AGNOSTIC SMOOTHED ONLINE REGRESSION WITH ADVERSARIAL RESPONSES

Anonymous authors
Paper under double-blind review

ABSTRACT

We study smoothed online prediction with bounded adversarial responses. This widely studied framework bridges i.i.d. sampling and adversarial covariate selection through a smoothness parameter $\mathsf{C}_{\mathsf{cov}}$, which bounds conditional covariate densities relative to a fixed, unknown base measure. We propose HEDGE-COVER, an information-theoretic algorithm that achieves sublinear regret $\widetilde{O}(\sqrt{\mathrm{Pdim}(\mathcal{F})\mathsf{C}_{\mathsf{cov}}T})$ for function classes with bounded pseudo-dimension. The algorithm aggregates a carefully constructed family of experts using HEDGE, with a prior that links regret to the number of disagreements between a consistent selector and a target function. We bound this number by exploiting covariate smoothness. This answers an open problem posed in Blanchard (2025) on the minimax optimal adaptive regret of the smoothed online regression problem. We establish a matching lower bound for the class of linear predictors. The main intricacy of the lower bound lies in explicitly constructing a challenging sequential covariate distribution supported on mutually orthogonal hyperplanes. This construction may be of independent technical interest. Finally, we revisit the well-specified setting and quantify the effect of response noise. For conditionally ν^2-subGaussian responses, we extend the existing lower bound under realizable responses by showing that the minimax expected regret is $\Omega((1 \vee \nu)\sqrt{(\mathsf{C}_{\mathsf{cov}} - 1)dT})$ for a function class of VC dimension d. A corresponding upper bound for ERM matches this dependence on ν.

1 INTRODUCTION

Online learning is a framework for sequential decision-making in which a learner iteratively makes predictions, observes feedback, and updates its strategy. Classical PAC learning (Vapnik & Chervonenkis, 1971; 1974) assumes i.i.d. samples, whereas online learning allows sequential, potentially adversarial covariates and responses. For binary classification under 0-1 loss, a class is online learnable if and only if it has finite Littlestone dimension (Littlestone, 1988; Shalev-Shwartz & Ben-David, 2014; Alon et al., 2021). This requirement is substantially stronger than finiteness of the VC dimension, which characterizes distribution-free PAC learnability for binary classification (Vapnik & Chervonenkis, 1971; 1974; Devroye et al., 1996; Mathiasen et al., 2026).

To accommodate dependence within the covariate sequence and interpolate between the i.i.d. and fully adversarial regimes, Rakhlin et al. (2011) first introduced the notion of a smoothed adversary in online learning. Subsequent work has developed and analyzed algorithms under several variants of smoothed online learning; see Block & Simchowitz (2022); Block et al. (2023); Haghtalab et al. (2022); Chen et al. (2026); Haghtalab et al. (2020; 2024); Raman et al. (2024); Wu et al. (2024); Blanchard (2025). In the recent literature, most work considers the case in which the base measure μ is known (or equivalently, uniform over its support) and the learner is able to sample from μ. See Section 1.1 for a detailed comparison with these results.

In the present work, we study the online prediction problem in which no assumptions except boundedness are made on the response values $\{y_t\}_{1 \le t \le T}$ selected by the adversary, and the base measure μ is only existential and is *not* announced to the learner a priori. We refer to the setting as the *agnostic* smoothed online learning setting. Arguably, removing knowledge of the base measure weakens the learner's information and generally makes the problem more realistic but also more challenging. In online regression with squared loss, at each round $t = 1, 2, \cdots, T$, the forecaster

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.

observes $x_t \in \mathbb{R}^d$, outputs a prediction $\widehat{y}_t \in \mathbb{R}$ and receives $y_t \in [-1, 1]$ from nature. The goal of the learner is to achieve low squared loss compared to a fixed function class baseline $\mathcal{F}$. Formally, the regret up to T is defined as

$$\mathsf{Regret}(T, \mathcal{F}) := \sum_{1 \le t \le T} (\widehat{y}_t - y_t)^2 - \inf_{f \in \mathcal{F}} \sum_{1 \le t \le T} (f(x_t) - y_t)^2. \tag{1.1}$$

We highlight that Blanchard (2025) studies a similar agnostic smoothed online prediction problem with arbitrarily adversarial responses and analyzes the *oblivious regret*

$$\mathsf{Regret}^{\mathsf{obl}}(T, \mathcal{F}) = \sum_{1 \le t \le T} \mathbb{E}(\widehat{y}_t - y_t)^2 - \inf_{f \in \mathcal{F}} \sum_{1 \le t \le T} \mathbb{E}(f(x_t) - y_t)^2 \le \mathbb{E}\,\mathsf{Regret}(T, \mathcal{F}). \tag{1.2}$$

Unless otherwise specified, we consider the regret defined in equation 1.1, which is sometimes called adaptive regret.

We next formalize a smoothness condition that caps how concentrated each X_t can be relative to a fixed base measure μ. An environment Env (i.e. the joint law of $(x_1, y_1, \cdots, x_T, y_T)$) is $\mathsf{C}_{\mathsf{cov}}$-smooth if there is a probability measure μ on $\mathbb{R}^d$ such that the conditional law of x_t, given $(x_1, \ldots, x_{t-1})$, has a Radon–Nikodym derivative with respect to μ at most $\mathsf{C}_{\mathsf{cov}}$.

Definition 1. There exists a probability measure μ on $\mathcal{X} \subseteq \mathbb{R}^d$ and a parameter $\mathsf{C}_{\mathsf{cov}} \ge 1$ such that for every round t, given the partial history $\mathcal{H}^x_{t-1} = \sigma(\{x_1, \ldots, x_{t-1}\})$, the conditional law of x_t, denoted by $P_t(\cdot \mid \mathcal{H}^x_{t-1})$, satisfies

$$\frac{\mathrm{d}P_t(\cdot \mid \mathcal{H}^x_{t-1})}{\mathrm{d}\mu}(x) \le \mathsf{C}_{\mathsf{cov}} \qquad \textit{for } \mu\textit{-a.e. } x \in \mathbb{R}^d.$$

Some works in the literature adopt the reciprocal parametrization $\sigma = \mathsf{C}_{\mathsf{cov}}^{-1}$ of the smoothness condition, while we stick to $\mathsf{C}_{\mathsf{cov}}$ in the present work. To quantify the benefit of covariate smoothness in online learning, we study the minimax optimal expected adaptive-benchmark regret. For a function class $\mathcal{F}$, we consider the following minimax regret

$$\mathfrak{R}(T, \mathcal{F}) = \mathfrak{R}(T, d, \mathsf{C}_{\mathsf{cov}}, \mathcal{F}) = \inf_{\pi} \sup_{\substack{\mathsf{Env}\text{ is} \\ \mathsf{C}_{\mathsf{cov}}-\text{smooth}}} \mathbb{E}_{\pi, \mathsf{Env}}\big(\mathsf{Regret}(T, \mathcal{F})\big),$$

where π ranges over all (possibly randomized) learners. We advance our understanding of the quantitative behavior of agnostic smoothed online regression by addressing the following question.

