UNIVERSALITY PROFILES: HOW MUCH TASK COVERAGE CAN A FIXED MODEL SUSTAIN?

Anonymous authors
Paper under double-blind review

ABSTRACT

We study *sustainable breadth*: for a *fixed* algorithm and a finite budget, how broad is the family of tasks on which it can succeed while remaining fast? For a frozen foundation model, an in-context learner, or a multi-task procedure, this is the question of how much coverage a single fixed model can sustain at a target error, a quantity observed to saturate in practice but left unquantified by existing theory. Rather than fixing a task class and optimizing over algorithms as in minimax theory, we fix the algorithm and quantify the complexity of the subset it can cover. We formalize this through the *universality profile*, the packing complexity of the algorithm's success set at a target accuracy or regret level. Our main converse shows that success on $\exp(B)$ distinguishable tasks requires the transcript to reveal at least $\Omega(B)$ information about the underlying task; once this information is upper bounded, one obtains a quantitative breadth–convergence frontier. In online prediction with expert advice, the classical minimax law $\Theta(\sqrt{T \log N})$ becomes an explicit universality law with budget $\log N$, and varying only the prior support of Hedge yields different algorithmic breadth on the same ambient class. In a packed K-armed bandit identification family with $\exp(\Theta(K))$ environments, the framework yields a sharp lower bound $T \gtrsim K/\Delta^2$ together with a matching, up to logarithmic factors, uniform-sampling upper bound. A local linear-bandit analysis turns the same proof skeleton into a two-sided structured frontier, with a lower bound that holds for every algorithm and a least-squares achievability guarantee at the scale set by full-parameter estimation. On a fixed expert class, we moreover characterize the success set of a given algorithm explicitly, showing that the profile ranges over a whole spectrum of values as the algorithm varies while the ambient minimax hardness stays put. A finite-feature kernel witness shows that capacity bounds breadth in a nonparametric setting, and real-data checks on digit and CIFAR-10 classification make the same breadth–budget law visible outside simulation: saturation at the class packing, a phase-transition knee, and coverage bounded by the capacity of a frozen representation. The point is not to replace minimax hardness, but to add an algorithm-centric axis that measures the task coverage achieved by a particular procedure.

1 INTRODUCTION

How many distinguishable tasks can one *fixed* model solve at a target error? This is the coverage question behind foundation models, in-context learners, and multi-task procedures, whose few-shot breadth is observed to saturate in practice, and it applies to any fixed procedure, central across machine learning, reinforcement learning, and decision making, where one seeks algorithms that are simultaneously *broad* and *fast*. What is missing is a quantitative language for *sustainable breadth*: for a *fixed* algorithm and a finite budget, how much breadth is actually sustainable?

That question is not answered by minimax theory, which fixes a class and asks what the *best possible* algorithm achieves. We instead fix the algorithm and quantify the packing complexity of the set of tasks on which it succeeds at accuracy ϵ within budget n: the *universality profile*, a measure of *algorithmic breadth* rather than class hardness. Two algorithms can face the same class yet cover provably different subfamilies (Section 5); success-set complexity is a reusable axis, orthogonal to class hardness, certified in explicit examples and paired with standard information bounds.

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.

Once breadth is phrased this way, a sharp principle emerges: succeeding on $\exp(B)$ distinguishable tasks forces the transcript to reveal $\Omega(B)$ bits about the task, making the theory rate–distortion-like (error is distortion, success-set complexity is rate) (Shannon, 1959; Yu, 1997; Tsybakov, 2009) and clarifying why PAC-Bayes and universal coding pay logarithmic penalties for broad coverage.

Contributions. We make three contributions.

- **An algorithm-centric complexity measure of task coverage.** We define the *universality profile* $U_{\mathcal{A}}(\epsilon, n)$ as the packing complexity of the success set of a fixed algorithm after budget n, an algorithm-dependent complexity measure that quantifies coverage of a procedure rather than hardness of a class.
- **A quantitative rate–distortion-style frontier linking coverage and convergence.** Broad success sets force mutual information with a packed task index, giving a generic route from breadth targets to sample- or horizon-lower bounds once a model-specific information budget is certified.
- **Two-sided frontiers on canonical models with observable phase transitions.** On a fixed expert class we characterize the comparator success set of a given algorithm explicitly, so the profile provably ranges while ambient minimax hardness stays put; bandit identification yields a near-sharp breadth–horizon frontier; and a structured linear-bandit model admits both a lower bound holding for all algorithms and a least-squares achievability guarantee. A finite-feature kernel witness shows capacity bounds breadth in a nonparametric setting. Synthetic experiments make the predicted trade-offs, including their sharp transitions, visible in finite regimes.

2 Related Work

The closest antecedents come from minimax lower bounds, metric entropy, PAC-Bayes, universal coding (the minimum description length, MDL, principle), and interactive complexity (Yu, 1997; Tsybakov, 2009; McAllester, 1999; Seeger, 2002; Catoni, 2007; Rissanen, 1984; Barron et al., 1998; Wolpert & Macready, 1997; Mannor & Tsitsiklis, 2004; Bubeck & Cesa-Bianchi, 2012; Cesa-Bianchi & Lugosi, 2006).

Interactive complexity and class-level lower bounds. Recent work sharpens class-level interactive lower bounds: the Decision-Estimation Coefficient (DEC) Foster et al. (2021; 2023) characterizes minimax regret through a class-level complexity measure, interactive Fano arguments Chen et al. (2024) give lower bounds that hold against any algorithm, and saddle-point analyses Kirschner et al. (2024); Appel & Kosoy (2025) sharpen worst-case regret in structured environments. All of these fix a class and study the *optimal* risk on it, taking an infimum over policies; none tracks how large a subfamily a specific algorithm covers at a fixed budget, which is the quantity we measure.

PAC-Bayes and information-theoretic generalization. PAC-Bayesian bounds McAllester (1999); Seeger (2002); Catoni (2007) and information-theoretic bounds Hellström et al. (2025) are algorithm-dependent: the posterior Q is specific to a learning algorithm, and the $\mathrm{KL}(Q\|P)$ penalty reflects how far its output departs from the prior. The present work is a packing-number analogue: instead of a KL cost paid by the posterior, we measure the log-packing number of the success set, where KL captures how much the algorithm has learned and log-packing how many distinguishable tasks it can solve.

In-context learning and foundation-model coverage. A distinct literature studies why *frozen* models generalize to new tasks without weight updates (Xie et al., 2022; Ahn et al., 2023; Bai et al., 2023; Wies et al., 2023; Zhang et al., 2024), asking whether a model can *represent or learn* a family of inner algorithms or priors. Our object is the complementary coverage question of how large a family of tasks a fixed procedure can cover at a target error, which those works do not bound. The framework is agnostic to how a procedure was trained: it takes the procedure as given and quantifies its coverage, the natural language for comparing frozen models of any origin.

3 PROBLEM SETUP AND UNIVERSALITY

3.1 TASK FAMILIES AND ALGORITHMS

Let $\{P_\theta : \theta \in \Theta\}$ be a family of data-generating mechanisms. In the passive setting, $\mathcal{A}$ observes n i.i.d. samples $X_{1:n} \sim P_\theta^{\otimes n}$ and outputs $\widehat{\theta}$; in interactive settings, the transcript Z_n includes actions and observations generated adaptively. We write $P_\theta^{\mathcal{A}}$ for the law of Z_n under $\mathcal{A}$ on task θ (the product law $P_\theta^{\otimes n}$ in the passive case), and ρ for a task-distance proxy (estimation error, excess risk, or regret).

3.2 CHOOSING THE TASK DISTANCE

The profile depends on the task metric ρ: require that $\rho(\widehat{\theta}, \theta) \le \epsilon$ encode a concrete performance guarantee and that 2ϵ-separated tasks be genuinely distinguishable. Standard choices are $\|\widehat{\theta} - \theta\|$ (passive), regret over comparators (experts), and normalized Hamming error (bandits); metric sensitivity is in Appendix B.

Definition 1 (ϵ-success and success sets)**.** *Fix $\epsilon > 0$ and $\delta \in (0,1)$. We say that $\mathcal{A}$ is* ϵ-successful *on task θ at sample size n if*

$$\mathbb{P}_{Z_n \sim P_\theta^{\mathcal{A}}}\Big(\rho(\widehat{\theta}, \theta) \le \epsilon\Big) \ge 1 - \delta. \tag{1}$$

The corresponding success set is

$$\mathsf{Succ}_{\mathcal{A}}(\epsilon, n) := \{\theta \in \Theta : \mathcal{A} \text{ is } \epsilon\text{-successful on } \theta \text{ at } n\}. \tag{2}$$

Remark 1 (The confidence parameter)**.** *$\delta \in (0,1)$ is the confidence level, with an explicit effect: success at level $1-\delta$ contributes a factor $(1-\delta)$ on $\log M$ plus an additive $-\log 2$ in Lemma 1, so the frontier constant degrades as $\frac{1}{1-\delta}$ while the rate is unchanged; the union bounds of Theorems 4–5 add a $\log(1/\delta)$ only when per-task failures are forced. Instantiations take $\delta = 1/4$.*

3.3 UNIVERSALITY AS SUCCESS-SET COMPLEXITY

To quantify how broad $\mathsf{Succ}_{\mathcal{A}}(\epsilon, n)$ is, we use a packing-based notion that applies to both discrete and continuous task spaces.

Notation. Throughout, $O(\cdot)$, $\Omega(\cdot)$, and $\Theta(\cdot)$ hide universal positive constants, and $\lesssim, \gtrsim$ mean the same as $O(\cdot), \Omega(\cdot)$; logarithmic factors are reported explicitly when they matter. For a metric space $(\mathcal{S}, \rho)$ and $\delta > 0$, the *δ-packing number* $\mathcal{M}(\delta, \mathcal{S}, \rho)$ is the maximum cardinality of a subset of $\mathcal{S}$ whose points are pairwise ρ-distance greater than δ.

Definition 2 (Universality profile)**.** *Let $\mathcal{M}(\delta, \mathcal{S}, \rho)$ denote the δ-packing number of a set $\mathcal{S} \subseteq \Theta$ under ρ. Define the* universality profile *of $\mathcal{A}$ by*

$$U_{\mathcal{A}}(\epsilon, n) := \log \mathcal{M}(2\epsilon, \mathsf{Succ}_{\mathcal{A}}(\epsilon, n), \rho). \tag{3}$$

The profile $U_{\mathcal{A}}(\epsilon, n)$ answers: after n samples, how many *distinguishable* tasks can $\mathcal{A}$ solve to accuracy ϵ? In contrast to minimax analyses that fix the task family and optimize over algorithms, eq. (3) is algorithm-centric and directly operationalizes "universality".

Remark 2 (Relation to minimax packings)**.** *When the success set equals the whole class, the profile reduces to the standard packing view; the new content appears when $\mathsf{Succ}_{\mathcal{A}}(\epsilon, n) \subsetneq \Theta$.*

3.4 A UNIFYING TEMPLATE: UNIVERSALITY DOMAINS

The definition above measures universality with respect to a *task domain* Θ equipped with a metric ρ. The same template extends to other domains over which a guarantee must hold uniformly, which we use in Section 5 for comparators.

Remark 3 (Universality domains: a unifying template)**.** *Let $\mathcal{D}$ index objects over which a guarantee must hold uniformly (environments, comparators, or hypotheses) with metric $\rho_{\mathcal{D}}$ and performance functional Perf; the profile over $\mathcal{D}$ is $U_{\mathcal{A}}^{\mathcal{D}}(\epsilon, n) = \log \mathcal{M}(2\epsilon, \{u \in \mathcal{D} : \mathsf{Perf}(\mathcal{A}, u; n) \le \epsilon\}, \rho_{\mathcal{D}})$, recovering eq. (3) for $\mathcal{D} = \Theta$. In experts the comparator set with the discrete metric makes packing reduce to cardinality, and the comparator profile of Section 5 is exactly this object.*

Proposition 1 (Feature capacity bounds breadth (nonparametric witness)). *Consider Gaussian regression $Y = f(x)+\xi$, $\xi \sim \mathcal{N}(0,1)$, with tasks in an RKHS ball $B_H(R)$ and $\rho(\widehat{f}, f) = \|\widehat{f} - f\|_{L^2}$. Let $\mathcal{A}_m$ be least squares restricted to the span V_m of the first m eigenfunctions. Then the success set of $\mathcal{A}_m$ is contained in an ϵ-tube around V_m, so $U_{\mathcal{A}_m}(\epsilon, n) \le \Theta(m \log(R/\epsilon))$ independent of the ambient class; for eigenvalue decay $\lambda_j \asymp j^{-\beta}$ the full ball has packing entropy $\asymp (R/\epsilon)^{2/\beta}$, so capacity sharply limits which subfamily a fixed learner covers (proof in Appendix B).*

4 A Rate–Distortion Analogy: A Generic Converse

We now turn the profile into a lower bound in two steps: broad success forces mutual information, and a model-specific bound on that information yields a convergence constraint.

4.1 From universality to mutual information

Let $\Theta_0 = \{\theta_1, \dots, \theta_M\} \subseteq \mathsf{Succ}_{\mathcal{A}}(\epsilon, n)$ be a 2ϵ-packing set under ρ, and let $\widetilde{\Theta}$ be a random variable uniform on Θ_0. Let Z_n be the algorithm transcript generated under $\widetilde{\Theta}$.

Lemma 1 (Universality implies mutual information). *If $\mathcal{A}$ is ϵ-successful on every $\theta \in \Theta_0$, then*

$$I(\widetilde{\Theta}; Z_n) \ge (1-\delta)\log M - \log 2. \tag{4}$$

The proof follows a standard reduction from estimation to multi-way testing: on a 2ϵ-packing set, any ϵ-accurate estimator induces a decoder with error probability bounded away from one, and Fano's inequality yields eq. (4) (Yu, 1997; Tsybakov, 2009).

4.2 An information budget per sample

We upper bound $I(\widetilde{\Theta}; Z_n)$ by the per-sample information available about $\widetilde{\Theta}$ under the algorithm.

