Smooth Online Learning is as Easy as Statistical Learning

Adam Block ABLOCK@MIT.EDU
Yuval Dagan DAGAN@MIT.EDU
Noah Golowich NZG@MIT.EDU
Alexander Rakhlin RAKHLIN@MIT.EDU
MIT

Editors: Po-Ling Loh and Maxim Raginsky

Abstract

Much of modern learning theory has been split between two regimes: the classical *offline* setting, where data arrive independently, and the *online* setting, where data arrive adversarially. While the former model is often both computationally and statistically tractable, the latter requires no distributional assumptions. In an attempt to achieve the best of both worlds, previous work proposed the smooth online setting where each sample is drawn from an adversarially chosen distribution, which is *smooth*, i.e., it has a bounded density with respect to a fixed dominating measure. Existing results for the smooth setting were known only for binary-valued function classes and were computationally expensive in general; in this paper, we fill these lacunae. In particular, we provide tight bounds on the minimax regret of learning a nonparametric function class, with nearly optimal dependence on both the horizon and smoothness parameters. Furthermore, we provide the first oracle-efficient, no-regret algorithms in this setting. In particular, we propose an oracle-efficient improper algorithm whose regret achieves optimal dependence on the horizon and a proper algorithm requiring only a *single* oracle call per round whose regret has the optimal horizon dependence in the classification setting and is sublinear in general. Both algorithms have exponentially worse dependence on the smoothness parameter of the adversary than the minimax rate. We then prove a lower bound on the oracle complexity of any proper learning algorithm, which matches the oracle-efficient upper bounds up to a polynomial factor, thus demonstrating the existence of a statistical-computational gap in smooth online learning. Finally, we apply our results to the contextual bandit setting to show that if a function class is learnable in the classical setting, then there is an oracle-efficient, no-regret algorithm for contextual bandits in the case that contexts arrive in a smooth manner.

Keywords: Online Learning, Smoothed Analysis, Oracle Complexity

1. Introduction

Modern learning theory has primarily focused on two regimes: batch and sequential settings. In the former, data are independent and learning is easy while in the latter, Nature has the power to adversarially choose data to make learning as difficult as possible. Much of the empirical success in machine learning has derived from assuming that independence is satisfied and applying the Empirical Risk Minimization (ERM) principle in the batch (offline) setting, for which algorithms such as gradient descent have been highly successful even with complex function classes. However, in many settings independence is likely to fail, and the sequential regime is attractive due to the minimal assumptions made on the data generating process. Unfortunately, it suffers from poor algorithmic efficiency and an inability to learn some of the most basic function classes.

A typical example of the gap in difficulty of learning in these two settings is furnished by threshold functions on the unit interval. Classical theory tells us that in the batch setting, due to the combinatorial simplicity of the function class, a simple ERM approach efficiently and optimally learns the class of thresholds; in contradistinction, an adversarial data generation process precludes sequential (online) learning entirely (Littlestone, 1988). One way to escape the difficulty of adversarial learning is to apply the technique of *smoothed analysis* introduced in a now-famous paper by Spielman and Teng (2004), which focused on solving linear programs with the simplex algorithm. In this regime, one analyzes worst case inputs that are perturbed by a small amount of stochastic noise. Smoothed analysis in the setting of online learning was first introduced in Rakhlin et al. (2011), where the authors showed non-constructively that thresholds again become learnable in this setting. More recently, a series of papers (Haghtalab et al., 2020, 2021) has demonstrated that the stochastic perturbation has beneficial effects in far greater generality than the class of thresholds; in fact, any classification task that is possible in the batch setting is also statistically tractable in the smoothed online setting.

In the more modern formulation of the smoothed paradigm studied in Haghtalab et al. (2020, 2021), instead of choosing an input that is then perturbed, the adversary chooses a distribution that is restricted to be sufficiently anti-concentrated so as to not put too much mass on the set of "hard" instances. This anti-concentration (referred to as *σ-smoothness*; Definition 1) is quantified by a parameter $\sigma \leq 1$ that governs how far from independent the adversary can be. When $\sigma = 1$, we are in the batch setting where the data arrive i.i.d.; as σ tends to 0, the adversary is given more and more power to choose bad instances, with the limit of $\sigma = 0$ being entirely adversarial.

While we provide a rigorous definition of the problem setting below, we outline the broad strokes here. We consider learning over the course of T rounds where, at each round, Nature reveals a context x_t sampled in a σ-smooth manner, the learner reveals a prediction $\widehat{y}_t$, and then Nature reveals y_t. Given a loss function, the objective of the learner is to minimize regret to the best predictor in some function class $\mathcal{F}$. When $\mathcal{F}$ is binary-valued and has finite VC dimension d (i.e. is learnable in the batch setting), Haghtalab et al. (2021) proved that $O\left(\sqrt{dT \log\left(T/\sigma\right)}\right)$ regret is achievable, albeit with an inefficient algorithm. Two natural questions arise: can we extend these results to nonparametric, real-valued classes? and, more importantly, are there efficient algorithms that can achieve comparable regret? In this paper, we answer both questions. With regard to the first question, the natural extension of the covering-based argument in Haghtalab et al. (2021) would yield suboptimal dependence on σ in the nonparametric regime; instead, we obtain a nonconstructive proof through careful application of combinatorial inequalities and an adaptation of the coupling lemma of Haghtalab et al. (2020).

For the question of practical algorithms, we need to more carefully consider what we mean by efficiency. Certainly the algorithm of Haghtalab et al. (2021) is not efficient as it requires constructing an ε-net of $\mathcal{F}$ as a first step, which is exponential in the VC dimension of $\mathcal{F}$. A natural choice is to look to the batch setting and try to leverage the success of ERM-based approaches to achieve reasonable runtime, as is done in Kalai and Vempala (2005); Hazan and Koren (2016); this approach is supported further by the empirical success of various heuristics for ERM (Goodfellow et al., 2016). As such, we suppose the learner has access to an *ERM oracle* (Definition 2) that can efficiently optimize some loss over our function class $\mathcal{F}$ given as input a dataset; we analyze the time complexity of our algorithms in terms of the number of calls to this oracle.

In our algorithmic results, we distinguish between *proper* and *improper* learners. Proper learners are required, before seeing x_t, to output a hypothesis $\widehat{f}_t \in \mathcal{F}$ that is used to produce the prediction $\widehat{y}_t := \widehat{f}_t(x_t)$, whereas improper learners can make any prediction $\widehat{y}_t$ based on knowledge of x_t. There are many settings in which proper (as opposed to improper) online learning may be desirable for downstream applications, such as learning in games (Daskalakis and Golowich, 2021) and reinforcement learning with function approximation (Foster et al., 2021). For proper learners, our oracle time complexity results are optimal up to a polynomial factor in all parameters; the question of optimality for improper learning remains an interesting open question.

We now briefly describe our key contributions (which are summarized in Table 1):

- In Section 3 we give tight upper bounds on the statistical rates of learning a real-valued function class in the smoothed online setting without regard to computational efficiency, while extending and providing a new proof to the case of binary classification treated in (Haghtalab et al., 2021, Theorems 3.1 & 3.2). Our bounds are tight both in their dependence on T and σ, up to logarithmic factors, showing that the regret in the smooth setting is only a factor $\log(T/\sigma)$ away from that in the i.i.d. setting (Theorem 3). In the process of doing this, we provide in Lemma 14 a more general and much simpler proof of the key technical step of coupling from Haghtalab et al. (2021).

- In Section 4 we present an *improper* algorithm based on the relaxation method of Rakhlin et al. (2012) with tight dependence on the horizon T but suboptimal dependence on σ: in particular, the regret in the smooth setting scales as $\sigma^{-1/2}$ times that in the i.i.d. setting (Theorem 7). Our algorithm is efficient in the sense that it requires only $O\left(\sqrt{T}\log T\right)$ oracle calls per round in general and only 2 oracle calls per round in the classification setting. We then show in Proposition 8 that the polynomial dependence on σ is not an artifact of our analysis, but rather inherent to the method.

- In Section 5 we present a *proper* algorithm based on Follow the Perturbed Leader (FTPL) that exhibits optimal dependence on T for classification and suboptimal dependence on T in general (Theorem 10). Further, the algorithm requires only 1 oracle call per round. We use a Gaussian white noise with intensity approximated by μ as our perturbation, which allows for optimization without enumeration of experts; to establish correctness, we overcome a major technical hurdle introduced by the complicated dependence structure of this perturbation.

- In Section 6 we show that the suboptimality of the FTPL learner is inherent for oracle-efficient algorithms: in particular, we provide a lower bound based on the method of Hazan and Koren (2016) that demonstrates that any *proper* algorithm with access to an ERM oracle requires $\widetilde{\Omega}\left(\sigma^{-\frac{1}{2}}\right)$ time. Combined with the upper bound of $\log 1/\sigma$ of Theorem 3, this implies that there exists an exponential statistical-computational gap in smoothed online learning.

- Finally, in Appendix A, we apply our results to the problem of Contextual Bandits. We show in Theorem 13 that whenever a function class $\mathcal{F}$ is learnable offline, we can get vanishing regret in the smooth contextual bandit setting with an oracle-efficient algorithm.

In an independent and concurrent work, Haghtalab et al. (2022) established several results similar to our own. In particular, they provided an analysis of an *improper* algorithm similar to our oracle-efficient improper learner from Section 4, as well as computational lower bounds similar to those presented in our Section 6. Additionally, they presented an oracle-efficient *proper* learner in

[1] This bound was shown in Haghtalab et al. (2021), though with slightly different log factors.

Reference	Algorithm	Iterations	Oracle calls	Total time
Theorem 3	Non-constructive, *proper*	**Cls**[1]: $\varepsilon^{-2} d \log(1/\sigma)$ **Reg**: $\varepsilon^{-2} \log(1/\sigma)$	NA	NA
Theorem 7	Relaxation-based, *improper*	**Cls**: $\varepsilon^{-2} d \sigma^{-1}$ **Reg**: $\varepsilon^{-2} \sigma^{-1}$	**Cls**: $\varepsilon^{-2} d \sigma^{-1}$ **Reg**: $\varepsilon^{-3} \sigma^{-3/2}$	**Cls**: $\varepsilon^{-4} d^2 \sigma^{-3}$ **Reg**: $\varepsilon^{-4} \sigma^{-5/2}$
Theorem 10	FTPL, *proper*	**Cls**: $\varepsilon^{-2} d \sigma^{-1}$ **Reg**: $\varepsilon^{-3} \sigma^{-1}$	**Cls**: $\varepsilon^{-2} d \sigma^{-1}$ **Reg**: $\varepsilon^{-3} \sigma^{-1}$	**Cls**: $\varepsilon^{-4} d^2 \sigma^{-5/2}$ **Reg**: $\varepsilon^{-7} \sigma^{-3}$
Theorems 11 & 52, Corollaries 12 & 53	Computational **lower bound** for any *proper* alg.	NA	**Cls**: $\max\left\{\sigma^{-1/2}\zeta^2, \varepsilon^{-2} d\right\}$ with ζ-approx. oracle	**Cls**: $\max\left\{\sigma^{-1/2}, \varepsilon^{-2} d\right\}$ with exact oracle
Theorem 3.2 from Haghtalab et al. (2021)	Statistical **lower bound** for any algorithm	**Cls**: $\varepsilon^{-2} d \log(1/\sigma)$ **Reg**: $\varepsilon^{-2} \log(1/\sigma)$	NA	NA

Table 1: Overview of our main results. For each algorithm/lower bound, the number of iterations T after which the algorithm achieves regret $\leq \varepsilon T$ is shown for two cases: (a) **Cls:** the case of binary classification for a class $\mathcal{F}$ of VC dimension d; (b) **Reg:** the case of regression for a class $\mathcal{F}$ with scale-sensitive VC dimension bounded as $\mathsf{vc}(\mathcal{F}, \alpha) \lesssim \alpha^{-p}$, for some $0 < p < 2$ (our results extend to the case of $p \geq 2$, which may be found in the theorem statements). We take $L = 1$ and suppress logarithmic factors where possible.

the case of binary classification that, while based on the principle of FTPL, has a substantially different analysis than our own in Section 5; note that our algorithm also applies for general, real-valued function classes in the nonparametric regime.

We discuss further related work in Appendix B.

2. Problem Setup and Notation

In this section we formally define the smoothed online learning setting. We then introduce some concepts and notation we use throughout the paper.

Miscellaneous notation. For distributions p, q on a measure space $\mathcal{X}$, we write $p \ll q$ if p is absolutely continuous with respect to q. For a positive integer m, let $[m] = \{1, 2, \ldots, m\}$. For expressions f, g we say $f \lesssim g$ if there is some universal constant C such that $f \leq Cg$. We also use $f = O(g)$ to signify the same thing.

2.1. Smoothed Online Learning

We consider the setting of smoothed online learning, following Haghtalab et al. (2021). Consider a space $\mathcal{X}$ of covariates equipped with some sigma-algebra. Let $\mathcal{F} \subset [-1, 1]^{\mathcal{X}}$ be a function class and $\ell : [-1, 1] \times [-1, 1] \to [0, 1]$ be an L-Lipschitz, convex loss function for some $L > 0$. Fix some $T \in \mathbb{N}$ denoting the number of rounds of learning. We consider the following learning setting, making a distinction between *improper learning* and *proper learning*:

1. For each $t \in [T]$, nature samples $x_t \sim p_t$ for some distribution p_t that may depend in any way on the past samples x_s for $s < t$ and the algorithm's past predictions. Nature also chooses $y_t \in [-1, 1]$ adversarially (perhaps depending on x_t) in a similar manner.

2. The learner makes a (possibly random) prediction as follows:

   - **Improper learner:** The learner observes x_t and makes a prediction $\widehat{y}_t \in [-1, 1]$.
   - **Proper learner:** The learner chooses a hypothesis $\widehat{f}_t \in \mathcal{F}$, and its prediction is defined as $\widehat{y}_t := \widehat{f}_t(x_t)$; the learner then observes x_t.

3. Nature reveals $y_t \in [-1, 1]$ to the learner, and the learner incurs loss $\ell(\widehat{y}_t, y_t)$.

With no restrictions on p_t, the above online learning problem has been extensively studied, and essentially tight rates are known (Rakhlin et al., 2015b; Block et al., 2021); furthermore, exponential lower bounds are known for oracle-efficient proper learning algorithms (Hazan and Koren, 2016). To circumvent these lower bounds, we consider the smooth setting. Our fundamental assumption (Definition 1 below) is that there is a distinguished distribution μ on $\mathcal{X}$, accessable to the learner through efficient sampling, so that the adversary is constrained to choose covariates x_t according to some distribution with bounded Radon-Nikodym derivative with respect to μ. This assumption, which follows that of Haghtalab et al. (2020, 2021), has been used extensively as well in the smoothed analysis of local search algorithms (Manthey, 2020) and discrete optimization problems (Beier and Vöcking, 2004). We emphasize the assumption that in all cases the learner has access to μ through efficient sampling; note that a typical example to keep in mind is that μ is uniform on some set, so the efficient sampling assumption is not very restrictive.

Definition 1 (Adaptive smooth distributions) *Let p, μ be probability measures on a set $\mathcal{X}$. We say that p is σ-*smooth *with respect to μ if $p \ll \mu$ and*

$$\operatorname{ess\,sup} \frac{dp}{d\mu} \leq \frac{1}{\sigma}.$$

Let $\mathfrak{P}(\sigma, \mu)$ denote the class of all distributions p that are σ-smooth with respect to μ. We denote this class simply by $\mathfrak{P}$ when σ, μ are clear from the context. For any $T \in \mathbb{N}$, we let $\mathfrak{P}_T(\sigma, \mu)$ denote the space of joint distributions $\mathscr{D}$ on $x_1, \ldots, x_T \in \mathcal{X}$ satisfying the following property: letting p_t denote the law of x_t conditional on x_s for all $s < t$, for all $t \in [T]$, $p_t \in \mathfrak{P}(\sigma, \mu)$ almost surely. Similarly, we let $\widetilde{\mathfrak{P}}_T(\sigma, \mu)$ denote the space of joint distributions $\mathscr{D}$ on $(x_1, y_1), \ldots, (x_T, y_T) \in \mathcal{X} \times [-1, 1]$ such that if p_t is the law of x_t conditional on (x_s, y_s) for $s < t$, then $p_t \in \mathfrak{P}(\sigma, \mu)$ almost surely. Note that no constraints are placed on the distribution of y_t.

For $\sigma = 1$ we recover the notion of the data $x_1, \ldots, x_T$ being sampled iid from μ; as σ tends to zero, the notion of σ-smoothness becomes weaker and thus we consider σ-smoothness as interpolating between the favorable situation of i.i.d. data and the unfavorable adversarial situation.

2.2. Minimax value

The goal of the learner is to minimize expected regret to the best function in $\mathcal{F}$, defined as

$$\mathbb{E}\left[\mathrm{Reg}_T(\mathcal{F})\right] = \mathbb{E}\left[\sum_{t=1}^{T} \ell(\widehat{y}_t, y_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right]$$

with the expectation taken over both the sampling of $x_t \sim p_t$ and the learner's possibly randomized predictions. For any function class, we consider the minimax regret for proper learners to be the value $\mathcal{V}_T^{\text{prop}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$, defined to be equal to the following expression:

$$\left\langle \inf_{q_t\in\Delta(\mathcal{F})} \sup_{p_t\in\mathfrak{P}(\sigma,\mu)} \mathbb{E}_{x_t\sim p_t} \sup_{y_t\in[-1,1]} \mathbb{E}_{\widehat{f}_t\sim q_t} \right\rangle_{t=1}^{T} \left[\sum_{t=1}^{n} \ell(\widehat{f}_t(x_t), y_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right]$$

where $\langle\cdot\rangle_{t=1}^{T}$ denotes iterated application of the enclosed operators. Similarly, we define the minimax regret for an improper learner to be $\mathcal{V}_t^{\text{improp}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$, defined to be equal to

$$\left\langle \sup_{p_t\in\mathfrak{P}(\sigma,\mu)} \mathbb{E}_{x_t\sim p_t} \inf_{q_t\in\Delta([-1,1])} \sup_{y_t\in[-1,1]} \mathbb{E}_{\widehat{y}_t\sim q_t} \right\rangle_{t=1}^{T} \left[\sum_{t=1}^{n} \ell(\widehat{y}_t, y_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right].$$

It is straightforward to see that $\mathcal{V}_T^{\text{prop}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \geq \mathcal{V}_T^{\text{improp}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$.

2.3. ERM oracle model

To capture the notion of computational efficiency, we consider the following *ERM oracle model*:

Definition 2 (ERM oracle) *For $\zeta > 0$, a* ζ-approximate (weighted) empirical risk minimization (ERM) oracle *takes as input a sequence $(x_1, y_1), \ldots, (x_m, y_m) \in \mathcal{X} \times [-1, 1]$ of data, a sequence $w_1, \ldots, w_m \in \mathbb{R}$ of weights, and a sequence $\ell_1, \ldots, \ell_m$ of $[-1, 1]$-valued loss functions and outputs some $\widehat{f} \in \mathcal{F}$ satisfying*

$$\sum_{i=1}^{m} w_i \ell_i(\widehat{f}(x_i), y_i) \leq \inf_{f\in\mathcal{F}} \sum_{i=1}^{m} w_i \ell_i(f(x_i), y_i) + \zeta \cdot \sum_{i=1}^{m} |w_i|. \tag{1}$$

We remark that while Definition 2 allows for an arbitrary sequence $\ell_1, \ldots, \ell_m$ of loss functions, all of our algorithms will set each ℓ_i equal to either ℓ (the given loss function of the learning problem) or ℓ^{Id}, where $\ell^{\text{Id}}(\widehat{y}, y) := \widehat{y}$.

We will measure the computation cost of algorithms via the following two metrics: first, the *number of calls to the ERM oracle*, and second, the *total computation time*. To define the latter, we must specify the manner in which our algorithm interacts with the ERM oracle: in particular, we assume that there is a certain region of memory on which the algorithm lists tuples (x_i, y_i, w_i). Listing (or modifying) any such tuple takes unit time, as does calling the ERM oracle, which performs the optimization in (1) in unit time. This convention mirrors that of oracle machines in complexity theory (Arora and Barak, 2009). We refer to the case that $\zeta = 0$ the *exact ERM oracle model*, and the case that $\zeta > 0$ the *approximate ERM oracle model*.

2.4. Statistical complexities

For a space $\mathcal{X}$, an $n \in \mathbb{N}$, and a function class $\mathcal{F} \subset [-1,1]^{\mathcal{X}}$, the Rademacher complexity, $\mathcal{R}_n(\mathcal{F})$, and Gaussian complexity, $\mathcal{G}_n(\mathcal{F})$, conditional on $x_1, \dots, x_n$, are defined as follows:

$$\mathcal{R}_n(\mathcal{F}) := \mathbb{E}_\varepsilon \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^n \varepsilon_i f(x_i) \right], \qquad \mathcal{G}_n(\mathcal{F}) := \mathbb{E}_\gamma \left[\sup_{f \in \mathcal{F}} \sum_{i=1}^n \gamma_i f(x_i) \right],$$

where the ε_i are i.i.d. Rademacher random variables, and the γ_i are i.i.d. standard normal random variables. It is a well-known fact that there is a universal constant C so that $\frac{1}{C} \cdot \mathcal{R}_n(\mathcal{F}) \leq \mathcal{G}_n(\mathcal{F}) \leq C \log n \cdot \mathcal{R}_n(\mathcal{F})$ for all $\mathcal{F}$ and n.

Furthermore, we consider the notion of scale-sensitive VC dimension from Kearns and Schapire (1994); Bartlett et al. (1996) that characterizes learnability of $\mathcal{F}$ in a batch setting. For any $\alpha > 0$ and points $x_1, \dots, x_m \in \mathcal{X}$, we say that $\mathcal{F}$ is shattered by the x_i at scale α with witness $s_1, \dots, s_m \in \mathbb{R}$ if for all $(\varepsilon_1, \dots, \varepsilon_m) \in \{\pm 1\}^m$ there is an $f_\varepsilon \in \mathcal{F}$ such that

$$\varepsilon_i (f_\varepsilon(x_i) - s_i) \geq \frac{\alpha}{2} \qquad \text{for all } i$$

We define the VC dimension of $\mathcal{F}$ at scale α, denoted by $\mathsf{vc}(\mathcal{F}, \alpha)$ to be the largest m such that there exists a shattering set of size m. We let $\mathsf{vc}(\mathcal{F}) = \lim_{\alpha \downarrow 0} \mathsf{vc}(\mathcal{F}, \alpha)$ denote the VC dimension of $\mathcal{F}$.

3. Minimax Value

In this section, we provide tight bounds on $\mathcal{V}_T^{\text{prop}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$ without regard to oracle efficiency. While our proof is nonconstructive, we emphasize that our results in Section 6 demonstrate that any proper algorithm based on ERM oracle calls which achieves the optimal dependence on σ cannot be computationally efficient. Our results show that $\mathcal{V}_T^{\text{prop}}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$ is always a $\text{poly}(\log(T/\sigma))$ factor away from the optimal statistical rates achievable in the batch setting. We now present our bound:

Theorem 3 *Let $\mathcal{F} : \mathcal{X} \to [-1,1]$ be a real-valued function class and denote by $\mathsf{vc}(\mathcal{F}, \delta)$ the scale-sensitive VC dimension of $\mathcal{F}$ at scale $\delta > 0$. Then, for some $c > 0$*

$$\mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \lesssim L \log^{3/2}(T) \cdot \log\left(\frac{T}{\sigma}\right) \inf_{\alpha > 0} \left\{ T\alpha + \sqrt{T} \int_\alpha^2 \sqrt{\mathsf{vc}\left(\mathcal{F}, c\delta\right)} d\delta \right\}$$

In particular, if $\mathsf{vc}(\mathcal{F}, \delta) \leq d \log\left(\frac{1}{\delta}\right)$ for all $\delta > 0$, then

$$\mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \lesssim L \log^{\frac{3}{2}}(T) \log\left(\frac{T}{\sigma}\right) \sqrt{dT}$$

and if $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ for some $p < \infty$, then

$$\mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \lesssim L \log^{\frac{3}{2}}(T) \log\left(\frac{T}{\sigma}\right) T^{\max\left(\frac{1}{2}, 1 - \frac{1}{p}\right)}$$

Our result extends (Haghtalab et al., 2021, Theorem 3.1) to the cases of real-valued and nonparametric function classes. In that paper, in order to prove their regret bounds for smoothed online classification, the authors introduced the clever approach of coupling, showing that if a distribution p is smooth with respect to the uniform distribution on a discrete set, then in expectation we may pretend the data comes independently from the uniform distribution. In Appendix C, we generalize their result in Lemmas 14 and 24 with a dramatically simpler proof[2]. Namely, for any $k \in \mathbb{N}$ we construct a coupling between $(x_1, \dots, x_T) \sim \mathscr{D} \in \mathfrak{P}_T(\sigma, \mu)$ and $\{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}} \sim \mu^{\otimes kT}$ such that $\{x_1, \dots, x_T\} \subset \{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}}$ with probability at least $1 - e^{-\sigma k}$.