Question 1. In the adversarial setting, can one achieve sublinear (adaptive) regret for smoothed online learning without prior knowledge of the base measure? If so, what is the minimax-optimal regret rate?

We also revisit the well-specified setting, a special case of agnostic smoothed online learning in which the responses satisfy $\mathbb{E}[y_t \mid x_t] = f^\star(x_t)$ for some $f^\star \in \mathcal{F}$. In this setting, empirical risk minimization (ERM) is a natural procedure that does not require knowledge of the base measure. Block et al. (2024) establish an upper bound on the expected cumulative estimation error of ERM in terms of the Wills functional (Wills, 1973; Hadwiger, 1975; Vitale, 1996; Mourtada, 2025), while Blanchard (2025) proves a matching worst-case lower bound for VC classes under realizable responses. Taken together, these results show that, when $\mathsf{C}_{\mathsf{cov}}$ is bounded away from 1, ERM is minimax-optimal in its dependence on $\mathsf{C}_{\mathsf{cov}}$, d, and T, up to logarithmic factors. However, they do not provide a sharp characterization of the dependence on the response-noise level.

We now summarize our main contributions in this paper.

1. In the adversarial response setting, we propose HEDGE-COVER, an algorithm that requires no knowledge of the base measure and achieves sublinear regret $\widetilde{O}(\sqrt{\mathrm{Pdim}(\mathcal{F})\mathsf{C}_{\mathsf{cov}}T})$, as established in Theorem 2.1. Blanchard (2025) proposes R-COVER, which achieves an oblivious regret bound of $O(\sqrt{\mathsf{C}_{\mathsf{cov}}dT})$ for regression (see equation 1.2), but leave open whether sublinear adaptive regret is attainable in this setting. Our result resolves this open question by providing the first sublinear adaptive regret guarantee for regression with adversarial responses and an unknown base measure.

2. We complement the upper bound for HEDGE-COVER by constructing a challenging smoothed environment that yields the lower bound $\mathfrak{R}(T) = \Omega(T \wedge \sqrt{d(\mathsf{C}_{\mathsf{cov}} - 1)T})$ for $d \geq 2$ and $\mathsf{C}_{\mathsf{cov}} > 1$. The proof develops a systematic method for constructing hard environments, which may be of independent theoretical interest.
3. In the well-specified setting, we quantify the effect of response noise. For conditionally ν^2-subGaussian responses, we prove that there exists a class of VC dimension d for which every learner incurs expected regret $\Omega\big(T \wedge \sqrt{(\mathsf{C}_{\mathsf{cov}} - 1)(1 \vee \nu^2)dT}\big)$. The bound recovers the known rate in Blanchard (2025) under realizable ($\nu = 0$) responses and reveals a missing noise-dependent term in the ERM analysis of Block et al. (2024). A minor modification of their argument yields an upper bound with matching dependence on ν. Finally, we show that the Wills-functional upper bound needs not be class-wise tight.

1.1 RELATED WORK

Online prediction with a smooth environment. We review the closest results in smoothed online learning in Table 1 and list the work on online regression with a known base measure as a comparison. In the online classification setting, Blanchard (2025) introduces R-COVER, a recursive-covering

Table 1: Upper and lower regret bounds for regression under $\mathsf{C}_{\mathsf{cov}}$-smooth covariates.

	Adversarial responses	**Well-specified responses**
Upper bound	**Unknown μ, oblivious.** R-COVER: $O\big(\log^3 T\sqrt{\mathsf{C}_{\mathsf{cov}}dT}\big)$, for $\mathsf{fat}_{\mathcal{F}}(r) \leq d\log(1/r)$. (Blanchard, 2025, Theorem 4)	
	Unknown μ, adaptive. HEDGE-COVER: $\widetilde{O}\big(\sqrt{\mathrm{Pdim}(\mathcal{F})\mathsf{C}_{\mathsf{cov}}T}\big)$. Theorem 2.1	**Unknown μ, estimation error.** ERM: $\widetilde{O}\big(\sqrt{\mathsf{C}_{\mathsf{cov}}dT}\big)$. (Block et al., 2024, Theorem 8)
	Known μ, adaptive. $O\big(\log^{3/2} T\log(\mathsf{C}_{\mathsf{cov}}T)\sqrt{dT}\big)$. (Block et al., 2022, Theorem 3)	
Lower bound	**Unknown μ, adaptive.** $\Omega\Big(\sqrt{(\mathsf{C}_{\mathsf{cov}} - 1)dT} \wedge T\Big)$ for all algorithms. Theorem 3.1	**Unknown μ, adaptive.** $\Omega\Big(\sqrt{(\mathsf{C}_{\mathsf{cov}} - 1)dT} \wedge T\Big)$ for all algorithms. Realizable responses: (Blanchard, 2025, Theorem 2) Noise-dependent extension: Theorem 3.3

algorithm that provides the first sublinear regret guarantee for agnostic smoothed online learning without prior knowledge of μ. Blanchard (2025) also establishes a minimax lower bound under realizable responses for a function class of VC dimension d. We establish a lower bound of the same order for the linear function class when $d \geq 2$. For this class, we further characterize a sharp phase transition: the minimax regret scales as $\sqrt{T}$ when $d \geq 2$ and $\mathsf{C}_{\mathsf{cov}} > 1$, but is logarithmic in T when $d = 1$ or $\mathsf{C}_{\mathsf{cov}} = 1$.

Online linear regression. The central ingredient in our lower bound for agnostic regression (Theorem 3.1) over a function class satisfying $\mathsf{VC}(\mathcal{F}) = d$ is an explicit construction of an unknown discrete base measure μ. The construction applies even to linear predictors. Online linear regression has a long history (Vovk, 1997; Azoury & Warmuth, 2001; Cesa-Bianchi et al., 2006). More recently, by exploiting a connection with self-normalized martingales, Chen et al. (2026) establish a regret upper bound of $O(\sqrt{d\mathsf{C}_{\mathsf{cov}}T\log T})$. As a minor technical contribution, we sharpen this bound to

$O(\sqrt{d\mathsf{C}_{\mathsf{cov}}T})$ by using an argument that avoids direct decoupling, which matches the worst regret even for the logarithmic term.