Lemma 2 (Mutual information upper bound via pairwise KL). *Let $\widetilde{\Theta}$ be uniform on $\Theta_0 = \{\theta_1, \dots, \theta_M\}$. For any algorithm $\mathcal{A}$ and transcript Z_n,*

$$I(\widetilde{\Theta}; Z_n) \le \frac{1}{M^2} \sum_{i,j=1}^{M} D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| P_{\theta_j}^{\mathcal{A}}). \tag{5}$$

Proof sketch. With $\overline{P}^{\mathcal{A}} = \frac{1}{M}\sum_j P_{\theta_j}^{\mathcal{A}}$ the mixture, $I(\widetilde{\Theta}; Z_n) = \frac{1}{M}\sum_i D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| \overline{P}^{\mathcal{A}})$; convexity of KL in its second argument gives $D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| \overline{P}^{\mathcal{A}}) \le \frac{1}{M}\sum_j D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| P_{\theta_j}^{\mathcal{A}})$, and averaging over i yields the bound.

Remark 4 (From pairwise KL to a per-round information budget). *Lemma 2 certifies a budget $I(\widetilde{\Theta}; Z_n) \le n \cdot \mathsf{Info}(\epsilon)$ by bounding the average pairwise transcript KL by $n\kappa(\epsilon)$: in passive settings $\kappa(\epsilon)$ is an average one-sample KL, and in interactive settings Lemma 3 gives an additive decomposition in time.*

Lemma 3 (Interactive information decomposition). *Let $Z_n = (A_{1:n}, Y_{1:n})$ be an interactive transcript generated by an algorithm $\mathcal{A}$, where A_t may depend on Z_{t-1}. For any θ, θ', the KL divergence between transcript distributions satisfies*

$$D_{\mathrm{KL}}(P_\theta^{\mathcal{A}} \| P_{\theta'}^{\mathcal{A}}) = \mathbb{E}_\theta \left[\sum_{t=1}^{n} D_{\mathrm{KL}}(P_\theta(Y_t \mid A_t, Z_{t-1}) \,\|\, P_{\theta'}(Y_t \mid A_t, Z_{t-1})) \right]. \tag{6}$$

Lemma 3 is the bandit mechanism: despite adaptive actions, information about θ accumulates additively over rounds through conditional divergences.

4.3 Main theorem

The next theorem is the paper's formal backbone: model-agnostic and purely information-theoretic, it becomes a concrete lower bound once a model certifies a transcript-information budget.

Theorem 1 (Universality–convergence frontier)**.** *Fix $\epsilon > 0$, $n \in \mathbb{N}$, and $\delta \in (0,1)$. For any algorithm $\mathcal{A}$ and any task-distance ρ,*

$$U_{\mathcal{A}}(\epsilon, n) \le \frac{1}{1-\delta} I(\widetilde{\Theta}; Z_n) + \frac{1}{1-\delta} \log 2, \tag{7}$$

where $\widetilde{\Theta}$ is uniform on a maximal 2ϵ-packing $\Theta_0 \subseteq \mathsf{Succ}_{\mathcal{A}}(\epsilon, n)$ and Z_n is the transcript under $\widetilde{\Theta}$. Consequently, if $\mathcal{A}$ is ϵ-successful on a 2ϵ-packing of size $\exp(B)$ and a model certifies an information budget $I(\widetilde{\Theta}; Z_n) \le n \cdot \mathsf{Info}(\epsilon)$, then necessarily

$$n \ge \Omega\left(\frac{B}{\mathsf{Info}(\epsilon)}\right). \tag{8}$$

Theorem 1 is a reusable template rather than a rearrangement of its constituent lemmas: it reduces the lower-bound problem in each new model to certifying a single transcript-information budget $\mathsf{Info}(\epsilon)$, after which the task geometry (the packing) and the per-sample KL control the frontier, and the remaining work is largely geometric. This reduction, not the Fano step itself, is the reusable contribution.

Corollary 1 (Passive-estimation specialization)**.** *In the passive setting with n i.i.d. samples $Z_n = X_{1:n}$, let $\Theta_0 \subseteq \mathsf{Succ}_{\mathcal{A}}(\epsilon, n)$ be a 2ϵ-packing of size M and assume the induced family $\{P_\theta : \theta \in \Theta_0\}$ satisfies*

$$\frac{1}{M^2} \sum_{i,j=1}^{M} D_{\mathrm{KL}}(P_{\theta_i} \| P_{\theta_j}) \le \kappa. \tag{9}$$

Then $U_{\mathcal{A}}(\epsilon, n) \le \frac{1}{1-\delta} n\kappa + \frac{1}{1-\delta} \log 2$.

Remark 5 (Interpretation as rate–distortion)**.** *Define $R(\epsilon) := \sup_{\mathcal{A}} U_{\mathcal{A}}(\epsilon, n)$ at fixed n. With an information budget $I(\widetilde{\Theta}; Z_n) \le n \cdot \mathsf{Info}(\epsilon)$, Theorem 1 makes $R(\epsilon)$ linear in n: a rate–distortion structure in which ϵ is the distortion and $R(\epsilon)$ the maximum representable task complexity (Shannon, 1959). Two features distinguish this from the mutual-information rate $I(Z;W)$ of information-theoretic generalization bounds (Xu & Raginsky, 2017; Nokleby & Beirami, 2021): our rate is algorithm-specific (the packing entropy of a fixed procedure's success set, not a class-level quantity), and its units differ (a geometric count of tasks, charged against rather than equal to the transcript information).*

5 Experts: Breadth Separates from Minimax Hardness

Our first flagship witness is online prediction with expert advice. The point of this section is not that experts are the most complicated setting, but that they make the algorithm-centric distinction maximally transparent: the ambient minimax hardness is fixed by N, yet the breadth achieved by a particular algorithm can vary substantially on that same class.

5.1 Setting

There are N experts; each round the learner chooses $p_t \in \Delta^N$, incurs expected loss $\langle p_t, \ell_t \rangle$, and observes the losses $\ell_t \in [0,1]^N$. The regret to the best expert is

$$\mathrm{Reg}_T := \sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \min_{i \in [N]} \sum_{t=1}^{T} \ell_{t,i}. \tag{10}$$

Proposition 2 (Exponential weights induce a universality penalty)**.** *Fix a prior $\pi \in \Delta^N$ on N experts and run exponential weights (Hedge) with learning rate $\eta > 0$. Then for every loss sequence $\ell_{1:T} \in [0,1]^{T \times N}$ and every expert i,*

$$\mathbb{E}\left[\sum_{t=1}^{T} \langle p_t, \ell_t \rangle - \sum_{t=1}^{T} \ell_{t,i}\right] \le \frac{\log(1/\pi_i)}{\eta} + \frac{\eta T}{8}. \tag{11}$$

Optimizing over η yields a regret cost of order $\sqrt{T \log(1/\pi_i)}$, showing that a universality budget $\log(1/\pi_i)$ induces a worst-case regret penalty scaling as $\sqrt{T \log(1/\pi_i)}$ (Cesa-Bianchi & Lugosi, 2006).

5.2 A SHARP BREADTH WITNESS ON A FIXED CLASS

The next proposition is classical, but it supplies the ambient baseline: minimax regret depends only on the class size N.

Proposition 3 (Minimax regret scales with $\log N$)**.** *There exist universal constants $c, C > 0$ such that for all $N \geq 2$ and $T \geq 1$,*

$$c\sqrt{T \log N} \leq \inf_{\mathcal{A}} \sup_{\ell_{1:T} \in [0,1]^{T \times N}} \mathbb{E}[\mathrm{Reg}_T(\mathcal{A})] \leq C\sqrt{T \log N}. \tag{12}$$

Definition 3 (Comparator-universality profile for experts)**.** *For an online algorithm $\mathcal{A}$ and a subset $S \subseteq [N]$, let $\mathrm{Reg}_T(\mathcal{A}; S)$ be the worst-case expected regret over all loss sequences against the best expert in S, and define the* comparator-universality profile *as*

$$U_{\mathcal{A}}^{\mathrm{cmp}}(\epsilon, T) := \log \max \{|S| : S \subseteq [N],\ \mathrm{Reg}_T(\mathcal{A}; S) \leq \epsilon\}. \tag{13}$$

Corollary 2 (Universality forces regret)**.** *There exists a universal constant $c > 0$ such that for any online algorithm $\mathcal{A}$ and any $\epsilon > 0$,*

$$\epsilon \geq c\sqrt{T \cdot U_{\mathcal{A}}^{\mathrm{cmp}}(\epsilon, T)}. \tag{14}$$

Proposition 4 (Breadth and hardness can separate on the same expert class)**.** *Fix an expert class $[N]$ and horizon T with $\log N < 2T$. Let $S \subseteq [N]$ have size m, and run Hedge with prior π^S uniform on S and zero on $[N] \setminus S$. Then, with*

$$\epsilon_m := \sqrt{\frac{T \log m}{2}},$$

the resulting algorithm $\mathcal{A}_S$ satisfies

$$U_{\mathcal{A}_S}^{\mathrm{cmp}}(\epsilon_m, T) = \log m.$$

In particular, on the same ambient expert class $[N]$, the algorithmic breadth can vary from $\log m$ to $\log N$ as the prior support changes, while the class-level minimax hardness remains $\Theta(\sqrt{T \log N})$ by Proposition 3.

Interpretation. At full breadth ($B = \log N$) the frontier reproduces the minimax law $\epsilon \asymp \sqrt{TB}$ (Proposition 3, Corollary 2). Proposition 4 is the cleanest algorithm-centric witness: with the ambient class fixed, varying only the prior support changes breadth from $\log m$ to $\log N$, which minimax regret alone does not distinguish. The frontier *law* itself is recoverable by applying the class-level minimax bound to subclasses: restrict to a comparator set of size $\exp(B)$ and invoke Proposition 3. What class-level packing cannot see is the *assignment*: which profile a fixed algorithm attains, and the strictness of its success set, as certified by Theorem 2. That assignment is the algorithm-centric content (extended discussion in Appendix B).

Theorem 2 (Strict comparator success sets)**.** *Fix the expert class $[N]$ and horizon T with $\log N < 2T$, and take confidence $\delta = 1/4$. For $m \in [N]$, let $\mathcal{A}_m$ be Hedge with prior uniform on a support $S_m \subset [N]$ of size m, and let $\epsilon_m = \sqrt{T \log m/2}$.*

1. *At tolerance ϵ_m, the comparator sets on which $\mathcal{A}_m$ is ϵ_m-good are exactly the subsets of its support:*

$$\mathrm{Reg}_T(\mathcal{A}_m; S') \leq \epsilon_m \quad \Longleftrightarrow \quad S' \subseteq S_m. \tag{15}$$

   In particular, every comparator set containing an expert outside S_m provably incurs worst-case regret at least $T > \epsilon_m$, so the success region is a strict *subfamily of the ambient comparator domain whenever $m < N$, and $U_{\mathcal{A}_m}^{\mathrm{cmp}}(\epsilon_m, T) = \log m$.*

2. *The ambient minimax regret $\inf_{\mathcal{A}} \sup_{\ell} \mathbb{E}[\mathrm{Reg}_T(\mathcal{A})] = \Theta(\sqrt{T \log N})$ is a single number, independent of the algorithm and of m. It bounds what is* possible *on the whole class but does not determine which subfamily a* fixed *algorithm covers, nor does it certify that a particular procedure fails beyond its support. Indeed the same ambient value is consistent with profiles ranging from $\log m$ to $\log N$ as the prior support varies.*

Proof sketch. $S' \subseteq S_m$ gives regret $\leq \epsilon_m$ by Proposition 4; for $j \in S' \setminus S_m$, the sequence with $\ell_{t,j} = 0, \ell_{t,i} = 1$ $(i \in S_m)$ forces regret $\geq T > \epsilon_m$, since zero prior mass keeps Hedge off j. Part (ii) follows because the minimax value depends only on N; details in Appendix A.7. The zero-mass prior is the clean limit of a continuous family of positive-prior separations: with strictly positive masses π_j, expert j becomes ϵ-admissible exactly when $\log(1/\pi_j) \leq 2\epsilon^2/T$, so a prior with effective support m_{eff} (mass $\asymp 1/m$ on m experts, exponentially small elsewhere) attains profile $\log m_{\text{eff}}$; the separation persists for every non-uniform prior and collapses only at the uniform prior (Appendix B).

6 Structured Interactive Witnesses

The second pillar of the paper is interactive. We start from a fully explicit Bernoulli identification family where both sides of the frontier can be computed, and then show that the same proof skeleton survives in a structured linear-bandit model.

6.1 A fully explicit Bernoulli bandit family

Consider a K-armed stochastic bandit where the environment is indexed by a binary vector $\theta \in \{-1, +1\}^K$. In environment θ, arm i yields i.i.d. rewards

$$Y_t \mid (A_t = i, \theta) \sim \text{Bernoulli}\left(\frac{1}{2} + \theta_i \Delta\right), \qquad \Delta \in \left(0, \frac{1}{4}\right]. \tag{16}$$

An algorithm $\mathcal{A}$ selects actions $A_t \in [K]$ adaptively, and after T rounds outputs an estimate $\widehat{\theta} \in \{-1, +1\}^K$. We evaluate performance using the normalized Hamming loss

$$\rho(\widehat{\theta}, \theta) := \frac{1}{K} \sum_{i=1}^{K} \mathbf{1}\{\widehat{\theta}_i \neq \theta_i\}. \tag{17}$$

Theorem 3 (Universality cost for bandit identification)**.** *Fix $\epsilon \in (0, 1/4]$ and $\Delta \in (0, 1/4]$. There exists a subset $\Theta_0 \subseteq \{-1, +1\}^K$ that is a 2ϵ-packing under ρ with $\log|\Theta_0| \geq c_0 K$ such that the following holds. For any (possibly adaptive) bandit algorithm $\mathcal{A}$, if $\mathcal{A}$ is ϵ-successful on every $\theta \in \Theta_0$ at horizon T, then necessarily*

$$T \geq c_1 \frac{\log|\Theta_0|}{\Delta^2} \geq c_0 c_1 \frac{K}{\Delta^2}. \tag{18}$$

In particular, $\Theta_0 \subseteq \mathsf{Succ}_{\mathcal{A}}(\epsilon, T)$ implies $U_{\mathcal{A}}(\epsilon, T) \geq \log|\Theta_0|$ for this task family, so the lower bound can be read as $T \gtrsim U_{\mathcal{A}}(\epsilon, T)/\Delta^2$.

Proof sketch. A Varshamov–Gilbert construction gives a 2ϵ-packing with $\log|\Theta_0| = \Omega(K)$ (Tsybakov, 2009), and the interactive chain rule bounds the pairwise transcript KL by $\mathbb{E}_\theta[\sum_{t=1}^{T} D_{\text{KL}}(\text{Ber}(\frac{1}{2} + \theta_{A_t}\Delta) \| \text{Ber}(\frac{1}{2} + \theta'_{A_t}\Delta))] \lesssim T\Delta^2$, since each round contributes 0 or $O(\Delta^2)$; Lemma 1 then gives $T \gtrsim \log|\Theta_0|/\Delta^2$. Full details are in Appendix A.8.