Curiously, our proof of Theorem 3 is quite different from that of (Haghtalab et al., 2021, Theorem 3.1), although we still use the coupling. In that paper, the authors apply the coupling to analyze covering numbers with respect to $L^2(\mu)$ (i.e. in the i.i.d. sense); the natural extension of this technique would be to apply chaining (Dudley, 1967) and bound $\mathcal{V}_T^{prop}(\mathcal{F}, \mathscr{D})$ by $\mathbb{E}_\mu\left[\mathcal{R}_T(\mathcal{F})\right]$. Unfortunately, Proposition 8 in the sequel shows that such a bound is not possible without a suboptimal, polynomial dependence on σ. We instead go by a different approach, which is based on the observation that the sequential fat-shattering dimension (see Definition 19) is bounded above by the scale-sensitive VC dimension times a logarithmic factor in the domain size (Lemma 21). Though the true domain $\mathcal{X}$ may be infinite, we show that it is possible adapt the coupling lemma of Haghtalab et al. (2021) to bound an “effective” domain size. In combination with the non-constructive bounds using distribution-dependent sequential Rademacher complexity from Rakhlin et al. (2011); Block et al. (2021), this provides a tight characterization of the minimax regret’s dependence on the horizon T, the complexity of the function class, and the smoothness parameter σ. A full proof is in Appendix C. Combining Theorem 3 with (Haghtalab et al., 2021, Theorem 3.2), we have a complete characterization of the statistical rates of smooth online learning. We further note that our proof applied to the results of Rakhlin and Sridharan (2015) immediately extends to nonconstructively show that the dependence of the minimax value on T for squared loss are the expected “fast rates” from Rakhlin et al. (2017). Unfortunately, our efficient algorithms below do not provably achieve these rates; resolving this disparity is an interesting future direction with applications to the study of contextual bandits, as described in Appendix A.

Finally, before we proceed to consider oracle-efficient algorithms, we note that the requirement that the learner has access to μ cannot be dropped without a substantial loss in the regret. In particular, we have the following result:

Proposition 4 *There exists a function class $\mathcal{F} : [0, 1] \to \{\pm 1\}$ with $\mathsf{vc}(\mathcal{F}) = 1$ and an adversary that is σ-smooth with respect to some unknown μ such that, no matter how the learner chooses $\widehat{y}_t$, it holds for $T \le \frac{1}{\sigma}$ that*

$$\mathbb{E}\left[\mathrm{Reg}_T\right] \ge \frac{T}{2}.$$

Proposition 4 is proved by letting $\mathcal{F}$ be thresholds on the unit interval and allowing the μ with respect to which the adversary is σ-smooth adapt to the data sequence. The details can be found in Appendix C. In particular, the result shows that for the learner to be able to ensure regret scaling as in Theorem 3, he needs to have access to μ in some way. Critically, the size of the set of possible μ in the lower bound of Proposition 4 is growing exponentially with T; indeed, if we know that our

[2] Already in the case of discrete distributions, the proof of (Haghtalab et al., 2021, Theorem 2.1) first demonstrates their claim for uniform measures and then proceeds to apply convex analysis for the general case. They then claim that Choquet’s Theorem (Choquet and Meyer, 1963, Corollaire 8) implies the general case when $\mathcal{X}$ is not discrete; however, it is not *a priori* obvious that $\mathfrak{P}(\sigma, \mu)$ is compact in the relevant topology.

adversary is σ-smooth with respect to some $\mu \in \mathscr{P}$, a finite class of distributions, then we can use Hedge (Freund and Schapire, 1997) to aggregate predictions assuming smoothness with respect to each $\mu \in \mathscr{P}$ and add an additive term of size $O\left(\sqrt{\log(|\mathscr{P}|)T}\right)$ to our regret.

4. Relaxations and Oracle-Efficient Algorithms

In[3] the previous section, we derived sharp bounds for the minimax regret in the smoothed online setting, with sharp dependence on the key parameter σ. A natural next step is to design an algorithm that achieves these bounds. One possibility, suggested in Haghtalab et al. (2021), constructs a $\frac{1}{\sqrt{T}}$-net on $\ell \circ \mathcal{F}$ with respect to $L^2(\mu)$ and runs Hedge (Freund and Schapire, 1997) on the resulting covering. Unfortunately, in the nonparametric case, after optimizing δ this approach yields suboptimal rates, corresponding to one-step discretization. Ideally, an algorithm achieving optimal regret would construct nets at multiple scales and aggregate the resulting predictions in some way. While there has been some progress on how to do this (Cesa-Bianchi and Lugosi, 1999; Gaillard and Gerchinovitz, 2015; Daskalakis and Golowich, 2021), optimal rates are not yet achievable; in any case, relying on the construction of δ-nets is inefficient as these can be exponentially large.

Thus in order to bring the smoothed online learning paradigm from the world of theory into that of practice, we need more efficient algorithms. Presently, we describe an oracle-efficient improper learning procedure. As we shall see, the algorithm has regret with optimal dependence on the horizon, T, but suboptimal dependence on σ. We will improve the dependence on σ with a proper learning algorithm in the following section, at the cost of worse dependence on T in general. Here, we leverage the relaxation approach, studied in Rakhlin et al. (2012).

Definition 5 *Fix $T \in \mathbb{N}$. A sequence of real valued functions $\mathbf{Rel}_T(\mathcal{F}|x_1, \dots, x_t) : \mathcal{X}^t \to \mathbb{R}$, $t \leq T$, is a relaxation if for any $x_{1:T} \in \mathcal{X}$, we have the following two properties:*

$$\sup_{p_t \in \mathfrak{P}(\sigma,\mu)} \mathbb{E}_{x'_t \sim p_t} \inf_{q_t \in \Delta([-1,1])} \sup_{y'_t \in [-1,1]} \left[\mathbb{E}_{\widehat{y}_t \sim q_t}[\ell(\widehat{y}, y'_t)] + \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x'_t, y'_t)\right] \tag{2}$$

$$\leq \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_{t-1}, y_{t-1})$$

$$-\inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t) \leq \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_T, y_T)$$

As established in (Rakhlin et al., 2012, Proposition 1), a relaxation gives rise to both an algorithm and an associated regret bound; indeed, any q_t guaranteeing (2) at each time t yields a regret at most $\mathbf{Rel}_T(\mathcal{F})$; the challenge, of course, is to define the relaxation. We have the following result:

Proposition 6 *Suppose that $\mathscr{D} \in \mathfrak{P}(\sigma, \mu)$. Then, for any function class $\mathcal{F}$ and L-Lipschitz, convex loss function ℓ, and any $k \in \mathbb{N}$,*

$$\mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_t, y_t) = \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - \sum_{s=1}^{t} \ell(f(x_s), y_s)\right] + (T-t)^3 e^{-\sigma k} \tag{3}$$

[3] In an earlier version of the paper, we used a slightly different relaxation with a worse rate. With a minor modification, we get a quadratic improvement in the dependence of the regret on σ. While our improvement was independent of other work, we note that Haghtalab et al. (2022) present the same final relaxation and analysis.

is a relaxation, where the expectation is over independent $x_{s,j} \sim \mu$ and Rademacher random variables $\varepsilon_{s,j}$ for $s > t$.

We provide a proof in Appendix D that uses the minimax theorem, symmetrization and Lemma 14. Applying (Rakhlin and Sridharan, 2014, Lemma 5.1) to reduce to deterministic predictions, Proposition 6 gives rise to an algorithm that plays

$$\widehat{y}_t = \operatorname*{argmin}_{\widehat{y} \in [-1,1]} \sup_{y_t \in [-1,1]} \ell(\widehat{y}, y_t) + \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f)\right]$$

where we drop the additive constant from (3) because it does not depend on f and we let $L_t(f) = \sum_{s=1}^{t} \ell(f(x_s), y_s)$. After optimizing the resulting regret bound with respect to k, we get that the regret of this algorithm is $O\left(\mathcal{R}_{T\frac{\log T}{\sigma}}(\mathcal{F})\right)$, which has an optimal dependence on T up to logarithmic factors, but is suboptimal with respect to σ. Note that while the supremum inside of the expectation in (3) can be solved with an ERM oracle by letting $\ell_s(f(x_{s,j}), y_s) = f(x_{s,j})$ for $x > t$, we still require a costly integration in order to find $\widehat{y}_t$. We can compute this expectation by sampling from μ and applying concentration but this approach requires many calls to the ERM oracle. Motivated by the random playout idea in Rakhlin et al. (2012), we propose a much more efficient algorithm:

Theorem 7 *Suppose that $\mathscr{D} \in \mathfrak{P}(\sigma, \mu)$, $\mathcal{F} : \mathcal{X} \to [-1,1]$ is a function class, and ℓ is an L-Lipschitz, convex loss function. At each time t, for $1 \leq j \leq k$, sample $x_{t+1,j}, \dots, x_{T,j} \sim \mu$ and $\varepsilon_{t+1,j}, \dots, \varepsilon_{T,j}$ independently. After observing x_t, predict*

$$\widehat{y}_t = \operatorname*{argmin}_{\widehat{y} \in [-1,1]} \sup_{y_t \in [-1,1]} \left\{ \ell(\widehat{y}, y_t) + \sup_{f \in \mathcal{F}} \left[6L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right\} \tag{4}$$

Then the expected regret against any smooth adversary is

$$\mathbb{E}\left[\mathrm{Reg}_T\right] \leq 6L\mathbb{E}_\mu\left[\mathcal{R}_{kT}(\mathcal{F})\right] + \sqrt{T} + T^3 e^{-\sigma k}$$

Moreover, this regret can be achieved with $O\left(\sqrt{T}\log T\right)$ calls to the ERM oracle per round in general and only 2 calls per round in the special case that $\ell(\widehat{y}, y) = \frac{1-\widehat{y}y}{2}$. In particular, when $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ for some $p > 0$, we have

$$\mathbb{E}\left[\mathrm{Reg}_T\right] \lesssim L\left(\frac{T\log(T)}{\sigma}\right)^{\max\left(\frac{1}{2}, 1 - \frac{1}{p}\right)}$$

Theorem 7 is proved in Appendix D. To understand why (4) is oracle-efficient, note that we can discretize the interval $[-1,1]$ at scale $\frac{1}{L\sqrt{T}}$ to produce a set S of size $2L\sqrt{T}$. For each $\widehat{y} \in S$, we can exhaustively search S for the optimal y_t with $O(\sqrt{T})$ calls to the (value of the) ERM oracle. Because the problem is convex in $\widehat{y}$ (due to the convexity of ℓ), we can run zeroth order optimization as in Agarwal et al. (2011) to find $\widehat{y}_t$ up to $\frac{1}{\sqrt{T}}$ error in $O\left(\sqrt{T}\log T\right)$ calls to the oracle. If the losses are linear, then the problem is convex in y_t and is thus extremized on the boundary; further leveraging the linear loss allows the problem to be solved in 2 oracle calls per round. Note that in the case of y being binary-valued, we can think of $\ell(\widehat{y}, y) = \frac{1-\widehat{y}y}{2}$ as the indicator loss when guessing $\widehat{y} \in \{\pm 1\}$ and thus, for classification, we can get optimal regret with 2 oracle calls per round.

While Theorem 3 demonstrates that regret can depend on σ only through a logarithmic factor, our relaxation-based algorithm has a polynomial dependence on σ. Unfortunately, this polynomial dependence cannot in general be eliminated for any relaxation relying on the classical Rademacher complexity. To see this, note that the regret of any algorithm is bounded below by $\mathcal{V}_T^{improp}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$. The following proposition shows that the value is in turn bounded below by a polynomial factor of σ.

Proposition 8 *For any $\sigma \leq 1$ and $0 < p < 2$, there exists a measure μ and a function class $\mathcal{F}$ satisfying $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ such that for all $T \gtrsim \frac{1}{\sigma} \log\left(\frac{1}{\sigma}\right)$ with ℓ the absolute loss,*

$$\sigma^{-\frac{p}{4}} \mathbb{E}_\mu \left[\mathcal{R}_T(\mathcal{F})\right] \lesssim \mathcal{V}_T^{improp}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$$

We construct a measure such that the learning problem is easy when the population distribution is μ by having μ concentrate a lot of mass on a distinguished point x^* on which all the functions in $\mathcal{F}$ agree; thus, a sample from μ will include many copies of x^*, which incur no regret. We then consider an i.i.d. adversary and let p_t be uniform over a set of points that shatters $\mathcal{F}$ at scale $\sqrt{\sigma}$; in this way, we make it so a sample from p_t will incur high regret and the gap between the performance on samples from p_t and μ is relatively large. A complete proof can be found in Appendix C. Note that Proposition 8 is not in conflict with Theorem 3 because the example described above makes $\mathcal{R}_T(\mathcal{F})$ polynomially small in σ for a carefully designed μ; in essence, the separation is created by making the Rademacher complexity much smaller than expected based on the complexity of the function class $\mathcal{F}$.

In the improper procedure (4), we have our first efficient algorithm for the smoothed online learning setting that works for a generic function class and achieves an optimal regret dependence on the horizon T. There are three drawbacks to Theorem 7. First, $\widehat{y}_t$ is *improper*. Second, our dependence on σ is significantly worse than the optimal statistical dependence explored in Section 3. Third, while the algorithm is efficient, we may hope to have an algorithm that makes only 1 oracle call per round in general. We address these issues in the following section.

5. Follow the Perturbed Leader and Oracle-Efficient Proper Learning

In the previous section, we provided an improper oracle-efficient algorithm that achieves optimal dependence on the horizon T, but is improper and requires more than one oracle call per round; here we demonstrate that a proper learner can have similar regret in some situations with only 1 oracle call per time step. In the following section, we will show that our algorithm's regret is optimal up to a polynomial factor for any oracle-efficient algorithm.

In Rakhlin et al. (2012), the authors make use of the connection between relaxations, random playout, and the Follow the Perturbed Leader (FTPL) style algorithms pioneered in Kalai and Vempala (2005) to make the relaxation approach more efficient in some cases. We expand upon this approach, using Theorem 7 as a starting point. Indeed, the prediction $\widehat{y}_t$ in (4) is cosmetically very similar to that of FTPL, were we recall that the FTPL approach introduces a noise process $\omega(f)$ and, at each time step, sets

$$f_t \in \operatorname*{argmin}_{f \in \mathcal{F}} L_{t-1}(f) + \eta\omega(f)$$

for some real-valued parameter η. The perturbation $\eta\omega(f)$ acts to regularize the predictions; typically, the noise ω is independent across functions, with the classic example being exponential noise in Kalai and Vempala (2005). On the other hand, up to a sign, the supremum in (4) returns the optimal value of $L_t(f) + \eta\omega(f)$ with appropriate values of η and letting ω be the Rademacher process. It is natural to wonder, then, if the min-max problem that is the source of the extra oracle calls is really necessary; we show below that it is not in the sense that FTPL provides an efficient, proper algorithm.

Were we to apply existing FTPL results, using independent perturbations for each f, we would require the enumeration of representative "experts," which would preclude the desired oracle-efficiency. Above, we saw that a Rademacher process perturbation is motivated by the relaxations of the previous section, but analysis is much easier with a Gaussian process. We first treat the case of binary classification:

Theorem 9 *Suppose that $\mathcal{F} : \mathcal{X} \to [-1, 1]$ is a function class and ℓ a loss function that is Lipschitz in both arguments. Suppose further that we are in the smoothed online learning setting where x_i are drawn from a distribution that is σ-smooth with respect to some distribution μ on $\mathcal{X}$. Let*

$$\hat{\omega}_{t,n}(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \gamma_{t,i} f(Z_{t,i})$$

where the $Z_{t,i} \sim \mu$ are independent and the $\gamma_{t,i}$ are indpendent standard normal random variables. Suppose that $\zeta \geq 0$ and consider the algorithm which uses the approximate ERM oracle to choose f_t according to[4]

$$L_{t-1}(f_t) + \eta\hat{\omega}_{t,n}(f_t) \leq \inf_{f \in \mathcal{F}} L_{t-1}(f) + \hat{\omega}_{t,n}(f) + \zeta \tag{5}$$

and let $\widehat{y}_t = f_t(x_t)$. If $\mathcal{F}$ and y_t are binary valued, $\mathsf{vc}(\mathcal{F}) \leq d$, then for appropriate choices of n and η[5]

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim \sqrt{\frac{dT \log T}{\sigma}} + \zeta T$$

More generally, if we let

$$\hat{\omega}_{t,n}(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \gamma_{t,i} \ell(f(Z_{t,i}), y_{t,i})$$

with $y_{t,i}$ drawn uniformly from $\varepsilon\mathbb{Z} \cap [-1, 1]$ and $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ for some $p < 2$, then for appropriate choices of the parameters[6]*, if f_t is chosen according to (5),*

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim T^{\frac{3}{4}} \sigma^{-\frac{1}{4}} \log\left(\frac{T}{\sigma}\right) + \zeta T$$

In order to improve the regret for the case of real-valued labels, we introduce a second, stabilizing perturbation. The following result bounds the regret of the resulting algorithm:

[4] Note that we have not included the total weight multiplying ζ in (5), as in (1); thus we are technically using a $\frac{\zeta}{T + \log(1/\delta) \cdot n}$-approximate ERM oracle with probability $1 - O(\delta)$.

[5] Specified in Proposition 40 in Appendix E.3

[6] Given in Corollary 44 in Appendix E.3.

Theorem 10 *Suppose that $\mathcal{F} : \mathcal{X} \to [-1, 1]$ is a function class and $\ell : [-1, 1] \times [-1, 1] \to [0, 1]$ is a loss function that is L-Lipschitz in* both *arguments. Suppose further that we are in the smooth online learning setting where x_t is chosen from a distribution that is σ-smooth with respect to some μ. Fix $\varepsilon > 0$ and consider the following two processes:*

$$\hat{\omega}_{t,m}(f) = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \gamma_{t,i} f(Z_{t,i}) \qquad \hat{\omega}'_{t,n}(f) = \sum_{j=1}^{n} \gamma'_{t,j} \ell(f(Z'_{t,j}), y'_{t,j})$$

where $\gamma_{t,i}, \gamma'_{t,j}$ are independent standard normal random variables, $Z_{t,i}, Z'_{t,j} \sim \mu$, and $y'_{t,j}$ are independent and uniform on $\varepsilon\mathbb{Z} \cap [-1, 1]$. Suppose that f_t is chosen according to

$$L_{t-1}(f_t) + \eta\hat{\omega}_{t,m}(f_t) + \hat{\omega}_{t,n}(f_t) \leq \inf_{f \in \mathcal{F}} L_{t-1}(f) + \eta\hat{\omega}_{t,m}(f) + \hat{\omega}_{t,n}(f) + \zeta$$

If there is some $p < 2$ such that $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ then for appropriate choices of the parameters[7]*, we have:*

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim \frac{T^{\frac{2}{3}} \log T}{\sigma^{\frac{1}{3}}} + \zeta T$$

If $2 \leq p < \infty$, then there are appropriate choices of parameters such that $\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] = o(T)$.

In Appendix E.3 we provide a slightly more general form of the above regret bounds, including in the case when the labels are smooth with respect to a known measure. Further, we give the precise dependence of our regret bounds on n and η; the optimal values of these parameters are polynomial in $L, T, 1/\sigma$, and the complexity of the function class. Interestingly, we can still achieve regret with the same dependence on T by setting the parameters independently of σ and L; this is useful for applications where we can assume σ-smoothness but do not know what σ is.

The proofs of Theorems 9 and 10 proceed similarly. First, we apply a variant of the classic "Be-the-Leader" approach (Cesa-Bianchi and Lugosi, 2006, Lemma 3.1), which leads to a regret decomposition into a perturbation term and a stability term. The perturbation term is easily controlled with classical empirical process theory. For the stability term, we further decompose the regret into a term quantifying the difference in losses of f_t and f_{t+1} on a tangent sequence and another quantifying the dependence of f_{t+1} on x_t, y_t. For the former term, we prove a novel anti-concentration inequality for the infimum of a Gaussian process, which may be of independent interest, and apply this inequality to control the Wasserstein distance between f_t and f_{t+1}. We bound the latter term in the case of linear loss in a similar way as Haghtalab et al. (2022) did for the corresponding term in their algorithm. This suffices for the binary labels case, but to extend to the more general setting, we use a discretization scheme to reduce to the case that the labels are also chosen in a smooth manner with respect to some distribution; we then reduce this setting to the case of linear loss and apply our earlier bound. The details can be found in Appendix E.

The stability estimate was a significant technical challenge due to the complex dependence structure of $\hat{\omega}_{t,n}$ accross $\mathcal{F}$; most regret bounds for FTPL-style algorithms are simplified by independent perturbations. To the best of our knowledge, Theorem 10 constitutes the first proof of an FTPL regret bound where the algorithm uses a generic Gaussian process as the perturbation.

[7] Outlined in Corollary 43 in Appendix E.3.

6. Computational Lower Bounds

Comparing the results of Theorem 3 and Theorem 10, we notice that the requirement of oracle efficiency incurs an exponential loss in the regret's dependence on σ. In this section, we show that this exponential gap is necessary for any oracle-efficient algorithm.

Theorem 11 *Fix any $T \in \mathbb{N}$ and $\sigma \in (0, 1]$. In the ERM oracle model, any randomized algorithm cannot guarantee expected regret smaller than $\frac{T}{200}$ against a σ-smooth online adversary over T rounds and any binary $\mathcal{F}$ with $|\mathcal{F}| \leq 1/\sigma$ in total time smaller than $\widetilde{O}(1/\sqrt{\sigma})$.*

Theorem 11 is proved in Appendix F by constructing a family of function classes on a space $\mathcal{X}$ of size $1/\sigma$ and noting that a worst-case adaptive adversary is σ-smooth in this setting. The construction then mirrors that in Hazan and Koren (2016), which reduces from Aldous' problem (Aldous, 1983); the main difference being that Definition 2 allows for negative weights in the ERM oracle, which complicates the proof.

As an immediate corollary of Theorem 11, we obtain the following regret lower bound for computationally efficient algorithms in the ERM oracle model, i.e., those whose total time after T time steps is $\mathrm{poly}(T)$:

Corollary 12 *Fix any $\alpha \geq 1$, $\varepsilon < 1/200, \sigma \in (0, 1]$, and $d \geq \log 1/\sigma$. Any algorithm whose total time in the ERM oracle model over T rounds is bounded as T^α requires that $T \geq \widetilde{\Omega}\left(\max\left\{\frac{d}{\varepsilon^2}, \sigma^{-\frac{1}{2\alpha}}\right\}\right)$ to achieve regret εT for classes $\mathcal{F}$ of VC dimension d against a σ-smooth adversary.*

Corollary 12 and Theorem 3 show that there is an exponential statistical-computational gap for smoothed online learning in the ERM oracle model: for general classes $\mathcal{F}$, it is possible to achieve regret proportional to $\log(1/\sigma)$, but the regret must be polynomial in $1/\sigma$ if the algorithm is required to be oracle-efficient. While Theorem 11 and Corollary 12 get a lower bound only on the total computation time, as opposed to the number of oracle calls, we provide analogous results obtaining lower bounds on the number of oracle calls with an *approximate* ERM oracle in Theorem 52 and Corollary 53 in the appendix. In particular, we show that any algorithm with $T^{O(1)}$ oracle calls with a $1/T^{O(1)}$-approximate ERM oracle needs $T \geq \max\{d, 1/\varepsilon, 1/\sigma\}^{\Omega(1)}$ to obtain sublinear regret against binary classes of VC dimension d.

Acknowledgments

AB acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 1122374. NG is supported by a Fannie & John Hertz Foundation Fellowship and an NSF Graduate Fellowship. AR acknowledges support from the ONR through awards N00014-20-1-2336 and N00014-20-1-2394, from the NSF through awards DMS-2031883 and DMS-1953181, and from the ARO through award W911NF-21-1-0328.

References

Scott Aaronson. Lower bounds for local search by quantum arguments. *SIAM Journal on Computing*, 35(4):804–824, 2006.

Jacob Abernethy, Chansoo Lee, Abhinav Sinha, and Ambuj Tewari. Online linear optimization via smoothing. In *Conference on Learning Theory*, pages 807–823. PMLR, 2014.

Jacob Abernethy, Chansoo Lee, and Ambuj Tewari. Fighting bandits with a new kind of smoothness. *arXiv preprint arXiv:1512.04152*, 2015.

Alekh Agarwal, Dean P Foster, Daniel J Hsu, Sham M Kakade, and Alexander Rakhlin. Stochastic convex optimization with bandit feedback. *Advances in Neural Information Processing Systems*, 24:1035–1043, 2011.

Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization oracle. In *Conference on Learning Theory*, pages 18–29. PMLR, 2019.

David Aldous. Minimization algorithms and random walk on the d-cube. *The Annals of Probability*, 11(2):403–413, 1983.

Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

David Arthur and Sergei Vassilvitskii. Worst-case and smoothed analysis of the icp algorithm, with an application to the k-means method. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 153–164, 2006.

Peter L Bartlett, Philip M Long, and Robert C Williamson. Fat-shattering and the learnability of real-valued functions. *journal of computer and system sciences*, 52(3):434–452, 1996.

Rene Beier and Berthold Vöcking. Typical properties of winners and losers in discrete optimization. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '04, page 343–352, New York, NY, USA, 2004. Association for Computing Machinery.

Adam Block, Yuval Dagan, and Alexander Rakhlin. Majorizing measures, sequential complexities, and online learning. In Mikhail Belkin and Samory Kpotufe, editors, *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pages 587–590. PMLR, 15–19 Aug 2021. URL https://proceedings.mlr.press/v134/block21a.html.

Shant Boodaghians, Rucha Kulkarni, and Ruta Mehta. Smoothed efficient algorithms and reductions for network coordination games. In Thomas Vidick, editor, *11th Innovations in Theoretical Computer Science Conference, ITCS 2020, January 12-14, 2020, Seattle, Washington, USA*, volume 151 of *LIPIcs*, pages 73:1–73:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

Nicolo Cesa-Bianchi and Gábor Lugosi. On prediction of individual sequences. *The Annals of Statistics*, 27(6):1865–1895, 1999.

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

Nicolo Cesa-Bianchi, Alex Conconi, and Claudio Gentile. On the generalization ability of on-line learning algorithms. *IEEE Transactions on Information Theory*, 50(9):2050–2057, 2004.

Gustave Choquet and Paul-André Meyer. Existence et unicité des représentations intégrales dans les convexes compacts quelconques. In *Annales de l'institut Fourier*, volume 13, pages 139–154, 1963.

Alon Cohen and Tamir Hazan. Following the perturbed leader for online structured learning. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, page 1034–1042. JMLR.org, 2015.

Constantinos Daskalakis and Noah Golowich. Fast rates for nonparametric online learning: From realizability to learning in games. *arXiv preprint arXiv:2111.08911*, 2021.