Hybrid online learning. Hybrid online learning (Lazaric & Munos, 2009) studies the setting in which covariates are sampled i.i.d. from an unknown distribution, while the responses may be chosen adversarially. Note that this corresponds to the special case of smoothed online learning with $\mathsf{C}_{\mathsf{cov}} = 1$. For an unrestricted response adversary, the best-known results leave a gap between statistical and computational guarantees. Even in this special case, it remains open whether statistical optimality and computational efficiency can be achieved simultaneously, suggesting a possible statistical–computational trade-off in smoothed online learning. Specifically, statistically optimal algorithms are typically computationally intractable for natural hypothesis classes (Lazaric & Munos, 2009; Wu et al., 2022), whereas known ERM-oracle-efficient algorithms attain statistically suboptimal regret rates (Wu et al., 2024). Under the additional structural assumption that, at each round, the adversary selects a possibly different labeling function from a fixed and known class, Okoroafor et al. (2026) develop an oracle-efficient learning algorithm whose regret is near-optimal up to logarithmic factors and the dependence on the adversary's labeling class.

1.2 Notations

Let $\mathbb{N}_+ = \{1, 2, \ldots\}$ and $[n] = \{1, \ldots, n\}$ for $n \in \mathbb{N}_+$. For a finite set A, let $|A|$ denote its cardinality, and let $\mathbb{1}_{\mathsf{E}}$ denote the indicator of an event E. For sets $A, B \subseteq \mathcal{X}$, their symmetric difference is denoted by $A \Delta B$. Uppercase letters denote random variables and the corresponding lowercase letters denote their realizations. Let $\mathcal{H}_t^x := \sigma(X_1, \ldots, X_t)$ denote the covariate history and let $\mathcal{H}_t := \sigma(X_s, \widehat{Y}_s, Y_s : 1 \le s \le t)$ denote the completed interaction history. We abbreviate $P_t(\cdot) := \mathcal{L}(X_t \mid \mathcal{H}_{t-1}^x)$ and $P_t(A) := \mathbb{P}(X_t \in A \mid \mathcal{H}_{t-1}^x)$. For probability measures P and μ, $P \ll \mu$ denotes absolute continuity, $\mathrm{d}P/\mathrm{d}\mu$ denotes the Radon–Nikodym derivative, and $\mu^{\otimes k}$ denotes the k-fold product measure. Finally, we use standard asymptotic notation: $a \lesssim b$ and $a \gtrsim b$ mean that the corresponding inequality holds up to a universal positive constant, while $a \asymp b$ means that both relations hold. The notation $\widetilde{O}(\cdot)$ suppresses polylogarithmic factors in T.

2 Hedge-Cover algorithm

Before formally introducing the algorithm Hedge-Cover, we introduce the necessary shorthand notation and define a selection procedure similar to the realizable device in Blanchard (2025) as follows. The learner first discretizes the range $[-1, 1]$ with a grid $\mathcal{N}_m = \{-1 + 2m^{-1}j\}$ where $j \in [m]$. With a slight abuse of notation, we use $\mathcal{N}_m \circ \mathcal{F} = \{\mathsf{proj}(f(x), \mathcal{N}_m) : f \in \mathcal{F}\}$. For a trajectory $\mathsf{D} = \{(x_i, \widetilde{y}_i)\}_{1 \le i \le n}$, where $x_i \in \mathcal{X}$ and $\widetilde{y}_i \in \mathcal{N}_m$, a selector $\mathsf{Sel} : (\mathcal{X} \times \mathcal{N}_m)^\infty \to \mathcal{N}_m \circ \mathcal{F}$ maps to a (non-zero) function in $\mathcal{F}$ if D is realizable.

Definition 2 (Selector)**.** Fix a well-ordering of $\mathcal{N}_m \circ \mathcal{F}$. Given a trajectory $\mathsf{D} = \{(x_i, \widetilde{y}_i)\}_{i=1}^n$, where $x_i \in \mathcal{X}$ and $\widetilde{y}_i \in \mathcal{N}_m$, define $\mathsf{Sel}(\mathsf{D})$ to be the first function $f_\star \in \mathcal{N}_m \circ \mathcal{F}$ in this ordering satisfying $f_\star(x_i) = \widetilde{y}_i$ for every $i \in [n]$. If no such function exists, set $\mathsf{Sel}(\mathsf{D}) \equiv 0$, the zero function.

To apply the Hedge algorithm (Cesa-Bianchi et al., 2006), we introduce hypotheses of the form $\boldsymbol{h} = (\mathcal{T}_{\boldsymbol{h}}, \mathsf{map}_{\boldsymbol{h}})$, where $\mathcal{T}_{\boldsymbol{h}} \subseteq [T]$ and $\mathsf{map}_{\boldsymbol{h}} : \mathcal{T}_{\boldsymbol{h}} \to \mathcal{N}_m$. Each hypothesis specifies a set of times $\mathcal{T}_{\boldsymbol{h}}$ at which predictions are prescribed and a mapping $\mathsf{map}_{\boldsymbol{h}}$ that assigns a value in $\mathcal{N}_m$ to each such time. At the remaining times, predictions are determined by the selector. Both the selector Sel and the hypotheses implicitly depend on the discretization $\mathcal{N}_m$. Let $\mathcal{H}$ denote the set of all such hypotheses. Since $|\mathcal{N}_m| = m$, a direct calculation gives $|\mathcal{H}| = \sum_{k=0}^{T} \binom{T}{k} m^k = (m+1)^T$. Before describing Hedge-Cover, we emphasize that it is an information-theoretic algorithm; its computational aspects are beyond the scope of this work.

Hedge-Cover proceeds in the following steps for each round.

- We initialize a prior weight $p_{0,\boldsymbol{h}}$ for each hypothesis $\boldsymbol{h} \in \mathcal{H}$ and set its cumulative loss $L_{0,\boldsymbol{h}}$ to zero. Each hypothesis also maintains its own trajectory, initialized as $\mathsf{D}_{0,\boldsymbol{h}} = \emptyset$. The prior is chosen so that hypotheses with fewer prescribed prediction times incur a smaller complexity penalty in the regret bound.