Theorem 4 (A matching upper bound by uniform sampling)**.** *There exists an algorithm that samples each arm m times (so $T = Km$) and outputs $\widehat{\theta}_i = \text{sign}(\widehat{\mu}_i - 1/2)$. For any $\delta \in (0, 1)$, choosing $m \gtrsim \Delta^{-2} \log(K/\delta)$ ensures $\mathbb{P}_\theta(\widehat{\theta} = \theta) \geq 1 - \delta$ simultaneously for all $\theta \in \{-1, +1\}^K$, and hence yields ϵ-success for any $\epsilon > 0$.*

Interpretation. Together, Theorems 3–4 give a near-sharp interactive frontier: covering $\exp(\Theta(K))$ environments needs $T \gtrsim K/\Delta^2$, and uniform sampling attains it up to logs: a two-sided witness rather than a one-sided lower-bound example (Mannor & Tsitsiklis, 2004; Bubeck & Cesa-Bianchi, 2012).

6.2 A structured linear-bandit witness

The same reduction extends to structured models via a plug-in recipe (metric, packing, KL). The next proposition records it for a Gaussian linear bandit.

Proposition 5 (Local universality lower bound for linear bandit identification)**.** *Consider a stochastic linear bandit with unknown parameter $\theta^\star \in \mathbb{R}^d$ and action set $\mathcal{X} \subseteq \mathbb{R}^d$ satisfying $\|x\|_2 \le 1$ for all $x \in \mathcal{X}$. At round t, the learner chooses $X_t \in \mathcal{X}$ and observes $Y_t = \langle X_t, \theta^\star \rangle + \xi_t$ with $\xi_t \sim \mathcal{N}(0,1)$. Let $\rho(\widehat{\theta}, \theta^\star) = \|\widehat{\theta} - \theta^\star\|_2$. There exist universal constants $c_0, c_1 > 0$ such that for any $\epsilon \in (0, 1]$, there exists a 2ϵ-packing $\Theta_0 \subseteq \mathbb{R}^d$ with $\log|\Theta_0| \ge c_0 d$ such that: for any algorithm $\mathcal{A}$, if $\mathcal{A}$ is ϵ-successful on every $\theta \in \Theta_0$ at horizon T, then*

$$T \ge c_1 \frac{d^2}{\epsilon^2}. \tag{19}$$

Proposition 5 follows from an isotropic Varshamov–Gilbert packing and a Gaussian-relaxation bound on the transcript information, which gives $I(\widetilde{\Theta}; Z_T) \lesssim T\epsilon^2/d$ (Appendix A.9). The isotropic packing is the crux: a fixed algorithm can at most probe the isotropically distributed parameter in one direction per round, so each round reveals only $O(\epsilon^2/d)$ bits, and adaptivity cannot beat this because the packing carries no direction to exploit (Wainwright, 2019; Maiti et al., 2026). (The scale d^2/ϵ^2 is the full-parameter ℓ_2 *estimation* rate under unit-norm actions, not the d/ϵ^2 prediction rate; Appendix A.9.) The budget follows from a Gaussian prior $\widetilde{\Theta}_G \sim \mathcal{N}(0, (4\epsilon^2/d) I_d)$ whose posterior covariance obeys $\Sigma_t \preceq (4\epsilon^2/d) I_d$ for *every* adaptive action, so $I(\widetilde{\Theta}_G; Z_T) \le 2\epsilon^2 T/d$; adaptivity cannot help because the isotropic packing has no preferred direction, and the radial direction measures the norm, constant over the packing. The next theorem matches this at scale d^2/ϵ^2.

Theorem 5 (Achievability by isotropic exploration and least squares)**.** *In the setting of Proposition 5, take the action set to be the unit sphere $\mathcal{X} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$. Let the algorithm sample $X_1, \dots, X_T$ i.i.d. uniformly on $\mathcal{X}$ and output the least-squares estimate $\widehat{\theta} = V_T^{-1} \sum_{t=1}^T X_t Y_t$, where $V_T = \sum_{t=1}^T X_t X_t^\top$. Then there is a universal constant $C > 0$ such that for every $\theta^\star$ with $\|\theta^\star\|_2 \le 2\epsilon$ and every $\delta \in (0,1)$,*

$$\mathbb{P}_{\theta^\star}\Big(\|\widehat{\theta} - \theta^\star\|_2 > \epsilon\Big) \le \delta \qquad \textit{whenever} \qquad T \ge C \frac{d^2}{\epsilon^2} \log \frac{d}{\delta}. \tag{20}$$

In particular, at horizon $T = \Theta(d^2 \epsilon^{-2} \log(d/\delta))$ the scheme is ϵ-successful uniformly on the packed family Θ_0 of Proposition 5.

Theorem 5 is proved in Appendix A.10 via Gram-matrix concentration ($\lambda_{\min}(V_T) \ge T/(2d)$ with high probability), the Gaussian structure of the least-squares error, and a χ^2_d tail bound; the error distribution is independent of $\theta^\star$, so the guarantee is uniform over the local family. Together the two bounds are tight up to logarithmic factors: the lower bound $T = \Omega(d^2/\epsilon^2)$ matches the achievability $T = O(d^2\epsilon^{-2}\log(d/\delta))$, so the isotropic (D-optimal) design is minimax-optimal for full-parameter ℓ_2 estimation on this local family. A target-aware design that already knows the support of $\theta^\star$ reaches d/ϵ^2 only by forgoing the uniform guarantee and concentrating on a single direction; this is the price of broad coverage, not a slack in the theory (full discussion in Appendix B).

7 Synthetic Validation of Algorithmic Breadth

The empirical goal is measurability, not benchmark superiority: does the profile track an observable breadth–budget trade-off in finite regimes? We report synthetic studies, each averaged over 10 seeds with 95% confidence intervals, with breadth proxy $\log(\#\{\text{packed environments solved at the target accuracy}\})$ (protocols in Appendix C).

7.1 E1: experts breadth separation

E1 varies only the prior support of Hedge on a fixed expert class: regret to the best supported expert grows monotonically with $\log m$ (Figure 1, left), matching Proposition 4.

7.2 E2: Bernoulli bandit coverage frontier

E2 works with the packed Bernoulli family of Section 6: uniform sampling and a Thompson-style identification rule both rapidly increase breadth with $T\Delta^2$ and saturate near $\log 64$, while a greedy proxy stays near zero breadth (Figure 1, middle), matching Theorems 3–4; the transition is robust to the success metric and threshold (Appendix C).

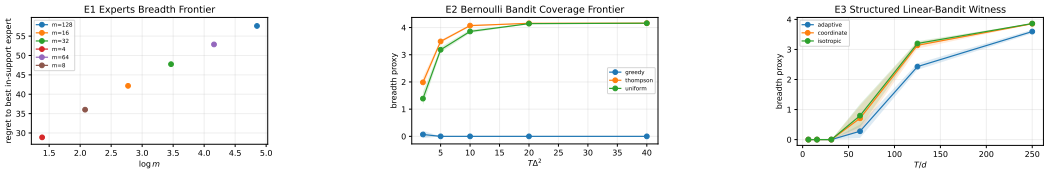


Figure 1: Synthetic validation of algorithmic breadth. Left: E1 with $N = 128$ and $T = 4000$ shows regret increasing monotonically with the breadth proxy $\log m$. Middle: E2 with $K = 16$ and $\Delta = 0.1$ shows uniform sampling approaching full packed-family coverage ($\log 64 \approx 4.16$) as $T\Delta^2$ grows, while a greedy proxy remains near zero breadth. Right: E3 with $d = 16$ and $\epsilon = 0.35$ shows a structured coverage transition toward $\log 48 \approx 3.87$ on a packed sign family, with isotropic exploration turning on slightly earlier than a balanced coordinate design. Shaded regions are 95% confidence intervals over 10 seeds.

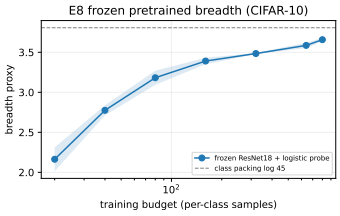


Figure 2: Frozen pretrained-representation breadth (CIFAR-10, 5 seeds): a logistic probe on frozen ImageNet-ResNet18 features covers more of the 45 class-pair tasks as the budget grows, saturating below the class packing $\log 45$ (details in Appendix C).

7.3 E3: Structured Linear-Bandit Witness

The packed local linear-bandit family shows the same transition: breadth rises sharply once the normalized horizon T/d crosses the predicted scale (Figure 1, right); a two-stage adaptive design does not beat isotropic exploration, consistent with the d^2/ϵ^2 lower bound of Proposition 5 (full phase-transition surface in Appendix B).

Real-data checks confirm the profile is observable outside simulation (Appendix B, Figures 3–6): on digits, breadth saturates at the class packing $\log 45$, and frozen representations, from PCA features to an ImageNet-pretrained ResNet on CIFAR-10, cover exactly as much as their capacity allows, saturating at or just below the packing limit. Figure 2 shows the largest-scale check.

8 Connections and Discussion

The paper is best read as a bridge between three ideas: PAC-Bayes and universal coding price breadth through KL or redundancy penalties (McAllester, 1999; Seeger, 2002; Catoni, 2007; Rissanen, 1984; Barron et al., 1998); interactive learning charges breadth to the information a policy extracts from its trajectory (Chen et al., 2024); and DEC/saddle-point analyses characterize class hardness, which $U_{\mathcal{A}}(\epsilon, n)$ complements by measuring the breadth of a fixed algorithm (Foster et al., 2021; Kirschner et al., 2024; Appel & Kosoy, 2025). Two comparisons sharpen the position. The interactive Fano unification of Chen et al. (2024) isolates an estimation factor (fractional covering) atop exploration costs; our linear-bandit factor-d is exactly such a factor, so a DEC-style class complexity and a per-policy profile could compose into a joint breadth–convergence view. Conditional-MI frameworks (Haghifam et al., 2021; Futami & Fujisawa, 2026) likewise decompose algorithm-dependent budgets roundwise; ours is coarser but model-agnostic (cf. the rate–distortion Bayes-risk line (Nokleby & Beirami, 2021; Nokleby et al., 2016)).

Limitations. The framework is primarily a converse: it does not itself identify frontier-achieving algorithms or resolve computation, and it depends on the metric ρ. Extending it to distribution-weighted universality is a natural next step.

AI USE STATEMENT

Generative AI tools were used to assist with drafting and editing portions of the text and with the implementation of the synthetic and real-data experiments. All AI-assisted content, including mathematical statements, claims, proofs, and experimental code, was reviewed and verified by the authors. We take responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

Reference implementations of all experiments and figure-generation scripts are provided in the supplementary code. The appendix contains complete proofs of all theoretical claims, the exact experimental protocols (packing constructions, success thresholds, decoding rules, and seed counts), and the full results CSVs, so that every figure can be reproduced from the stated procedures.

REFERENCES

Kwangjun Ahn, Xiang Cheng, Hadi Daneshmand, and Suvrit Sra. Transformers learn to implement preconditioned gradient descent for in-context learning. In *Advances in Neural Information Processing Systems*, volume 36, pp. 45614–45650, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/8ed3d610ea4b68e7afb30ea7d01422c6-Abstract-Conference.html.

Alexander Appel and Vanessa Kosoy. Regret bounds for robust online decision making. In *Proceedings of the 38th Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pp. 64–146. PMLR, 2025. URL https://proceedings.mlr.press/v291/appel25a.html.

Yu Bai, Fan Chen, Huan Wang, Caiming Xiong, and Song Mei. Transformers as statisticians: Provable in-context learning with in-context algorithm selection. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://arxiv.org/abs/2306.04637.

Andrew R. Barron, Jorma Rissanen, and Bin Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, 44(6):2743–2760, 1998. doi: 10.1109/18.720554.

Raghav Bongole, Tobias J. Oechtering, and Mikael Skoglund. Generalizing the Fano inequality further. *arXiv preprint arXiv:2601.12027*, 2026. URL https://arxiv.org/abs/2601.12027.

Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012. doi: 10.1561/2200000024.

Olivier Catoni. *PAC-Bayesian Supervised Classification: The Thermodynamics of Statistical Learning*, volume 56 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. IMS, 2007.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. ISBN 9780521841085.

Fan Chen, Dylan J. Foster, Yanjun Han, Jian Qian, Alexander Rakhlin, and Yunbei Xu. Assouad, Fano, and Le Cam with interaction: A unifying lower bound framework and characterization for bandit learnability. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/8a23a95e26d016711c0d70f79ade3c95-Abstract-Conference.html. NeurIPS 2024 Spotlight.

Dylan J. Foster, Sham M. Kakade, Jian Qian, and Alexander Rakhlin. The statistical complexity of interactive decision making. *arXiv preprint arXiv:2112.13487*, 2021. doi: 10.48550/arXiv.2112.13487. URL https://arxiv.org/abs/2112.13487.

Dylan J. Foster, Noah Golowich, and Yanjun Han. Tight guarantees for interactive decision making with the decision-estimation coefficient. In *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 3969–4043. PMLR, 2023. URL https://proceedings.mlr.press/v195/foster23b.html.

Futoshi Futami and Masahiro Fujisawa. Information-theoretic generalization bounds for sequential decision making. *arXiv preprint arXiv:2605.12190*, 2026. URL https://arxiv.org/abs/2605.12190.

Mahdi Haghifam, Gintare Karolina Dziugaite, Shay Moran, and Daniel M. Roy. Towards a unified information-theoretic framework for generalization. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://arxiv.org/abs/2111.05275.

Fredrik Hellström, Giuseppe Durisi, Benjamin Guedj, and Maxim Raginsky. Generalization bounds: Perspectives from information theory and PAC-Bayes. *Foundations and Trends in Machine Learning*, 18(1):1–223, 2025. doi: 10.1561/2200000112. arXiv:2309.04381.

Johannes Kirschner, Seyed Alireza Bakhtiari, Kushagra Chandak, Volodymyr Tkachuk, and Csaba Szepesvári. Regret minimization via saddle point optimization. *arXiv preprint arXiv:2403.10379*, 2024. doi: 10.48550/arXiv.2403.10379. URL https://arxiv.org/abs/2403.10379.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020. doi: 10.1017/9781108571401. URL https://tor-lattimore.com/downloads/book/book.pdf.