Luc Devroye, Gábor Lugosi, and Gergely Neu. Prediction by random-walk perturbation. In *Conference on Learning Theory*, pages 460–473. PMLR, 2013.

Miroslav Dudík, Nika Haghtalab, Haipeng Luo, Robert E Schapire, Vasilis Syrgkanis, and Jennifer Wortman Vaughan. Oracle-efficient online learning and auction design. In *2017 ieee 58th annual symposium on foundations of computer science (focs)*, pages 528–539. IEEE, 2017.

Richard M Dudley. The sizes of compact subsets of hilbert space and continuity of gaussian processes. *Journal of Functional Analysis*, 1(3):290–330, 1967.

Michael Etscheid and Heiko Röglin. Smoothed analysis of local search for the maximum-cut problem. *ACM Trans. Algorithms*, 13(2), mar 2017. ISSN 1549-6325.

Xavier Fernique. Regularité des trajectoires des fonctions aléatoires gaussiennes. In *Ecole d'Eté de Probabilités de Saint-Flour IV—1974*, pages 1–96. Springer, 1975.

Dylan Foster and Alexander Rakhlin. Beyond ucb: Optimal and efficient contextual bandits with regression oracles. In *International Conference on Machine Learning*, pages 3199–3210. PMLR, 2020.

Dylan J Foster, Alexander Rakhlin, David Simchi-Levi, and Yunzong Xu. Instance-dependent complexity of contextual bandits and reinforcement learning: A disagreement-based perspective. *arXiv preprint arXiv:2010.03104*, 2020.

Dylan J Foster, Sham M Kakade, Jian Qian, and Alexander Rakhlin. The statistical complexity of interactive decision making. *arXiv preprint arXiv:2112.13487*, 2021.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

Pierre Gaillard and Sébastien Gerchinovitz. A chaining algorithm for online nonparametric regression. In *Conference on Learning Theory*, pages 764–796. PMLR, 2015.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis of online and differentially private learning. *arXiv preprint arXiv:2006.10129*, 2020.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *arXiv preprint arXiv:2102.08446*, 2021.

Nika Haghtalab, Yanjun Han, Abhishek Shetty, and Kunhe Yang. Oracle-efficient online learning for beyond worst-case adversaries. *arXiv preprint arXiv:2202.08549*, 2022.

James Hannan. 4. approximation to rayes risk in repeated play. In *Contributions to the Theory of Games (AM-39), Volume III*, pages 97–140. Princeton University Press, 2016.

Elad Hazan and Tomer Koren. The computational power of optimization in online learning. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 128–141, 2016.

David P. Helmbold and Manfred K. Warmuth. Learning permutations with exponential weights. In *COLT*, 2007.

Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Michael J Kearns and Robert E Schapire. Efficient distribution-free learning of probabilistic concepts. *Journal of Computer and System Sciences*, 48(3):464–497, 1994.

Victor Klee and George J Minty. How good is the simplex algorithm. *Inequalities*, 3(3):159–175, 1972.

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

Bodo Manthey. *Smoothed Analysis of Local Search*, pages 285–308. Cambridge University Press, United States, December 2020.

Shahar Mendelson. Rademacher averages and phase transitions in glivenko-cantelli classes. *IEEE transactions on Information Theory*, 48(1):251–263, 2002.

Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pages 993–1019, Princeton, NJ, USA, 12–14 Jun 2013. PMLR. URL https://proceedings.mlr.press/v30/Rakhlin13.html.

Alexander Rakhlin and Karthik Sridharan. Statistical learning and sequential prediction. *Book Draft*, 2014.

Alexander Rakhlin and Karthik Sridharan. Online nonparametric regression with general loss functions. *arXiv preprint arXiv:1501.06598*, 2015.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. *Advances in neural information processing systems*, 24:1764–1772, 2011.

Alexander Rakhlin, Ohad Shamir, and Karthik Sridharan. Relax and randomize: From value to algorithms. *Advances in neural information processing systems*, 25, 2012.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning via sequential complexities. *J. Mach. Learn. Res.*, 16(1):155–186, 2015a.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Sequential complexities and uniform martingale laws of large numbers. *Probability Theory and Related Fields*, 161(1-2):111–153, 2015b.

Alexander Rakhlin, Karthik Sridharan, and Alexandre B Tsybakov. Empirical entropy, minimax regret and minimax risk. *Bernoulli*, 23(2):789–824, 2017.

Tim Roughgarden, editor. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2021.

Mark Rudelson and Roman Vershynin. Combinatorics of random processes and sections of convex bodies. *Annals of Mathematics*, pages 603–648, 2006.

Norbert Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13 (1):145–147, 1972.

Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and trends in Machine Learning*, 4(2):107–194, 2011.

Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41(1):247–261, 1972.

David Simchi-Levi and Yunzong Xu. Bypassing the monster: A faster and simpler optimal algorithm for contextual bandits under realizability. *Mathematics of Operations Research*, 2021.

Daniel A Spielman and Shang-Hua Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 81–90, 2004.

Nathan Srebro, Karthik Sridharan, and Ambuj Tewari. Smoothness, low noise and fast rates. *Advances in neural information processing systems*, 23, 2010.

Vladimir Nikolaevich Sudakov. Gaussian random processes and measures of solid angles in hilbert space. In *Doklady Akademii Nauk*, volume 197, pages 43–45. Russian Academy of Sciences, 1971.

Eiji Takimoto and Manfred K. Warmuth. Path kernels and multiplicative updates. In *Proceedings of the 15th Annual Conference on Computational Learning Theory*, COLT '02, page 74–89, Berlin, Heidelberg, 2002. Springer-Verlag.

Vladimir Vapnik and Alexey Chervonenkis. *Theory of pattern recognition*. Nauka, Moscow, 1974.

Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Appendix A. Applications to Contextual Bandits

We apply our results to the study of contextual bandits. A series of recent papers (Foster and Rakhlin, 2020; Simchi-Levi and Xu, 2021; Foster et al., 2020) has focused on reducing the contextual bandit framework to that of online learning, with Foster and Rakhlin (2020) introducing an efficient and optimal reduction, SquareCB, that turns an online regression oracle into a fast, no-regret contextual bandit algorithm. One of the key advantages of this reduction is the fact that the learner "only" has to design algorithms that have small regret in the full-information setting, thought to be an easier task than one that requires a careful balance of exploration and exploitation. Unfortunately, there is still a dearth of oracle-efficient online algorithms with provably good regret in general, limiting the broader application of these results. In Simchi-Levi and Xu (2021), the authors use a similar reduction, but with an offline regression oracle, for which there are many practical algorithms; unfortunately, the result requires the contexts to arrive in i.i.d. fashion, unlike the more general result of Foster and Rakhlin (2020). Here, we show that whenever a function class is learnable in the offline setting, we can still use the SquareCB reduction to get a no-regret algorithm that is efficient with respect to an ERM oracle in the smooth contextual bandit setting.

We consider the setting described in Foster et al. (2020) with the modification that contexts arrive in a σ-smooth manner. Formally, at each $1 \leq t \leq T$, Nature selects a σ-smooth distribution p_t, and samples $x_t \sim p_t$, then samples a loss function ℓ_t independently from some distribution depending on x_t. The learner selects an action $a_t \in [K]$ and observes $\ell_t(a_t)$. We are given a function class $\mathcal{F} : \mathcal{X} \times [K] \to [0,1]$ and suppose that there is some unknown $f^* \in \mathcal{F}$ such that $\mathbb{E}\left[\ell_t(a) | x_t = x\right] = f^*(x,a)$ for all $x \in \mathcal{X}$ and $a \in [K]$. The goal is to minimize regret to the best policy induced by $\mathcal{F}$, where for any $f \in \mathcal{F}$, we define $\pi_f(x) = \operatorname{argmin}_{a \in [K]} f(x,a)$, i.e., we wish to minimize

$$\mathrm{Reg}_{CB}(T) = \sum_{t=1}^{T} \ell_t(a_t) - \ell_t(\pi_{f^*}(x_t))$$

We have the following result:

Theorem 13 *Suppose we are in the σ-smooth Contextual Bandit setting described above. If we run SquareCB with the relaxation-induced online regressor from (4), we can achieve*

$$\mathbb{E}\left[\mathrm{Reg}_{CB}(T)\right] \leq 12 \frac{K \log T}{\sqrt{\sigma}} \sqrt{T \mathcal{R}_T(\mathcal{F})}$$

with $O\left(\sqrt{T} \log T\right)$ calls to the ERM oracle per round. If we instead instantiate SquareCB with the FTPL algorithm from Theorem 10 and $\mathcal{R}_T(\mathcal{F}) = o(T)$, then $\mathbb{E}\left[\mathrm{Reg}_{CB}(T)\right] = o(T)$ as well with only 1 call to the ERM oracle per round.

To prove Theorem 13, we observe that if the contexts arrive in a σ-smooth manner with respect to μ, then the context-action pairs can be taken to be $\frac{\sigma}{K}$-smooth with respect to $\mu \otimes \mathrm{Unif}([K])$. We then apply (Foster and Rakhlin, 2020, Theorem 1). The details and precise rates in the case of the FTPL instantiation can be found in Appendix G.

Note that the regret bound does not have optimal rates with respect to either T or K. In order to recover optimal rates with respect to T, we would need to find an algorithm that exhibits fast rates with square loss in the smoothed-online setting. This is an interesting further direction in its own right, in addition to the practical implications on better rates for efficient algorithms for contextual bandits.

Appendix B. Related Work

Here we describe some recent of the recent literature and how it relates to our work.

Smoothed Analysis. Smoothed analysis was first introduced in Spielman and Teng (2004), where it was proposed as an explanation for the gap between theoretical lower bounds and excellent empirical performance of the simplex algorithm (Klee and Minty, 1972). Since then, smoothed analysis has been applied to analyze the performance of algorithms for many other problems which are known to be hard in the worst-case, such as the k-means algorithm for clustering (Arthur and Vassilvitskii, 2006), the flip algorithm for finding a local max-cut (Etscheid and Röglin, 2017), and more generally better-response algorithms for finding Nash equilibria in network coordination games (Boodaghians et al., 2020) (see also Roughgarden (2021) for a more comprehensive overview).

In the context of learning theory, Rakhlin et al. (2011) gave a nonconstructive proof demonstrating its utility for the specific case of threshold functions, while Haghtalab et al. (2020, 2021) proved that the minimax regret of binary classification in smoothed online learning is governed by VC dimension.

Online Learning. The optimal statistical rates attainable by online learning algorithms was shown to be characterized by sequential complexity measures of the function class in Rakhlin et al. (2015b); Rakhlin and Sridharan (2013); Rakhlin et al. (2015a). This characterization was extended to the case of constrained adversaries (including the special case of smoothed adversaries) in Rakhlin et al. (2011). Several subsequent papers have established further refined regret bounds Block et al. (2021); Rakhlin and Sridharan (2015). The profusion of publications relating to algorithmic questions about online learning is too large to enumerate here, but notable relevant work includes Hazan and Koren (2016), which provides lower bounds on oracle-efficiency and Rakhlin et al. (2012) which introduces a general framework for constructing algorithms.

Follow The Perturbed Leader. Our proper learning algorithm is motivated by Follow the Perturbed Leader (FTPL) Kalai and Vempala (2005); Hannan (2016). FTPL has been successful for many problems, including learning from experts (Kalai and Vempala, 2005), multi-armed bandits (Abernethy et al., 2015), and online structured learning (Cohen and Hazan, 2015), which includes as special cases problems such as online shortest path (Takimoto and Warmuth, 2002) and online learning of permutations (Helmbold and Warmuth, 2007). There is a similar diversity in methods of proving regret bounds for FTPL style algorithms, including potential-based analysis (Abernethy et al., 2014; Cohen and Hazan, 2015) and relaxation methods (Rakhlin et al., 2011). A common approach, which we adopt, is to show that the algorithm is stable (Kalai and Vempala, 2005; Agarwal et al., 2019; Devroye et al., 2013; Agarwal et al., 2011). One of the primary advantages of our FTPL approach is the fact that we do not generate independent noise for each function in our class. In (Dudík et al., 2017), the authors present an FTPL-style algorithm which aims to do something similar, mitigating the computational burden by sharing randomness between functions. Their method, however, is very different from ours in that they rely on their new notions of *admissability* and *implementability* of a matrix to transform low-dimensional independent noise into a more structured form; in contradistinction, we directly use a Gaussian Process on the function class to ensure stability of our algorithm.

Contextual Bandits. There is a rich history of studying contextual bandits. Most relevant to our work is the series of papers Foster and Rakhlin (2020); Simchi-Levi and Xu (2021); Foster et al.

(2020) which provides a reduction from contextual bandits to an online learning oracle. See these papers for further references.

Appendix C. Proofs from Section 3

C.1. Proofs Related to the Coupling

We first extend (Haghtalab et al., 2021, Theorem 2.1) by providing a simpler and more general proof of the coupling between $\mathscr{D} \in \mathfrak{P}_T(\sigma, \mu)$ and independent random variables drawn according to μ. While we use a slightly different version (Lemma 24) in the proof of Theorem 3, the following lemma is both simpler for exposition and is used in the proofs of the results in Section 4.

Lemma 14 *Suppose that $\mathscr{D} \in \mathfrak{P}_T(\sigma, \mu)$. Then for any T there exists a measure Π with random variables $(x_t, Z_t^j)_{\substack{1 \le t \le T \\ 1 \le j \le k}}$ satisfying the following properties:*

1. *x_t is distributed according to $p_t(\cdot | x_1, \dots, x_{t-1})$ induced by $\mathscr{D}$.*
2. *$\{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}}$ are iid according to μ*
3. *With probability at least $1 - Te^{-\sigma k}$, we have $x_t \in \{Z_t^j\}_{1 \le j \le k}$ for all t*

Proof We construct the coupling recursively. For any t, suppose that Z_s^j, x_s has been constructed for $s < t$. If $t = 0$ then this is the empty set. Now, sample Z_t^j iid according to μ. Let $\pi_t^j = \sigma \frac{dp_t}{d\mu}(Z_t^j)$. Note that $\pi_t^j \le 1$ by the assuption of σ-smoothness. Construct the random set $S_t \subset [k]$ by adding j to S_t iwth probability π_t^j independently for each $1 \le j \le k$. If S_t is nonempty, then sample x_t uniformly from S_t. Otherwise, sample x_t independently from p_t. We now show that this process exhibits the desired properties.

It is clear form the construction that Z_t^j are iid according to μ. To verify that $x_t \in \{Z_t^j\}$, we note that for any t, j, we have

$$\mathbb{P}(Z_t^j \in S_t) = \mathbb{E}_\mu\left[\sigma \frac{dp_t}{d\mu}(Z_t^j)\right] = \sigma \tag{6}$$

Because the Z_t^j are added to S_t independently, the probability that S_t is empty is $(1 - \sigma)^k$. Thus, by a union bound, the probability that there exists some $t \le T$ such that any S_t is empty is bounded by $T(1 - \sigma)^k \le Te^{-\sigma k}$.

Finally, to see that x_t are distributed according to p_t, let $A \subset \mathcal{X}$ be measurable and χ_A denote the indicator for A. We compute:

$$\begin{aligned}
\mathbb{P}(Z_t^j \in A | Z_t^j \in S_t) &= \frac{\mathbb{P}\left(Z_t^j \in A \text{ and } Z_t^j \in S_t\right)}{\mathbb{P}\left(Z_t^j \in S_t\right)} \\
&= \frac{\mathbb{P}\left(Z_t^j \in A \text{ and } Z_t^j \in S_t\right)}{\sigma} \\
&= \frac{\mathbb{E}_\mu\left[\chi_A \sigma \frac{dp_t}{d\mu}(Z_t^j)\right]}{\sigma} \\
&= \mathbb{E}_{p_t}\left[\mathbf{1}_A\right] = p_t(A)
\end{aligned}$$

where the second equality follows from (6), the third equality follows from the construction of S_t, and the penultimate equality following from the definnition of the Radon-Nikodym derivative. The result follows. ■

We further note that the coupling in Lemma 14 is optimal with respect to the dependence on k in the third requirement, as seen in the following proposition.

Proposition 15 *For any $\sigma < 1$ and non-atomic measure μ on $\mathcal{X}$, there exists a measure on $\mathcal{X}$, p such that p is σ-smooth with respect to μ and the following property holds. For any coupling Π which has random variables Z^j for $1 \leq j \leq k$ and X such that $Z^j \sim \mu$ are independent and $X \sim p$, the probability that $X \in \{Z^j\}$ is bounded below by $1 - (1 - \sigma)^k$.*

Proof Given μ, let $A \subset \mathcal{X}$ denote a measurable set such that $\mu(A) = \sigma$. Let p be a measure on $\mathcal{X}$ such that $\frac{dp}{d\mu} = \frac{1}{\sigma}\chi_A$. Then p is σ-smooth with respect to μ. Note that if $Z^j \sim \mu$ then with probability $1 - \sigma$, $Z^j \notin A$. Thus with probability $(1 - \sigma)^k$, none of Z^j are in A. Thus with probability at least $(1 - \sigma)^k$, $X \notin \{Z^j | 1 \leq j \leq k\}$ if $X \sim p$. The result follows. ■

C.2. Preliminaries on Distribution-Dependent Sequential Rademacher Complexity

In this section, we recall the definition of the distribution-dependent sequential Rademacher complexity from Rakhlin et al. (2011) and how it relates to the minimax regret. To begin, we formally construct a measure $\rho_{\mathscr{D}}$ used in the definition of distribution-dependent sequential Rademacher complexity from Rakhlin et al. (2011).

Throughout, we follow Rakhlin et al. (2015b, 2011) and introduce as our basic object in analyzing sequential complexities a tree. Specifically, we consider complete binary trees $\mathbf{z}$ of depth T with each vertex of the tree labelled by some element of $\mathcal{X}$. We associate each $\varepsilon \in \{\pm 1\}^T$ to a path in the tree from the root to a leaf, where the path is constructed recursively by beginning at the root and at each level going to the left if $\varepsilon_{t-1} = 1$ and to the right otherwise. For a given tree $\mathbf{z}$, we denote by $\mathbf{z}_t(\varepsilon)$ the label of the t^{th} vertex along the path ε.

Let $\mathscr{D}$ be the joint distribution of $z_1, \dots, z_T \in \mathcal{Z}$. Define $p_t(\cdot, | z_1, \dots, z_{t-1})$ as the distribution under $\mathscr{D}$ of z_t, given z_s for $s < t$. We recursively construct the measure $\rho_{\mathscr{D}}$ on pairs of binary trees as follows. We first construct the roots of each tree by sampling $\mathbf{z}_0(\varepsilon), \mathbf{z}'_0(\varepsilon) \sim p_0$ independently. Suppose we have $\mathbf{z}_{1:t-1}, \mathbf{z}'_{1:t-1}$ already constructed. For any $s < t$, let

$$\chi_s(\varepsilon) = \begin{cases} \mathbf{z}_s(\varepsilon) & \varepsilon_s = 1 \\ \mathbf{z}'_s(\varepsilon) & \varepsilon_s = -1 \end{cases}$$

then sample $\mathbf{z}_t(\varepsilon), \mathbf{z}'_t(\varepsilon)$ independently from $p_t(\cdot | \chi_1(\varepsilon), \dots, \chi_{t-1}(\varepsilon))$. In this way, we can recursively construct the measure $\rho_{\mathscr{D}}$.

With the definition of $\rho_{\mathscr{D}}$ completed, we can now define the key notion of complexity.

Definition 16 (Definition 2 from Rakhlin et al. (2011)) *Given a space $\mathcal{Z}$, a function class $\mathcal{F} \subset [-1, 1]^{\mathcal{Z}}$, and a joint distribution $\mathscr{D}$, let $\rho_{\mathscr{D}}$ be the measure on an ordered pair of binary trees of depth T with values in $\mathcal{Z}$, defined above. Then, we define the distribution-dependent sequential Rademacher complexities as*

$$\mathcal{R}_T^{seq}(\mathcal{F}, \mathscr{D}) = \mathbb{E}_{(\mathbf{z}, \mathbf{z}') \sim \rho_{\mathscr{D}}} \mathbb{E}_{\varepsilon} \left[\sup_{f \in \mathcal{F}} \sum_{t=1}^{T} \varepsilon_t f(\mathbf{z}_t(\varepsilon)) \right]$$

If $\mathfrak{P}$ is a class of distributions $\mathscr{D}$, we define

$$\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P}) = \sup_{\mathscr{D} \in \mathfrak{P}} \mathcal{R}_T^{seq}(\mathcal{F}, \mathscr{D})$$

for any class of distributions $\mathfrak{P}$.

Intuitively, depending on the nature of the class $\mathfrak{P}$, $\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P})$ interpolates between the classical batch Rademacher complexity (if we force $\mathscr{D}$ to be iid) and the fully adversarial sequential Rademacher complexity from Rakhlin et al. (2015b). In the special case that $\mathfrak{P} = \mathfrak{P}(\sigma, \mu)$, we see that we are much closer to the classical Rademacher complexity than to the fully adversarial analogue. Indeed, using Lemma 24, which is an extension of the coupling result contained in Lemma 14 above, we can bound the distribution-dependent sequential Rademacher complexity by that of the classical Rademacher complexity:

Lemma 17 *Let $\mathcal{F} \subset [-1,1]^{\mathcal{X}}$ be a function class. Then, for any $k \in \mathbb{N}$,*

$$\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \leq \left(\frac{4}{\sigma} \log T \right) \mathbb{E}_\mu \left[\mathcal{R}_{kT}(\mathcal{F}) \right] + 2T^2 e^{-\sigma k}$$

In particular, in the case that $\mathsf{vc}(\mathcal{F}) \leq d$, we have:

$$\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \lesssim \sqrt{T d \log\left(\frac{T}{\sigma} \right)}$$

and in the case that $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$, we have:

$$\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \lesssim \left(T \log\left(\frac{T}{\sigma} \right) \right)^{\max\left(\frac{1}{2}, 1 - \frac{1}{p}\right)}$$

Proof Let A be the high probability event in Lemma 24 below, i.e., the event that $x_t \in \left\{ Z_t^j \right\}_{1 \leq j \leq k}$ for all t. We have for any $\mathscr{D} \in \mathfrak{P}(\sigma, \mu)$,

$$\begin{aligned}
\mathcal{R}_T^{seq}(\mathcal{F}, \mathscr{D}) &= \mathbb{E}_{(\mathbf{z}, \mathbf{z}') \sim \rho_{\mathscr{D}}} \mathbb{E}_\varepsilon \left[\sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t f(\mathbf{z}_t(\varepsilon)) \right] \\
&= \mathbb{E}_\Pi \mathbb{E}_\varepsilon \left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t f(\mathbf{z}_t(\varepsilon)) \right] + \mathbb{E}_\Pi \mathbb{E}_\varepsilon \left[\chi_{A^c} \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t f(\mathbf{z}_t(\varepsilon)) \right] \\
&\leq \mathbb{E}_\Pi \mathbb{E}_\varepsilon \left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t f(\mathbf{z}_t(\varepsilon)) + \sum_{j \text{ such that } Z_t^j \neq \mathbf{z}_t(\varepsilon)} \mathbb{E}_{\varepsilon_{t,j}} \left[\varepsilon_{t,j} f(Z_t^j) \right] \right] + 2T^2 e^{-\sigma k} \\
&\leq \mathbb{E}_\Pi \mathbb{E}_\varepsilon \left[\sup_{f \in \mathcal{F}} \sum_{j=1}^k \sum_{t=1}^T \varepsilon_{t,j} f(Z_t^j) \right] + 2T^2 e^{-\sigma k} \\
&\leq 2T^2 e^{-\sigma k} + \mathbb{E}_\mu \left[\mathcal{R}_{kT}(\mathcal{F}) \right]
\end{aligned}$$

where Π is the coupling in Lemma 24, the first inequality follows because ε_t is mean zero, the second inequality follows by Jensen's, and the lastfollows by definition of Rademacher complexity. Setting $k = \frac{2}{\sigma} \log T$ concludes the proof. ■

As is the case in both the fully adversarial and classical regimes, we see that $\mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P})$ is determined up to constants by $\mathcal{R}_T^{seq}(\mathcal{F}, \mathfrak{P})$:

Proposition 18 (Theorem 3 and Lemma 20 from Rakhlin et al. (2011)) *For any $\mathcal{F}$, we have*

$$\mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P}(\sigma, \mu)) \le 2 \sup_{\mathscr{D} \in \widetilde{\mathfrak{P}}_T(\sigma,\mu)} \mathcal{R}_T^{seq}(\ell \circ \mathcal{F}, \mathscr{D}).$$

where we recall from Definition 1 that $\widetilde{\mathfrak{P}}_T(\sigma, \mu)$ is the class of distributions on (x_t, y_t) such that the x_t are chosen in a σ-smooth way and the y_t are adversarial. In the special case where ℓ is absolute loss, we also have

$$\sup_{\mathscr{D} \in \widetilde{\mathfrak{P}}_T(\sigma,\mu)} \mathcal{R}_T^{seq}(\ell \circ \mathcal{F}, \mathscr{D}) \le \mathcal{V}_T^{prop}(\mathcal{F}, \mathfrak{P}(\sigma, \mu))$$

In the statement of Proposition 18, $\ell \circ \mathcal{F}$ denotes the class of functions in $[0, 1]^{\mathcal{X} \times [-1,1]}$ of the form $(x, y) \mapsto \ell(f(x), y)$, for $f \in \mathcal{F}$. By Proposition 18, it suffices to provide upper and lower bounds on $\mathcal{R}_T^{seq}(\ell \circ \mathcal{F}, \widetilde{\mathfrak{P}}_T(\sigma, \mu))$, which is significantly more tractable than working with the iterated operators involved in $\mathcal{V}_T^{\text{prop}}(\mathcal{F}, \mathfrak{P})$.