Algorithm 1: HEDGE-COVER

Input: Time horizon T, function class $\mathcal{F}$

Set $m \leftarrow 2T$;

// HEDGE initialization

for $\boldsymbol{h} \in \mathcal{H}$ **do**

  $p_{0,\boldsymbol{h}} \propto \dfrac{1}{(|\mathcal{T}_{\boldsymbol{h}}|+1)^2(m+1)^{|\mathcal{T}_{\boldsymbol{h}}|}\binom{T}{|\mathcal{T}_{\boldsymbol{h}}|}}, \quad \sum_{\boldsymbol{h}\in\mathcal{H}} p_{0,\boldsymbol{h}} = 1$

  $L_{0,\boldsymbol{h}} \leftarrow 0$, $\mathsf{D}_{0,\boldsymbol{h}} \leftarrow \emptyset$;

for $t \in [T]$ **do**

  Observe x_t;

  for $\boldsymbol{h} \in \mathcal{H}$ **do**

    $\widehat{y}_{t,\boldsymbol{h}} = \begin{cases} \mathsf{map}_{\boldsymbol{h}}(t) & \text{if } t \in \mathcal{T}_{\boldsymbol{h}} \\ \left[\mathsf{Sel}(\mathsf{D}_{t-1,\boldsymbol{h}})\right](x_t) & \text{if } t \notin \mathcal{T}_{\boldsymbol{h}} \end{cases}$

  $\widehat{y}_t \leftarrow$ HEDGE-PREDICTION$(\{\widehat{y}_{t,\boldsymbol{h}}\})$;

  Observe y_t;

  $\{L_{t,\boldsymbol{h}}\}_{h\in\mathcal{H}} \leftarrow$ HEDGE-UPDATE(y_t);

  for $\boldsymbol{h} \in \mathcal{H}$ **do**

    $\mathsf{D}_{t,\boldsymbol{h}} \leftarrow \mathsf{D}_{t-1,\boldsymbol{h}} \bigcup \{(\boldsymbol{x}_t, \widehat{y}_{t,\boldsymbol{h}})\}$;

return $\{\widehat{y}_t\}_{1\le t\le T}$

- At round t, after observing x_t, each hypothesis $\boldsymbol{h}$ predicts $\mathsf{map}_{\boldsymbol{h}}(t)$ if $t \in \mathcal{T}_{\boldsymbol{h}}$; otherwise, it applies the selector Sel to its trajectory $\mathsf{D}_{t-1,\boldsymbol{h}}$ and evaluates the selected function at x_t. Thus, each hypothesis follows a consistent selector, with prescribed predictions at the times in $\mathcal{T}_{\boldsymbol{h}}$.
- The learner aggregates these predictions using HEDGE-PREDICTION, assigning weights proportional to $p_{0,\boldsymbol{h}}\exp(-\eta L_{t-1,\boldsymbol{h}})$. After observing y_t, it adds $(\widehat{y}_{t,\boldsymbol{h}} - y_t)^2$ to each hypothesis's cumulative loss and appends $(x_t, \widehat{y}_{t,\boldsymbol{h}})$ to that hypothesis's trajectory.

The role and the reason for our choice of the initialization $\{p_{0,\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathcal{H}}$ will be explained shortly after the proof sketch in Remark 2.

Definition 3 (pseudo dimension)**.** For a function class $\mathcal{F}$ consisting of functions $\mathcal{X} \to [-1, 1]$, its pseudo dimension is defined as

$$\mathrm{PDim}(\mathcal{F}) = \mathsf{VC}(\{(x, r) : r \le f(x), f \in \mathcal{F}\}).$$

Pseudo dimension extends VC dimension to real-valued function classes by allowing a distinct real threshold at each shattered point (Pollard, 2012). We provide an upper bound on the expected regret as follows. The formal high probability version of the theorem is stated in Theorem B.5 in Appendix B.

Theorem 2.1. *There exists an algorithm such that, for each function class $\mathcal{F}$ with* $\mathrm{Pdim}(\mathcal{F}) \le d$ *and each* $\mathsf{C}_{\mathsf{cov}}$*-smooth environment* Env*,*

$$\mathbb{E}\big[\mathsf{Regret}(T, \mathcal{F})\big] = \widetilde{O}\big(\sqrt{\mathsf{C}_{\mathsf{cov}} dT} \wedge T\big). \tag{2.1}$$

Remark 1. We measure the complexity of $\mathcal{F}$ by its pseudo dimension, as in Haghtalab et al. (2022). This assumption is stronger than the scale-dependent fat-shattering conditions considered by Block et al. (2022) and Blanchard (2025): finite pseudo dimension bounds the fat-shattering dimension uniformly over all positive scales, whereas the latter conditions permit it to diverge as the scale vanishes. Our analysis uses this uniform control to bound the growth of discretized hypotheses and their disagreement sets. Accordingly, our adaptive regret guarantee addresses the unknown base-measure problem for classes of finite pseudo dimension. Extending it to the broader classes allowed by fat-shattering growth conditions remains open.

Proof Sketch of Theorem 2.1

Step 1: Covering. For sufficiently large T, we fix a finite cover of the discretized function class $\mathcal{M}_1 \subseteq \mathcal{N}_m \circ \mathcal{F}$ such that, for each $f \in \mathcal{N}_m \circ \mathcal{F}$, there exists $g \in \mathcal{M}_1$ such that

$$\mu\{x : f(x) \ne g(x)\} \le \varepsilon \tag{2.2}$$

with $\log|\mathcal{M}_1| = O(d\log(T\varepsilon^{-1}))$. This cover is used only in the analysis.

Step 2: From a function to a hypothesis $\boldsymbol{h}$. Given the preceding class $\mathcal{M}_1$, for any $f_j \in \mathcal{M}_1$, we consider the following set:

$$\mathsf{B}(j,t) := \big\{x \in \mathcal{X} : \mathsf{Sel}(\{x_i, f_j(x_i)\}_{1\le j\le t-1})(x) \neq f_j(x)\big\}.$$

As $\mathcal{M}_1 \subseteq \mathcal{N}_m \circ \mathcal{F}$, the selector Sel cannot be trivially set to 0, as f_j certainly satisfies the definition. Though there is apparently no guarantee that the selector can be the same as f_j, the observed covariates $\{x_1, \cdots, x_{t-1}\}$ cannot belong to $\mathsf{B}(j,t)$. From the set, one can construct a good enough hypothesis $\boldsymbol{h} = (\mathcal{T}_{\boldsymbol{h}} := \{t : x_t \in \mathsf{B}(j,t)\}, \mathsf{map}_{\boldsymbol{h}}(t) = f_j(x_t))$. Such a construction ensures that $\widehat{y}_{t,\boldsymbol{h}} = f_j(x_t)$ for every t.