Arnab Maiti, Yunbei Xu, and Kevin Jamieson. On the power of adaptivity for ε-best arm identification in linear bandits. In *Proceedings of Thirty Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 4911–4968. PMLR, 2026. URL https://proceedings.mlr.press/v336/maiti26a.html.

Shie Mannor and John N. Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5:623–648, 2004. URL https://www.jmlr.org/papers/volume5/mannor04b/mannor04b.pdf.

David A. McAllester. PAC-Bayesian bounds. In *Proceedings of the 12th Annual Conference on Computational Learning Theory*, pp. 164–172. ACM, 1999.

Matthew Nokleby and Ahmad Beirami. Information-theoretic Bayes risk lower bounds for realizable models. *arXiv preprint arXiv:2111.04579*, 2021. URL https://arxiv.org/abs/2111.04579.

Matthew Nokleby, Ahmad Beirami, and Robert Calderbank. Rate-distortion bounds on Bayes risk in supervised learning. In *IEEE International Symposium on Information Theory*, 2016. URL https://arxiv.org/abs/1605.02268.

Jorma Rissanen. Universal coding, information, prediction, and estimation. *IEEE Transactions on Information Theory*, 30(4):629–636, 1984. doi: 10.1109/TIT.1984.1056936.

Matthias Seeger. PAC-Bayesian generalisation error bounds. In *Advances in Neural Information Processing Systems*, volume 15, pp. 603–610, 2002.

Claude E. Shannon. Coding theorems for a discrete source with a fidelity criterion. In *IRE National Convention Record*, volume 4, pp. 142–163, 1959.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009.

Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge University Press, 2019. doi: 10.1017/9781108627771.

Noam Wies, Yoav Levine, and Amnon Shashua. The learnability of in-context learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/73950f0eb4ac0925dc71ba2406893320-Abstract-Conference.html.

David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997. doi: 10.1109/4235.585893.

Sang Michael Xie, Aditi Raghunathan, Percy Liang, and Tengyu Ma. An explanation of in-context learning as implicit Bayesian inference. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2111.02080.

Aolin Xu and Maxim Raginsky. Information-theoretic analysis of generalization capability of learning algorithms. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://arxiv.org/abs/1705.07809.

Bin Yu. Assouad, Fano, and Le Cam. In David Pollard, Erik Torgersen, and Grace L. Yang (eds.), *Festschrift for Lucien Le Cam: Research Papers in Probability and Statistics*, pp. 423–435. Springer, 1997.

Ruiqi Zhang, Spencer Frei, and Peter L. Bartlett. Trained transformers learn linear models in-context. *Journal of Machine Learning Research*, 25(49):1–55, 2024. URL https://jmlr.org/papers/v25/23-1042.html.

A PROOFS

A.1 PROOF OF LEMMA 1

Let $\widetilde{\Theta}$ be uniform on $\Theta_0 = \{\theta_1, \ldots, \theta_M\}$ and let Z_n be the transcript. Define an estimator $\widehat{\Theta} = \widehat{\Theta}(Z_n)$ by mapping $\widehat{\theta}$ to the nearest element of Θ_0 under ρ, breaking ties arbitrarily:

$$\widehat{\Theta}(Z_n) \in \arg\min_{\theta \in \Theta_0} \rho(\widehat{\theta}(Z_n), \theta). \tag{21}$$

Because Θ_0 is a 2ϵ-packing, whenever $\rho(\widehat{\theta}, \widetilde{\Theta}) \le \epsilon$ we must have $\widehat{\Theta} = \widetilde{\Theta}$. For each $\theta \in \Theta_0$,

$$\mathbb{P}_\theta(\widehat{\Theta} \neq \theta) \le \mathbb{P}_\theta(\rho(\widehat{\theta}, \theta) > \epsilon) \le \delta, \tag{22}$$

using the ϵ-success definition in the main text. Averaging over the uniform prior on Θ_0 gives $\mathbb{P}(\widehat{\Theta} \neq \widetilde{\Theta}) \le \delta$. By Fano's inequality,

$$\mathbb{P}(\widehat{\Theta} \neq \widetilde{\Theta}) \ge 1 - \frac{I(\widetilde{\Theta}; Z_n) + \log 2}{\log M}, \tag{23}$$

so $I(\widetilde{\Theta}; Z_n) \ge (1 - \delta) \log M - \log 2$.

A.2 PROOF OF LEMMA 2

Let $\overline{P}^{\mathcal{A}} := \frac{1}{M} \sum_{j=1}^{M} P_{\theta_j}^{\mathcal{A}}$ be the mixture distribution over transcripts. A standard identity yields

$$I(\widetilde{\Theta}; Z_n) = \frac{1}{M} \sum_{i=1}^{M} D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| \overline{P}^{\mathcal{A}}). \tag{24}$$

By convexity of KL in its second argument,

$$D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| \overline{P}^{\mathcal{A}}) \le \frac{1}{M} \sum_{j=1}^{M} D_{\mathrm{KL}}(P_{\theta_i}^{\mathcal{A}} \| P_{\theta_j}^{\mathcal{A}}). \tag{25}$$

Averaging over i gives the claim.

A.3 PROOF OF LEMMA 3

Let $Z_t = (A_{1:t}, Y_{1:t})$. Since the algorithm is the same under θ and θ', the conditional distribution of A_t given Z_{t-1} is identical under the two environments. Therefore, the chain rule yields

$$D_{\mathrm{KL}}(P_\theta^{\mathcal{A}}(Z_n) \| P_{\theta'}^{\mathcal{A}}(Z_n)) = \sum_{t=1}^{n} \mathbb{E}_\theta[D_{\mathrm{KL}}(P_\theta(Y_t \mid A_t, Z_{t-1}) \,\|\, P_{\theta'}(Y_t \mid A_t, Z_{t-1}))], \tag{26}$$

since the A_t terms contribute zero KL.

A.4 PROOF OF THEOREM 1

By definition, there exists a maximal 2ϵ-packing $\Theta_0 \subseteq \mathsf{Succ}_{\mathcal{A}}(\epsilon, n)$ of size $M = \exp(U_{\mathcal{A}}(\epsilon, n))$. Apply Lemma 1 to obtain $I(\widetilde{\Theta}; Z_n) \ge (1 - \delta) \log M - \log 2$. Rearranging gives

$$U_{\mathcal{A}}(\epsilon, n) = \log M \le \frac{1}{1 - \delta} I(\widetilde{\Theta}; Z_n) + \frac{1}{1 - \delta} \log 2, \tag{27}$$

which is eq. (7). For the sample-complexity form, if $\mathcal{A}$ is ϵ-successful on a 2ϵ-packing of size $\exp(B)$, then $B \le U_{\mathcal{A}}(\epsilon, n)$ (the packing is contained in a maximal one); combining this with eq. (7) and the budget $I(\widetilde{\Theta}; Z_n) \le n \cdot \mathsf{Info}(\epsilon)$ yields $B \le \frac{1}{1-\delta} n \, \mathsf{Info}(\epsilon) + \frac{1}{1-\delta} \log 2$, i.e. $n \ge \Omega(B / \mathsf{Info}(\epsilon))$.

A.5 Experts instantiation: proof of Corollary 2

Let $B := U_{\mathcal{A}}^{\mathrm{cmp}}(\epsilon, T)$ and let S be a subset with $|S| = \exp(B)$ such that $\mathrm{Reg}_T(\mathcal{A}; S) \le \epsilon$. Consider the online learning problem restricted to the expert set S. By definition of minimax regret and Proposition 3 (applied with $N = |S|$), we have

$$\epsilon \ge \sup_{\ell_{1:T} \in [0,1]^{T \times N}} \mathbb{E}[\mathrm{Reg}_T(\mathcal{A}; S)] \ge c\sqrt{T \log |S|} = c\sqrt{T \cdot B}, \tag{28}$$

which is the stated inequality.

A.6 Proof of Proposition 4

Let π^S be uniform on S and zero on $[N] \setminus S$, and let $\mathcal{A}_S$ be Hedge with this prior. For each $i \in S$, Proposition 2 gives

$$\mathrm{Reg}_T(\mathcal{A}_S; \{i\}) \le \frac{\log m}{\eta} + \frac{\eta T}{8}. \tag{29}$$

Choosing $\eta = \sqrt{8 \log m / T}$ yields

$$\mathrm{Reg}_T(\mathcal{A}_S; S) \le \sqrt{\frac{T \log m}{2}}. \tag{30}$$

Hence

$$U_{\mathcal{A}_S}^{\mathrm{cmp}}(\epsilon_m, T) \ge \log |S| = \log m, \tag{31}$$

where $\epsilon_m = \sqrt{T \log m / 2}$.

To see that no larger comparator set is admissible at tolerance $\epsilon < T$, fix any expert $j \notin S$ and consider the loss sequence

$$\ell_{t,j} = 0, \qquad \ell_{t,i} = 1 \text{ for all } i \in S, \qquad t = 1, \dots, T. \tag{32}$$

Since $\pi_j^S = 0$, the multiplicative-weights update keeps zero mass on expert j at every round, so the learner incurs loss 1 at each time while expert j incurs loss 0. Therefore any comparator set containing an expert outside S has worst-case regret at least T. Moreover,

$$\epsilon_m = \sqrt{\frac{T \log m}{2}} < T, \tag{33}$$

because $m \le N$ and $\log N < 2T$. Therefore every admissible comparator set at tolerance ϵ_m must be contained in S, so

$$U_{\mathcal{A}_S}^{\mathrm{cmp}}(\epsilon_m, T) \le \log m. \tag{34}$$

Combining the lower and upper bounds gives

$$U_{\mathcal{A}_S}^{\mathrm{cmp}}(\epsilon_m, T) = \log m. \tag{35}$$

Finally, Proposition 3 shows that the minimax hardness of the ambient expert class $[N]$ remains $\Theta(\sqrt{T \log N})$, independent of the chosen prior.

A.7 Proof of Theorem 2

Fix $m \in [N]$ and let $\mathcal{A}_m$ be Hedge with uniform prior on S_m and learning rate $\eta = \sqrt{8 \log m / T}$.

Forward direction of (i). For $S' \subseteq S_m$, the best expert in S' is at least as good as the best expert in S_m on every loss sequence, so

$$\mathrm{Reg}_T(\mathcal{A}_m; S') \le \mathrm{Reg}_T(\mathcal{A}_m; S_m) \le \sqrt{\frac{T \log m}{2}} = \epsilon_m, \tag{36}$$

where the second inequality is Proposition 4.

Reverse direction of (i). Fix $j \in S' \setminus S_m$ and consider the loss sequence

$$\ell_{t,j} = 0, \qquad \ell_{t,i} = 1 \text{ for all } i \in S_m, \qquad t = 1, \dots, T, \tag{37}$$

with losses of experts in $[N] \setminus (S_m \cup \{j\})$ arbitrary. Since $\pi_j^{S_m} = 0$, the multiplicative-weights update assigns zero mass to j at every round, so the learner's expected loss is 1 at each time while the best expert in S' (namely j) incurs loss 0. Hence $\mathrm{Reg}_T(\mathcal{A}_m; S') \geq T$. As in Proposition 4, $\epsilon_m = \sqrt{T \log m/2} < T$ because $m \leq N$ and $\log N < 2T$. Therefore S' is not ϵ_m-admissible, establishing the equivalence.

Profile identity. By the definition of the comparator-universality profile, $U^{\mathrm{cmp}}_{\mathcal{A}_m}(\epsilon_m, T) = \log \max\{|S| : \mathrm{Reg}_T(\mathcal{A}_m; S) \leq \epsilon_m\}$. By the equivalence, the admissible comparator sets are exactly the subsets of S_m, whose maximum size is m, so $U^{\mathrm{cmp}}_{\mathcal{A}_m}(\epsilon_m, T) = \log m$.

Part (ii). The ambient minimax regret $\inf_{\mathcal{A}} \sup_{\ell} \mathbb{E}[\mathrm{Reg}_T(\mathcal{A})]$ equals $\Theta(\sqrt{T \log N})$ by Proposition 3, independently of m. Part (i) shows that the profiles of $\mathcal{A}_m$ range over $\{\log m : m \in [N]\}$ on the fixed class $[N]$, so the profile is not determined by the ambient minimax value.

A.8 Bandit instantiation: proofs for Theorems 3–4

A packing set. By a standard Varshamov–Gilbert argument (see, e.g., (Tsybakov, 2009)), there exists a subset $\Theta_0 \subseteq \{-1, +1\}^K$ such that $|\Theta_0| \geq \exp(c_0 K)$ and for all distinct $\theta, \theta' \in \Theta_0$,

$$\rho(\theta, \theta') = \frac{1}{K} d_H(\theta, \theta') \geq 2\epsilon. \tag{38}$$

This ensures Θ_0 is a 2ϵ-packing under ρ.

A KL upper bound for any adaptive algorithm. Fix $\theta, \theta' \in \Theta_0$. Let $P_\theta^{\mathcal{A}}$ denote the distribution over the full transcript $Z_T = (A_{1:T}, Y_{1:T})$ induced by algorithm $\mathcal{A}$ in environment θ. By the chain rule for KL divergence and the bandit conditional independence,

$$D_{\mathrm{KL}}(P_\theta^{\mathcal{A}} \| P_{\theta'}^{\mathcal{A}}) = \mathbb{E}_\theta \left[\sum_{t=1}^{T} D_{\mathrm{KL}}(P_\theta(Y_t \mid A_t, Z_{t-1}) \,\|\, P_{\theta'}(Y_t \mid A_t, Z_{t-1})) \right]. \tag{39}$$

Conditioned on $(A_t = i, Z_{t-1})$, Y_t is Bernoulli with mean $1/2 + \theta_i \Delta$ under θ, and $1/2 + \theta'_i \Delta$ under θ'. Therefore each summand is either 0 (if $\theta_i = \theta'_i$) or equals $D_{\mathrm{KL}}(\mathrm{Ber}(1/2+\Delta) \| \mathrm{Ber}(1/2-\Delta))$ (if $\theta_i \neq \theta'_i$), which is bounded by a constant multiple of Δ^2 for $\Delta \leq 1/4$. In particular, for all $\Delta \leq 1/4$,

$$D_{\mathrm{KL}}(\mathrm{Ber}(1/2 + \Delta) \| \mathrm{Ber}(1/2 - \Delta)) \leq 16\Delta^2. \tag{40}$$

To verify this bound, write $p = 1/2 + \Delta$ and note that

$$D_{\mathrm{KL}}(\mathrm{Ber}(p) \| \mathrm{Ber}(1-p)) = 2\Delta \log\left(\frac{1+2\Delta}{1-2\Delta}\right). \tag{41}$$

Using $\log\left(\frac{1+x}{1-x}\right) \leq \frac{2x}{1-x}$ for $x \in (0,1)$ and $x = 2\Delta \leq 1/2$ yields $D_{\mathrm{KL}}(\mathrm{Ber}(p) \| \mathrm{Ber}(1-p)) \leq \frac{8\Delta^2}{1-2\Delta} \leq 16\Delta^2$. Since the total number of rounds is T,

$$D_{\mathrm{KL}}(P_\theta^{\mathcal{A}} \| P_{\theta'}^{\mathcal{A}}) \leq 16T\Delta^2. \tag{42}$$

Proof of Theorem 3. Assume $\mathcal{A}$ is ϵ-successful on every $\theta \in \Theta_0$ and let $\widetilde{\Theta}$ be uniform on Θ_0. Define

$$\kappa = \frac{1}{|\Theta_0|^2} \sum_{\theta, \theta' \in \Theta_0} D_{\mathrm{KL}}(P_\theta^{\mathcal{A}} \| P_{\theta'}^{\mathcal{A}}) \leq 16T\Delta^2. \tag{43}$$

By Lemma 2, we have $I(\widetilde{\Theta}; Z_T) \leq \kappa \leq 16T\Delta^2$. On the other hand, Lemma 1 with $\delta = 1/4$ (Remark 1) gives $I(\widetilde{\Theta}; Z_T) \geq (3/4) \log |\Theta_0| - \log 2$. Combining and rearranging yields $T \geq c_1 \log |\Theta_0| / \Delta^2$ for a universal c_1.