In the proof below, we will also need a sequential analogue of $\mathsf{vc}(\mathcal{F}, \alpha)$:

Definition 19 (Definition 7 from Rakhlin et al. (2015b)) *We say that a $\mathcal{X}$-valued binary tree of depth T, $\mathbf{x}$, is shattered by $\mathcal{F}$ at scale $\delta \ge 0$ if there exists an $\mathbb{R}$-valued binary tree $\mathbf{s}$ of depth T such that for all $\varepsilon \in \{\pm 1\}^T$, there exists an $f_\varepsilon \in \mathcal{F}$ such that*

$$\varepsilon_t \left(f(\mathbf{x}_t(\varepsilon)) - \mathbf{s}_t(\varepsilon) \right) \ge \frac{\alpha}{2}$$

Define the sequential fat-shattering dimension of $\mathcal{F}$, $\mathsf{fat}_\delta(\mathcal{F})$ as the maximal T such that there exists a tree of depth T shattering $\mathcal{F}$ at scale δ.

Finally, we require a structural result showing that worst-case sequential Rademacher complexity contracts with Lipschitz loss functions:

Lemma 20 (Lemma 13 from Rakhlin et al. (2015b)) *Let $\mathcal{F}$ be a function class with values in $[-1, 1]$ and let ℓ be L-Lipschitz. Then,*

$$\sup_{\mathscr{D} \in \Delta(\mathcal{X}^{\times T})} \mathcal{R}_T^{seq}(\ell \circ \mathcal{F}, \mathscr{D}) \lesssim L \log^{\frac{3}{2}}(T) \sup_{\mathscr{D} \in \Delta(\mathcal{X}^{\times T})} \mathcal{R}_T^{seq}(\mathcal{F}, \mathscr{D})$$

Because the supremum in Lemma 20 is taken over all distributions on the product space $\mathcal{X}^{\times T}$, the above distribution-dependent sequential Rademacher complexities are reduced to the adversarial sequential Rademacher complexities of Rakhlin et al. (2015b). In the following section, we show that on small domains, we can control $\mathsf{fat}_\delta(\mathcal{F})$ by $\mathsf{vc}(\mathcal{F}, \delta)$.

C.3. Sequential and Batch Complexities

In this section, we prove the following lemma, which bounds the sequential fat-shattering dimension by the scale-sensitive VC dimension when the domain is small:

Lemma 21 *Let $\mathcal{F}$ be a function class from $\mathcal{X}$ to $[-1,1]$ and let $\mathsf{fat}_\delta(\mathcal{F})$ denote the sequential fat-shattering dimension of $\mathcal{F}$ (Definition 19). Then, for any $\alpha > 0$,*

$$\mathsf{fat}_\delta(\mathcal{F}) \lesssim \mathsf{vc}\left(\mathcal{F}, c\alpha\delta\right) \log^{1+\alpha}\left(\frac{C\left|\mathcal{X}\right|}{\mathsf{vc}\left(\mathcal{F}, c\delta\right)\delta}\right)$$

In order to prove this result, we require a generalization of the Sauer-Shelah lemma (Sauer, 1972; Shelah, 1972). We first define covering numbers with respect to the $\sup$ norm:

Definition 22 *Let $\mathcal{F}$ be a class of functions on $\mathcal{X}$. A set S of functions on $\mathcal{X}$ is a δ covering if for all $f \in \mathcal{F}$, there exists a $s_f \in S$ such that*

$$\sup_{x \in \mathcal{X}} \left| s_f(x) - f(x) \right| \leq \delta$$

We let $N(\mathcal{F}, \delta)$ to be the minimal size of a δ-covering of $\mathcal{F}$.

In order to bound $\mathsf{fat}_\delta(\mathcal{F})$ by $\mathsf{vc}(\mathcal{F}, \delta)$, we first recall a result that bounds $N(\mathcal{F}, \delta)$ by $\mathsf{vc}(\mathcal{F}, \delta)$:

Theorem 23 (Theorem 4.4 from Rudelson and Vershynin (2006)) *Let $\mathcal{F}$ be a function class on $\mathcal{X}$, a finite set, to $[-1,1]$. Then for any $\alpha > 0$, there are constants $c, C > 0$ such that*

$$\log N(\mathcal{F}, \delta) \lesssim \mathsf{vc}(\mathcal{F}, c\alpha\delta) \log^{1+\alpha}\left(\frac{C\left|\mathcal{X}\right|}{\mathsf{vc}(\mathcal{F}, c\delta)\delta}\right)$$

The above theorem is an intermediate result, so our bound will come down to comparing $\mathsf{fat}_\delta(\mathcal{F})$ to the covering numbers. We can now provide the main proof in the section.

Proof [(Lemma 21)] We first note that $2^{\mathsf{fat}_\delta(\mathcal{F})} \leq N\left(\mathcal{F}, \frac{\delta}{3}\right)$. To see this, let $d = \mathsf{fat}_\delta(\mathcal{F})$ and let $\mathbf{x}$ denote a depth d tree that shatters $\mathcal{F}$ at scale δ with witness tree $\mathbf{s}$. Let S be a $\frac{\delta}{2}$ net for $\mathcal{F}$. For each $\varepsilon \in \{\pm 1\}^d$, let f_ε be the function that realizes the shattering on path ε. If $v_{f_\varepsilon} \in \mathcal{F}$ is the projection into S, then we note that the function $\varepsilon \mapsto v_{f_\varepsilon}$ is injective. Indeed, if there are two different $\varepsilon, \varepsilon'$ mapping to the same $v \in S$, then there is some t such that $\varepsilon_t = -\varepsilon_t'$ but $\varepsilon_s = \varepsilon_s'$ for $s < t$. Thus $\mathbf{x}_t(\varepsilon) = \mathbf{x}_t(\varepsilon')$ and $\mathbf{s}_t(\varepsilon) = \mathbf{s}_t(\varepsilon')$. We know, however, that

$$\left| f_\varepsilon(\mathbf{x}_t(\varepsilon)) - f_{\varepsilon'}(\mathbf{x}_t(\varepsilon')) \right| \leq \left| f_\varepsilon(\mathbf{x}_t(\varepsilon)) - v(\mathbf{x}_t(\varepsilon)) \right| + \left| v(\mathbf{x}_t(\varepsilon)) - f_{\varepsilon'}(\mathbf{x}_t(\varepsilon)) \right| \leq \frac{2\delta}{3}$$

Thus we have by the shattering assumption,

$$\begin{aligned}
\frac{\delta}{2} &\leq \varepsilon_t \left(f_\varepsilon(\mathbf{x}_t(\varepsilon) - \mathbf{s}_t(\mathbf{x}_t(\varepsilon))) \right) \\
&= -\varepsilon_t' \left(f_\varepsilon(\mathbf{x}_t(\varepsilon) - \mathbf{s}_t(\mathbf{x}_t(\varepsilon))) \right) \\
&\leq -\varepsilon_t' \left(f_{\varepsilon'}(\mathbf{x}_t(\varepsilon) - \varepsilon_t' \frac{2\delta}{3} - \mathbf{s}_t(\mathbf{x}_t(\varepsilon))) \right) \\
&\leq -\frac{\delta}{2} + \frac{2\delta}{3} \\
&\leq \frac{\delta}{6}
\end{aligned}$$

Thus we have a contradiction and the mapping is injective. But this means then that $2^d \leq N\left(\mathcal{F}, \frac{\delta}{2}\right)$ as desired. By Theorem 23, for any $\alpha > 0$, we have

$$\begin{aligned}\mathsf{fat}_\delta(\mathcal{F}) &\lesssim \log\left(2^{\mathsf{fat}_\delta(\mathcal{F})}\right) \lesssim \log\left(N\left(\mathcal{F}, \frac{\delta}{3}\right)\right) \\ &\lesssim \mathsf{vc}\left(\mathcal{F}, c\alpha\delta\right) \log^{1+\alpha}\left(\frac{C\left|\mathcal{X}\right|}{\mathsf{vc}\left(\mathcal{F}, c\alpha\delta\right)\delta}\right)\end{aligned}$$

as desired. ■

We are now ready to prove the main results from Section 3.

C.4. Proof of Theorem 3

By Proposition 18, it suffices to control the distribution-dependent sequential Rademacher complexity of Definition 16, specialized to the case that $\mathcal{Z} = \mathcal{X} \times [-1, 1]$, and $\mathscr{D} \in \mathfrak{P}_T(\sigma, \mu)$. We first adapt Lemma 14 to construct a coupling with $\varepsilon, \rho_{\mathscr{D}}$ and independent samples from μ; this will allow us to move from sequential Rademacher complexity to standard Rademacher complexity. The lemma is again a variant of the coupling in Haghtalab et al. (2021), albeit simpler to describe and preserving independence between ε and Z_t^j. We have the following lemma:

Lemma 24 *Suppose that $\mathscr{D} \in \widetilde{\mathfrak{P}}_T(\sigma, \mu)$. Then for any T there exists a measure Π with random variables $(\varepsilon_{1:T}, \mathbf{z}(\varepsilon), \mathbf{z}'(\varepsilon), Z_t^j, Z_t^{j'})_{\substack{1 \leq t \leq T \\ 1 \leq j \leq k}}$ satisfying the following properties, where we write $\mathbf{z}(\varepsilon) = (\mathbf{x}(\varepsilon), \mathbf{y}(\varepsilon)), \mathbf{z}'(\varepsilon) = (\mathbf{x}'(\varepsilon), \mathbf{y}'(\varepsilon))$ to separate the $\mathcal{X}$- and $[-1, 1]$-components of $\mathbf{z}(\varepsilon) \in \mathcal{Z}$:*

1. *$\varepsilon_{1:T}$ are iid Rademacher random variables.*
2. *$(\mathbf{z}, \mathbf{z}')$ is distributed according ρ.*
3. *$\{Z_t^j, Z_t^{j'}\}$ are iid according to μ*
4. *$\{\varepsilon, Z_t^j, Z_t^{j'}\}$ are independent*
5. *With probability at least $1 - 2T(1 - \sigma)^k$, $\mathbf{x}_t(\varepsilon) \in \{Z_t^j\}_{1 \leq j \leq k}$ for all t*

Proof Given $\mathscr{D}$, let:

- $p_t(\cdot|(x_1, y_1), \dots, (x_{t-1}, y_{t-1}))$ denote the distribution of x_t under $\mathscr{D}$ given (x_s, y_s) for $s < t$ (since $\mathscr{D} \in \widetilde{\mathfrak{P}}_T(\sigma, \mu)$, p_t is σ-smooth with respect to μ a.s.);
- $q_t(\cdot|(x_1, y_1), \dots, (x_{t-1}, y_{t-1}), x_t)$ denote the distribution of y_t under $\mathscr{D}$ given (x_s, y_s) for $s < t$ and x_t.

We construct the coupling recursively. For any t, suppose that $Z_s^j, Z_s^{j'}, \varepsilon_s, \mathbf{z}_s(\varepsilon)$ has been constructed for $s < t$. If $t = 0$ then this is the empty set. Now, sample $Z_t^j, Z_t^{j'}$ iid according to μ and ε_t a Rademacher random variable. Let $\pi_t^j = \sigma \frac{dp_t}{d\mu}(Z_t^j)$. Note that $\pi_t^j \leq 1$ by the assumption of σ-smoothness. As in the proof of Lemma 14, construct the random set S_t by adding each Z_t^j to S_t

with probability π_t^j. If S_t is nonempty, then sample $\mathbf{x}_t(\varepsilon)$ independently from S_t uniformly at random; if S_t is empty, sample $\mathbf{x}_t(\varepsilon)$ from $p_t(\cdot|\mathbf{z}_1(\varepsilon), \dots, \mathbf{z}_{t-1}(\varepsilon))$. Then sample $\mathbf{y}_t(\varepsilon)$ independently from $q_t(\cdot|\mathbf{z}_1(\varepsilon), \dots, \mathbf{z}_{t-1}(\varepsilon), \mathbf{x}_t(\varepsilon))$, and set $\mathbf{z}_t(\varepsilon) = (\mathbf{x}_t(\varepsilon), \mathbf{y}_t(\varepsilon))$.

We may construct the $\mathbf{z}_t'(\varepsilon)$ similarly by constructing a set S_t' in the same way, using $Z_t^{j'}$ instead of Z_t^j. Finally sample ε_t independently.

It is clear from the construction that ε_t, $t \in [T]$, are independent Rademacher random variables. Similarly, it is clear that $Z_t^j, Z_t^{j'}$ are iid according to μ and independent of ε. The remainder of the properties are proved in the same way as in Lemma 14. ■

We are now ready to prove the theorem.

Proof (Theorem 3) By Theorem 18, it suffices to bound the distribution-dependent sequential Rademacher complexity. Let Π be the coupling in Lemma 24 and let A denote the event that $\mathbf{x}_t(\varepsilon) \in \{Z_t^j\}_{1 \le j \le k}$ for all t (we continue to write $\mathbf{z}_t(\varepsilon) = (\mathbf{x}_t(\varepsilon), \mathbf{y}_t(\varepsilon))$). Then we compute for a fixed k,

$$\begin{aligned}
&\sup_{\mathscr{D} \in \mathfrak{P}} \mathbb{E}_{\rho_{\mathscr{D}}, \varepsilon}\left[\sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right] \\
&\le \mathbb{E}_{\Pi, \varepsilon}\left[\sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right] \\
&\le \mathbb{E}_{\Pi, \varepsilon}\left[\chi_{A^c} \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right] + \mathbb{E}_{\Pi, \varepsilon}\left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right] \\
&\le 2T^2 e^{-\sigma k} + \mathbb{E}_{\Pi, \varepsilon}\left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right]
\end{aligned}$$

By the tower property of expectations, denoting by $\mathcal{F}|_{\{Z_t^j\}}$ the restriction of $\mathcal{F}$ to the set of all Z_t^j, we have

$$\begin{aligned}
\mathbb{E}_{\Pi, \varepsilon}\left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon))\right] &= \mathbb{E}_{Z_t^j \overset{iid}{\sim} \mu}\left[\mathbb{E}_{\Pi}\left[\chi_A \sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon)) \middle| \{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}}\right]\right] \\
&= \mathbb{E}_{Z_t^j \overset{iid}{\sim} \mu}\left[\mathbb{E}_{\Pi}\left[\chi_A \sup_{f \in \mathcal{F}|_{\{Z_t^j\}}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon)) \middle| \{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}}\right]\right] \\
&\lesssim L \log^{\frac{3}{2}}(T) \sup_{\{Z_t^j\}_{\substack{1 \le t \le T \\ 1 \le j \le k}}} \sup_{\mathbf{x}} \mathbb{E}_{\varepsilon}\left[\sup_{f \in \mathcal{F}|_{\{Z_t^j\}}} \sum_{t=1}^T \varepsilon_t f(\mathbf{x}_t(\varepsilon))\right]
\end{aligned}$$

where the last inner supremum is over all $\mathbf{x}$ such that $\mathbf{x}$ is a $\{Z_t^j\}$-labelled binary tree of depth T; the last inequality follows, then, from Lemma 20. Let $\mathsf{fat}_\delta(\mathcal{F})$ denote the sequential fat-shattering dimension in Definition 19. We may apply (Block et al., 2021, Corollary 10 and Proposition 15), which bounds the worst-case sequential Rademacher complexity by the sequential fat-shattering dimension to get

$$\sup_{\substack{\{Z_t^j\}_{1 \leq t \leq T} \\ 1 \leq j \leq k}} \sup_{\mathbf{x}} \mathbb{E}_\varepsilon \left[\sup_{f \in \mathcal{F}|_{\{Z_t^j\}}} \sum_{t=1}^T \varepsilon_t f(\mathbf{x}_t(\varepsilon)) \right] \lesssim \sup_{\substack{\{Z_t^j\}_{1 \leq t \leq T} \\ 1 \leq j \leq k}} \inf_{\alpha > 0} \left\{ \alpha T + \sqrt{T} \int_\alpha^1 \sqrt{\mathsf{fat}_\delta \left(\mathcal{F}_{\{Z_t^j\}} \right)} d\delta \right\}$$

By Lemma 21, we have for any $\beta > 0$,

$$\mathsf{fat}_\delta \left(\mathcal{F}_{\{Z_t^j\}} \right) \lesssim \mathsf{vc}(\mathcal{F}, c\beta\delta) \log^{1+\beta} \left(\frac{C \left| \{Z_t^j\}_{\substack{1 \leq t \leq T \\ 1 \leq j \leq k}} \right|}{\mathsf{vc}\left(\mathcal{F}, c\beta\delta\right) \delta} \right) \lesssim \mathsf{vc}(\mathcal{F}, c\beta\delta) \log^{1+\beta} \left(\frac{CkT}{\mathsf{vc}\left(\mathcal{F}, c\beta\delta\right) \delta} \right)$$

independent of the realization of Z_t^j. Thus we have

$$\begin{aligned} &\mathbb{E}_{Z_t^j \overset{iid}{\sim} \mu} \left[\mathbb{E}_\Pi \left[\chi_A \sup_{f \in \mathcal{F}|_{\{Z_t^j\}}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon)) \middle| \{Z_t^j\}_{\substack{1 \leq t \leq T \\ 1 \leq j \leq k}} \right] \right] \\ &\qquad \lesssim L \log^{\frac{3}{2}}(T) \inf_{\alpha > 0} \left\{ \alpha T + \sqrt{T} \int_\alpha^1 \sqrt{\mathsf{vc}(\mathcal{F}, c\beta\delta) \log^{1+\beta} \left(\frac{CkT}{\mathsf{vc}\left(\mathcal{F}, c\beta\delta\right) \delta} \right)}) d\delta \right\} \end{aligned}$$

Putting everything together, we have

$$\begin{aligned} &\sup_{\mathscr{D} \in \mathfrak{P}} \mathbb{E}_{\rho_{\mathscr{D}}, \varepsilon} \left[\sup_{f \in \mathcal{F}} \sum_{t=1}^T \varepsilon_t \ell(f(\mathbf{x}_t(\varepsilon)), \mathbf{y}_t(\varepsilon)) \right] \\ &\qquad \lesssim 2T^2 e^{-\sigma k} + L \log^{\frac{3}{2}}(T) \inf_{\alpha > 0} \left\{ \alpha T + \sqrt{T \log^{1+\beta} \left(\frac{3kT}{\mathsf{vc}\left(\mathcal{F}, c\beta\alpha\right) \alpha} \right)} \int_\alpha^1 \sqrt{\mathsf{vc}(\mathcal{F}, c\beta\delta)} d\delta \right\} \end{aligned}$$

Setting $k = \frac{2 \log T}{\sigma}$ and $\beta = 1$ concludes the proof. ■

C.5. Proof of Proposition 4

Let $\mathcal{F}$ be the class of thresholds on the unit interval, i.e.,

$$\mathcal{F} = \{x \mapsto \text{sign}(x - \theta) | \theta \in [0, 1]\}$$

It is well-known that $\mathsf{vc}(\mathcal{F}) = 1$. Consider an adversary that sets $x_1 = 0$, $x_2 = 1$, $y_1 = -1$, $y_2 = 1$ and for all $t > 2$, sets

$$x_t = x_{t-1} - y_{t-1} 2^{-(t-2)}$$

and the y_t are independent Rademacher random variables. Note that by construction, the adversary is realizable with respect to $\mathcal{F}$ in the sense that for any realization of the (x_t, y_t), there is some $f \in \mathcal{F}$ with $f(x_t) = y_t$ for all t. Also by construction, we see that the expected number of mistakes in T rounds is $\frac{T}{2}$. For fixed T, let

$$\mathscr{P}_T = \left\{ \frac{1}{T} \sum_{t=1}^{T} \delta_{x_t} \right\}$$

be the set of empirical distributions generated by the contexts over all realizations of $x_1, \ldots, x_T$. Note that for each $\mu \in \mathscr{P}_T$, the support has size T and thus the adversary constructed above is $\left(\frac{1}{T}\right)$-smooth with respect to some $\mu \in \mathscr{P}_T$. The result follows by noting that if $T \leq \frac{1}{\sigma}$ then the adversary is σ-smooth with respect to some $\mu \in \mathscr{P}_T$.

Appendix D. Proofs from Section 4

D.1. Proofs Related to Relaxations

Proof (Proposition 6) It suffices to prove (2) as the other property follows immediately from the construction. For the sake of convenience, we denote

$$L_t(f) = \sum_{s=1}^{t} \ell(f(x_t), y_t)$$

We begin by noting that (Rakhlin and Sridharan, 2014, Lemma 5.1) tells us that due to the convexity of ℓ in the first argument, it suffices to replace distributions q_t over $[-1, 1]$ with values $\widehat{y}_t \in [-1, 1]$. In particular,

$$\begin{aligned} &\inf_{q_t \in \Delta([-1,1])} \sup_{y_t \in [-1,1]} \mathbb{E}_{q_t}\left[\ell(\widehat{y}_t, y_t)\right] + \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f)\right] \\ &\qquad = \inf_{\widehat{y}_t \in [-1,1]} \sup_{y_t \in [-1,1]} \ell(\widehat{y}_t, y_t) + \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f)\right] \end{aligned}$$

Now, omitting the feasible set for $\widehat{y}_t, y_t$ to ease the notational load, we plug in our relaxation:

$$\begin{aligned}
&\inf_{\widehat{y}_t} \sup_{y_t} \ell(\widehat{y}_t, y_t) + \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f)\right] \\
&\quad = \inf_{\widehat{y}_t} \sup_{y_t} \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + \ell(\widehat{y}_t, y_t) - \ell(f(x_t), y_t)\right] \\
&\quad \leq \inf_{\widehat{y}_t} \sup_{y_t} \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + \partial\ell(\widehat{y}_t, y_t)(\widehat{y}_t - f(x_t))]\right] \\
&\quad \leq \inf_{\widehat{y}_t} \sup_{y_t} \sup_{g_t \in [-L,L]} \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + g_t(\widehat{y}_t - f(x_t))\right] \\
&\quad = \inf_{\widehat{y}_t} \max_{g_t \in \{-L,L\}} \mathbb{E}_{\mu,\varepsilon}\left[g_t \widehat{y}_t + \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) - g_t f(x_t)\right]
\end{aligned}$$

where we let $\partial\ell$ denote a subgradient of ℓ with respect to the first argument. The first inequality follows by convexity, the second inequality follows by Lipschitzness, and the last equality follows because the inner expectation is convex as a function of g_t and so obtains its maximum on the boundary. Let d_t denote a distribution on $\{-L, L\}$; the y_t vanishes because it only appeared in the $\partial\ell(\widehat{y}_t, y_t)$ and this was bounded by g_t. Then by the minimax theorem, we have

$$\begin{aligned}
&\inf_{\widehat{y}_t} \max_{g_t \in \{-L,L\}} \mathbb{E}_{\mu,\varepsilon}\left[g_t \widehat{y}_t + \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) - g_t f(x_t)\right] \\
&\quad = \sup_{d_t} \inf_{\widehat{y}_t} \mathbb{E}_{g_t \sim d_t} \mathbb{E}_{\mu,\varepsilon}\left[g_t \widehat{y}_t + \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) - g_t f(x_t)\right] \\
&\quad \leq \sup_{d_t} \mathbb{E}_{g_t \sim d_t} \mathbb{E}_{\mu,\varepsilon}\left[\inf_{\widehat{y}_t} \mathbb{E}_{g_t' \sim d_t}[g_t' \widehat{y}_t] + \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) - g_t f(x_t)\right] \\
&\quad \leq \sup_{d_t} \mathbb{E}_{\mu,\varepsilon} \mathbb{E}_{g_t \sim d_t}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + (\mathbb{E}_{g_t' \sim d_t}[g_t'] - g_t) f(x_t)\right] \\
&\quad \leq \sup_{d_t} \mathbb{E}_{\mu,\varepsilon} \mathbb{E}_{g_t, g_t' \sim d_t}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + \varepsilon_t(g_t' - g_t) f(x_t)\right] \\
&\quad \leq \sup_{d_t} \mathbb{E}_{\mu,\varepsilon} \mathbb{E}_{g_t \sim d_t}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2\varepsilon_t g_t f(x_t)\right] \\
&\quad = \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t)\right]
\end{aligned}$$

Now, it would seem that we are done, but note that $x_t \sim p_t$ while $x_s \sim \mu$ for $s > t$. We thus apply $\sup_{p_t \in \mathfrak{P}_t} \mathbb{E}_{x_t \sim p_t}$ to all of the preceding equations and, adding back in the additive constant, we have shown

$$\sup_{p_t \in \mathfrak{P}} \mathbb{E}_{x_t \sim p_t} \inf_q \sup_{y_t} \left[\mathbb{E}_{\widehat{y}_t \sim q}[\ell(\widehat{y}, y)] + \mathbf{Rel}_T(\mathcal{F} | x_1, y_1, \dots, x_t, y_t) \right]$$

$$\leq \sup_{p_t \in \mathfrak{P}} \mathbb{E}_{x_t \sim p_t} \mathbb{E}_{\mu, \varepsilon} \left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right] + (T-t)^3 e^{-\sigma k}$$

Now, applying the coupling Π from Lemma 14, we have

$$\mathbb{E}_{x_t \sim p_t} \mathbb{E}_{\mu, \varepsilon} \left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

$$= \mathbb{E}_{x_t \sim \Pi} \mathbb{E}_{\mu, \varepsilon} \left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

$$= \mathbb{E}_{x_t \sim \Pi} \mathbb{E}_{\mu, \varepsilon} \left[\chi_{A^c} \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

$$+ \mathbb{E}_{x_t \sim \Pi} \mathbb{E}_{\mu, \varepsilon} \left[\chi_{A} \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

where A is the event that $x_t \in \{Z_t^j\}$ for $1 \leq j \leq k$ and χ_A is the indicator. For the first term, we have

$$\mathbb{E}_{x_t \sim \Pi} \mathbb{E}_{\mu, \varepsilon} \left[\chi_{A^c} \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

$$\leq \mathbb{P}(A^c)(n-t+1) \leq (n-t+1)^2 e^{-\sigma k}$$

For the second term, we have

$$\mathbb{E}_{x_t \sim \Pi} \mathbb{E}_{\mu, \varepsilon} \left[\chi_{A} \sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + 2L\varepsilon_t f(x_t) \right]$$

$$\leq \mathbb{E}_{\Pi, \mu, \varepsilon} \chi_A \sup_{f \in \mathcal{F}} \left[2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) + \sum_{j=1}^{k} \varepsilon_{t,j} f(Z_t^j) \right]$$

$$= \mathbb{E}_{\mu, \varepsilon} \left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) \right]$$

Putting this back together, we have

$$\begin{aligned}
&\sup_{p_t \in \mathfrak{P}_t} \mathbb{E}_{x_t \sim p_t} \inf_q \sup_{y_t} \left[\mathbb{E}_{\widehat{y}_t \sim q}[\ell(\widehat{y}, y)] + \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_t, y_t) \right] \\
&\qquad \leq \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) \right] + (T-t+1)^2 e^{-\sigma k} + (T-t)^3 e^{-\sigma k} \\
&\qquad \leq \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) \right] + (T-t+1)^3 e^{-\sigma k} \\
&\qquad = \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_{t-1}, y_{t-1})
\end{aligned}$$

as desired. Thus we have an admissable relaxation. ■

Proof (Theorem 7) It suffices to show the following claim:

$$\sup_{p_t \in \mathfrak{P}} \mathbb{E}_{x_t \sim p_t} \left[\sup_{y_t \in [-1,1]} \ell(\widehat{y}_t, y_t) + \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{s=t+1}^{T} \sum_{j=1}^{k} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) + (T-t)^3 e^{-\sigma k} \right] \right]$$

$$\leq \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_{t-1}, y_{t-1})$$

Indeed, if this is the case, then $\widehat{y}_t$ is admissable with respect to $\mathbf{Rel}_T(\mathcal{F}|\cdot)$, for which we already have a regret bound in Proposition 6. To prove the stated claim, we have

$$\begin{aligned}
&\sup_{y_t} \mathbb{E}_{\mu,\varepsilon} \left[\ell(\widehat{y}_t, y_t) + \sup_{f \in \mathcal{F}} \left[2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right] \\
&\qquad \leq \mathbb{E}_{\mu,\varepsilon} \left[\sup_{y_t} \ell(\widehat{y}_t, y_t) + \sup_{f \in \mathcal{F}} \left[2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right] \\
&\qquad = \mathbb{E}_{\mu,\varepsilon} \left[\inf_{\widehat{y}} \sup_{y_t} \ell(\widehat{y}, y_t) + \sup_{f \in \mathcal{F}} \left[2L \sum_{s=t+1}^{n} \sum_{j=1}^{k} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right]
\end{aligned}$$

where the inequality follows by Jensen's and the equality follow from the construction of $\widehat{y}_t$ in (4). Now we may apply the proof of Proposition 6 with the expectation with respect to μ, ε taking place outside of the minimax operation. This shows that

$$\begin{aligned}
&\mathbb{E}_{\mu,\varepsilon} \left[\inf_{\widehat{y}} \sup_{y_t} \ell(\widehat{y}, y_t)] + \sup_{f \in \mathcal{F}} \left[2L \sum_{s=t+1}^{n} \sum_{j=1}^{k} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right] + (T-t)^3 e^{-\sigma k} \\
&\qquad \leq \mathbb{E}_{\mu,\varepsilon}\left[\sup_{f \in \mathcal{F}} 2L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_{t-1}(f) \right] + (T-t+1)^3 e^{-\sigma k} \\
&\qquad = \mathbf{Rel}_T(\mathcal{F}|x_1, y_1, \dots, x_{t-1}, y_{t-1})
\end{aligned}$$

as desired and the claim holds.