Step 3: Hedge bound. For a sufficiently small η, HEDGE guarantees that, for any hypothesis $\boldsymbol{h}$, we can relate the cumulative squared loss induced by Algorithm 1 to that of the predictions by $\boldsymbol{h}$:

$$\begin{aligned}\sum_{1\le t\le T}(\widehat{y}_t - y_t)^2 - \sum_{1\le t\le T}(f_j(x_t) - y_t)^2 &= \sum_{1\le t\le T}(\widehat{y}_t - y_t)^2 - \sum_{1\le t\le T}(\widehat{y}_{t,\boldsymbol{h}} - y_t)^2 \\ &\le -\frac{1}{\eta}\log(p_{0,\boldsymbol{h}}) \lesssim \eta^{-1}(1 + |\mathcal{T}_{\boldsymbol{h}}|)\log(T),\end{aligned} \tag{2.3}$$

as $m \asymp T$. See Lemma C.1 for the details.

Step 4: Upper bound on $|\mathcal{T}_{\boldsymbol{h}}|$. From Steps 1–3, it suffices to show that, under the smooth environment, the cardinality of $\mathcal{T}_{\boldsymbol{h}} = \sum_{1\le t\le T}\mathbb{1}\{x_t \in \mathsf{B}(j,t)\}$ is of order $\widetilde{O}(\sqrt{d\mathsf{C}_{\mathsf{cov}}T})$. This is arguably the most technically challenging step, where we use the decoupling techniques and martingale arguments. In order to strengthen the result to its high-probability version, we also invoke a Vapnik-Chervonenkis inequality. This step is formalized in Proposition 2.4.

Step 5: Putting everything together. To relate the left-hand side of equation 2.3 to regret, it remains to consider (1) the discretization error and (2) the bad event in equation 2.2. The former is at most $O(1)$ as $m \asymp T$, and the latter can be controlled by carefully choosing ε in terms of T. Precisely, we choose $\varepsilon \asymp (\mathsf{C}_{\mathsf{cov}}T\log T)^{-1}$ to accommodate the extra $T\log T$ terms that occur in the decoupling step.

□

Remark 2 (Initialization $p_{0,\boldsymbol{h}}$)**.** The initialization of the distribution over $\mathcal{H}$ is designed to make the aggregation penalty depend on the number of corrections required by a hypothesis. As shown in Step 3 of the proof sketch, the regret is controlled, modulo approximation and discretization errors, by the negative log initialized weight $\log(1/p_{0,\boldsymbol{h}})$ of a suitable hypothesis $\boldsymbol{h}$. For each $\tau \in \{0,\dots,T\}$, there are $\binom{T}{\tau}m^\tau$ hypotheses satisfying $|\mathcal{T}_{\boldsymbol{h}}| = \tau$. The factor $\binom{T}{\tau}(m+1)^\tau$ in the denominator accounts for this multiplicity, while the additional factor $(\tau+1)^2$ keeps the normalizing constant bounded uniformly in T. Consequently, when $m \asymp T$, we have $\log(1/p_{0,\boldsymbol{h}}) = O(1 + \tau\log(T))$. For the hypothesis constructed to track $f_j \in \mathcal{M}_1$, τ equals the number of rounds on which the selector disagrees with f_j. Thus, a sufficiently tight bound on this count τ yields the desired aggregation regret bound.

2.1 DECOUPLING

As in much of the smoothed online learning literature, a key step in our analysis relies on the following decoupling technique, first introduced by Haghtalab et al. (2024, Theorem 2.1).

Lemma 2.2 (Theorem 2.1 (Haghtalab et al., 2024))**.** *Let $\mathcal{X}$ be a Borel subset of Euclidean space, and let $P_t = \mathcal{L}(X_t \mid \mathcal{H}^x_{t-1})$, $1 \le t \le T$, be $\mathsf{C}_{\mathsf{cov}}$-smooth with respect to a fixed probability measure μ. For each $k \in \mathbb{N}_+$ there is a coupling Π of $(X_t, Z_1^{(t)}, \dots, Z_k^{(t)})_{1\le t\le T}$ such that:*

(a.) *The covariates $(X_1,\dots,X_T)$ have the prescribed sequential law.*

(b.) *All the variables $Z_i^{(t)}$, $i \in [k]$, $t \in [T]$, are i.i.d. with law μ.*

(c.) *Conditional on $X_1,\dots,X_{t-1}$, the variables $\{Z_i^{(s)} : s \ge t,\ i \in [k]\}$ are i.i.d. with law μ.*

(d.) *For each $t \le T$, with probability at least $1 - t(1 - \mathsf{C}_{\mathsf{cov}}^{-1})^k$, simultaneously for all $s \le t$,*

$$X_s \in \{Z_1^{(s)}, \dots, Z_k^{(s)}\}.$$

We will also need the following non-asymptotic inequality for classes of functions whose complexity is described through subgraphs. There are multiple ways to prove results stated in Claim B.4. Here we detail the approach via the classical Vapnik-Chervonenkis inequality (Anthony & Shawe-Taylor, 1993). A left-tail bound is proved and the constants have recently been improved in Portier (2026). We only need the inequality in equation 2.4.

Lemma 2.3 (Theorem 2.1 (Anthony & Shawe-Taylor, 1993); Theorem 1 (Portier, 2026))**.** *Let $Z_1, \cdots, Z_n \overset{i.i.d.}{\sim} P$ and $\mathcal{C}$ be a countable class of measurable subsets of $\mathcal{X}$. We have, for all $\zeta > 0$,*

$$\mathbb{P}\Big(\sup_{C \in \mathcal{C}} \frac{P_n(C) - P(C)}{\sqrt{P_n(C)}} > \zeta \Big) \le \Pi_{\mathcal{C}}(2n) \exp\big(- \mathsf{c} n \zeta^2 \big). \tag{2.4}$$

Now we are ready to prove a result holding for a general countable set class.

Proposition 2.4. For some base measure μ, $x_1, \cdots, x_T \sim P_t(\cdot)$ satisfy the $\mathsf{C}_{\mathsf{cov}}$-smoothness condition, as described in Definition 1. Let $\mathcal{A}$ be a countable set class. With probability at least $1 - \delta$, the following events hold simultaneously:

1. every pair of sets $A_1, A_2 \in \mathcal{A}$ such that $\mu(A_1 \Delta A_2) \le \varepsilon := \frac{\mathsf{c}_1}{\mathsf{C}_{\mathsf{cov}} T \log(T\delta^{-1})}$ satisfies
$$\sum_{1 \le t \le T} \mathbb{1}\Big\{ x_t \in A_1 \Delta A_2 \Big\} \lesssim \log\Big(\delta^{-1} \Pi_{\mathcal{A}}(2T\mathsf{C}_{\mathsf{cov}} \log(T\delta^{-1})) \Big); \tag{2.5}$$
2. simultaneously for any $1 \le t' \le T$ and set $A \in \mathcal{A}$, the intersection $A \cap \{x_1, \cdots, x_{t'}\} = \emptyset$ implies that
$$\sum_{1 \le t \le t'} P_t(A) \lesssim \log\Big(\delta^{-1} \Pi_{\mathcal{A}}(2T\mathsf{C}_{\mathsf{cov}} \log(T\delta^{-1})) \Big) + \log\Big(\frac{|\mathcal{A}'|}{\delta} \Big), \tag{2.6}$$
where $\mathcal{A}'$ is a finite ε-proper covering of $\mathcal{A}$.