Proof of Theorem 4. Let $\widehat{\mu}_i$ be the empirical mean of arm i based on m independent pulls. By Hoeffding's inequality, for each fixed i,

$$\mathbb{P}_\theta\big(|\widehat{\mu}_i - (1/2 + \theta_i\Delta)| \geq \Delta/2\big) \leq 2\exp(-m\Delta^2/2). \tag{44}$$

Taking a union bound over $i \in [K]$, if $m \gtrsim \Delta^{-2}\log(K/\delta)$ then with probability at least $1-\delta$ we have $\mathrm{sign}(\widehat{\mu}_i - 1/2) = \theta_i$ for all i, i.e., $\widehat{\theta} = \theta$.

A.9 Linear bandit instantiation: proof of Proposition 5

We consider the Gaussian linear bandit model in the statement of Proposition 5. Let $\epsilon \in (0, 1]$.

A packing set. By a standard Varshamov–Gilbert argument (see, e.g., (Tsybakov, 2009)), there exists a subset $\mathcal{V} \subseteq \{-1, +1\}^d$ with $|\mathcal{V}| \geq \exp(c_0 d)$ such that for all distinct $v, v' \in \mathcal{V}$, $d_H(v, v') \geq d/4$. Averaging the code over the sign-flip group if needed, we may take it symmetric and isotropic: for $\widetilde{\Theta}$ uniform on the packing, $\mathbb{E}[\widetilde{\Theta}] = 0$ and $\mathbb{E}[\widetilde{\Theta}\widetilde{\Theta}^\top] = (4\epsilon^2/d)\, I_d$. Define $\Theta_0 := \{(2\epsilon/\sqrt{d})v : v \in \mathcal{V}\}$. Then for distinct $\theta, \theta' \in \Theta_0$,

$$\|\theta - \theta'\|_2 = \frac{2\epsilon}{\sqrt{d}}\|v - v'\|_2 = \frac{2\epsilon}{\sqrt{d}} \cdot 2\sqrt{d_H(v, v')} \geq 2\epsilon, \tag{45}$$

so Θ_0 is a 2ϵ-packing under $\rho(\widehat{\theta}, \theta) = \|\widehat{\theta} - \theta\|_2$, and $\log|\Theta_0| \geq c_0 d$.

A transcript information upper bound. Let $\widetilde{\Theta}_G \sim \mathcal{N}(0, (4\epsilon^2/d)I_d)$ be a Gaussian prior matching the isotropic second moment of the packing. For the Gaussian linear channel with adaptive unit-norm actions, the posterior covariance obeys the rank-one update

$$\Sigma_t = \Sigma_{t-1} - \frac{\Sigma_{t-1}X_t X_t^\top \Sigma_{t-1}}{1 + X_t^\top \Sigma_{t-1} X_t}, \tag{46}$$

so $\Sigma_t \preceq \Sigma_0 = (4\epsilon^2/d)I_d$ in the Loewner order and $\lambda_{\max}(\Sigma_{t-1}) \leq 4\epsilon^2/d$. Hence

$$I(\widetilde{\Theta}_G; Z_T) = \frac{1}{2}\sum_{t=1}^{T}\log(1 + X_t^\top \Sigma_{t-1} X_t) \leq \frac{T}{2}\log\Big(1 + \frac{4\epsilon^2}{d}\Big) \leq \frac{2\epsilon^2 T}{d}. \tag{47}$$

Reduction to the discrete packing. Quantize the Gaussian prior to the packing by nearest-neighbor projection, $\widetilde{\Theta}_q = \pi(\widetilde{\Theta}_G)$. Since Θ_0 is a maximal 2ϵ-packing of the sphere of radius 2ϵ, its Voronoi cells are near-congruent, so the isotropic Gaussian projects to a law that is $(1 + o(1))$-uniform on Θ_0 (the standard covering/packing uniformity, (Wainwright, 2019)). By data processing and the continuity of mutual information on the finite alphabet Θ_0,

$$I(\widetilde{\Theta}; Z_T) \leq I(\widetilde{\Theta}_q; Z_T) + o(d) \leq I(\widetilde{\Theta}_G; Z_T) + o(d) \leq \frac{2\epsilon^2 T}{d} + o(d). \tag{48}$$

Conclusion. Assume $\mathcal{A}$ is ϵ-successful on every $\theta \in \Theta_0$. Lemma 1 with $\delta = 1/4$ (Remark 1) gives $I(\widetilde{\Theta}; Z_T) \geq (3/4)\log|\Theta_0| - \log 2 \geq c_0' d$. Combining with the previous upper bound yields $c_0' d \leq 2\epsilon^2 T/d + o(d)$, i.e. $T \geq c_1 d^2/\epsilon^2$ for a universal constant c_1. This matches the isotropic achievability of Theorem 5 up to logarithmic factors, so the local linear bandit is a tight structured witness (cf. the fixed-design minimax rate (Wainwright, 2019) and the adaptivity results of (Maiti et al., 2026)).

Why d^2/ϵ^2, not d/ϵ^2. The scale d^2/ϵ^2 is forced by the unit-norm action constraint $\|X_t\|_2 \leq 1$. Under an isotropic unit-norm design, $X^\top X \approx (T/d)I_d$, so the least-squares error $\widehat{\theta} - \theta = V_T^{-1}\sum_t X_t\xi_t$ has covariance $V_T^{-1} \approx (d/T)I_d$ and squared ℓ_2 error $\mathbb{E}\|\widehat{\theta} - \theta\|_2^2 = \mathrm{tr}(V_T^{-1}) = d^2/T$; equating this to ϵ^2 gives $T = d^2/\epsilon^2$. This is the classical full-parameter ℓ_2 *estimation* rate (Wainwright, 2019). The d/ϵ^2 rate one might expect corresponds instead to (i) ℓ_2 *prediction* error, $\|\widehat{\theta} - \theta\|^2_{X^\top X/n} = \mathrm{tr}(V_T V_T^{-1})/n = d/n$, or (ii) designs with $\|x\| \approx \sqrt{d}$ (fixed-design regression with $X^\top X \approx TI$), which violate the unit-norm action constraint. Neither applies here: the profile is defined in terms of the parameter estimation error $\rho(\widehat{\theta}, \theta) = \|\widehat{\theta} - \theta\|_2$, under unit-norm actions, and both give the d^2/ϵ^2 scale.

A.10 Linear bandit instantiation: proof of Theorem 5

We use the unit-sphere action set and the isotropic design $X_1, \dots, X_T$ i.i.d. uniform on $\mathcal{X}$. Let $V_T = \sum_{t=1}^T X_t X_t^\top$ and $S_T = \sum_{t=1}^T X_t Y_t$, so that $\widehat{\theta} = V_T^{-1} S_T$. Substituting $Y_t = \langle X_t, \theta^\star \rangle + \xi_t$ gives

$$\widehat{\theta} - \theta^\star = V_T^{-1} \sum_{t=1}^T X_t \xi_t. \tag{49}$$

Since X_t is independent of ξ_t and the noise is $\mathcal{N}(0,1)$, conditionally on the design the error $\widehat{\theta} - \theta^\star$ is centered Gaussian with covariance V_T^{-1}; in particular its distribution does not depend on $\theta^\star$.

Gram-matrix concentration. Because X_t is uniform on the unit sphere, $\mathbb{E}[X_t X_t^\top] = \frac{1}{d} I_d$ and $\|X_t X_t^\top\|_2 \le 1$. By the matrix Bernstein bound, there is a universal constant $c > 0$ such that for every $\tau \in (0,1)$,

$$\mathbb{P}\left(\lambda_{\min}(V_T) \le \frac{(1-\tau)T}{d} \right) \le d \exp\left(-\frac{c\tau^2 T}{d} \right). \tag{50}$$

Choosing $\tau = 1/2$ and $T \ge C_0 d \log(2d/\delta)$ for a sufficiently large universal C_0, we obtain the event

$$\mathcal{E} := \left\{ \lambda_{\min}(V_T) \ge \frac{T}{2d} \right\}, \qquad \mathbb{P}(\mathcal{E}^c) \le \frac{\delta}{2}. \tag{51}$$

Error tail. On $\mathcal{E}$, writing the standardized error as $Z \sim \mathcal{N}(0, I_d)$, we have

$$\|\widehat{\theta} - \theta^\star\|_2^2 \le \frac{2d}{T} \|Z\|_2^2. \tag{52}$$

The Laurent–Massart χ^2_d tail bound yields

$$\mathbb{P}\Big(\|Z\|_2^2 \ge d + 2\sqrt{d \log(2/\delta)} + 2\log(2/\delta) \Big) \le \frac{\delta}{2}. \tag{53}$$

Hence, on $\mathcal{E}$ and with probability at least $1 - \delta/2$ over the noise,

$$\|\widehat{\theta} - \theta^\star\|_2^2 \le \frac{2d}{T} \Big(d + 2\sqrt{d \log(2/\delta)} + 2\log(2/\delta) \Big). \tag{54}$$

For $T \ge C \frac{d^2}{\epsilon^2} \log \frac{d}{\delta}$ with a sufficiently large universal constant C, the right-hand side is at most ϵ^2 and the total failure probability is at most δ. Since the error distribution is independent of $\theta^\star$, the bound holds uniformly for every $\theta^\star$ with $\|\theta^\star\|_2 \le 2\epsilon$, and in particular uniformly over the packed family Θ_0 of Proposition 5. This yields ϵ-success at horizon T, as claimed.

A.11 A more explicit passive-estimation specialization

In the passive setting with $Z_n = X_{1:n}$, a standard upper bound is

$$I(\widetilde{\Theta}; X_{1:n}) \le \frac{1}{M^2} \sum_{i,j=1}^M D_{\mathrm{KL}}(P_{\theta_i}^{\otimes n} \| P_{\theta_j}^{\otimes n}) = n \cdot \frac{1}{M^2} \sum_{i,j=1}^M D_{\mathrm{KL}}(P_{\theta_i} \| P_{\theta_j}), \tag{55}$$

which can be combined with Lemma 2 to obtain an information budget with $\mathsf{Info}(\epsilon)$ equal to the average pairwise one-sample KL on the packing.

A.12 Passive-estimation sanity check

The framework also recovers the usual passive-estimation scaling. For a function class $\mathcal{F}$ with L^2 metric, classical nonparametric lower bounds use packings whose size is governed by metric entropy (Tsybakov, 2009); in our notation, the maximal universality at accuracy ϵ is of order $\log \mathcal{M}(2\epsilon, \mathcal{F}, \|\cdot\|_{L^2})$, and eq. (8) yields the standard implication that smaller ϵ requires larger n.

Corollary 3 (Metric-entropy scaling under average-KL control)**.** *Suppose that for each $\epsilon > 0$ there exists a 2ϵ-packing $\Theta_0 \subseteq \mathcal{F}$ with $\log|\Theta_0| \geq c_0\epsilon^{-\alpha}$ and average one-sample KL at most κ_0, i.e.*

$$\frac{1}{|\Theta_0|^2}\sum_{f,g\in\Theta_0} D_{\mathrm{KL}}(P_f\|P_g) \leq \kappa_0,$$

where $c_0, \kappa_0 > 0$ are independent of ϵ. Then any algorithm that is ϵ-successful on every element of Θ_0 must use $n \geq c_1\epsilon^{-\alpha}$ for a constant $c_1 > 0$. Thus standard nonparametric families with entropy $\asymp \epsilon^{-\alpha}$ and bounded local KL recover the usual entropy-driven sample complexity.

A.13 Proof of Corollary 1

Under the stated average-KL condition, the previous inequality yields $I(\widetilde{\Theta}; X_{1:n}) \leq n\kappa$. Combining with Lemma 1, we obtain $(1-\delta)\log M - \log 2 \leq n\kappa$. Rearranging gives $\log M \leq \frac{1}{1-\delta}n\kappa + \frac{1}{1-\delta}\log 2$, which implies the claim since $U_{\mathcal{A}}(\epsilon, n) = \log M$ for a maximal packing.

A.14 Proof of Corollary 3

Apply Corollary 1 with the stated packing Θ_0. Since

$$\frac{1}{|\Theta_0|^2}\sum_{f,g\in\Theta_0} D_{\mathrm{KL}}(P_f\|P_g) \leq \kappa_0, \tag{56}$$

we obtain

$$\log|\Theta_0| \leq \frac{1}{1-\delta}n\kappa_0 + \frac{1}{1-\delta}\log 2. \tag{57}$$

Using the lower bound $\log|\Theta_0| \geq c_0\epsilon^{-\alpha}$ and rearranging yields

$$n \geq \frac{3c_0}{4\kappa_0}\epsilon^{-\alpha} - \frac{\log 2}{\kappa_0}. \tag{58}$$

For sufficiently small ϵ, the first term dominates the additive constant, so $n \geq c_1\epsilon^{-\alpha}$ for a constant $c_1 > 0$ depending only on c_0 and κ_0.