To prove the oracle efficiency claims, for a fixed δ, let S be a δ-discretization of $[-1, 1]$ of size $\frac{2}{\delta}$. If we solve the minimax problem (4) over S, then by the assumption of ℓ being L-Lipshitz, we the regret bound for our approximate solution is greater than that of the exact solution by at most an additive constant of $L\delta T$. In general, for any fixed $\widehat{y}$, we can optimize

$$\sup_{y_t \in S} \left\{ \ell(\widehat{y}, y_t) + \sup_{f \in \mathcal{F}} \left[6L \sum_{j=1}^{k} \sum_{s=t+1}^{T} \varepsilon_{s,j} f(x_{s,j}) - L_t(f) \right] \right\} \tag{7}$$

with $|S|$ calls to the ERM oracle. Now, note that (7) is convex in $\widehat{y}$ by the assumption of convexity in the first argument of ℓ. By a simple three point method from zeroth order optimization Agarwal et al. (2011), which we prove as Lemma 25 for the sake of completeness below, we can minimize (7) with respect to $\widehat{y} \in S$ with $O(\log |S|)$ evaluations of the supremum. Each evaluation of the supremum requires $O(|S|)$ calls to the ERM oracle so we require $O\left(|S| \log |S|\right)$ calls in total. Noting that $|S| \leq \frac{1}{\delta}$, we can get

$$\mathrm{Reg}_T(\widehat{y}) \leq 2L\mathfrak{R}_{kT}(\mathcal{F}) + T^3 e^{-\sigma k} + LT\delta$$

with $O\left(\frac{T}{\delta} \log\left(\frac{1}{\delta}\right)\right)$ calls to the ERM oracle. Setting $\delta = \frac{1}{L\sqrt{T}}$ recovers the bound in the theorem statement. The last statement, on linear losses, is proven in Lemma 26, where we give an explicit representation of the solution using only 2 oracle calls. Optimizing k concludes the proof. ■

Lemma 25 *Let $f : [0, 1] \to \mathbb{R}$ be a convex function and let $S \subset [0, 1]$. Then*

$$x_0 \in \operatorname*{argmin}_S f$$

can be found with $O(\log |S|)$ calls to a value oracle that returns $f(x)$ given input $x \in S$.

Proof Motivated by Agarwal et al. (2011), we describe the following recursive algorithm that shrinks S until it contains only one point, yet always includes the minimizer. Let $S_0 = S$. To construct S_{i+1} from S_i, order the points $x_1, \ldots, x_m \in S_i$ such that $x_i < x_j$ for all $i < j$. Let z_1, z_2, z_3 be the $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ quanatiles of S respectively and evaluate $f(z_1)$, $f(z_2)$, and $f(z_3)$ with three calls to the value oracle. There are several cases:

$\mathbf{f(z_1) > f(z_2) < f(z_3)}$ If the middle point is smaller than either point on the end, then be the convexity of f we know that the minimum must occcur for some x such that $z_1 < x < z_3$. In this case let S_{i+1} contain all the points $x \in S_i$ such that $z_1 < x < z_3$. Note that $|S_{i+1}| \leq \frac{1}{2} |S_i|$.

$\mathbf{f(z_1) < f(z_2) > f(z_3)}$ This case corresponds to the middle point being higher than the end points. This is not possible, however, as f is convex.

$\mathbf{f(z_1) < f(z_2) < f(z_3)}$ In this case, convexity assures us that the minimizer cannot be at any point $x \geq z_2$ and so we let S_{i+1} to be the set of all points $x \in S_i$ such that $x < z_2$. Note that $|S_{i+1}| \leq \frac{1}{2} |S_i|$.

$\mathbf{f(z_1) > f(z_2) > f(z_3)}$ This is the mirror image of the previous case and can be handled similarly.

$\mathbf{f}(\mathbf{z_1}) = \mathbf{f}(\mathbf{z_2}) = \mathbf{f}(\mathbf{z_3})$ In this case, convexity ensures that the minimizer must have $x \leq z_1$ or $x \geq z_3$ and so we let S_{i+1} be the set of $x \in S_i$ satisfying this constraint. Again, $|S_{i+1}| \leq \frac{1}{2}|S_i|$.

In any case, with three calls to the value oracle, we reduce the size of S_i by a factor of 2. Thus we can find x_0 in $O(\log|S|)$ calls as claimed. ■

Lemma 26 *Suppose we are in the situation of Theorem 7 and the loss $\ell(\widehat{y}, y) = \frac{1-\widehat{y}y}{2}$. Then the problem (4) can be solved with two calls to the ERM oracle. In fact, $\widehat{y}_t$ is given by*

$$\frac{1}{2}\sup_{f\in\mathcal{F}}\left[6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f(x_{s,j}) - L_{t-1}(f) - \ell(f(x_t),1)\right]$$
$$-\frac{1}{2}\sup_{f\in\mathcal{F}}\left[6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f(x_{s,j}) - L_{t-1}(f) - \ell(f(x_t),-1)\right]$$

Proof We are trying to minimize with respect to $\widehat{y} \in [-1, 1]$.

$$\sup_{y_t\in S}\left\{\ell(\widehat{y}, y_t) + \sup_{f\in\mathcal{F}}\left[6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f(x_{s,j}) - L_t(f)\right]\right\} \tag{8}$$

Independent of $\widehat{y}$, if ℓ is linear in y, then the expresion over which we are taking the supremum in (8) is convex in y and thus $y_t \in \{\pm 1\}$. For the sake of simplicity, suppose that

$$\sup_{f\in\mathcal{F}}\left[6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f(x_{s,j}) - L_t(f)\right]$$

is maximized by functions f_+ and f_- depending on if $y_t = 1$ or $y_t = -1$ (in the general case, we could take a sequence of functions attaining the supremum). Let

$$a_+ = 6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f_+(x_{s,j}) - L_t(f_+)$$
$$a_- = 6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f_-(x_{s,j}) - L_t(f_-)$$

Then the solution to (4) is given by

$$\min_{\widehat{y}\in[-1,1]}\max\left(\frac{1-\widehat{y}}{2} + a_+, \frac{1+\widehat{y}}{2} + a_-\right)$$

The maximum is taken over two linear functions of $\widehat{y}$ with opposite slope and so the minimax result is where they intersect, assuming they intersect somewhere in $[-1, 1]$, which they do if $|a_+ - a_-| \leq$

1. Note that

$$\begin{aligned}
a_+ &= 6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f_+(x_{s,j}) - L_t(f_+) \\
&= 6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f_+(x_{s,j}) - L_{t-1}(f_+) - \ell(f_+(x_t),1) + \ell(f_+(x_t),-1) - \ell(f_+(x_t),-1) \\
&\geq 6L\sum_{j=1}^{k}\sum_{s=t+1}^{T}\varepsilon_{s,j}f_-(x_{s,j}) - L_t(f_-) - \ell(f_+(x_t),1) + \ell(f_+(x_t),-1) \\
&\geq a_- - 1
\end{aligned}$$

By symmetry, $|a_- - a_+| \leq 1$. Thus,

$$\widehat{y}_t = \frac{a_+ - a_-}{2} \in [-1,1]$$

solves the minimax problem with two calls to the ERM oracle. ■

D.2. Proof of Proposition 8

In order to prove Proposition 8, we consider the following setting. Fix an $\alpha > 0$, suppose that $\mathsf{vc}(\mathcal{F},\alpha) = m$ and let $x_1,\ldots,x_m$ shatter $\mathcal{F}$ at scale α. We let p_t be uniform on $x_1,\ldots,x_m$ and let $\mu = (1-\sigma)\delta_{x^*} + \sigma p_t$, where x^* is a distinguished point satisfying $f(x^*) = 0$ for all $f \in \mathcal{F}$. Note that p_t is σ-smooth with respect to μ. We compare the expected Rademacher complexity sampling n points according to μ to that when sampling according to p_t. We require two lemmata.

Lemma 27 *Suppose we are in the above setting and $X_1,\ldots,X_T$ are sampled independently according to μ. Then, with probability at least $1-\delta$, the number of indices i such that $X_i \neq x^*$ is at most $2\sigma T$ if $T \geq \frac{3}{\sigma}\log\left(\frac{1}{\delta}\right)$.*

Proof Let $Y_i = \mathbf{1}[X_i \neq x^*]$. Then Y_i are independent Bernoulli random variables with parameter σ and the number of such indices is the sum of Y_i. Applying Chernoff's inequality, we have

$$\mathbb{P}\left(\sum_{i=1}^{T} Y_i \geq 2\sigma T\right) \leq e^{-\frac{\sigma T}{3}}$$

The assumption of T large enough concludes the proof. ■

Lemma 28 *Suppose that we are in the setting described above and $X_1,\ldots,X_T$ are sampled according to p_t. Suppose that $T \geq 8m\log\left(\frac{m}{\delta}\right)$. Then with probability at least $1-\delta$, for each $1 \leq j \leq m$, there are at least $\frac{T}{2m}$ indices i such that $X_i = x_j$.*

Proof Fix j and let $Y_i = \mathbf{1}[X_i = x_j]$. Then the Y_i are independent Bernoulli random variables with parameter $\frac{1}{m}$. Letting S_T denote the sum of the Y_i, which is the desired number of indices, we may apply Chernoff's inequality to get

$$\mathbb{P}\left(S_T \leq \frac{1}{2}\frac{T}{m}\right) \leq e^{-\frac{T}{8m}}$$

and so with probability at least $1 - \frac{\delta}{m}$, there are at least $\frac{T}{2m}$ indices i such that $X_i = x_j$. Applying a union bound concludes the proof. ■

We may now adapt an argument from Mendelson (2002) and (Srebro et al., 2010, Lemma A.2) to lower bound the Rademacher complexity according to p_t:

Lemma 29 *Suppose that we are in the above setting and suppose that $T \geq 8m \log(2m)$. Then,*

$$\mathbb{E}_{p_t}\left[\mathcal{R}_T(\mathcal{F})\right] \geq \frac{\alpha}{8}\sqrt{\mathsf{vc}(\mathcal{F}, \alpha) T}$$

Proof Let χ_A denote the indicator of the high probability event A from Lemma 28. Consider any fixed choice of $X_1, \dots, X_T$ so that the event A holds. For each $j \in [m], k \in [T/(2m)]$, let $\phi(j, m) \in [T]$ denote the kth smallest value of i so that $X_i = x_j$ (that all such $\phi(j, m)$ exist is guaranteed by A). Furthermore let $\Phi \subset [T]$ denote the image of ϕ, so that $|\Phi| = T/2$. Next, for any $\varepsilon \in \{-1, 1\}^n$, define $f^\varepsilon := \arg\max_{f \in \mathcal{F}} \sum_{j=1}^m \sum_{k=1}^{T/(2m)} \varepsilon_{\phi(j,m)} f(x_j)$.[8]

$$\begin{aligned}
\mathbb{E}_{p_t}\left[\mathcal{R}_T(\mathcal{F})\right] &= \mathbb{E}_{p_t}\left[\mathbb{E}_\varepsilon\left[\sup_{f \in \mathcal{F}} \sum_{i=1}^T \varepsilon_i f(X_i) \middle| X_1, \dots, X_T\right]\right] \\
&\geq \mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sup_{f \in \mathcal{F}} \sum_{i=1}^T \varepsilon_i f(X_i) \middle| X_1, \dots, X_T\right]\right] \qquad (9) \\
&\geq \mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sum_{i=1}^T \varepsilon_i f^\varepsilon(X_i) \middle| X_1, \dots, X_T\right]\right] \\
&\geq \mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sum_{j=1}^m \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{jk} f^\varepsilon(x_j) \middle| X_1, \dots, X_T\right]\right] + \mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sum_{i \notin \Phi} \varepsilon_i f^\varepsilon(X_i) \middle| X_1, \dots, X_n\right]\right] \\
&\geq \mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sup_{f \in \mathcal{F}} \sum_{j=1}^m \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} f(x_j) \middle| X_1, \dots, X_T\right]\right] \qquad (10) \\
&= \mathbb{P}(A) \mathbb{E}_\varepsilon\left[\sup_{f \in \mathcal{F}} \sum_{j=1}^m \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} f(x_j)\right]
\end{aligned}$$

where (9) follows by Jensen's inequality coupled with the fact that the ε_i are mean zero, and (10) follows since $\{\varepsilon_i : i \in \Phi\}$ are independent of $\{\varepsilon_i : i \notin \Phi\}$ (so that $\mathbb{E}_{p_t}\left[\chi_A \mathbb{E}_\varepsilon\left[\sum_{i \notin \Phi} \varepsilon_i f^\varepsilon(X_i) \middle| X_1, \dots, X_T\right]\right] =$

[8] If the argmax does not exist, we may instead consider a sequence of functions that approximates the argmax to arbitrarily small precision.

0), and by definition of f^ε. By the triangle inequality and symmetry of the ε_i, $i \in [T]$, we have

$$\begin{aligned}
\mathbb{E}_\varepsilon \left[\sup_{f \in \mathcal{F}} \sum_{j=1}^{m} \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} f(x_j) \right] &\geq \frac{1}{2} \mathbb{E}_\varepsilon \left[\sup_{f, f' \in \mathcal{F}} \sum_{j=1}^{m} \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} (f(x_j) - f'(x_j)) \right] \\
&\geq \frac{1}{2} \mathbb{E}_\varepsilon \left[\sum_{j=1}^{m} \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} (f_\varepsilon(x_j) - f'_\varepsilon(x_j)) \right]
\end{aligned}$$

where $f_\varepsilon, f'_\varepsilon$ are chosen so that

$$\operatorname{sign}\left(\sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} \right) (f_\varepsilon(x_j) - s_j) \geq \frac{\alpha}{2} \qquad \operatorname{sign}\left(\sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} \right) \left(f'_\varepsilon(x_j) - s_j\right) \leq -\frac{\alpha}{2}$$

for some $s_1, \ldots, s_m \in \mathbb{R}$. Note that there exist such $f_\varepsilon, f'_\varepsilon \in \mathcal{F}$ by the assumption that $x_1, \ldots, x_m$ shatter $\mathcal{F}$ at scale α. We thus have

$$\begin{aligned}
\mathbb{E}_\varepsilon \left[\sum_{j=1}^{m} \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} (f_\varepsilon(x_j) - f'_\varepsilon(x_t)) \right] &\geq \frac{1}{2} \mathbb{E}_\varepsilon \left[\sum_{j=1}^{m} \left| \sum_{k=1}^{\frac{T}{2m}} \varepsilon_{\phi(j,k)} \right| \alpha \right] \\
&\geq \frac{m}{2} \alpha \mathbb{E}_\varepsilon \left[\left| \sum_{i=1}^{\frac{T}{2m}} \varepsilon_i \right| \right] \\
&\geq \frac{m}{2} \alpha \sqrt{\frac{T}{4m}} \qquad (11) \\
&= \frac{\alpha}{4} \sqrt{Tm},
\end{aligned}$$

where (11) follows from Khintchine's inequality. By Lemma 28, $\mathbb{P}(A) \geq \frac{1}{2}$. Thus, putting everything together, we have

$$\mathbb{E}_{p_t} \left[\mathcal{R}_T(\mathcal{F}) \right] \geq \frac{\alpha}{8} \sqrt{mT}$$

We finally recall that $m = \mathsf{vc}(\mathcal{F}, \alpha)$ and conclude the proof. ■

The last thing we need to do is provide an upper bound on $\mathbb{E}_\mu \left[\mathcal{R}_T(\mathcal{F}) \right]$. We can do this using chaining and Lemma 27.

Lemma 30 *Suppose we are in the setting above and suppose that $T \geq \frac{3}{\sigma} \log T$. Then there is an absolute constant C such that*

$$\mathbb{E}_\mu \left[\mathcal{R}_T(\mathcal{F}) \right] \leq C\sqrt{\sigma T}$$

Proof Let

$$A = \{ |\{ i | X_i \neq x^* \}| \leq 2\sigma T \}$$

and let χ_A denote the indicator for this event. By Lemma 27, $\mathbb{P}(A) \geq 1 - \frac{1}{T}$. Note that as $\mathcal{F}$ is uniformly bounded by 1, we always have the trivial upper bound of $\mathcal{R}_T(\mathcal{F}) \leq T$ on any data set. We can thus compute

$$\mathbb{E}_\mu \left[\mathcal{R}_T(\mathcal{F}) \right] = \mathbb{E}_\mu \left[\chi_A \mathcal{R}_T(\mathcal{F}) \right] + (1 - \mathbb{P}(A))T \leq \mathbb{E}_\mu \left[\chi_A \mathcal{R}_T(\mathcal{F}) \right] + 1$$

By the definition of the event A, we have:

$$\begin{aligned}
\mathbb{E}_\mu\left[\chi_A \mathcal{R}_T(\mathcal{F})\right] &= \mathbb{E}_\mu\left[\chi_A \mathbb{E}_\varepsilon\left[\sup_{f\in\mathcal{F}} \sum_{X_i\neq x^*} \varepsilon_i f(X_i) + \sum_{X_i = x^*} \varepsilon_i f(X^*) \middle| X_1, \dots, X_T\right]\right] \\
&= \mathbb{E}_\mu\left[\chi_A \mathbb{E}_\varepsilon\left[\sup_{f\in\mathcal{F}} \sum_{X_i\neq x^*} \varepsilon_i f(X_i) \middle| X_1, \dots, X_T\right]\right] \\
&\leq \sup_{X_i} \mathbb{E}_\varepsilon\left[\sup_{f\in\mathcal{F}} \sum_{i=1}^{2\sigma T} \varepsilon_i f(X_i)\right]
\end{aligned}$$

We may now apply (Rudelson and Vershynin, 2006) to get that

$$\sup_{X_i} \mathbb{E}_\varepsilon\left[\sup_{f\in\mathcal{F}} \sum_{i=1}^{2\sigma T} \varepsilon_i f(X_i)\right] \leq C\sqrt{2\sigma T} \int_0^1 \sqrt{\mathsf{vc}(\mathcal{F}, \alpha)} d\alpha$$

Because $\mathsf{vc}(\mathcal{F}, \alpha) \leq C\alpha^{-p}$ for some $p < 2$, the result follows. ■

We are now ready to prove the main bound.

Proof (Proposition 8) Let $\alpha = \sqrt{\sigma}$ and set $\mathcal{F}, \mu, p_t, \mathcal{X}$ as above. By Lemma 29 and Lemma 30, we have

$$\frac{\mathbb{E}_{p_t}[\mathcal{R}_T(\mathcal{F})]}{\mathbb{E}_\mu[\mathcal{R}_T(\mathcal{F})]} \geq \frac{c\sqrt{\sigma \mathsf{vc}(\mathcal{F}, \sqrt{\sigma}) T}}{C\sqrt{\sigma T}} = c\sqrt{\mathsf{vc}(\mathcal{F}, \sqrt{\sigma})} \geq c\sigma^{-\frac{p}{4}}$$

where the last inequality follows from the assumption on the complexity of $\mathcal{F}$. We may now apply the lower bound in Proposition 18 with $\mathscr{D}$ just independent copies of p_t. The result follows. ■

Appendix E. Proof of Theorem 10

In this section we prove Theorem 10, which gives a regret bound for the follow-the-perturbed-leader (FTPL) algorithm (5) with respect to general classes $\mathcal{F}$ for convex, Lipschitz loss functions. This result bounds the regret of an FTPL style algorithm by a stability term and a term corresponding to the size of the perturbation. In particular, we bound the stability term by controlling the Wasserstein distance between the laws of $\widehat{y}_t$ and $\widehat{y}_{t+1}$. The techniques involved are of independent interest as we develop a novel Gaussian anti-concentration inequality that applies even when the labels are not assumed smooth. We begin by stating and proving the relevant variant of the BTL lemma. Then, in Appendix E.1 we provide the stability bound, using our Gaussian anti-concentration approach. We continue in Appendix E.2 by controlling the final stability term in the below decomposition, in the special case of linear loss. Finally, we conclude the proof in Appendix E.3 by extending from linear loss to general loss in the case of smooth labels and applying a discretization approach to recover full generality.

We consider the smoothed online setting with distribution μ. More specifically, we consider the following setting: for some parameter $n \in \mathbb{N}$, for each time step $t \in [T+1]$, consider points

$X_{t,1}, \ldots, X_{t,n} \in \mathcal{X}$, and define

$$\forall f \in \mathcal{F} : \qquad \hat{\omega}_{t,n}(f) := \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} \gamma_{t,i} \cdot f(X_{t,i}), \tag{12}$$

where $\gamma_{t,1}, \ldots, \gamma_{t,n}$ are i.i.d. standard normal random variables. We define $\hat{\mu}_{t,n}$ to be the distribution $\hat{\mu}_{t,n} := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_{t,i}}$, where $\delta_{X_{t,i}}$ denotes the point mass at $X_{t,i}$. In what follows we will consider iterates f_t, $1 \le t \le T+1$, satisfying, for some $\zeta > 0$,

$$L_{t-1}(f_t) + \eta \cdot \hat{\omega}_{t,n}(f_t) \le \operatorname*{argmin}_{f \in \mathcal{F}} L_{t-1}(f) + \eta \cdot \hat{\omega}_{t,n}(f) + \zeta. \tag{13}$$

We begin with a classic regret decomposition based on the well-known "Be-the-Leader" Lemma (Kalai and Vempala, 2005; Cesa-Bianchi and Lugosi, 2006). We first prove a related, auxiliary result that allows us to deal with different perturbations at each time step:

Lemma 31 *Suppose f_t, for $t \in [T]$, is defined as in (13), for any (adaptively chosen) sequence $(x_1, y_1), \ldots, (x_T, y_T) \in \mathcal{X} \times [-1, 1]$. Then it holds that*

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{t+1}(x_t), y_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right] \le \zeta \cdot (T+1) + 2 \cdot \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}\right].$$

Proof We first use induction on $T \ge 0$ to show the following statement:

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{t+1}(x_t), y_t)\right] \le \mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{T+1}(x_t), y_t) + \hat{\omega}_{T+1,n}(f_{T+1})\right] + \zeta \cdot T + \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right]. \tag{14}$$

To establish the base case $T = 0$, we note that $0 \le \mathbb{E}\left[\hat{\omega}_{1,n}(f_1)\right] + \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right]$, which follows because $\mathbb{E}\left[\sup_{f \in \mathcal{F}} -\hat{\omega}_{1,n}(f)\right] = \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right] \ge 0$ (by symmetry of the process $\hat{\omega}_{1,n}$).