We denote the joint events described in Proposition 2.4 as E.

Corollary 2.5. *If the event E occurs, then for fixed predictable events $\{A_t\} \subseteq \mathcal{X}$ (i.e. A_t is $\mathcal{H}_{t-1}^x$ measurable) satisfying $A_t \cap \{x_1, \cdots x_{t-1}\} = \emptyset$, we have*

$$\sum_{1 \le t \le T} P_t(A_t) \lesssim \sqrt{\mathsf{C}_{\mathsf{cov}} T \log\Big(1 + \frac{T}{\mathsf{C}_{\mathsf{cov}}}\Big) \Big[\log\Big(\delta^{-1} \Pi_{\mathcal{A}}(2T\mathsf{C}_{\mathsf{cov}} \log(T\delta^{-1})) \Big) + \log\Big(\frac{|\mathcal{A}'|}{\delta} \Big) \Big]} + \mathsf{C}_{\mathsf{cov}} \log\Big(1 + \frac{T}{\mathsf{C}_{\mathsf{cov}}}\Big). \tag{2.7}$$

3 Lower bound for online prediction in a smooth environment

Theorem 3.1. *For sufficiently large $T \in \mathbb{N}^+$, there exists an adversarial environment Env and a function class with finite pseudo dimension $\mathrm{Pdim}(\mathcal{F}) \le d$ such that for any learner π, the expected regret satisfies*

$$\mathbb{E}\,\mathsf{Regret}(T, \mathcal{F}) \gtrsim \begin{cases} \sqrt{d(\mathsf{C}_{\mathsf{cov}} - 1)T} & \text{if} \quad \mathsf{C}_{\mathsf{cov}} > 1 \text{ and } d \ge 2; \\ d \log(1 + T/d) & \text{if} \quad \mathsf{C}_{\mathsf{cov}} = 1 \text{ or } d = 1. \end{cases} \tag{3.1}$$

Remark 3 (Rate comparison)**.** We now give a detailed comparison between our rate and existing work. First, our results characterize a phase transition between $\mathsf{C}_{\mathsf{cov}} = 1$ (i.e., when the covariates are drawn i.i.d.) and $\mathsf{C}_{\mathsf{cov}} > 1$. It is well known that logarithmic regret is attainable in the classical PAC learning setting. Second, in our construction of a hard environment, we focus on the linear predictor

class. Somewhat surprisingly, the linear predictor class with $\mathcal{X} \subseteq \mathbb{R}^d$ already captures the worst-case difficulty among function classes with pseudo-dimension $\mathrm{Pdim}(\mathcal{F}) \leq d$. Finally, our lower bound rules out logarithmic dependence on the smoothness parameter $\mathsf{C}_{\mathsf{cov}}$ in the agnostic setting. This contrasts with the setting in which one knows and can sample from μ, where logarithmic dependence is possible, aligned with the observations by Block & Simchowitz (2022) and Wu et al. (2023).

As the idea to prove the lower bound will also be used in the proof of the lower bound in Theorem 3.3 later, we sketch some proof heuristic as follows. The complete proof is relegated to Appendix D.

Proof Idea By choosing the (truncated) linear predictor $\mathcal{F}_{\mathsf{lin}} := \{x \mapsto (\theta^{\mathsf{T}} x \vee -1) \wedge 1 : x \in \mathcal{X}\}$, we explicitly construct a hard sequential law by constructing $d/2$ trees of depth $\Theta(d^{-1/2}(\mathsf{C}_{\mathsf{cov}} T)^{1/2})$, when $\mathsf{C}_{\mathsf{cov}} > 1$ and $d = 2$. Each visited level of a two-dimensional tree contributes at least $1/2$ expected regret for any learner. The corner cases $d = 1$ or $\mathsf{C}_{\mathsf{cov}} = 1$ correspond to the case where each x_t lies on the same line or, at each round, $\{x_t\}_{1 \leq t \leq T}$ are independent. In the practically relevant regime $d \geq 2$ and $\mathsf{C}_{\mathsf{cov}} > 1$, we further establish the sharp rate $\mathfrak{R}(T) = \Theta(\sqrt{d\mathsf{C}_{\mathsf{cov}} T})$, with no hidden logarithmic factor. The upper bound is achieved by the unregularized VAW estimator, improving the existing result by a factor of $\sqrt{\log T}$ and thereby establishing the optimality of the unregularized VAW procedure. □

As the lower bound instance focuses on $\mathcal{F}_{\mathsf{lin}}$, we revisit the VAW algorithm following Chen et al. (2026) and avoid the extra logarithmic factor on T. In other words, at least for the linear predictor class, for the regime $d \geq 2, \mathsf{C}_{\mathsf{cov}} > 1$, the regret rate stated in Theorem 3.1 is optimal even with respect to logarithmic factor.

Proposition 3.2. For every $\delta \in (0, 1)$, with probability at least $1 - \delta$ over the covariate sequence $\{x_t\}_{1 \leq t \leq T}$, the VAW predictor satisfies

$$\sum_{1 \leq t \leq T} (\widehat{y}_t - y_t)^2 - \inf_{\theta \in \mathbb{R}^d} \sum_{1 \leq t \leq T} (\theta^{\mathsf{T}} x_t - y_t)^2 \lesssim \sqrt{d\mathsf{C}_{\mathsf{cov}} T} + \log(1/\delta), \tag{3.2}$$

simultaneously for every outcome sequence $y_1, \ldots, y_T \in [-1, 1]$. Consequently, the VAW predictor is optimal, up to universal constants, for all $d \geq 2$, $\mathsf{C}_{\mathsf{cov}} > 1$, and T.

The proof proceeds largely as in Chen et al. (2026, Theorem 8) by carefully avoiding the rejection sampling and coupling lemma originally proposed in Haghtalab et al. (2024). The rejection sampling step leads to the maximum of T geometric random variables $J_t \overset{i.i.d.}{\sim} \mathsf{Geom}(\mathsf{C}_{\mathsf{cov}}^{-1})$, thus for each round, one must sample $\Theta(\mathsf{C}_{\mathsf{cov}} \log(T))$ times. Instead, we use the smoothness by relating the law of the subsequence $\boldsymbol{x}_I$ to its i.i.d. dominating measure counterpart $\mu_{\otimes |I|}$.