A.15 Proof of Proposition 1 (full statement: Proposition 6)

Success-set containment. Since $\mathcal{A}_m$ outputs estimates in V_m, whenever $\|\widehat{f} - f\|_{L^2} \leq \epsilon$ we have $\mathrm{dist}_{L^2}(f, V_m) \leq \|f - \widehat{f}\|_{L^2} \leq \epsilon$; together with $\|f\|_{L^2} \leq R$ this gives $\mathsf{Succ}_{\mathcal{A}_m}(\epsilon, n) \subseteq S_m$.

Packing of the constrained set. Let $S_m = \{f : \|f\|_{L^2} \leq R,\ \mathrm{dist}_{L^2}(f, V_m) \leq \epsilon\}$. Write $f = g + h$ with $g = \mathrm{proj}_{V_m} f \in V_m$ and $\|h\|_{L^2} \leq \epsilon$. For the upper bound, an ϵ-net of the m-dimensional ball $\{g \in V_m : \|g\|_{L^2} \leq R\}$ has size at most $(cR/\epsilon)^m$ by the volume ratio in $\mathbb{R}^m$, and the ball $\{h : \|h\|_{L^2} \leq \epsilon\}$ is covered by a single ϵ-ball, so $\mathcal{M}(2\epsilon, S_m, L^2) \leq (cR/\epsilon)^m$. For the lower bound, the ball $\{g \in V_m : \|g\|_{L^2} \leq R/2\}$ is contained in S_m and has a 2ϵ-packing of size at least $(c'R/\epsilon)^m$. Hence $\log\mathcal{M}(2\epsilon, S_m, L^2) = \Theta(m\log(R/\epsilon))$, and the containment gives $U_{\mathcal{A}_m}(\epsilon, n) \leq \Theta(m\log(R/\epsilon))$.

RKHS-ball entropy. For a Mercer kernel with $\lambda_j \asymp j^{-\beta}$, the RKHS ball $B_H(R) = \{\sum_j a_j\phi_j : \sum_j a_j^2/\lambda_j \leq R^2\}$ is, in the coefficient coordinates (where the L^2 metric becomes the ℓ^2 metric), an ellipsoid with semiaxes $R\lambda_j^{1/2} \asymp Rj^{-\beta/2}$. The metric entropy of such an ellipsoid is classical; it satisfies $\log\mathcal{M}(2\epsilon, B_H(R), L^2) \asymp (R/\epsilon)^{2/\beta}$ (Tsybakov, 2009). For completeness, the matching lower bound follows from the volume of the truncated ellipsoid: choosing $J = (cR/\epsilon)^{2/\beta}$, the first J coordinates have semiaxes $Rj^{-\beta/2} \geq c\epsilon$, and the volume ratio gives $\log\mathcal{M}(2\epsilon, B_H(R), L^2) \gtrsim \sum_{j=1}^{J}\log(Rj^{-\beta/2}/\epsilon) = \Theta((R/\epsilon)^{2/\beta})$.

B Additional Results and Discussion

B.1 Metric sensitivity

Remark 6 (Metric sensitivity). *The profile is defined relative to a fixed task metric, so one may ask how the conclusions depend on that choice. Suppose ρ' is bi-Lipschitz equivalent to ρ: there exist constants $c, C > 0$ with $c\,\rho'(s,t) \le \rho(s,t) \le C\,\rho'(s,t)$ for all s, t. Then a 2ϵ-packing under ρ is a $(2\epsilon/C)$-packing under ρ' and a 2ϵ-packing under ρ' is a $(2c\epsilon)$-packing under ρ, so the log-packing numbers satisfy $\log \mathcal{M}(2\epsilon, \mathcal{S}, \rho) \le \log \mathcal{M}(2\epsilon/C, \mathcal{S}, \rho')$ and symmetrically. Consequently, at tolerances within a constant factor, the log-profile differs by at most an absolute constant and the ordering of algorithms by breadth is unchanged: the framework is robust to reasonable choices of ρ up to constants. Concretely, in the Bernoulli family the normalized Hamming error on the environment index and the ℓ_1 distance on the reward means are equivalent up to the gap Δ and the arm count K, so choosing between them changes constants, not qualitative conclusions. The same holds for the observed phase transitions: replacing ρ by an equivalent metric shifts the knee of the breadth–budget curve by only an $O(1)$ factor in the budget, leaving the qualitative coverage transition intact. In high-dimensional or nonparametric settings the practical choice is guided by the same principle: pick ρ so that $\rho \le \epsilon$ encodes the downstream performance requirement; any two such choices in the same bi-Lipschitz class yield equivalent profiles up to constants.*

B.2 Feature capacity bounds breadth: a kernel witness

To show that the framework is not limited to finite or linear structures, consider the nonparametric Gaussian regression setting that is the linear analogue of a random-feature or truncated-network learner. Let k be a Mercer kernel on $\mathcal{X}$ with L^2-orthonormal eigenfunctions $\phi_1, \phi_2, \dots$ and eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots > 0$ under the input distribution, and let $V_m = \operatorname{span}\{\phi_1, \dots, \phi_m\}$. For $m \in \mathbb{N} \cup \{\infty\}$, let $\mathcal{A}_m$ be the least-squares estimator restricted to V_m, which for $m < \infty$ is the linear model with the first m features and for $m = \infty$ is full kernel ridge regression. The task family is the RKHS ball $B_H(R) = \{f \in \mathcal{H}_k : \|f\|_{\mathcal{H}_k} \le R\}$.

Proposition 6 (Feature count bounds sustainable breadth). *Consider Gaussian regression $Y = f(x) + \xi$, $\xi \sim \mathcal{N}(0,1)$, with task distance $\rho(\widehat{f}, f) = \|\widehat{f} - f\|_{L^2}$ and tasks in $B_H(R)$.*

1. *For the m-feature estimator $\mathcal{A}_m$, any task it estimates to accuracy ϵ must lie within L^2-distance ϵ of V_m, i.e. $\mathsf{Succ}_{\mathcal{A}_m}(\epsilon, n) \subseteq S_m$ where $S_m = \{f : \|f\|_{L^2} \le R,\ \operatorname{dist}_{L^2}(f, V_m) \le \epsilon\}$. The packing entropy of S_m satisfies*

$$\log \mathcal{M}(2\epsilon, S_m, L^2) = \Theta\big(m \log(R/\epsilon)\big), \tag{59}$$

   so the universality profile is bounded by feature capacity: $U_{\mathcal{A}_m}(\epsilon, n) \le \Theta(m \log(R/\epsilon))$, independent of the ambient class.

2. *If the eigenvalues decay polynomially, $\lambda_j \asymp j^{-\beta}$ for some $\beta > 0$, the RKHS ball itself has packing entropy*

$$\log \mathcal{M}(2\epsilon, B_H(R), L^2) \asymp (R/\epsilon)^{2/\beta}, \tag{60}$$

   so the full-kernel estimator can in principle sustain far more breadth than $\mathcal{A}_m$ at the same accuracy. The gap between the two profiles is an algorithm-centric constraint that the ambient minimax value, a single number for the class, does not express.

Interpretation. The finite-feature witness is the nonparametric echo of the experts separation: on the same ambient class, an algorithm's capacity (here, the number of features) sharply limits which subfamily it covers, and this limit is invisible to class-level minimax. The packing bound is a volume-ratio argument: an ϵ-tube around V_m is covered by $(cR/\epsilon)^m$ balls (the m-dimensional ball's volume ratio) times a single ball for the ϵ-sized orthogonal tail, giving $\Theta(m \log(R/\epsilon))$; no assumption on the noise level or regularization is needed for the profile bound itself. Its sample-complexity side is the classical passive rate: in Gaussian regression with feature covariance $\Sigma = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$, least squares estimates the m-dimensional coefficient vector to ℓ_2 accuracy ϵ with $n \gtrsim \operatorname{tr}(\Sigma^{-1})/\epsilon^2$ samples, which is $\Theta(m/\epsilon^2)$ for well-conditioned features and matches standard practice. The m^2/ϵ^2 rate of the interactive model in Section 6 is a distinct phenomenon, arising because bounded actions cap the per-coordinate information; the two rates describe the same packing-entropy object under different observation models. A proof of the profile claim is in Appendix Section A.15.

B.3 THE EXPERTS WITNESSES IN DETAIL

Remark 7 (Sensitivity to the prior's support)**.** *Theorem 2 uses a prior with zero mass outside S_m, and its certificate is exact: any comparator set containing an out-of-support expert fails. The strictness degrades smoothly as outside experts acquire positive prior mass. Suppose each expert $j \notin S_m$ carries prior mass $\pi_j > 0$. By Proposition 2, at the tuned learning rate the regret of Hedge against expert j satisfies $\mathrm{Reg}_T(\mathcal{A}_m; \{j\}) \le \sqrt{T \log(1/\pi_j)/2}$, so expert j becomes ϵ-admissible as soon as $\log(1/\pi_j) \le 2\epsilon^2/T$. Since the regret against any comparator set S' is at most the regret against its best-supported expert, the success set at tolerance ϵ is*

$$\Big\{S' \subseteq [N] : \min_{j \in S' \setminus S_m} \log(1/\pi_j) \le 2\epsilon^2/T\Big\},$$

and the profile interpolates continuously from $\log m$ (as $\pi_j \to 0$) toward $\log N$ (at the uniform prior) as the outside mass grows. This quantifies exactly how the strict certificate depends on the prior's support, and shows the separation of Theorem 2 is sharp precisely for committed (zero-mass) strategies. Under the uniform prior $\pi_i = 1/N$ every algorithm attains the same profile $\log N$ and the separation collapses; a meaningful separation persists exactly when the prior is non-uniform, with its strength governed by how concentrated the prior is on its support. The zero-mass case is therefore the clean limit of a continuous family of positive-prior separations, not a degenerate artefact. The characterization is unchanged under other multiplicative-weights variants (AdaHedge, Prod, or online mirror descent with entropic regularization): all keep zero mass on zero-prior experts and attain the same regret order at their tuned rates, so the separation is about the prior's support rather than the exponentiation rule, with the learning-rate schedule affecting only the constant in $\epsilon_m = \sqrt{T \log m/2}$.

Remark 8 (What is, and is not, recoverable from minimax)**.** *Whether the framework yields a bound not obtainable from standard minimax packings has a precise answer in this section. The frontier law itself, $\epsilon \ge c\sqrt{T\, U^{\mathrm{cmp}}_{\mathcal{A}}(\epsilon, T)}$ of Corollary 2,* is *recoverable by applying the class-level minimax bound to subclasses: restricting to a comparator set of size $\exp(B)$ and invoking Proposition 3 reproduces it. What the class-level packing cannot see is the* assignment*: which profile a fixed algorithm attains, and the strictness of its success set. Theorem 2 exhibits a fixed ambient class on which the profile provably ranges from $\log m$ to $\log N$ as the prior support varies while the ambient minimax value stays put, and on which a specific algorithm provably fails on every comparator set outside its support. This is the algorithm-centric content that the minimax packing of the ambient class does not express.*

Relation to known expert-learning bounds. At full breadth ($B = \log N$) the frontier reproduces the classical law $\epsilon \asymp \sqrt{T \log N}$, matching the minimax regret of the class by Proposition 3. The added content at intermediate breadth is that the frontier is now *assigned to an algorithm*: it is the guarantee of a fixed Hedge instance rather than the minimax optimum over all algorithms, and it comes with a provable strictness certificate (Theorem 2). Two algorithms on the same class, Hedge with support m and Hedge with support N, face the identical ambient minimax value $\Theta(\sqrt{T \log N})$ yet have profiles differing by $\log(N/m)$; the class-level bound, which takes an infimum over algorithms, does not expose this axis.

Remark 9 (A game-theoretic reading)**.** *Proposition 4 and Theorem 2 admit a max–min interpretation. The expert setting is a two-player game in which the learner commits to a strategy and the adversary then chooses the loss sequence. The comparator-universality profile of a fixed algorithm asks how large a comparator set the learner can serve while keeping worst-case regret at most ϵ; it is the value of the game restricted to the learner's committed strategy (for $\mathcal{A}_m$, a strategy supported on S_m). The ambient minimax regret $\Theta(\sqrt{T \log N})$ is the value of the unrestricted game, in which the learner may choose the strategy after seeing the class. The separation between $\log m$ and $\log N$ is then the statement that a committed support strictly reduces the value of the game, and minimax theory computes only the unrestricted value. This is the precise sense in which the profile measures a committed strategy's breadth rather than the class's intrinsic difficulty.*

B.4 BANDIT IDENTIFICATION IN CONTEXT

Relation to known identification bounds. The horizon law $T \gtrsim K/\Delta^2$ of Theorem 3 coincides, up to constants, with the classical sample complexity of K-armed bandit identification at separation Δ (Mannor & Tsitsiklis, 2004; Bubeck & Cesa-Bianchi, 2012): it is the minimax rate of the family,

and the framework recovers it from the universality profile without an arm-by-arm argument. What the profile adds is the coverage reading: $U_{\mathcal{A}}(\epsilon, T)$ counts how many environments a *fixed* algorithm solves, so two algorithms on the same packed family can be ranked directly. In the synthetic study of Section 7, uniform sampling attains breadth $\log 64$ once $T\Delta^2$ is large enough, while a greedy proxy stays near zero breadth on the same family, an algorithm-level distinction that the family's minimax value does not express. A finer instance-dependent analysis, weighting each environment by the number of arms on which its reward mean actually differs, would parallel instance-optimal best-arm identification, where sample complexity is governed by distribution-dependent information quantities rather than the worst-case gap; we leave this refinement open.

B.5 The linear-bandit frontier in context

The price of broad coverage: universal versus targeted estimation. Together, Proposition 5 and Theorem 5 establish a frontier for the local linear bandit that is tight up to logarithmic factors: the information-theoretic lower bound $T = \Omega(d^2/\epsilon^2)$ matches the isotropic least-squares achievability $T = O(d^2\epsilon^{-2}\log(d/\delta))$. The scale d^2/ϵ^2 is the price of broad coverage. An algorithm with no prior knowledge of which coordinate of $\theta^\star$ is large must allocate observations uniformly across all d directions; reducing each coordinate's error to $\epsilon/\sqrt{d}$ costs $O(d/\epsilon^2)$ samples per coordinate and $O(d^2/\epsilon^2)$ in total. By contrast, a target-aware oracle that knows the support of $\theta^\star$ focuses all samples on a single direction and achieves $T = O(1/\epsilon^2)$, forgoing the uniform guarantee. This is the design-of-experiments reading: for uniform ℓ_2 accuracy ϵ, the D-optimal design is isotropic at scale d^2/ϵ^2, and adaptivity cannot beat it, because the isotropic packing offers no direction to exploit (Wainwright, 2019; Maiti et al., 2026). The profile makes this tension legible: the universality profile of the isotropic scheme saturates at $\log|\Theta_0| = \Theta(d)$ exactly at the horizon where uniform coverage first becomes feasible, whereas a target-aware scheme reaches d/ϵ^2 only by collapsing its success set to a single direction. The same estimation factor is isolated in DEC-style analyses through fractional coverings (Foster et al., 2021; 2023); the profile attributes that cost to a fixed algorithm's isotropic exploration rather than to the class.