Now assume that (14) holds at some step $T-1$, namely that

$$\mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_{t+1}(x_t), y_t)\right] \le \mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_{T}(x_t), y_t) + \hat{\omega}_{T,n}(f_{T})\right] + \zeta \cdot (T-1) + \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right]. \tag{15}$$

By definition of f_T in (13), we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_T(x_t), y_t) + \hat{\omega}_{T,n}(f_T)\right] &\le \mathbb{E}\left[\inf_{f \in \mathcal{F}} \sum_{t=1}^{T-1} \ell(f(x_t), y_t) + \hat{\omega}_{T,n}(f)\right] + \zeta \\
&= \mathbb{E}\left[\inf_{f \in \mathcal{F}} \sum_{t=1}^{T-1} \ell(f(x_t), y_t) + \hat{\omega}_{T+1,n}(f)\right] + \zeta \qquad (16) \\
&\le \mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_{T+1}(x_t), y_t) + \hat{\omega}_{T+1,n}(f_{T+1})\right] + \zeta, \qquad (17)
\end{aligned}$$

where (16) follows because, conditioned on $(x_1, y_1), \ldots, (x_{T-1}, y_{T-1})$, the process $\hat{\omega}_{T,n}$ is drawn independently, as is the process $\hat{\omega}_{T+1,n}$, and both have the same conditional distribution. From (15) and (17) we have that

$$\mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_{t+1}(x_t), y_t)\right] \leq \mathbb{E}\left[\sum_{t=1}^{T-1} \ell(f_{T+1}(x_t), y_t) + \hat{\omega}_{T+1,n}(f_{T+1})\right] + \zeta \cdot T + \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right].$$

Adding $\mathbb{E}[\ell(f_{T+1}(x_T), y_T)]$ to both sides establishes (14), thus completing the inductive hypothesis.

To complete the proof of the lemma, we note that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{T+1}(x_t), y_t) + \hat{\omega}_{T+1,n}(f_{T+1})\right] \\
\leq &\mathbb{E}\left[\inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t) + \hat{\omega}_{T+1,n}(f)\right] + \zeta \\
\leq &\mathbb{E}\left[\inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t) + \sup_{f' \in \mathcal{F}} \hat{\omega}_{T+1,n}(f')\right] + \zeta \\
= &\mathbb{E}\left[\inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right] + \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{T+1,n}(f)\right] + \zeta,
\end{aligned}$$

which implies, combined with (14) and the fact that $\hat{\omega}_{1,n}$ and $\hat{\omega}_{T+1,n}$ are identically distributed, that

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{t+1}(x_t), y_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right] \leq \zeta \cdot (T+1) + 2 \cdot \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}\right].$$

■

Using Lemma 31, we get a decomposition of the regret into a stability term and a perturbation size term. The stability term is further decomposed for the future analysis.

Lemma 32 *Let f_t be defined as in (13) and let $(x_1, y_1), \ldots, (x_T, y_T)$ be any sequence of elements in $\mathcal{X} \times [-1, 1]$. Let $(x'_1, y'_1), \ldots, (x'_T, y'_T)$ be a tangent sequence, meaning that for all $1 \leq t \leq T$, (x'_t, y'_t) is independent and identically distributed as (x_t, y_t) conditioned on (x_s, y_s) for $s < t$. Then we may upper bound the expected regret by the following expression:*

$$2\eta\mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}(f)\right] + \sum_{t=1}^{T} \mathbb{E}\left[\ell(f_t(x'_t), y'_t) - \ell(f_{t+1}(x'_t), y'_t)\right] + \sum_{t=1}^{T} \mathbb{E}\left[\ell(f_{t+1}(x'_t), y'_t) - \ell(f_{t+1}(x_t), y_t)\right] \tag{18}$$

Proof By Lemma 31, we have

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell(f_{t+1}(x_t), y_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t)\right] \leq \zeta \cdot (T+1) + 2 \cdot \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{1,n}\right]$$

Adding and subtracting $\ell(f_t(x_t), y_t)$ from both sides and rearranging yields

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \le \mathbb{E}\left[\sum_{t=1}^{T} \ell(f_t(x_t), y_t) - \ell(f_{t+1}(x_t), y_t)\right] + 2\eta\mathbb{E}\left[\sup_{f\in\mathcal{F}} \hat{\omega}_{1,n}(f)\right] + \zeta T$$

Now, note that f_t is independent of (x_t, y_t) be construction, so $\mathbb{E}\left[\ell(f_t(x_t), y_t)\right] = \mathbb{E}\left[\ell(f_t(x_t'), y_t')\right]$. Adding and subtracting $\ell(f_{t+1}(x_t'), y_t')$ yields

$$\mathbb{E}\left[\ell(f_t(x_t), y_t) - \ell(f_{t+1}(x_t), y_t)\right] = \mathbb{E}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t')\right] + \mathbb{E}\left[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t)\right]$$

Applying linearity of expectation concludes the proof. ■

The classic decomposition in (18) allows for the control of each term independently. For the first, empirical process theory allows us to control $\mathbb{E}\left[\sup \hat{\omega}_{1,n}(f)\right]$ by the entropy of $\mathcal{F}$. The last term, called "generalization error" in Haghtalab et al. (2022), can be controlled in the case of linear loss by appealing to standard uniform deviations bounds; this is done in Appendix E.2. The key term is the middle one, whose control guarantees that f_t and f_{t+1} are close in an appropriate sense. We now present this bound.

E.1. Regret Bound Using the Wasserstein Distance

In this section we provide our bound on the middle term of (18). In particular, for any t, we show that $\mathbb{E}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t')\right]$ is small. We leverage the fact that ℓ is Lipschitz in the first coordinate and use this fact along with the smoothness of x_t' to reduce to showing that $||f_t - f_{t+1}||_{L^2(\mu)}$ is small in expectation over the perturbation.

We first argue that it suffices to consider $\mathcal{F}$ such that

$$\inf_{f\in\mathcal{F}} ||f||_{L^2(\mu)} \ge \frac{2}{3} \tag{19}$$

Indeed, take any $\mathcal{F}$ and any μ. Enlarge $\mathcal{X}$ to $\mathcal{X} \cup \{x^*\}$, where x^* is a new point such that $f(x^*) = 1$ for all $f \in \mathcal{F}$. Let $\widetilde{\mu} = \frac{1}{3}\mu + \frac{2}{3}\delta_{x^*}$. Then if p_t is σ-smooth with respect to μ then it is $\frac{\sigma}{3}$-smooth with respect to $\widetilde{\mu}$. Moreover, $||f||_{L^2(\widetilde{\mu})} \ge \frac{2}{3}$ for all $f \in \mathcal{F}$, and since all f take the same value on x^*, an ERM oracle for the original class clearly yields an ERM oracle for the new class with domain $\mathcal{X} \cup \{x^*\}$ (the oracle can simply ignore all points of the form (x^*, y)). Thus, at the cost of shrinking σ by a factor of 3, we will suppose this lower bound throughout this section.

In Lemma 33 below, we show that if f_t, f_{t+1} are defined with respect to a common noise process $\omega(\cdot)$, then they are close with high probability. In the lemma, we consider an arbitrary Lipschitz loss function ℓ, and define $L_t(f) := \sum_{s=1}^{t} \ell(f(x_s), y_s)$.

Lemma 33 *Fix any $\zeta > 0$, $t \in \mathbb{N}$, ℓ as above, and an arbitrary sequence $(x_1, w_1), \dots, (x_{t-1}, w_{t-1}) \in \mathcal{X}\times\mathbb{R}$. Let ω denote a Gaussian process on a separable class $\mathcal{F}$ with covariance $\Sigma_{fg} = \mathbb{E}_{X\sim\mu}\left[f(X)g(X)\right]$ for some measure μ on $\mathcal{X}$ and, by abuse of notation, let $\omega(f)$ denote a single sample from this process. Suppose that f_t satisfies*

$$L_{t-1}(f_t) + \eta\omega(f_t) \le \inf_{f\in\mathcal{F}} L_{t-1}(f) + \eta\omega(f) + \zeta,$$

and for each $(x, w) \in \mathcal{X} \times [-1,]$, *there is some* $f_{t+1,x,w}$ *such that*

$$L_{t,x,w}(f_{t+1,x,w}) + \eta\omega(f_{t+1,x,w}) \leq \inf_{f \in \mathcal{F}} L_{t,x,w}(f) + \eta\omega(f) + \zeta.$$

Suppose further that $f_t, f_{t+1,x,w}$ *are measurable with respect to the* σ*-algebra generated by* ω. [9] *Then,*

$$\mathbb{P}\left(\sup_{(x,w)\in\mathcal{X}\times[-1,1]} ||f_t - f_{t+1,x,w}||_{L^2(\mu)} > \alpha\right) \leq \frac{8(L+2\zeta)^2}{\alpha^4\eta^2 \inf_{f\in\mathcal{F}} ||f||^6_{L^2(\mu)}} + \frac{4(L+2\zeta)}{\alpha^2\eta \inf_{f\in\mathcal{F}} ||f||^4_{L^2(\mu)}} \mathbb{E}\left[\sup_{f\in\mathcal{F}} \omega(f)\right]. \tag{20}$$

Proof As $\mathcal{F}$ is seperable, it suffices to take a countable dense subset and assume that $\mathcal{F}$ is countable. By assumption we can write $f_t = f_t(\omega)$ for some measurable function $f_t(\cdot)$ (where measurability is with respect to the product topology on $\mathbb{R}^{\mathcal{F}}$).

Let

$$A_t = \left\{g \in \mathcal{F} |\, ||g - f_t||_{L^2(\mu)} > \alpha \text{ and } L_{t-1}(g) + \eta\omega(g) \leq L_{t-1}(f_t) + \eta\omega(f_t) + 2L + \zeta\right\}$$

Note that for any g for which $L_{t-1}(g) + \eta\omega(g) > L_{t-1}(f_t) + \eta\omega(f_t) + 2L + \zeta$ and for any $(x, w) \in \mathcal{X} \times [-1, 1]$,

$$\begin{aligned} L_{t,x,w}(f_t) + \eta\omega(f_t) &= L_{t-1}(f_t) + \eta\omega(f_t) + \ell(f_t(x), w) \\ &< L_{t-1}(g) + \eta\omega(g) + \ell(f_t(x), w) - 2L - \zeta \\ &= L_t(g) + \eta\omega(g) + \ell(f_t(x), w) - \ell(g(x), w) - 2L - \zeta \\ &\leq L_t(g) + \eta\omega(g) - \zeta, \end{aligned}$$

where the final inequality follows because $|\ell(f_t(x), w) - \ell(g(x), w)| \leq 2L$ as ℓ is L-Lipschitz. Suppose that $g \not\in A_t$. Then either $||g - f_t||_{L^2(\mu)} \leq \alpha$ or, using the above display, for all x, w, $L_{t,x,w}(f_t) + \eta\omega(f_t) + \zeta < L_{t,x,w}(g) + \eta\omega(g)$, meaning that $f_{t+1,x,w}$ cannot be equal to g for any choice of x, w. Hence, the event that $\sup_{(x,w)\in\mathcal{X}\times[-1,1]} \|f_t - f_{t+1,x,w}\|_{L^2(\mu)} > \alpha$ implies that for some (x, w), $f_{t+1,x,w} \in A_t$. Thus, it suffices to bound the probability that A_t is nonempty.

Let $\mathcal{D}_\alpha(f) := \{g \in \mathcal{F} |\, ||g - f||_{L^2(\mu)} > \alpha\}$. As $\mathcal{F}$ is assumed countable, we have

$$\begin{aligned} &\mathbb{P}(|A_t| = 0) \\ &\geq \sum_{f\in\mathcal{F}} \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} L_{t-1}(g) + \eta\omega(g) - (2L + 4\zeta) \geq L_{t-1}(f) + \eta\omega(f) - \zeta\right) \\ &= \sum_{f\in\mathcal{F}} \mathbb{E}_y\left[\mathbb{P}\left[f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} L_{t-1}(g) + \eta\omega(g) - (2L + 4\zeta) \geq y - \zeta | L_{t-1}(f) + \eta\omega(f) = y\right]\right], \end{aligned} \tag{21}$$

where in (21) the expectation is over the distribution of $y = L_{t-1}(f) + \eta\omega(f)$. We now fix an f and let, for all $g \in \mathcal{F}$,

$$\Omega_t(g) = L_{t-1}(g) + \eta\omega(g).$$

[9] Here we write $L_{t,x,w}(f) = \sum_{s=1}^{t-1} \ell(f(x_s), w_s) + \ell(f(x), w)$.

Note that the process Ω_t is a Gaussian process and, conditioning on $\Omega_t(f) = y$ remains a Gaussian process. Then conditioned on $\Omega_t(f) = y$, Ω_t has mean

$$m_{f,y}(g) = L_{t-1}(g) + \frac{\mathbb{E}_{X\sim\mu}\left[f(X)g(X)\right]}{\left|\left|f\right|\right|_{L^2(\mu)}^2}\left(y - L_{t-1}(f)\right)$$

and covariance Σ^f; critically, Σ^f does not depend on y. Let

$$\gamma(g) = \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}^2}\mathbb{E}_{X\sim\mu}[f(X)g(X)] \qquad\qquad \beta(g) = \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}} - \gamma(g)$$

Then we have

$$m_{f,y+\frac{4(L+2\zeta)}{\alpha^2}} = m_{f,y} + \gamma. \tag{22}$$

Now suppose that $\left|\left|g - f\right|\right|_{L^2(\mu)} > \alpha$, i.e., $g \in \mathcal{D}_\alpha(f)$. Then

$$\mathbb{E}_{X\sim\mu}\left[f(X)g(X)\right] = \frac{\left|\left|f\right|\right|_{L^2(\mu)}^2 + \left|\left|g\right|\right|_{L^2(\mu)}^2}{2} - \frac{1}{2}\left|\left|f - g\right|\right|_{L^2(\mu)}^2 \le 1 - \frac{\alpha^2}{2}$$

by $\left|\left|f\right|\right|_\infty \le 1$ for all $f \in \mathcal{F}$. Thus for all such g,

$$\beta(g) = \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}^2}\left(1 - \mathbb{E}_{X\sim\mu}[f(X)g(X)]\right) \ge \frac{2(L+2\zeta)}{\left|\left|f\right|\right|_{L^2(\mu)}^2} \ge 2(L+2\zeta).$$

We now fix y and note that

$$\begin{aligned}
&\mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) - 2(L+2\zeta) \ge y - \zeta \mid \Omega_t(f) = y\right)\\
&\ge \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) - \beta(g) \ge y - \zeta \mid \Omega_t(f) = y\right)\\
&= \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) - \beta(g) - \gamma(g) + \gamma(g) \ge y - \zeta \mid \Omega_t(f) = y\right)\\
&= \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) + \gamma(g) \ge y - \zeta + \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2}(\mu)^2} \mid \Omega_t(f) = y\right)\\
&= \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) \ge y - \zeta + \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}^2} \mid \Omega_t(f) = y + \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}^2}\right)
\end{aligned}$$

where the inequality follows from the control of χ_B by β, the second equality follows from $\gamma + \beta = \frac{4(L+2\zeta)}{\alpha^2\left|\left|f\right|\right|_{L^2(\mu)}^2}$, and the last equality follows from the fact that a Gaussian process is determined only by its covariance and mean (in particular, we are using (22)).

Note that $L_{t-1}(f) + \eta\omega(f)$ is a Gaussian random variable with mean $L_{t-1}(f)$ and variance $\eta^2\left|\left|f\right|\right|_{L^2(\mu)}^2$. Denote by $q_f(y)$ the density of this distribution with respect to the Lebesgue measure on $\mathbb{R}$. Now, we compute

$$\begin{aligned}
&\mathbb{P}(|A_t| = 0)\\
&= \sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) - 2(L + 2\zeta) \geq y - \zeta \mid \Omega_t(f) = y\right) dy\\
&\geq \sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta + \frac{4(L+2\zeta)}{\alpha^2 \left|\left|f\right|\right|_{L^2(\mu)}^2} \mid \Omega_t(f) = y + \frac{4(L+2\zeta)}{\alpha^2 \left|\left|f\right|\right|_{L^2(\mu)}^2}\right) dy\\
&= \sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy\\
&+ \sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta + \frac{4(L+2\zeta)}{\alpha^2 \left|\left|f\right|\right|_{L^2(\mu)}^2} \mid \Omega_t(f) = y + \frac{4(L+2\zeta)}{\alpha^2 \left|\left|f\right|\right|_{L^2(\mu)}^2}\right) dy\\
&- \sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy.
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy\\
&= \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \mid \Omega_t(f) = y\right) dy = \mathbb{P}\left(f_t(\omega) = f\right),
\end{aligned}$$

where the first equality above follows since, conditioned on $\Omega_t(f) = y$, $f_t(\omega)$ implies that $\inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq \inf_{g \in \mathcal{F}} \Omega_t(g) \geq y - \zeta$. Hence

$$\sum_{f \in \mathcal{F}} \int_{-\infty}^{\infty} q_f(y) \cdot \mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g \in \mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy = \sum_{f \in \mathcal{F}} \mathbb{P}(f_t(\omega) = f) = 1.$$

For the second and third terms, using that

$$q_f(z) = \frac{1}{\sqrt{2\pi \cdot \eta^2 \left|\left|f\right|\right|_{L^2(\mu)}^2}} \cdot \exp\left(-\frac{1}{2} \cdot \frac{(z - L_{t-1}(f))^2}{\eta^2 \left|\left|f\right|\right|_{L^2(\mu)}^2}\right),$$

we observe that

$$q_f(y) - q_f\left(y - \frac{4(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}\right) = q_f(y)\left(1 - \exp\left(\frac{(y-L_{t-1}(f))^2}{2\eta^2 ||f||^2_{L^2(\mu)}} - \frac{\left(y - L_{t-1}(f) - \frac{4(L+2\zeta)}{\alpha^2||f||^2_{L^2(\mu)}}\right)^2}{2\eta^2 ||f||^2_{L^2(\mu)}}\right)\right)$$

$$\leq \frac{q_f(y)}{2\eta^2 ||f||^2_{L^2(\mu)}}\left(\left(y - L_{t-1}(f) - \frac{4(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}\right)^2 - (y - L_{t-1}(f))^2\right)$$

$$= \frac{q_f(y)}{2\eta^2 ||f||^2_{L^2(\mu)}}\left(\frac{16(L+2\zeta)^2}{\alpha^4 ||f||^4_{L^2(\mu)}} - \frac{8(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}(y - L_{t-1}(f))\right)$$

Thus we have

$$- \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} q_f(y)\cdot\mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta + \frac{4(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}} \mid \Omega_t(f) = y + \frac{4(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}\right) dy$$

$$+ \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} q_f(y)\cdot\mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy$$

$$= \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} \left(q_f(y) - q_f\left(y - \frac{4(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}\right)\right)\mathbb{P}\left(f_t(\omega) = f \text{ and } \inf_{g\in\mathcal{D}_\alpha(f)} \Omega_t(g) \geq y - \zeta \mid \Omega_t(f) = y\right) dy$$

$$\leq \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} \frac{q_f(y)}{2\eta^2 ||f||^2_{L^2(\mu)}}\left(\frac{16(L+2\zeta)^2}{\alpha^4 ||f||^4_{L^2(\mu)}} - \frac{8(L+2\zeta)}{\alpha^2 ||f||^2_{L^2(\mu)}}(y - L_{t-1}(f))\right)\mathbb{P}\left(f_t(\omega) = f | \Omega_t(f) = y\right) dy$$

$$\leq \frac{8(L+2\zeta)^2}{\alpha^4\eta^2 \inf_{f\in\mathcal{F}} ||f||^6_{L^2(\mu)}} \sum_{f\in\mathcal{F}} \mathbb{P}(f_t = f)$$
$$- \frac{4(L+2\zeta)}{\alpha^2\eta^2} \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} \frac{y - L_{t-1}(f)}{||f||^4} \mathbb{P}\left(f_t(\omega) = f | \Omega(f_t) = y\right) q_f(y) dy$$

$$= \frac{8(L+2\zeta)^2}{\alpha^4\eta^2 \inf_{f\in\mathcal{F}} ||f||^6_{L^2(\mu)}} - \frac{4(L+2\zeta)}{\alpha^2\eta^2} \sum_{f\in\mathcal{F}} \int_{-\infty}^{\infty} \frac{y - L_{t-1}(f)}{||f||^4} \mathbb{P}\left(f_t(\omega) = f | \Omega(f_t) = y\right) q_f(y) dy$$

$$\overset{(*)}{\leq} \frac{8(L+2\zeta)^2}{\alpha^4\eta^2 \inf_{f\in\mathcal{F}} ||f||^6_{L^2(\mu)}} + \frac{4(L+2\zeta)}{\alpha^2\eta \inf_{f\in\mathcal{F}} ||f||^4_{L^2(\mu)}} \mathbb{E}\left[\sup_{f\in\mathcal{F}} \omega(f)\right],$$

where the first inequality uses the previous computation, the second follows by linearity, and the last equality follows because $f_t \in \mathcal{F}$ is distinct. To see that inequality $(*)$ holds, note that, by definition, $(y - L_{t-1}(f)) \overset{d}{=} \eta\omega(f)$ and thus,

$$\frac{1}{\eta}\sum_{f\in\mathcal{F}}\int_{-\infty}^{\infty}\frac{y - L_{t-1}(f)}{||f||^4}\mathbb{P}\left(f_t(\omega) = f|\Omega(f_t) = y\right)q_f(y)dy = \mathbb{E}\left[\frac{\omega(f_t)}{||f_t||^4}\right]$$

We now have

$$\begin{aligned}
-\mathbb{E}\left[\frac{\omega(f_t)}{||f_t||^4}\right] &\leq -\mathbb{E}\left[\inf_{f\in\mathcal{F}}\frac{\omega(f)}{||f||^4}\right]\\
&\overset{(a)}{=} \left|\mathbb{E}\left[-\inf_{f\in\mathcal{F}}\frac{\omega(f)}{||f||^4}\right]\right|\\
&\overset{(b)}{\leq} \mathbb{E}\left[\left|-\inf_{f\in\mathcal{F}}\frac{\omega(f)}{||f||^4}\right|\right]\\
&\overset{(c)}{=} \mathbb{E}\left[\sup_{f\in\mathcal{F}}\frac{|\omega(f)|}{||f||^4}\right]\\
&\overset{(d)}{\leq} \frac{1}{\inf_{f\in\mathcal{F}}||f||^4}\mathbb{E}\left[\sup_{f\in\mathcal{F}}|\omega(f)|\right]\\
&\overset{(e)}{\leq} \frac{1}{\inf_{f\in\mathcal{F}}||f||^4}\mathbb{E}\left[\sup_{f\in\mathcal{F}}\omega(f)\right]
\end{aligned}$$

Note that by Jensen's inequality,

$$\mathbb{E}\left[\inf_{f\in\mathcal{F}}\frac{\omega(f)}{||f||^4}\right] \leq \inf_{f\in\mathcal{F}}\mathbb{E}\left[\frac{\omega(f)}{||f||^4}\right] = 0$$

and so (a) holds. Then (b) follows from Jensen's inequality, (c) follows from the symmetry of the Gaussian, (d) follows by linearity, and (e) follows from the Sudakov-Fernique inequality (Sudakov, 1971; Fernique, 1975) applied to the contraction $|\cdot|$. The result follows. ■

Note that Lemma 33 holds for an arbitrary measure μ on $\mathcal{X}$ and applies even in the case where x_t, w_t are adversarially chosen. To apply this result, we choose μ to be the empirical measure on the perturbation samples $X_{t,i}$. We consider two cases. First, we suppose that we are in a classification setting, where we get better rates. We then prove the more general setting where $\mathcal{F}$ is real-valued.

Lemma 34 *Suppose that we are in the smoothed online setting, f_t is chosen so as to satisfy (13), and the empirical distribution $\hat{\mu}_{t,n}$ satisfies*

$$\sup_{f,f'\in\mathcal{F}}\left|||f - f'||^2_{L^2(\mu)} - ||f - f'||^2_{L^2(\hat{\mu}_{t,n})}\right| \leq \Delta.$$

Suppose further that for all $f \in \mathcal{F}$ and all $x \in \mathcal{X}$, $f(x) \in \{\pm 1\}$. Then

$$\mathbb{E}\left[\ell(f_t(x_t), y_t) - \ell(f_{t+1}(x_t'), y_t')\right] \leq \frac{30(L + 2\zeta)^3\log\eta}{\sigma\eta}\mathbb{E}\left[1 + \sup_{f\in\mathcal{F}}\hat{\omega}_{t,n}(f)\right] + \frac{2L\Delta}{\sigma}$$

Proof Let $(x_t', y_t') \in \mathcal{X} \times [-1,1]$ be a sample distributed independently and identically to (x_t, y_t) conditioned on $(x_1, y_1), \dots, (x_{t-1}, y_{t-1}), f_1, \dots, f_{t-1}$. Since f_t is selected independently of (x_t, y_t), it is immediate that

$$\mathbb{E}[\ell(f_t(x_t), y_t)] = \mathbb{E}[\ell(f_t(x_t'), y_t')]. \tag{23}$$

Therefore, it suffices to bound

$$\mathbb{E}[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t')].$$

Let us now fix any values of $S := \{(x_1, y_1), \dots, (x_{t-1}, y_{t-1}), f_1, \dots, f_{t-1}\}$. By Lemma 33, there is a joint distribution ν over (f_t', f_{t+1}'), so that, conditioned on S:

1. The marginal distribution of f_t', conditioned on S, is equal to the marginal distribution of f_t, conditioned on S.

2. The marginal distribution of f_{t+1}', conditioned on S, is equal to the marginal distribution of f_{t+1}', conditioned on S.