We now turn to the well-specified setting, starting with the following definitions.

Definition 4. An environment is well-specified with respect to a function class $\mathcal{F}$ if there exists an $\mathcal{H}_0$-measurable function $f^\star \in \mathcal{F}$ such that, for every $t \in [T]$, $\mathbb{E}[y_t \,|\, \mathcal{H}_{t-1}, x_t] = f^\star(x_t)$. The environment is realizable if $y_t = f^\star(x_t)$, and it is ν^2-subGaussian if $y_t = f^\star(x_t) + \eta_t$ with $\eta_t \,|\, \mathcal{H}_{t-1}, x_t$ being mean-zero and ν^2-subGaussian.

ERM is a natural algorithm for well-specified environments. At each round t, the learner chooses

$$\widehat{f}_t \in \operatorname*{argmin}_{f \in \mathcal{F}} \sum_{s=1}^{t-1} \big(f(x_s) - y_s\big)^2, \tag{3.3}$$

the minimizer of the empirical error on historical observations, and predicts $\widehat{y}_t = \widehat{f}_t(x_t)$.

Under realizable responses, Blanchard (2025) constructs a function class of VC dimension d for which the minimax expected regret is $\Omega(\sqrt{(\mathsf{C}_{\mathsf{cov}} - 1) dT})$. We refine this result by quantifying the additional difficulty caused by response noise. The following theorem extends their lower bound to conditionally ν^2-subGaussian responses and makes its dependence on the noise level explicit.

Theorem 3.3. *For any $d \in \mathbb{N}_+$, there exists a function class $\mathcal{F} : \mathcal{X} \to \{0, 1\}$ with $\mathsf{VC}(\mathcal{F}) = d$ such that, for any $\mathsf{C}_{\mathsf{cov}} > 1$ and any algorithm such that $\widehat{y}_t$ is some (possibly randomized) function*

of $\mathcal{H}_{t-1}$ and X_t, there is a $\mathsf{C}_{\mathsf{cov}}$-smooth ν^2-subGaussian environment such that, for any horizon $T \geq C[1 \vee (\mathsf{C}_{\mathsf{cov}} - 1)^{-1}]\nu^2 d$ with some sufficiently large constant $C > 0$,

$$\mathbb{E}[\mathsf{Regret}(T)] \gtrsim T \wedge \sqrt{(\mathsf{C}_{\mathsf{cov}} - 1)(1 \vee \nu^2)dT}.$$

Remark 4 (Comparison with linear classes)**.** When $d \geq 2$, the lower bound in Theorem 3.3 matches its counterpart for linear function classes established in Theorem 3.1, both scaling as $\sqrt{T}$ with the horizon. Thus, for any fixed $\mathsf{C}_{\mathsf{cov}} > 1$, imposing well-specification (or even realizability) does not improve the worst-case regret rate. The intuition is that, although well-specification rules out arbitrary response sequences, it leaves the adaptive covariate process sufficiently unconstrained to reproduce the same worst-case difficulty, while preserving realizability of the responses. When $d = 1$, Theorem 3.1 gives a logarithmic dependence on T for linear function classes, whereas Theorem 3.3 shows that a function class of VC dimension one can still exhibit a $\sqrt{T}$ dependence. This difference shows that the phase transition at $d = 1$ presented in Theorem 3.1 is not a consequence of dimension alone, but rather reflects additional structural properties of the one-dimensional linear class.

Remark 5 (Optimality of ERM)**.** Theorem 3.3, together with the upper bound in Block et al. (2024, Theorem 8), establishes the minimax optimality of ERM in its dependence on ν, $\mathsf{C}_{\mathsf{cov}}$, d, and T for function classes with finite VC dimension, up to logarithmic factors. To be specific, for any subGaussian environment, a minor technical correction to Theorem 8 of Block et al. (2024) (see Appendix D.4) shows that the expected cumulative estimation error of ERM satisfies[1]

$$\mathbb{E}\left[\sum_{t=1}^{T}\left(\widehat{f}_t(x_t) - f^\star(x_t)\right)^2\right] = \widetilde{O}\left(\sqrt{\mathsf{C}_{\mathsf{cov}}(1 \vee \nu^2)T\left(1 \vee \log \mathbb{E}_\mu\left[W_{2\mathsf{C}_{\mathsf{cov}}T\log T}(256 \cdot \mathcal{F})\right]\right)}\right), \tag{3.4}$$

where $\widetilde{O}$ hides logarithmic factors in T and $\mathsf{C}_{\mathsf{cov}}$, and $W_m(\mathcal{F})$ is the Wills functional defined in Definition 6 in Appendix D.3. In the well-specified setting, the expected cumulative error is no greater than the expected regret, but it is straightforward to show that the same upper bound still holds for the regret using Lemma 13 therein. The VC specialization of equation 3.4 follows directly from the fact that $\log W_m(\mathcal{F}) \lesssim \mathsf{VC}(\mathcal{F}) \cdot \log(m)$ for all $m \in \mathbb{N}_+$ (Block et al., 2024).

Although Theorem 3.3 gives a matching worst-case lower bound over function classes, it does not imply that the Wills-functional upper bound in equation 3.4 is sharp for every class. The following proposition shows that a function class can have large Wills complexity while every online ERM procedure has expected regret bounded independently of T, even when the covariate process is not smooth.

Proposition 3.4. Recall the Wills functional $W_m(\mathcal{F})$ from Definition 6. For any $K \in \mathbb{N}_+$, let $m \geq K\log(2K)$. There exists a function class $\mathcal{F}$ and a base measure μ on $\mathcal{X}$ with $\log \mathbb{E}_\mu[W_m(\mathcal{F})] \asymp K$, such that for any horizon T and any ν^2-subGaussian (possibly non-smooth) environment, any empirical risk minimizer as defined in equation 3.3 satisfies $\mathbb{E}[\mathsf{Regret}(T)] \lesssim (1 \vee \nu^2)K$.