Relation to known linear-bandit estimation. The lower bound $T \geq c\, d^2/\epsilon^2$ is an information-theoretic impossibility for any algorithm, while the achievable rate of Theorem 5, $T = O(d^2\epsilon^{-2}\log(d/\delta))$, matches the standard sample complexity of full-parameter ℓ_2 estimation in d dimensions (Lattimore & Szepesvári, 2020; Wainwright, 2019). The class-level minimax rate for this local family is of order d/ϵ^2 for a single environment (easy once you know which one you face), but an algorithm that must succeed on all $\exp(\Theta(d))$ environments simultaneously is forced to the d^2/ϵ^2 scale. As in the Bernoulli family, the framework additionally lets two algorithms on the same local family, isotropic least squares and a balanced coordinate design (Section 7), be ranked by their empirical breadth, a comparison that the class-level minimax rate does not make.

B.6 Real-data result: digits (E4)

To confirm that the profile is observable outside synthetic regimes, we run a lightweight real-data check. A single multinomial logistic-regression model is trained on the `sklearn` digits dataset (10 classes, 1797 samples) on a random subset of n training samples, the budget, and evaluated on all 45 binary class-pair tasks at target accuracy $\geq 90\%$. The breadth proxy is the logarithm of the number of pairs solved. Figure 3 shows the result averaged over 10 seeds: breadth rises monotonically from $\approx \log 2$ at $n = 15$ to its saturation value $\log 45 \approx 3.81$ as the budget grows, with a knee near $n \approx 120$. Two features match the theory: breadth saturates exactly at the ambient class packing $\log 45$ once the success set covers the whole task family (Remark 8), and the transition region is where the frontier predicts coverage becomes information-feasible. E4 is deliberately a sanity check: its aim is to confirm that the profile is a measurable quantity on non-synthetic data and that the predicted breadth–budget law, saturation at the class packing and a phase-transition knee, is visible there, not to benchmark large-model coverage.

B.7 Frozen-representation result (E5)

The finite-feature theory (Proposition 6) predicts that an algorithm's capacity bounds its breadth. E5 tests this on real data with *frozen* representations: a fixed, untrained representation, PCA learned

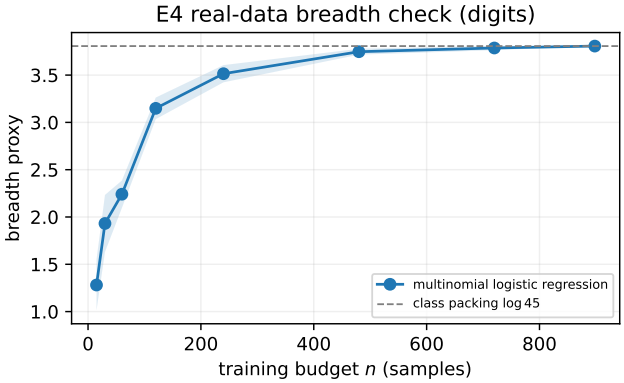


Figure 3: Real-data breadth check (digits, 10 seeds, 95% CI). A single multinomial logistic-regression model trained on n samples covers an increasing number of the 45 binary class-pair tasks at $\geq 90\%$ accuracy; breadth saturates at the class packing $\log 45 \approx 3.81$ with a knee near $n \approx 120$.

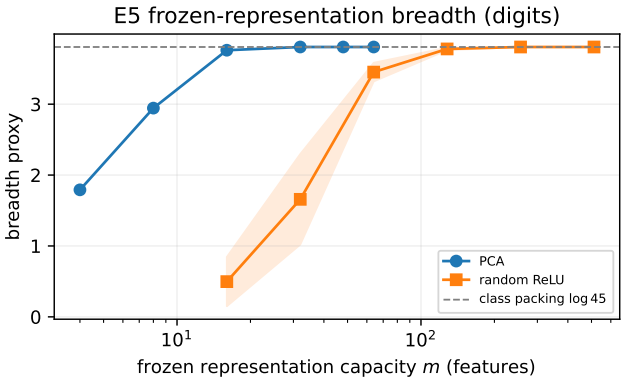


Figure 4: Frozen-representation breadth on digits. A frozen PCA or random-ReLU representation with a single multinomial logistic probe: breadth (log of the number of the 45 class-pair tasks solved at $\geq 90\%$ accuracy) rises with the representation capacity m and saturates at the class packing $\log 45 \approx 3.81$, matching the capacity-bounds-breadth prediction of Proposition 6. Shaded band: 95% CI over representation draws.

once on the train split or random ReLU features, is frozen, and a single multinomial logistic probe is trained on the frozen features for all 45 class-pair tasks. Figure 4 plots breadth against the number of features m (averaged over 10 representation draws for the random-ReLU curve). Breadth rises monotonically with capacity and saturates at the class packing $\log 45 \approx 3.81$ once the representation is rich enough ($m \gtrsim 32$ for PCA, $m \gtrsim 128$ for random ReLU). This is the empirical counterpart of Proposition 6: the frozen model covers exactly as much as its feature budget allows, a capacity–coverage trade-off that the ambient class's minimax value does not express. Like E4, this is a sanity check on a small-scale public dataset, not an evaluation of large models.

Role of the experiments. These experiments are deliberately synthetic and tightly coupled to the theory, except for the real-data checks E4, E5, and E8, whose role is to confirm the profile is observable beyond simulation. Their common role is to make algorithmic breadth observable and to show that the new complexity object tracks behavior not immediately visible from ambient class-level hardness alone. All synthetic experiments are lightweight NumPy simulations, E4 and E5 are lightweight `sklearn` fits, E8 additionally uses `torch`, and everything runs on a single CPU workstation.

Claim	Evidence
Profile measures the coverage of a fixed procedure	Def. 1, (3)
Broad success sets force mutual information	Lem. 1, Thm. 1
Breadth–budget frontier $n \gtrsim B/\mathsf{Info}(\epsilon)$	Thm. 1, Cor. 1
Strict success sets separate algorithms from class hardness	Prop. 4, Thm. 2
Near-sharp interactive identification frontier	Thms. 3–4
Two-sided structured frontier (linear bandit)	Prop. 5, Thm. 5
Feature capacity bounds breadth (kernel)	Prop. 6
Observable phase transitions	Figs. 1, 5, 3, 4

Table 1: Map of the paper's main claims to their supporting results and figures.

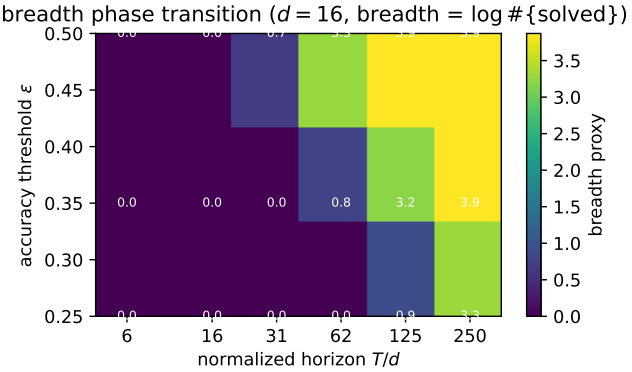


Figure 5: Breadth phase transition of the E3 witness ($d = 16$, isotropic least squares): the transition boundary moves to larger T/d as the threshold ϵ shrinks, matching $T/d \asymp d/\epsilon^2$ up to constants.

Remark 10 (Estimating the universality profile)**.** *In these instantiations the profile is estimated directly by Monte Carlo: fix a packed family Θ_0, run the algorithm on each task for several seeds, declare success when the task metric falls below ϵ, and report the logarithm of the number of successful environments (Section 7). This estimates the success-set cardinality on the chosen packing, which is exactly the profile. For a general system the same protocol applies with a held-out packing set, at the cost of sampling each task; the estimate carries the usual $O(1)$ slack from the packing and the failure tolerance δ. We do not claim an efficient estimator for the success set of an arbitrary black-box system; the protocol is meant for settings in which tasks can be sampled, as they can be here. Designing surrogate estimators that avoid running the algorithm on every task, for example adaptive probing guided by the predicted frontier or model-based surrogates of the success set, is a promising direction that we leave open.*

B.8 Claim–evidence map and covering-based characterizations

Remark 11 (Relation to covering- and quantile-based characterizations)**.** *The profile is itself a covering/geometric object, the packing complexity of an algorithm's success set, so it complements the mutual-information machinery that produces the frontier rather than depending on it. This also sharpens the relation to class-level characterizations. The decision-estimation coefficient (Foster et al., 2021; 2023) is a covering-based, worst-case quantity: it takes an infimum over policies and measures the intrinsic hardness of a class. The universality profile is the dual object restricted to a single procedure: it measures the covering complexity of the tasks that* a given *algorithm actually covers, at a fixed tolerance. Where DEC and interactive-Fano bounds (Chen et al., 2024) answer "how hard is the class for the best policy," the profile answers "how much of the class does this policy cover," a question that class-level covering numbers do not distinguish: two algorithms with identical DEC can have different profiles, as Theorem 2 shows. A quantile reading makes this concrete: the profile is the ϵ-covering complexity of the success set, i.e., the number of ϵ-separated tasks solved; sweeping the tolerance moves along a breadth–accuracy quantile curve, exactly as sweeping the*

distortion moves along a rate–distortion curve. Recent interactive-Fano extensions to tail-sensitive objectives use the same quantile/CVaR viewpoint (Bongole et al., 2026), and the profile can be seen as the success-set analogue of their risk quantiles.

B.9 Phase transitions, computational universality, and guidance

Remark 12 (Coverage transitions as a signature of the frontier)**.** *The synthetic studies show a sharp rise rather than a smooth increase: breadth stays near zero until the budget crosses a critical scale and then climbs to saturation (*$\log 64$ *in E2,* $\log 48$ *in E3). This is the empirical counterpart of the linear frontier* $n \gtrsim B/\mathsf{Info}(\epsilon)$*: for a finite packing, the number of environments a fixed algorithm can solve is roughly* $\exp(\Theta(n\,\mathsf{Info}))$ *capped at the family size, so a target coverage level is either information-feasible or not within a relatively narrow budget window. In identification tasks the threshold admits an early-stopping reading: it is the horizon at which a target coverage first becomes attainable, and stopping well before it forfeits coverage while going far beyond it returns little additional breadth.*

Remark 13 (Toward a computational universality)**.** *The frontier of Section 4 is purely information-theoretic: it counts samples or horizon, not compute. The template extends naturally by endowing the algorithm with a computation budget* τ *and defining the success set relative to both constraints,* $\mathsf{Succ}_{\mathcal{A}}(\epsilon, n, \tau)$*. The corresponding computational universality profile* $U_{\mathcal{A}}(\epsilon, n, \tau)$ *then measures breadth achievable within sample budget* n *and compute budget* τ*, and the frontier becomes a three-way trade-off between universality, convergence, and computation: broader coverage forces either more information or more compute. Appendix Section D formalizes this profile and records its basic properties (monotonicity in* τ*, saturation at the unconstrained profile, and a feature-count vignette linking compute to breadth). A fully quantitative theory requires a concrete model of computation and remains open; we flag it here as the natural next axis after the information-theoretic one. The experiments of Section 7 are lightweight and operate far from any computational constraint, so this extension does not affect them.*

When to use the universality lens. The framework is most informative when the success set is a strict subset of the ambient class, so that the profile lies strictly below the class's packing entropy and is not predicted by the minimax value (Remarks 8 and 6). When the success set equals the whole class, the profile collapses to the class packing and adds nothing beyond minimax; when the question is about a specific procedure, the strictness certificate of Theorem 2 is the decisive content. As a practical rule, use the profile to rank fixed procedures on a shared task family, where minimax is silent about which procedure covers which subfamily; fall back on minimax when the question is only about the best possible procedure on a class.

A practitioner's diagnostic. For a practitioner training a broad model, a foundation model, a multi-task learner, or an in-context learner, the profile yields a concrete diagnostic: estimate the Monte Carlo breadth $\widehat{U}$ on a packed evaluation family (Remark 10), compare it with the ambient class packing $\log|\Theta_0|$, and read the gap as untapped coverage capacity. Concretely, construct the evaluation family as a held-out set of tasks (a packing of the task space, or all K-choose-2 binary subtasks of a K-class benchmark as in E4–E5), fix the success threshold ϵ from the downstream accuracy requirement, and sweep the budget to locate the knee of $\widehat{U}(\cdot)$. The knee is the point where coverage becomes information-feasible: budgets below it forfeit coverage, budgets far above it add little. The residual gap between $\widehat{U}$ at saturation and $\log|\Theta_0|$ then indicates either a capacity bottleneck (if raising capacity raises the saturation level, as in E5) or a genuine modeling gap. A sharp phase transition (Remark 12) is the actionable signal throughout: it means a modest increase in budget can unlock large coverage gains, so capacity should be invested where the transition is steep rather than where coverage has already saturated.

C Experimental Protocol Details

This appendix records the exact synthetic protocols behind Section 7 and Figure 1. All experiments are implemented in Python with NumPy; plots are generated with Matplotlib. Unless otherwise stated, each curve is averaged over 10 random seeds and the reported uncertainty is a normal-approximation 95% confidence interval computed from the seed-level summaries.

C.1 SHARED CONVENTIONS

For each experiment we report a primary metric and one auxiliary diagnostic. The primary metric is the quantity plotted in the paper. The auxiliary diagnostic is retained in the CSV outputs to check that the plotted trend is not hiding a pathological failure mode. All experiments are lightweight CPU simulations; the synthetic studies E1–E3 do not use external datasets, while the real-data checks E4–E5 use the `sklearn` digits dataset and E8 uses CIFAR-10, both standard small-scale public benchmarks. In the bandit and linear-bandit studies, the packed environment family is regenerated independently for each seed by the rejection-sampling routines described in the corresponding subsections below, so the reported confidence intervals over seeds include the variability of the packing construction itself; we do not average over a single fixed packing. Across the runs, different random packings moved the transition locations only within the reported bands, so the qualitative phase transitions are not an artifact of any one construction. Reference implementations of all experiments and figure-generation scripts are provided in the accompanying code.