3. It holds that

$$\mathbb{P}_\nu\left(||f_t' - f_{t+1}'||_{L^2(\mu)} > \alpha\right) \le \frac{8(L+2\zeta)^2}{\alpha^4 \eta^2 \inf_{f \in \mathcal{F}} ||f||^6_{L^2(\hat{\mu}_{t,n})}} + \frac{4(L+2\zeta)}{\alpha^2 \eta \inf_{f \in \mathcal{F}} ||f||^4_{L^2(\mu)}} \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right]. \tag{24}$$

In particular, this joint distribution ν is constructed by setting f_t' to equal f_t from (13) and then defining f_{t+1}' so that

$$L_{t-1}(f_{t+1}') + \ell(f_{t+1}'(x_t), y_t) + \eta \hat{\omega}_{t,n}(f_{t+1}') \le \operatorname*{argmin}_{f \in \mathcal{F}} L_{t-1}(f) + \ell(f(x_t), y_t) + \eta \cdot \hat{\omega}_{t,n}(f) + \zeta.$$

Note that $\hat{\omega}_{t,n}$ has been used here as opposed to $\hat{\omega}_{t+1,n}$. Since $\hat{\omega}_{t,n}$ and $\hat{\omega}_{t+1,n}$ have the same distribution, the first two requirements of ν above are immediate. To see that the third holds, we note that, in the notation of Lemma 33, f_{t+1}' is exactly $f_{t+1,x,w}$ with $x = x_t, w = y_t$, and thus (24) is immediate from (20) (with the distribution μ set to $\hat{\mu}_{t,n}$ and the Gaussian process ω set to $\hat{\omega}_{t,n}$).

By the first two conditions above of the coupling ν and since (x_t', y_t') is drawn independently from (f_t', f_{t+1}'), it holds that $\mathbb{E}[\ell(f_t(x_t'), y_t')] = \mathbb{E}[\ell(f_t'(x_t'), y_t')]$ and $\mathbb{E}[\ell(f_{t+1}(x_t'), y_t')] = \mathbb{E}[\ell(f_{t+1}'(x_t'), y_t')]$.

Fix any $0 < \beta < \alpha$. By L-Lipschitzness of ℓ, the fact that $f_t', f_{t+1}' \in \{\pm 1\}$, and the fact that (x_t', y_t') are drawn independently from f_t', f_{t+1}', we have

$$\begin{aligned}
&\mathbb{E}_{\nu,\, x_t' \sim p_t,\, y_t'}\left[(\ell(f_t'(x_t'), y_t') - \ell(f_{t+1}'(x_t'), y_t')) \cdot \chi_{\beta \le ||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \le \alpha}\right] \qquad (25)\\
&\quad \le L \mathbb{E}_{\nu,\, x_t' \sim p_t}\left[\left|f_t'(x_t') - f_{t+1}'(x_t')\right| \cdot \chi_{\beta \le ||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \le \alpha}\right]\\
&\quad = L \mathbb{E}_\nu\left[\mathbb{E}_{x_t' \sim p_t}[(f_t'(x_t') - f_{t+1}'(x_t'))^2 \mid f_t', f_{t+1}'] \cdot \chi_{\beta \le ||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \le \alpha}\right]\\
&\quad \le \frac{L}{\sigma} \cdot \mathbb{E}_\nu\left[\mathbb{E}_{x_t' \sim \mu}[(f_t'(x_t') - f_{t+1}'(x_t'))^2 \mid f_t', f_{t+1}'] \cdot \chi_{\beta \le ||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \le \alpha}\right]\\
&\quad \le \frac{L \cdot (\alpha^2 + \Delta)}{\sigma} \mathbb{P}_\nu(||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} > \beta)
\end{aligned}$$

Set $S = \lceil \log \min\{\eta, 1/\Delta\} \rceil$ and let $\alpha_i = 2^{\frac{1-i}{2}}$. Then, noting that $||f||_{L^2(\hat{\mu}_n)} = 1$ for all $f \in \mathcal{F}$, we see, using (24),

$$
\begin{aligned}
&\mathbb{E}_{x_t' \sim p_t}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t')\right] \\
&= \mathbb{E}_{x_t' \sim p_t}\left[\ell(f_t'(x_t'), y_t') - \ell(f_{t+1}'(x_t'), y_t')\right] \\
&\leq \mathbb{E}_{x_t' \sim p_t}\left[\left(\ell(f_t'(x_t'), y_t') - \ell(f_{t+1}'(x_t'), y_t')\right) \cdot \chi_{||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \leq \alpha_S}\right] \\
&\quad + \sum_{i=0}^{S} \mathbb{E}_{x_t' \sim p_t}\left[\left(\ell(f_t'(x_t'), y_t') - \ell(f_{t+1}'(x_t'), y_t')\right) \cdot \chi_{\alpha_i < ||f_t' - f_{t+1}'||_{L^2(\hat{\mu}_{t,n})} \leq \sqrt{2}\alpha_i}\right] \\
&\leq \frac{L(\alpha_S^2 + \Delta)}{\sigma} + \sum_{i=0}^{S} \left(\frac{8(L+2\zeta)^2}{\alpha_i^4 \eta^2} + \frac{4(L+2\zeta)}{\alpha_i^2 \eta}\mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right]\right) \frac{L(\alpha_i^2 + \Delta)}{\sigma} \\
&\leq \frac{4L}{\sigma} \cdot \left(\frac{1}{\eta} + \Delta\right) + \sum_{i=0}^{S} \left(\frac{8(L+2\zeta)^2}{\alpha_i^2 \eta^2} + \frac{4(L+2\zeta)}{\eta}\mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right]\right) \frac{2L}{\sigma} \\
&\leq \frac{4L}{\sigma} \cdot \left(\frac{1}{\eta} + \Delta\right) + \sum_{i=0}^{S} \left(\frac{8(L+2\zeta)^2}{\eta} + \frac{4(L+2\zeta)}{\eta}\mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right]\right) \frac{2L}{\sigma} \\
&\leq \frac{30(L+2\zeta)^3 \log \eta}{\sigma \eta}\mathbb{E}\left[1 + \sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right] + \frac{2L\Delta}{\sigma},
\end{aligned}
$$

where the second inequality follows by the above argument (setting $\beta = 0$ for the first term) and from (24), the third inequality follows from $\Delta \leq \alpha_i^2$ for all $i \leq S$, the penultimate inequality follows from $\frac{1}{\alpha_i^2} \leq \frac{1}{\eta}$ for $i \leq S$ and the last inequality follows from $S \leq \log \eta$. The result follows from the above display and (23). ■

We now prove a more general result that has worse dependence on η.

Lemma 35 *Suppose that we are in the smoothed online setting, f_t is chosen so as to satisfy (13), and the empirical distribution $\hat{\mu}_{t,n}$ satisfies*

$$\sup_{f, f' \in \mathcal{F}} \left| ||f - f'||_{L^2(\mu)}^2 - ||f - f'||_{L^2(\hat{\mu}_{t,n})}^2 \right| \leq \Delta.$$

Suppose further that $\inf_{f \in \mathcal{F}} ||f||_{L^2(\hat{\mu}_{t,n})}^2 \geq 1/2$. Then

$$\mathbb{E}\left[\ell(f_t(x_t), y_t) - \ell(f_{t+1}(x_t'), y_t')\right] \leq \frac{1200(L+2\zeta)^3 \log \eta}{\sqrt{\sigma \eta}}\mathbb{E}\left[1 + \sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f)\right] + 4L \cdot \sqrt{\frac{\Delta}{\sigma}}$$

Proof Exactly as in the proof of Lemma 34, we introduce the independent sample (x_t', y_t'), as well as the coupling ν over (f_t', f_{t+1}'). In particular, (23) and (24) continue to hold. Next, we bound the expression in (25) as in the proof of Lemma 34 but this time applying Jensen's inequality:

$$\begin{aligned}
&\mathbb{E}_{x_t'\sim p_t}\left[(\ell(f_t(x_t'),y_t')-\ell(f_{t+1}(x_t'),y_t'))\chi_{\beta\leq||f_t-f_{t+1}||_{L^2(\mu_n)}\leq\alpha}\right]\\
&\qquad\leq L\mathbb{E}_{x_t'\sim p_t}\left[\left|f_t(x_t)-f_{t+1}(x_t)\right|\chi_{\beta\leq||f_t-f_{t+1}||_{L^2(\mu_n)}\leq\alpha}\right]\\
&\qquad\leq L\mathbb{P}\left(\sup_{x,y}||f_t-f_{t+1,x,y}||_{L^2(\hat{\mu}_n)}>\beta\right)\sqrt{\mathbb{E}_{x_t'\sim p_t}\left[(f_t(x_t')-f_{t+1}(x_t'))^2\chi_{||f_t-f_{t+1}||_{L^2(\mu_n)}\leq\alpha}\right]}\\
&\qquad\leq L\mathbb{P}\left(\sup_{x,y}||f_t-f_{t+1,x,y}||_{L^2(\hat{\mu}_n)}>\beta\right)\sqrt{\frac{\alpha^2+\Delta}{\sigma}}.
\end{aligned}$$

Setting $S=\lceil\log\min\{\sqrt{\eta},1/\sqrt{\Delta}\}\rceil$ and $\alpha_i=2^{1-i}$ for $0\leq i\leq S$, we have:

$$\begin{aligned}
&\mathbb{E}_{x_t'\sim p_t}\left[\ell(f_t(x_t'),y_t')-\ell(f_{t+1}(x_t'),y_t')\right]\\
&\leq\mathbb{E}_{x_t'\sim p_t}\left[\left(\ell(f_t(x_t'),y_t')-\ell(f_{t+1}(x_t'),y_t')\right)\cdot\chi_{||f_t-f_{t+1}||_{L^2(\hat{\mu}_n)}\leq\alpha_S}\right]\\
&\quad+\sum_{i=0}^{S}\mathbb{P}\left(||f_t-f_{t+1}||_{L^2(\hat{\mu}_n)}>\alpha_i\right)\mathbb{E}_{x_t\sim p_t}\left[\left(\ell(f_t(x_t'),y_t')-\ell(f_{t+1}(x_t'),y_t')\right)\cdot\chi_{||f_t-f_{t+1}||_{L^2(\hat{\mu}_n)}\leq2\alpha_i}\right]\\
&\leq L\cdot\sqrt{\frac{\alpha_S^2+\Delta}{\sigma}}+\sum_{i=0}^{S}\left(\frac{8(L+2\zeta)^2}{\alpha_i^4\eta^2\inf_{f\in\mathcal{F}}||f||_{L^2(\mu)}^6}+\frac{4(L+2\zeta)}{\alpha_i^2\eta\inf_{f\in\mathcal{F}}||f||_{L^2(\mu)}^4}\mathbb{E}\left[\sup_{f\in\mathcal{F}}\hat{\omega}_n(f)\right]\right)L\cdot\sqrt{\frac{\alpha_i^2+\Delta}{\sigma}}\\
&\leq L\cdot\sqrt{\frac{8(\Delta+1/\eta)}{\sigma}}+\sum_{i=0}^{S}\left(\frac{512(L+2\zeta)^2}{\alpha_i^3\eta^2}+\frac{64(L+2\zeta)}{\alpha_i\eta}\mathbb{E}\left[\sup_{f\in\mathcal{F}}\hat{\omega}_n(f)\right]\right)L\cdot\sqrt{\frac{2}{\sigma}}\\
&\leq L\cdot\sqrt{\frac{8(\Delta+1/\eta)}{\sigma}}+\sum_{i=0}^{S}\left(\frac{512(L+2\zeta)^2}{\sqrt{\eta}}+\frac{64(L+2\zeta)}{\sqrt{\eta}}\mathbb{E}\left[\sup_{f\in\mathcal{F}}\hat{\omega}_n(f)\right]\right)L\cdot\sqrt{\frac{2}{\sigma}}\\
&\leq\frac{1200(L+2\zeta)^3\log\eta}{\sqrt{\sigma\eta}}\mathbb{E}\left[1+\sup_{f\in\mathcal{F}}\hat{\omega}_n(f)\right]+4L\cdot\sqrt{\frac{\Delta}{\sigma}},
\end{aligned}$$

where we used the fact that $||f||_{L^2(\hat{\mu}_n)}\geq\frac{1}{2}$, $\alpha_i^2\geq\Delta$, and $\frac{1}{\alpha_i}\leq\sqrt{\eta}$ for all $i\leq S$. ■

Finally, we need to verify that $||\cdot||_{L^2(\mu)}$ and $||\cdot||_{L^2(\mu_n)}$ are close together, a key condition of Lemmas 34 and 35. The below standard result shows that this condition holds in high probability.

Lemma 36 *There is a constant $C>0$ so that the following holds. Consider any distribution μ over $\mathcal{X}$, suppose $x_1,\ldots,x_n\sim\mu$ are sampled independently, and define $\hat{\mu}_n:=\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}$. For any $\delta>0$, with probability at least $1-\delta$ over the x_i, we have*

$$\sup_{f,f'\in\mathcal{F}}\left|||f-f'||_{L^2(\mu)}^2-||f-f'||_{L^2(\hat{\mu}_n)}^2\right|\leq\frac{C}{\sqrt{n}}\cdot\left(\frac{1}{n}\mathcal{G}_n(\mathcal{F})+\sqrt{\frac{\log\left(\frac{1}{\delta}\right)}{n}}\right).$$

Proof Write $\mathcal{F}^2 = \{x \mapsto f(x)^2 : f \in \mathcal{F}\}$. Standard results in empirical processes, such as (Wainwright, 2019, Theorem 4.10) guarantee that with probability at least $1 - \delta$,

$$\sup_{f, f' \in \mathcal{F}} \left| \left|\left|f - f'\right|\right|_{L^2(\mu)}^2 - \left|\left|f - f'\right|\right|_{L^2(\mu_n)}^2 \right| \leq C \left(\frac{1}{n} \mathcal{R}_n(\mathcal{F}^2) + \sqrt{\frac{\log\left(\frac{1}{\delta}\right)}{n}} \right)$$

Noting that $\mathcal{F}$ has image in $[-1, 1]$ and thus the square is 2-Lipschitz, we may apply contraction and the bound of Rademacher complexity by Gaussian complexity to conclude the proof. ■

E.2. Bounding the generalization error

In this section, we bound the final term in (18). This term was called the "Generalization Error" in Haghtalab et al. (2022) and our control of this quantity follows a similar general approach as their Lemma 4.5. For our proof, we require the following variant of the coupling approach of Lemma 14:

Lemma 37 (Lemma 4.6 of Haghtalab et al. (2022)) *Fix a distribution μ on a set $\mathcal{X}$ and suppose that $p \in \mathfrak{P}(\sigma, \mu)$. Suppose that $X_1, \dots, X_m \sim \mu$ are iid. Then there is an external probability space with sample space Ω and measure ν which produces a sample $R \sim \nu$ so that the following holds. There is a measurable function $I : \mathcal{X}^m \times \Omega \to [m]$ so that, for some event $\mathcal{E} = \mathcal{E}(X_1, \dots, X_m, R)$ with $\Pr(\mathcal{E}) \geq 1 - (1 - \sigma)^m$, $(X_I | \mathcal{E}, (X_i)_{i \neq I}) \sim p$ (in words, conditioned on the event $\mathcal{E}$ and the value of any measurable function of $(x_i)_{i \neq I}$, X_I has conditional distribution p).*

We restrict our focus to linear loss $\ell(f(x), y) = yf(x)$. and provide the following bound:

Lemma 38 *Suppose that we are in the smoothed online setting. Fix any $t \in [T - 1]$ and suppose that f_{t+1} is chosen so as to satisfy (13), with the process $\hat{\omega}_{n,t}(\cdot)$ defined as in (12), and the parameters η, n satisfy $\eta/\sqrt{n} \geq L$. Furthermore, let (x_t', y_t') be an independent sample drawn from the conditional distribution of (x_t, y_t) given $\{(x_s, y_s)\}_{s \leq t-1}$ and $\{f_s\}_{s \leq t-1}$. Then, for some constant $c_0 \in (0, 1)$, it holds that*

$$\mathbb{E}[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t)] \leq 4L \cdot \frac{\log T}{c_0 \sigma n} \cdot \mathcal{R}_{c_0 \sigma n / (2 \log T)}(\mathcal{F}) + 2\zeta + \frac{2Ln\sigma}{T^2}.$$

Proof Fix any realization of $(x_1, y_1), \dots, (x_{t-1}, y_{t-1}), f_1, \dots, f_{t-1}$. Recalling the definition of smoothed adversary, let p_t denote the conditional distribution of x_t (which is the same as the conditional distribution of x_t') given $(x_1, y_1), \dots, (x_{t-1}, y_{t-1})$. Also let $q_t(\cdot | x_t)$ denote the conditional distribution of y_t given x_t (and conditioned on the fixed values of (x_s, y_s), $s < t$, which are omitted for clarity). Recall that we make no smoothness assumption on q_t. We denote the distribution of (x_t, y_t), where $x_t \sim p_t$ and $y_t \sim q_t(\cdot | x_t)$ as $p_t \odot q_t$. Furthermore let $\tilde{p}_t$ denote the conditional distribution of $(x_t, \operatorname{sign}(y_t))$ given $(x_1, y_1), \dots, (x_{t-1}, y_{t-1})$ (i.e., where $x_t \sim p_t$ and $y_t \sim q_t(\cdot | x_t)$). Set $c_0 := \Pr_{\gamma \sim \mathcal{N}(0,1)}(\gamma \geq 1) > 0$, where $\mathcal{N}(0, 1)$ is the standard normal distribution. Let $\tilde{\mu} \in \Delta(\mathcal{X} \times \{-1, 0, 1\})$ denote the product of μ and the distribution over $\{-1, 0, 1\}$ which puts mass c_0 on $1, -1$ and mass $1 - 2c_0$ on 0. For any measurable subset $\mathcal{A} \subset \mathcal{X}$ and and $b \in \{-1, 1\}$, we have, from σ-smoothness of p_t that for each $b \in \{-1, 1\}$,

$$\frac{\tilde{p}_t(\mathcal{A} \times \{b\})}{\tilde{\mu}(\mathcal{A} \times \{b\})} = \frac{p_t(\mathcal{A}) \cdot \Pr_{(x,y) \sim \tilde{p}_t}(y = b | x \in \mathcal{A})}{\mu(\mathcal{A}) \cdot c_0} \leq \frac{1}{c_0} \cdot \frac{p_t(\mathcal{A})}{\mu(\mathcal{A})} \leq \frac{1}{c_0 \sigma},$$

meaning that $\tilde{p}_t \in \mathfrak{P}(c_0\sigma, \tilde{\mu})$.

Define the function $\mathrm{thr} : \mathbb{R} \to \{-1, 0, 1\}$ as follows:

$$\mathrm{thr}(y) := \begin{cases} 1 & : y \geq 1 \\ 0 & : y \in (-1, 1) \\ -1 & : y \leq -1. \end{cases}$$

Recall the i.i.d. samples (x_i, γ_i), $i \in [n]$ defining the process $\hat{\omega}_n(\cdot)$ in (12); note that $(x_i, \mathrm{thr}(\gamma_i)) \sim \tilde{\mu}$ by the definition of $\tilde{\mu}$ and thr. For $i \in [n]$ set $z_i = (x_i, \mathrm{thr}(\gamma_i))$. Fix $m = 2\log T \cdot \frac{1}{c_0\sigma}$. We divide the i.i.d. sample $(z_1, \ldots, z_n)$ into n/m groups of m samples each: the first group consists of $(z_1, \ldots, z_m)$, the second consists of $(z_{m+1}, \ldots, z_{2m})$, and so on. By Lemma 37, for each group index $0 \leq j < n/m$, letting Ω_j denote the sample space of the external probability space in the statement (of Lemma 37) and $R_j \in \Omega_j$ denote the corresponding random variable, there is a function $I_j : (\mathcal{X} \times \{-1, 0, 1\})^m \times \Omega_j$ so that for some event $\mathcal{E}_j = \mathcal{E}_j(z_{jm+1}, \ldots, z_{jm+m}, R_j)$ occuring with probability at least $1 - (1 - c_0\sigma)^m$, letting $I_j = I_j(z_{jm+1}, \ldots, z_{jm+m}, R_j)$,

$$(z_I | \mathcal{E}_j, (z_{jm+i} : i \neq I)) \sim p_t.$$

In particular, we have applied Lemma 37 with μ set to $\tilde{\mu}$ and p set to $\tilde{p}_t$. Write $\mathcal{E} := \cap_{0 \leq j < n/m} \mathcal{E}_j$, so that $\Pr(\mathcal{E}) \geq 1 - (n/m) \cdot (1 - c_0\sigma)^m$. Let $\mathcal{I} \in [n]^{n/m}$ be the (random) vector defined as $\mathcal{I} = (I_0, \ldots, I_{n/m-1})$, and let $\bar{\mathcal{I}} \in [n]^{n-(n/m)}$ be the vector defined as $(i \in [n] : i \notin \mathcal{I})$. Since the individual groups $(z_1, \ldots, z_m), (z_{m+1}, \ldots, z_{2m}), \ldots$ are mutually independent, it follows that

$$((z_i)_{i \in \mathcal{I}} | \mathcal{E}, (z_i : i \in \bar{\mathcal{I}})) \sim \tilde{p}_t^{\otimes m}, \tag{26}$$

i.e., conditioned on all $z_i,\ i \in \bar{\mathcal{I}}$, the distribution of $z_i,\ i \in \mathcal{I}$ is i.i.d. according to $\tilde{p}_t$. Let us write the vector $(z_i)_{i \in \mathcal{I}}$ as $w \in (\mathcal{X} \times \{-1, 1\})^{n/m}$. Note that by independence of $\hat{\omega}_{n,t}$ across t, the distribution of (x_t, y_t) is independent of $z_1, \ldots, z_n, \mathcal{E}, \mathcal{I}$; further, $x_t \sim p_t$ and $y_t \sim q_t(\cdot | x_t)$.

For each $i \in [n]$, let $\hat{y}_i \in \mathbb{R}$ denote an independent sample from $q_t(\hat{y}_i | x_i)$ conditioned on $\mathrm{sign}(\hat{y}_i) = \mathrm{sign}(\gamma_i)$. Recalling the definition of the (random) index I_j above, we have $z_{I_j} = (x_{I_j}, \mathrm{thr}(\gamma_{I_j}))$. Recalling that $(x_{I_j}, \mathrm{thr}(\gamma_{I_j})) \sim \tilde{p}_t$ conditioned on $\mathcal{E}, (z_i : i \in \bar{\mathcal{I}})$ (which follows from (26)), which in particular means that $\mathrm{sign}(\gamma_{I_j}) = \mathrm{thr}(\gamma_{I_j}) \in \{-1, 1\}$, it follows that $(x_{I_j}, \hat{y}_{I_j})$ has the same distribution as (x_t, y_t) (namely, $p_t \odot q_t$) and both are independent, conditioned on $((z_i : i \neq I_j), \mathcal{E})$. In particular, conditioned on the event $\mathcal{E}$, the distributions of the following vectors in $(\mathcal{X} \times \mathbb{R})^{n+1}$ are the same:

$$((x_t, y_t), (x_{I_j}, \hat{y}_{I_j}), (z_i : i \neq I_j)) \overset{d}{=} ((x_{I_j}, \hat{y}_{I_j}), (x_t, y_t), (z_i : i \neq I_j)),$$

where $\overset{d}{=}$ denotes equality in distribution and the above notation means that the entries $(z_i : i \neq I_j)$ are concatenated to the other entries. Since the values of γ_i, $i \in [n]$ are independent and identically distributed conditioned on $(z_1, \ldots, z_n)$, it follows that conditioned on the event $\mathcal{E}$, the distributions of the following vectors are the same:

$$((x_t, y_t), (x_{I_j}, \hat{y}_{I_j}), ((x_i, \gamma_i) : i \neq I_j)) \overset{d}{=} ((x_{I_j}, \hat{y}_{I_j}), (x_t, y_t), ((x_i, \gamma_i) : i \neq I_j)).$$

Furthermore, note that under the event $\mathcal{E}$, we have that $\gamma_{I_j} = |\gamma_{I_j}| \cdot \mathrm{thr}(\gamma_{I_j}) = |\gamma_{I_j}| \cdot \mathrm{sign}(\hat{y}_{I_j})$ and $|\gamma_{I_j}| \geq 1$ (as $\mathrm{thr}(\gamma_{I_j}) \in \{-1, 1\}$ under the event $\mathcal{E}$). From (13), f_{t+1} is a ζ-approximate minimizer (among $f \in \mathcal{F}$) of

$$\begin{aligned} &\sum_{s=1}^{t} \ell(f(x_s), y_s) + \sum_{i=1}^{n} \frac{\eta \gamma_i}{\sqrt{n}} \cdot f(x_i) \\ =&y_t \cdot f(x_t) + \frac{\eta \cdot |\gamma_{I_j}|}{\sqrt{n}} \cdot \mathrm{sign}(\hat{y}_{I_j}) \cdot f(\hat{x}) + \sum_{i \neq I_j} \frac{\eta \gamma_i}{\sqrt{n}} \cdot f(x_i) + \sum_{s=1}^{t-1} y_s \cdot f(x_s). \end{aligned}$$

Since $\eta/\sqrt{n} \geq L$ by assumption, it follows from Lemma 39 that $\mathbb{E}[y_t \cdot f_{t+1}(x_t) \,|\, \mathcal{E}] \geq \mathbb{E}[\hat{y}_{I_j} \cdot f_{t+1}(x_{I_j}) \,|\, \mathcal{E}] - 2\zeta$.