4 Conclusion and Outlook

We first study the optimal regret in agnostic smoothed online learning with adversarial responses. By Theorem 2.1, Algorithm 1 achieves regret $\widetilde{O}\big(\sqrt{\mathrm{Pdim}(\mathcal{F})\mathsf{C}_{\mathsf{cov}}T}\big)$, matching the lower bound established in Theorem 3.1 up to logarithmic factors. Notably, this lower bound already holds for the linear prediction class $\mathcal{F}_{\mathsf{lin}}$ in $\mathbb{R}^d$. We then examine empirical risk minimization (ERM) in the well-specified setting as a natural benchmark for oracle-efficient learning. Theorem 3.3, however, shows that even in this setting, ERM cannot attain a regret rate faster than the minimax rate for adversarial responses. Together with the intricate but computationally inefficient R-Cover algorithm of Blanchard (2025), these findings suggest a possible gap between information-theoretic optimality and computational efficiency. Whether an oracle-efficient algorithm can attain the optimal regret rate in our setting, or whether a computational barrier precludes this, remains an important open question.

[1]Their result assumes smoothness conditional on the full history $\mathcal{H}_{t-1}$, whereas this paper imposes the weaker requirement of smoothness conditional only on the covariate history $\mathcal{H}^x_{t-1}$.

AI USE STATEMENT

In this work, we used generative AI tools for interpreting results and language polishing purposes during the preparation of this manuscript. We have not used generative AI tools to help develop theoretical models or conceptual frameworks, and providing critical ingredients for proving mathematical claims are not applicable to this work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI. If there are errors, the responsibility lies with the author.

REFERENCES

Noga Alon, Omri Ben-Eliezer, Yuval Dagan, Shay Moran, Moni Naor, and Eylon Yogev. Adversarial laws of large numbers and optimal regret in online classification. In *Proceedings of the 53rd annual ACM SIGACT symposium on theory of computing*, pp. 447–455, 2021.

Martin Anthony and John Shawe-Taylor. A result of vapnik with applications. *Discrete Applied Mathematics*, 47(3):207–217, 1993.

Katy S Azoury and Manfred K Warmuth. Relative loss bounds for on-line density estimation with the exponential family of distributions. *Machine learning*, 43(3):211–246, 2001.

Moïse Blanchard. Agnostic smoothed online learning. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pp. 1997–2006, 2025.

Adam Block and Max Simchowitz. Efficient and near-optimal smoothed online learning for generalized linear functions. *Advances in neural information processing systems*, 35:7477–7489, 2022.

Adam Block, Yuval Dagan, Noah Golowich, and Alexander Rakhlin. Smoothed online learning is as easy as statistical learning. In Po-Ling Loh and Maxim Raginsky (eds.), *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pp. 1716–1786. PMLR, 02–05 Jul 2022.

Adam Block, Max Simchowitz, and Alexander Rakhlin. Oracle-efficient smoothed online learning for piecewise continuous decision making. In Gergely Neu and Lorenzo Rosasco (eds.), *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 1618–1665. PMLR, 12–15 Jul 2023.

Adam Block, Alexander Rakhlin, and Abhishek Shetty. On the performance of empirical risk minimization with smoothed data. In *The Thirty Seventh Annual Conference on Learning Theory*, pp. 596–629. PMLR, 2024.

Nicolo Cesa-Bianchi, Gábor Lugosi, et al. *Prediction, learning, and games*. Cambridge university press Cambridge, 2006.

Fan Chen, Jian Qian, Alexander Rakhlin, and Nikita Zhivotovskiy. Self-normalized martingales and uniform regret bounds for linear regression. *arXiv preprint arXiv:2605.01628*, 2026.

Luc Devroye, László Györfi, and Gábor Lugosi. *A probabilistic theory of pattern recognition*, volume 31. springer New York, 1996.

Pierre Gaillard, Sébastien Gerchinovitz, Malo Huard, and Gilles Stoltz. Uniform regret bounds over $\mathbb{R}^d$ for the sequential linear regression problem with the square loss. In Aurélien Garivier and Satyen Kale (eds.), *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, volume 98 of *Proceedings of Machine Learning Research*, pp. 404–432. PMLR, 22–24 Mar 2019.

Hugo Hadwiger. Das will'sche funktional. *Monatshefte für Mathematik*, 79(3):213–221, 1975.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis of online and differentially private learning. *Advances in Neural Information Processing Systems*, 33:9203–9215, 2020.

Nika Haghtalab, Yanjun Han, Abhishek Shetty, and Kunhe Yang. Oracle-efficient online learning for smoothed adversaries. *Advances in Neural Information Processing Systems*, 35:4072–4084, 2022.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *Journal of the ACM*, 71(3):1–34, 2024.

Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021.

Alessandro Lazaric and Rémi Munos. Hybrid stochastic-adversarial on-line learning. In *COLT*, 2009.

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

Markus Engelund Mathiasen, Jian Qian, and Nikita Zhivotovskiy. An optimal agnostic pac algorithm. *arXiv preprint arXiv:2608.06363*, 2026.

Jaouad Mourtada. Universal coding, intrinsic volumes, and metric complexity. *Journal of the European Mathematical Society*, May 2025.

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

David Pollard. *Convergence of stochastic processes*. Springer Science & Business Media, 2012.

François Portier. A uniform relative deviation inequality for vc-subgraph classes. *arXiv preprint arXiv:2607.11719*, 2026.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. *Advances in neural information processing systems*, 24, 2011.

Vinod Raman, Unique Subedi, and Ambuj Tewari. Smoothed online classification can be harder than batch classification. *Advances in Neural Information Processing Systems*, 37:68120–68143, 2024.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Vladimir N. Vapnik and Alexey Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications*, 16(2):264–280, 1971. doi: 10.1137/1116025.

Vladimir N. Vapnik and Alexey Ya. Chervonenkis. *Theory of Pattern Recognition: Statistical Problems of Learning*. Nauka, Moscow, 1974. In Russian.

Richard A Vitale. The wills functional and gaussian processes. *The Annals of Probability*, 24(4): 2172–2178, 1996.

Volodya Vovk. Competitive on-line linear regression. *Advances in Neural Information Processing Systems*, 10, 1997.

Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge University Press, February 2019. ISBN 978-1-108-49802-9. Google-Books-ID: IluHDwAAQBAJ.

Jörg M Wills. Zur gitterpunktanzahl konvexer mengen. *Elemente der Mathematik*, 28:57–63, 1973.

Changlong Wu, Mohsen Heidari, Ananth Grama, and Wojciech Szpankowski. Expected worst case regret via stochastic sequential covering. *arXiv preprint arXiv:2209.04417*, 2022.

Changlong Wu, Ananth Grama, and Wojciech Szpankowski. Online learning in dynamically changing environments. In *The Thirty Sixth Annual Conference on Learning Theory*, pp. 325–358. PMLR, 2023.

Changlong Wu, Jin Sima, and Wojciech Szpankowski. Oracle-efficient hybrid online learning with unknown distribution. In *The Thirty Seventh Annual Conference on Learning Theory*, pp. 4992–5018. PMLR, 2024.