C.2 E1: EXPERTS BREADTH SEPARATION

The ambient expert count is $N \in \{64, 128, 256\}$, the horizon is $T \in \{1000, 4000, 16000\}$, and the support size is

$$m \in \{4, 8, 16, 32, 64, N\} \cap [N].$$

For a fixed configuration (N, T, m), the first m experts form the prior support. Losses are Bernoulli with means initialized at $1/2$, then modified so that one in-support expert has mean $1/2 - \mathrm{gap}$ with $\mathrm{gap} = 0.1$. When $m < N$, we also place one better expert just outside the support with mean $1/2 - 2\,\mathrm{gap}$ to expose brittleness of narrow coverage. Hedge is run with a prior uniform on the support and zero outside, with learning rate

$$\eta = \sqrt{\frac{8 \log(\max\{m, 2\})}{T}}.$$

The primary metric is regret to the best expert inside the support; the auxiliary metric is regret to the best expert in the full class. The main-text figure uses the representative setting $N = 128$ and $T = 4000$.

C.3 E2: BERNOULLI BANDIT COVERAGE FRONTIER

The number of arms is $K \in \{8, 16, 32\}$, the reward gap is $\Delta \in \{0.05, 0.1, 0.2\}$, and the horizon is $T \in \{200, 500, 1000, 2000, 4000\}$. For each (K, Δ) we sample a packed environment family of size 64 by rejection sampling sign vectors in $\{-1, +1\}^K$ with minimum Hamming distance $\max\{2, K/4\}$. In environment θ, arm i has Bernoulli mean $1/2 + \theta_i \Delta$.

We compare three algorithms. 'uniform' samples each arm $\lfloor T/K \rfloor$ times and predicts the sign of the empirical mean shift. 'greedy_proxy' initializes each arm once and then repeatedly samples the arm whose empirical mean is currently closest to $1/2$. 'thompson' maintains $\mathrm{Beta}(1, 1)$ posteriors and each round samples a posterior mean, pulling the arm whose sampled mean is closest to $1/2$; it is evaluated on the representative setting $K = 16, \Delta = 0.1$ only. Success is declared when normalized Hamming error is at most $\epsilon = 0.1$. The primary metric is the empirical breadth proxy

$$\log(\#\{\text{successful packed environments}\}),$$

and the auxiliary metric is the mean normalized Hamming error over the packed family. The main-text figure uses the representative setting $K = 16$ and $\Delta = 0.1$, plotted against the scaled budget $T\Delta^2$.

C.4 E3: STRUCTURED LINEAR-BANDIT WITNESS

The dimension is $d \in \{8, 16, 32\}$, the target threshold is $\epsilon \in \{0.25, 0.35, 0.5\}$, and the horizon is $T \in \{100, 250, 500, 1000, 2000, 4000\}$. For each (d, ϵ) we build a packed sign family of size 48 by rejection sampling sign vectors in $\{-1, +1\}^d$ with minimum Hamming distance $\max\{2, d/4\}$, then scaling them to

$$\theta = \frac{2\epsilon}{\sqrt{d}} v.$$

Observations follow the local Gaussian linear model

$$Y_t = \langle X_t, \theta \rangle + \xi_t, \qquad \xi_t \sim \mathcal{N}(0, 1).$$

We compare three least-squares protocols. 'isotropic_ls' uses i.i.d. Gaussian design vectors normalized to unit norm. 'coordinate_ls' uses a balanced axis-aligned design that cycles approximately evenly through coordinates with random signs. 'adaptive_ls' is a two-stage design: isotropic exploration for the first $T/2$ rounds, then focused sampling along the least-squares direction for the remaining $T/2$ rounds. All three decode with ridge-stabilized least squares using ridge parameter 10^{-6}. The primary metric is the empirical breadth proxy $\log(\#\{\text{successful packed environments}\})$, where success means Euclidean error at most ϵ; the auxiliary metric is the mean Euclidean estimation error over the packed family. The main-text figure uses the representative setting $d = 16$ and $\epsilon = 0.35$, where the coverage transition is visible and the adaptive design does not beat isotropic exploration.

C.5 E4: REAL-DATA BREADTH CHECK

The real-data check uses the `sklearn` digits dataset (10 classes, 1797 samples), split $50/50$ stratified into train and test. Tasks are the 45 binary class-pair classification problems. The algorithm is a single multinomial logistic-regression model (no feature engineering beyond pixel normalization to $[0, 1]$). For each budget $n \in \{15, 30, 60, 120, 240, 480, 720, 898\}$ and each of 10 seeds, we draw n training samples uniformly at random, fit the model, and evaluate pair accuracy on the test set restricted to the two classes of each pair. A pair is solved when its accuracy is at least 90%. The primary metric is the breadth proxy $\log(\#\{\text{solved pairs}\})$; the auxiliary metric is the mean pair accuracy. The results are in `results/e4_digits.csv`.

C.6 E5: FROZEN-REPRESENTATION BREADTH CHECK

The frozen-representation check uses the same digits data and split as E4, with the same 45 class-pair tasks and $\geq 90\%$ threshold. The representation is *frozen*: either PCA fit once on the train split (deterministic), or random ReLU features $z(x) = \max(0, Wx + b)$ with W and b drawn once and kept fixed. On the frozen features of dimension m, a single multinomial logistic probe is fit on the full train set and evaluated as in E4. For the random-ReLU representation we average the breadth proxy over 10 independent representation draws. The primary metric is the breadth proxy $\log(\#\{\text{solved pairs}\})$; the results are in `results/e5_frozen_rep.csv`.

C.7 E8: FROZEN PRETRAINED-REPRESENTATION BREADTH

E8 is the large-scale analogue of E5, using a frozen ImageNet-pretrained ResNet18 (`torchvision`) rather than a small PCA/ReLU feature. The final classification layer is removed and the 512-dimensional average-pooled output is used as a fixed representation of CIFAR-10 images (resized to 128×128, normalized with ImageNet statistics). On the 45 binary class-pair tasks, a single multinomial logistic probe is trained on the frozen features for a budget of n samples per class ($n \in \{20, 40, 80, 160, 320, 640, 800\}$, 5 seeds), and a pair is solved when its accuracy is at least 75% (a lower threshold than E4–E5, reflecting the frozen feature's discriminability ceiling). The breadth proxy rises from ≈ 2.2 to ≈ 3.66 as the budget grows, saturating below the class packing $\log 45 \approx 3.81$ because a few hard class pairs stay below the frozen feature's ceiling; this is the same capacity-bounds-breadth phenomenon as E5, now for a pretrained model. The results are in `results/e8_pretrained.csv`.

C.8 E6: METRIC AND THRESHOLD SENSITIVITY

E6 checks two robustness properties of the E2 witness (uniform sampling, $K = 16$, $\Delta = 0.1$). First, *metric sensitivity*: we measure breadth under two bi-Lipschitz-equivalent success criteria, normalized Hamming error ≤ 0.1 on the recovered sign vector and mean-vector ℓ_2 error ≤ 0.12 on the estimated reward means, and verify that the two breadth curves share the same phase-transition structure (the knee sits at the same $T\Delta^2$ scale, up to the $O(1)$ shift predicted by the bi-Lipschitz equivalence). Second, *threshold sensitivity*: we sweep the Hamming threshold $\epsilon_H \in \{0.05, 0.1, 0.2\}$ and observe the knee move right as ϵ_H shrinks, matching the frontier scaling $T \gtrsim B/\Delta^2$. The results are in `results/e6_metric_sensitivity.csv` and `results/e6_threshold_sensitivity.csv`.

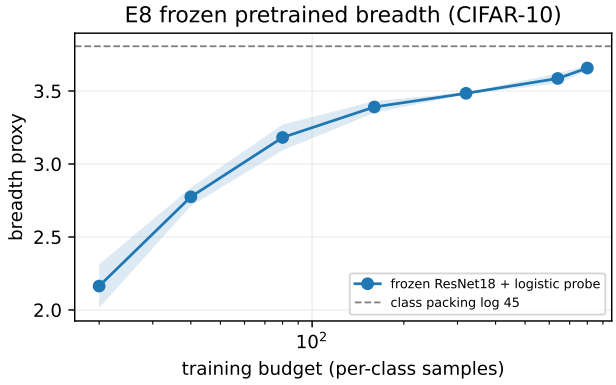


Figure 6: Frozen pretrained-representation breadth (CIFAR-10, 5 seeds). A single logistic probe on the frozen ImageNet-ResNet18 features covers an increasing number of the 45 class-pair tasks ($\geq 75\%$ accuracy) as the budget grows, saturating below the class packing $\log 45 \approx 3.81$ because a few hard pairs stay below the frozen feature's ceiling.

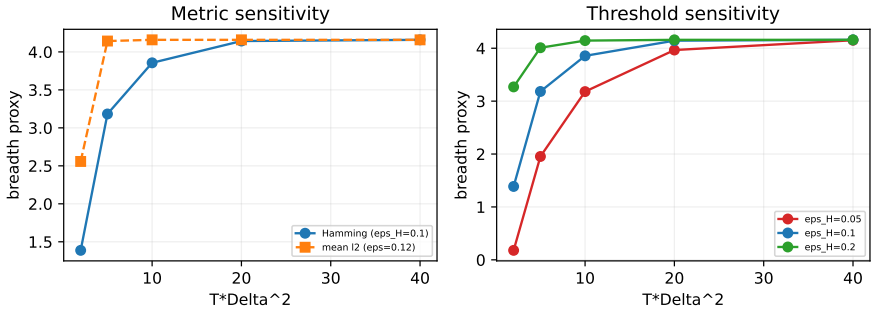


Figure 7: Robustness of the E2 witness ($K = 16$, $\Delta = 0.1$, uniform sampling). Left: breadth under normalized Hamming error ($\epsilon_H = 0.1$) vs. mean-vector ℓ_2 error ($\epsilon = 0.12$); the two bi-Lipschitz-equivalent metrics share the same phase-transition knee. Right: breadth under Hamming thresholds $\epsilon_H \in \{0.05, 0.1, 0.2\}$; the knee moves to larger $T\Delta^2$ as the threshold shrinks.

C.9 Figure-generation choices

The paper-facing figures are generated from the CSV summaries by selecting one representative setting per experiment: $(N, T) = (128, 4000)$ for E1, $(K, \Delta) = (16, 0.1)$ for E2, and $(d, \epsilon) = (16, 0.35)$ for E3. This keeps the main text focused on one clean witness per experiment; the main text carries the three-panel synthetic figure (Figure 1) and the real-data capacity curve of E8 (Figure 2), while the phase-transition heatmap of E3 (Figure 5), the digits budget curve of E4 (Figure 3), the frozen-representation capacity curve of E5 (Figure 4), and the metric- and threshold-sensitivity curves of E6 (Figure 7) appear in this appendix. The full parameter sweeps for each experiment remain in the CSV outputs `results/e1_experts.csv`, `results/e2_bandit.csv`, `results/e3_linear_bandit.csv`, `results/e4_digits.csv`, `results/e5_frozen_rep.csv`, `results/e6_metric_sensitivity.csv`, `results/e6_threshold_sensitivity.csv`, and `results/e8_pretrained.csv`, respectively.

D Computational Universality

This appendix formalizes the computational extension sketched in Remark 13: the compute budget becomes an explicit third resource alongside samples and horizon.

Definition 4 (Computational universality profile)**.** *Let $\mathcal{A}$ be an algorithm and let $\tau \geq 0$ bound the number of elementary operations $\mathcal{A}$ may use. For $\epsilon > 0$ and budget n, define the ϵ-success set under the compute constraint as the set of tasks θ on which some τ-compute-bounded run of $\mathcal{A}$ is ϵ-successful (Definition 1), and let $U_{\mathcal{A}}(\epsilon, n, \tau)$ be the packing complexity of this set. We write $U_{\mathcal{A}}(\epsilon, n, \infty)$ for the unconstrained profile.*

Proposition 7 (Basic properties of the computational profile)**.** *Let $\mathsf{comp}(\mathcal{A}, n, \epsilon)$ be the compute $\mathcal{A}$ actually uses on n samples at accuracy ϵ.*

1. Monotonicity. *The map $\tau \mapsto U_{\mathcal{A}}(\epsilon, n, \tau)$ is nondecreasing, and $U_{\mathcal{A}}(\epsilon, n, \tau) \leq U_{\mathcal{A}}(\epsilon, n, \infty)$.*
2. Saturation. *If $\tau \geq \mathsf{comp}(\mathcal{A}, n, \epsilon)$, then $U_{\mathcal{A}}(\epsilon, n, \tau) = U_{\mathcal{A}}(\epsilon, n, \infty)$, so the information-theoretic frontier of Theorem 1 applies unchanged once the compute bound is not binding.*
3. Feature vignette. *For the m-feature estimator $\mathcal{A}_m$ of Proposition 6, a run on n samples uses $\tau \asymp mn$ operations and attains $U_{\mathcal{A}_m}(\epsilon, n, \tau) \leq \Theta(m \log(R/\epsilon))$. Since $m = \Theta(\tau/n)$, breadth grows linearly in the compute-per-sample budget, while estimating the m coordinates to accuracy ϵ requires $n \gtrsim m^2/\epsilon^2$ samples, coupling the two budgets through the full-parameter rate.*

Discussion. Property (i) is immediate from the definition. Property (ii) holds because a non-binding compute constraint leaves the transcript unchanged, so Theorem 1 bounds the profile exactly as in the unconstrained case. The vignette shows that compute and samples interact nontrivially: breadth is limited both by capacity (packing entropy $m \log(R/\epsilon)$) and by the horizon needed to estimate m coordinates to accuracy ϵ (the rate m^2/ϵ^2), so sustaining a coverage level requires trading the two budgets. Making this trade-off fully quantitative requires a concrete model of computation and remains open.

E Additional Discussion: Universality and Model Selection

One generic way to approach the frontier is to maintain a mixture over a hierarchy of candidate models/priors and to perform model selection. In online learning this corresponds to expert aggregation; in statistical learning it corresponds to structural risk minimization. These procedures are known to incur a logarithmic penalty in the size of the candidate family, matching the universality dependence in many classical settings.