Further, letting $(x'_{t,1}, y'_{t,1}), \ldots, (x'_{t,n/m}, y'_{t,n/m})$ denote i.i.d. samples from the distribution of (x_t, y_t) (independent of (x_t, y_t)), it is immediate that for all $0 \leq j < n/m$,

$$\mathbb{E}[y'_{t,j} \cdot f_{t+1}(x'_{t,j}) \,|\, \mathcal{E}] = \mathbb{E}[y'_t \cdot f_{t+1}(x'_t) \,|\, \mathcal{E}].$$

Then it follows that

$$\begin{aligned} &\frac{n}{m} \cdot \mathbb{E}[y'_t \cdot f_{t+1}(x'_t) - y_t \cdot f_{t+1}(x_t) - 2\zeta \,|\, \mathcal{E}] \\ \leq& \mathbb{E}\left[\sum_{j=0}^{n/m-1} y'_{t,j} \cdot f_{t+1}(x'_{t,j}) - \sum_{j=0}^{n/m-1} \hat{y}_{I_j} \cdot f_{t+1}(x_{I_j}) \,|\, \mathcal{E}\right] \\ \leq& \mathbb{E}\left[\sup_{f \in \mathcal{F}} \sum_{j=0}^{n/m-1} y'_{t,j} \cdot f(x'_{t,j}) - \hat{y}_{I_j} \cdot f(x_{I_j}) \,|\, \mathcal{E}\right] \\ \leq& \mathbb{E}_{(\bar{x}_j, \bar{y}_j), (\bar{x}'_j, \bar{y}'_j) \sim p_t \odot q_t \,:\, 0 \leq j < n/m} \left[\sup_{f \in \mathcal{F}} \sum_{j=0}^{n/m-1} \bar{y}'_j \cdot f(\bar{x}'_j) - \bar{y}_j \cdot f(\bar{x}_j)\right] \qquad (27) \\ \leq& 2 \cdot \mathbb{E}_{(\bar{x}_j, \bar{y}_j) \sim p_t \odot q_t,\, \varepsilon_j \sim \mathrm{Unif}(\pm 1) \,:\, 0 \leq j < n/m} \left[\sup_{f \in \mathcal{F}} \sum_{j=0}^{n/m-1} \varepsilon_j \cdot \bar{y}_j \cdot f(\bar{x}_j)\right] \\ \leq& 2L \cdot \mathcal{R}_{n/m}(\mathcal{F}). \qquad (28) \end{aligned}$$

where (27) follows since, conditioned on $\mathcal{E}$, $(x'_{t,j}, y_{t,j'}), (x_{I_j}, \hat{y}_{I_j})$, $0 \leq j < n/m$ are all mutually independent (here we are using (26) as well as the definition of the labels $\hat{y}_i$). Furthermore, in (28) above, we are using the contraction inequality for Rademacher complexity.

By our choice of $m = 2 \log T \cdot \frac{1}{c_0 \sigma}$, we have that $\Pr(\mathcal{E}) \geq 1 - (n/m) \cdot (1 - c_0 \sigma)^m \geq 1 - (n/m) \cdot \exp(-c_0 \sigma m) \geq 1 - \frac{n}{mT^2} \geq 1 - \frac{n\sigma}{T^2}$. Then we see that

$$\mathbb{E}[y'_t \cdot f_{t+1}(x'_t) - y_t \cdot f_{t+1}(x_t)] \leq 2L \cdot \frac{m}{n} \cdot \mathcal{R}_{n/m}(\mathcal{F}) + 2\zeta + \frac{2Ln\sigma}{T^2}.$$

■

Lemma 39 *Fix $L \geq 1$, $\zeta \geq 0$. Consider random variables $x, x', x_1, \ldots, x_n \in \mathcal{X}$, $y, y' \in [-L, L]$, $y_1, \ldots, y_n \in \mathbb{R}$ drawn according to some distribution Q, and a constant $\gamma \geq L$. Suppose $h_1 \in \mathcal{F}$ is a function of $(x, y), (x', y'), (x_1, y_1), \ldots, (x_n, y_n)$ satisfying*

$$y \cdot h_1(x) + \gamma \cdot h_1(x') \cdot \operatorname{sign}(y') + \sum_{i=1}^n h_1(x_i) \cdot y_i \leq \min_{f \in \mathcal{F}} y \cdot f(x) + \gamma \cdot f(x') \cdot \operatorname{sign}(y') + \sum_{i=1}^n f(x_i) \cdot y_i + \zeta.$$

Suppose that the distribution of the $(n+2)$-tuples

$$((x, y), (x', y'), (x_1, y_1), \ldots, (x_n, y_n))$$

and

$$((x', y'), (x, y), (x_1, y_1), \ldots, (x_n, y_n))$$

are identical. Then $\mathbb{E}_Q[y \cdot h_1(x)] \geq \mathbb{E}_Q[y' \cdot h_1(x')] - 2\zeta$.

Proof Define h_2 as the function h_1 applied to the sequence $(x', y'), (x, y), (x_1, y_1), \ldots, (x_n, y_n)$, so that

$$y' \cdot h_2(x') + \gamma \cdot h_2(x) \cdot \operatorname{sign}(y) + \sum_{i=1}^n h_2(x_i) \cdot y_i \leq \min_{f \in \mathcal{F}} y' \cdot f(x') + \gamma \cdot f(x) \cdot \operatorname{sign}(y) + \sum_{i=1}^n f(x_i) \cdot y_i.$$

By definition of h_1, we have that

$$\begin{aligned} & y \cdot h_1(x) + \frac{\gamma}{|y'|} \cdot h_1(x') \cdot y' + \sum_{i=1}^n h_1(x_i) \cdot y_i \\ \leq & y \cdot h_2(x) + \frac{\gamma}{|y'|} \cdot h_2(x') \cdot y' + \sum_{i=1}^n h_2(x_i) \cdot y_i + \zeta. \end{aligned}$$

By definition of h_2, we have that

$$\begin{aligned} & y' \cdot h_2(x') + \frac{\gamma}{|y|} \cdot h_2(x) \cdot y + \sum_{i=1}^n h_2(x_i) \cdot y_i \\ \leq & y' \cdot h_1(x') + \frac{\gamma}{|y|} \cdot h_1(x) \cdot y + \sum_{i=1}^n h_1(x_i) \cdot y_i + \zeta. \end{aligned}$$

Adding the two previous displays and simplifying gives

$$\begin{aligned} & h_1(x) \cdot y \cdot (1 - \gamma/|y|) + h_1(x') \cdot y' \cdot (\gamma/|y'| - 1) \\ & \leq h_2(x) \cdot y \cdot (1 - \gamma/|y|) + h_2(x') \cdot y' \cdot (\gamma/|y'| - 1) + 2\zeta. \end{aligned}$$

Since (x, y) and (x', y') are exchangable, we have by definition of h_1, h_2 that for all constants $a, b \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{E}[h_2(x) \cdot y \cdot (1 - \gamma/|y|) + h_2(x') \cdot y' \cdot (\gamma/|y'| - 1) \mid |y| = a, |y'| = b] \\ & = \mathbb{E}[h_1(x') \cdot y' \cdot (1 - \gamma/|y'|) + h_1(x) \cdot y \cdot (\gamma/|y| - 1) \mid |y'| = a, |y| = b]. \end{aligned}$$

Combining the two above displays and rearranging gives

$$\begin{aligned}&\mathbb{E}[h_1(x)\cdot y\cdot(1-\gamma/a)+h_1(x')\cdot y'\cdot(\gamma/b-1)\mid |y|=a,|y'|=b]\\ \leq&\mathbb{E}[h_1(x')\cdot y'\cdot(1-\gamma/a)+h_1(x)\cdot y\cdot(\gamma/b-1)\mid |y|=b,|y'|=a]+2\zeta.\end{aligned}$$

Interchanging the roles of a, b, we get

$$\begin{aligned}&\mathbb{E}[h_1(x)\cdot y\cdot(1-\gamma/b)+h_1(x')\cdot y'\cdot(\gamma/a-1)\mid |y|=b,|y'|=a]\\ \leq&\mathbb{E}[h_1(x')\cdot y'\cdot(1-\gamma/b)+h_1(x)\cdot y\cdot(\gamma/a-1)\mid |y|=a,|y'|=b]+2\zeta.\end{aligned}$$

Exchangeability of (x,y) and (x',y') implies that $\Pr(|y|=a,|y'|=b)=\Pr(|y|=b,|y'|=a)$, and thus, by averaging the two above displays, we get

$$\begin{aligned}&\mathbb{E}[h_1(x)\cdot y\cdot(2-\gamma/a-\gamma/b)\mid \{|y|,|y'|\}=\{a,b\}]\\ &\leq\mathbb{E}[h_1(x')\cdot y'\cdot(2-\gamma/a-\gamma/b)\mid \{|y|,|y'|\}=\{a,b\}]+2\zeta.\end{aligned}$$

Using that $\gamma\geq L\geq\max\{a,b\}$ gives that

$$\mathbb{E}[h_1(x)\cdot y-h_1(x')\cdot y'\mid \{|y|,|y'|\}=\{a,b\}]\geq-2\zeta.$$

Taking expectation over $\{|y|,|y'|\}$ gives that $\mathbb{E}[h_1(x)\cdot y-h_1(x')\cdot y']\geq-2\zeta$, as desired. ■

E.3. Conclusion of Proof

We are now ready to start putting everything together. We first consider the case of binary labels.

Proposition 40 *Let $\mathcal{F}\subset\{\pm1\}^{\mathcal{X}}$ be a binary-valued function class and suppose that we are in the smoothed online learning setting, i.e., the conditional distribution of x_t given the history is σ-smooth with respect to some measure μ on $\mathcal{X}$. Let $\ell(\widehat{y},y)=-\widehat{y}y$ be indicator loss. For each $1\leq t\leq T$ and any n, define for any $f\in\mathcal{F}$,*

$$\hat{\omega}_{t,n}(f)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\gamma_i f(x_i)$$

where $x_i\sim\mu$ are independent ant $\gamma_i\sim N(0,1)$ are independent standard normal random variables. Let $f_t\in\mathcal{F}$ such that

$$L_{t-1}(f_t)+\eta\hat{\omega}_{t,n}(f_t)\leq\inf_{f\in\mathcal{F}}L_{t-1}(f)+\eta\hat{\omega}_{t,n}(f)+\zeta$$

Then, if $\mathsf{vc}(\mathcal{F})\leq d$, we have for $\eta=\sqrt{\frac{T\log(TL/\sigma)}{\sigma}}$ and $n=T/\sqrt{\sigma}$ that the regret satisfies:

$$\mathbb{E}\left[\mathrm{Reg}(f_t)\right]\lesssim\zeta T+\sqrt{\frac{Td\log(T/\sigma)}{\sigma}}$$

Proof By Hoeffding's inequality and (19) for some constant $C > 0$, as long as $n \geq C \log \frac{1}{\delta}$, an i.i.d. sample $x_1, \dots, x_n \sim \mu$ contains at least $n/2$ copies of x^* with probability $1 - \delta$, meaning that $\inf_{f \in \mathcal{F}} ||f||_{L^2(\hat{\mu}_n)} \geq \frac{1}{2}$ with probability $1 - \delta$; let this probability $1 - \delta$ event be denoted $\mathcal{E}_1$.

Further, for a sufficiently large constant $C > 0$, by Lemma 36, with probability $1 - \delta$ over the sample $x_1, \dots, x_n \sim \mu$, it holds that

$$\sup_{f, f' \in \mathcal{F}} \left| ||f - f'||_{L^2(\mu)}^2 - ||f - f'||_{L^2(\hat{\mu}_n)}^2 \right| \leq \Delta_n := C \cdot \left(\frac{1}{n} \mathcal{G}_n(\mathcal{F}) + \sqrt{\frac{\log(1/\delta)}{n}} \right). \tag{29}$$

Let this event (i.e., that (29) holds) be denoted $\mathcal{E}_2$.

Finally, it holds that with probability $1 - \delta$ over the sample $x_1, \dots, x_n \sim \mu$,

$$\left| \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_n(f) \right] - \frac{1}{\sqrt{n}} \mathcal{G}_n(\mathcal{F}) \right| \leq C \sqrt{\log\left(\frac{1}{\delta}\right)}. \tag{30}$$

Let this event (i.e., that (30) holds) be denoted $\mathcal{E}_3$.

The event $\mathcal{E} := \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3$ occurs with probability $1 - 3\delta$; taking $\delta = 1/T$, the contribution to expected regret on the complement of $\mathcal{E}$ is at most $6 = O(1)$. Thus, it suffices to bound regret in expectation conditioned on the event $\mathcal{E}$, which is what we proceed to do.

In particular, we use Lemma 32 to decompose the regret into three terms:

$$\mathbb{E}\left[\mathrm{Reg}_T(f_T) | \mathcal{E}\right] \leq \zeta T + 2\eta \left(\frac{1}{\sqrt{n}} \mathcal{G}_n(\mathcal{F}) + C\sqrt{\log T} \right) + T \max_{t \leq T} \mathbb{E}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t') \right]$$

$$+ T \max_{t \leq T} \mathbb{E}\left[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t) \right]$$

We can now apply Lemma 38 and Lemma 34 coupled with Lemma 36 to control Δ. In particular, we note that as we assume that $\mathsf{vc}(\mathcal{F}) \leq d$, we have

$$\mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f) \right] \lesssim \sqrt{d}, \qquad \frac{1}{\sqrt{n}} \mathcal{G}_n(\mathcal{F}) \lesssim \sqrt{d}$$

Then applying Lemmata 34 and 38 we may conclude that

$$\mathbb{E}\left[\sup_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f_t(x_t), y_t) - \ell(f(x_t), y_t) \right] \lesssim \zeta T + \eta\sqrt{d} + T \cdot \frac{(1+\zeta)^3 \log \eta}{\sigma \eta} \cdot \sqrt{d} + \frac{2T\Delta_n}{\sigma}$$

$$+ \frac{T \log T}{\sigma n} \mathcal{R}_{\frac{\sigma n}{\log T}}(\mathcal{F}) + \frac{n\sigma}{T} + \zeta T$$

Now, noting that for any $m \in \mathbb{N}$, the assumption that $\mathsf{vc}(\mathcal{F}) \leq d$ implies that

$$\frac{1}{m} \mathcal{R}_m(\mathcal{F}) \lesssim \sqrt{\frac{d}{m}}$$

we get

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim \zeta T + \eta\sqrt{d} + T \cdot \frac{(1+\zeta)^3 \log \eta}{\sigma \eta} \cdot \sqrt{d} + \frac{2T\Delta_n}{\sigma}$$

$$+ \sqrt{\frac{dT \log T}{\sigma n}} + \frac{n\sigma}{T} + \zeta T$$

We may choose $\eta = \sqrt{\frac{T \log(TL/\sigma)}{\sigma}}$ and $n = T/\sqrt{\sigma}$ to get a regret bound

$$\mathbb{E}\left[\mathrm{Reg}(f_t)\right] \lesssim \zeta T + \sqrt{\frac{Td \log(T/\sigma)}{\sigma}}$$

concluding the proof. ■

Note that Proposition 40 suffices to prove Theorem 9 in the case of binary values.

We now turn to the more challenging case of arbitrary labels. To understand the difficulty, note that Lemma 38 requires that the loss be linear. If we assume that the labels y_t are drawn in some smooth manner from a distribution $q_t(\cdot|x_t)$ so that the pair $(x_t, y_t) \sim \widetilde{p}_t$ with $\widetilde{p}_t$ being σ-smooth with respect to a distribution $\widetilde{\mu}$ on $\mathcal{X} \times [-1, 1]$, then we can reduce to the linear case by replacing $\mathcal{F} : \mathcal{X} \to [-1, 1]$ by $\ell \circ \mathcal{F} : \mathcal{X} \times [-1, 1] \to [-1, 1]$ with functions in $\ell \circ \mathcal{F}$ consisting of maps of the form $(x, y) \mapsto \ell(f(x), y)$ for any $f \in \mathcal{F}$. In the following result, we make use of this observation to bound the regret in the smoothed label setting, for arbitrary loss functions.

Proposition 41 *Let $\mathcal{F}$ be a function class mapping $\mathcal{X} \to [-1, 1]$ and suppose we are in the smoothed online learning setting with smooth labels, i.e., suppose that for all t, the adaptive adversary chooses a distribution $\widetilde{p}_t$ on $\mathcal{X} \times [-1, 1]$, σ smooth with respect to some distribution $\widetilde{\mu}$, and samples $(x_t, y_t) \sim \widetilde{p}_t$. Let $\ell : [-1, 1] \times [-1, 1] \to [-1, 1]$ be a loss function that is L-Lipschitz in the first argument. For each $1 \le t \le T$ and any n, let*

$$\hat{\omega}_{t,m}(f) = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \gamma_i f(x_i) \qquad\qquad \hat{\omega}'_{t,n}(f) = \sum_{j=1}^{n} \gamma'_j \ell(f(x'_j), y'_j)$$

where $x_i \sim \mu$ and $(x'_j, y'_j) \sim \widetilde{\mu}$ are independent and $\gamma_i, \gamma'_j \sim N(0, 1)$ are independent standard normal random variables. Let $f_t \in \mathcal{F}$ such that

$$L_{t-1}(f_t) + \eta \hat{\omega}_{t,m}(f_t) + \hat{\omega}_{t,n}(f) \le \inf_{f \in \mathcal{F}} L_{t-1}(f) + \eta \hat{\omega}_{t,n}(f) + \hat{\omega}_{t,n}(f) + \zeta$$

Then,

$$\begin{aligned}\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim & \left(\frac{L}{\sqrt{m}} \mathcal{G}_m(\mathcal{F}) + L\mathcal{G}_n(\mathcal{F}) + \sqrt{\log T} \right) \left(2\eta + T \frac{(1+\zeta)^3 \log \eta}{\sqrt{\eta\sigma}} \right) \\ & + \frac{L^2 T \log T}{\sigma n} \mathcal{R}_{\frac{\sigma n}{\log T}}(\mathcal{F}) + \frac{n\sigma}{T} + \zeta T \end{aligned} \tag{31}$$

In particular, if $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ for $p < 2$, we may choose $\eta = T^{2/3}\sigma^{-1/3}$ and $n = T/\sigma$ to get

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim T^{\frac{2}{3}} \sigma^{-\frac{1}{3}} \log\left(\frac{T}{\sigma}\right)$$

Proof The proof proceeds in a similar manner to that of Proposition 40 except we replace x by (x, y), $\mathcal{F}$ by $\ell \circ \mathcal{F}$ and ℓ by the identity. More formally, let $\ell \circ \mathcal{F} = \{(x, y) \mapsto \ell(f(x), y) | f \in \mathcal{F}\}$ and note that (29), (30), and $\inf_{f \in \mathcal{F}} ||\ell \circ f||_{L^2(\hat{\mu}_n)} \ge \frac{1}{2}$ all hold with probability at least $1 - 4\delta$, just as in the proof of the earlier proposition. Letting $\delta = T^{-2}$, let $\mathcal{E}$ denote the event that all of these hold, i.e., for all $1 \leq t \leq T$,

$$\sup_{f, f' \in \mathcal{F}} \left| \left|\left|\ell \circ f - \ell \circ f'\right|\right|_{L^2(\hat{\mu}_n)}^2 - \left|\left|\ell \circ f - \ell \circ f'\right|\right|_{L^2(\mu)}^2 \right| \leq \Delta_n := C \left(\frac{1}{n} \mathcal{G}_n(\ell \circ \mathcal{F}) + \sqrt{\frac{\log T}{n}} \right)$$

$$\sup_{f, f' \in \mathcal{F}} \left| \left|\left|f - f'\right|\right|_{L^2(\hat{\mu}_m)}^2 - \left|\left|f - f'\right|\right|_{L^2(\mu)}^2 \right| \leq \Delta_m$$

$$\left| \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,m}(f) \right] - \frac{1}{\sqrt{n}} \mathcal{G}_n(\mathcal{F}) \right| \leq C \sqrt{\log T}$$

$$\left| \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,n}(f) \right] - \frac{1}{\sqrt{n}} \mathcal{G}_n(\ell \circ \mathcal{F}) \right| \leq C \sqrt{\log T}$$

$$\inf_{f \in \mathcal{F}} \left|\left|\ell \circ f\right|\right|_{L^2(\hat{\mu}_n)} \geq \frac{1}{2}$$

The expected regret on the complement of $\mathcal{E}$ is at most $4T\delta \leq 4$ by boundedness of the loss. We may now apply Lemma 32 to get

$$\begin{aligned} \mathbb{E}\left[\mathrm{Reg}_T(f_t); \mathcal{E}\right] \lesssim \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,m}(f); \mathcal{E} \right] &+ \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}'_{t,n}(f); \mathcal{E} \right] \\ &+ T \max_{1 \leq t \leq T} \mathbb{E}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t'); \mathcal{E} \right] \\ &+ T \max_{1 \leq t \leq T} \mathbb{E}\left[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t); \mathcal{E} \right] \end{aligned}$$

The first two terms are bounded by the restriction to $\mathcal{E}$. To control the third term, we consider a coupling where $\hat{\omega}'_{t,n} = \hat{\omega}'_{t+1,n}$ and $\hat{\omega}_{t,m} = \hat{\omega}_{t+1,m}$ and note that by the independence of $\hat{\omega}'_{t,n}$ and $\hat{\omega}_{t,m}$, we may condition on the value of the former and let

$$\widetilde{L}_t(f) = L_t(f) + \hat{\omega}'_{t,n}(f)$$

We may then apply Lemma 35 to the resulting expression and get

$$T \max_{1 \leq t \leq T} \mathbb{E}\left[\ell(f_t(x_t'), y_t') - \ell(f_{t+1}(x_t'), y_t'); \mathcal{E} \right] \lesssim T \frac{(L + 2\zeta)^3 \log \eta}{\sqrt{\sigma \eta}} \mathbb{E}\left[\sup_{f \in \mathcal{F}} \hat{\omega}_{t,m}(f) \right] + 4LT \sqrt{\frac{\Delta_m}{\sigma}}$$

Note that this is further controlled using $\mathcal{E}$ to bound the expected supremum of $\hat{\omega}_{t,m}$.

To take care of the last term, we consider a coupling where $\hat{\omega}_{t,m} = \hat{\omega}_{t+1,m}$ but $\hat{\omega}'_{t,n}$ and $\hat{\omega}'_{t+1,n}$ are independent. We may now condition on $\hat{\omega}_{t,m}$ as in the previous paragraph and apply Lemma 38 to get

$$T \max_{1 \leq t \leq T} \mathbb{E}\left[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t); \mathcal{E} \right] \lesssim 4LT \frac{\log T}{\sigma n} \mathcal{R}_{\frac{\sigma n}{2 \log T}}(\ell \circ \mathcal{F}) + \zeta T + \frac{2Ln\sigma}{T}$$

To conclude, we apply contraction to note that for all $k \in \mathbb{N}$,

$$\mathcal{G}_m(\ell \circ \mathcal{F}) \leq L \mathcal{G}_m(\mathcal{F}) \qquad\qquad \mathcal{R}_m(\ell \circ \mathcal{F}) \leq L \mathcal{R}_m(\mathcal{F})$$

This proves the first statement.

To prove the second statement, note that if $\mathsf{vc}(\mathcal{F}, \delta) \ll \delta^{-2}$, then

$$\max\left(\mathcal{R}_k(\mathcal{F}), \mathcal{G}_k(\mathcal{F})\right) \lesssim \sqrt{k} \tag{32}$$

The result then follows by a direct computation. ■

While Proposition 41 attains no-regret, the assumption that the labels y_t are drawn in a smoothed manner is much stronger than desired. In order to mitigate this issue, we apply a discretization scheme.

Proposition 42 *Let $\mathcal{F}$ be a function class mapping $\mathcal{X} \to [-1,1]$ and suppose that we are in the smoothed online learning setting, where $x_t \sim p_t$ are drawn from a distribution that is σ-smooth with respect to μ. Suppose that $\ell : [-1,1] \times [-1,1] \to [-1,1]$ is a loss function that is L-Lipschitz in* both *arguments. Consider the following processes:*

$$\hat{\omega}_{t,m}(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{m} \gamma_i f(x_i) \qquad\qquad \hat{\omega}'_{t,n}(f) = \sum_{j=1}^{n} \gamma'_j \ell(f(x'_j), y'_j)$$

where $x_i, x'_j \sim \mu$, $\gamma_i \sim N(0,1)$ and y'_j are uniform on $[-1,1] \cap \varepsilon\mathbb{Z}$ for some fixed $\varepsilon > 0$. Suppose that f_t is chosen such that

$$L_{t-1}(f_t) + \eta\hat{\omega}_{t,m}(f_t) + \hat{\omega}_{t,n'}(f_t) \le \inf_{f\in\mathcal{F}} L_{t-1}(f) + \eta\hat{\omega}_{t,m}(f) + \hat{\omega}_{t,n'}(f) + \zeta$$

Then

$$\begin{aligned}\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim &\left(\frac{L}{\sqrt{m}}\mathcal{G}_m(\mathcal{F}) + L\mathcal{G}_n(\mathcal{F}) + \sqrt{\log T}\right)\left(2\eta + T\frac{(1+\zeta)^3 \log\eta}{\sqrt{\eta\sigma}}\right) \\ &+ \frac{L^2 T \log T}{\varepsilon\sigma n}\mathcal{R}_{\frac{\varepsilon\sigma n}{\log T}}(\mathcal{F}) + \frac{\varepsilon n \sigma}{T} + (\zeta + L\varepsilon)T\end{aligned}$$

Proof Let $S^\varepsilon = \varepsilon\mathbb{Z} \cap [-1,1]$ and, for any $y \in [-1,1]$, let y^ε be the projection of y into S^ε. By assumption, we have $|\ell(\cdot, y) - \ell(\cdot, y^\varepsilon)| \le L\varepsilon$. The key observation is that (x_t, y_t^ε) is $(\varepsilon\sigma/2)$-smooth with respect to $\mu \otimes \mathrm{Unif}(S^\varepsilon)$ by the fact that $|S^\varepsilon| \le 2/\varepsilon$. We may now apply Lemma 32 and note that the first term, the magnitude of the perturbation, is unchanged. For the second term, we note that

$$\mathbb{E}\left[\ell(f_t(x_t), y_t) - \ell(f_{t+1}(x'_t), y'_t)\right] \le \mathbb{E}\left[\ell(f_t(x_t), y_t^\varepsilon) - \ell(f_{t+1}(x'_t), (y'_t)^\varepsilon)\right] + 2L\varepsilon$$

which is in turn controlled by Lemma 35 by the same reasoning as the analogous statement in the proof of Proposition 41. To bound the generalization error, we note that, again by the Lipschitz assumption,

$$\mathbb{E}\left[\ell(f_{t+1}(x'_t), y'_t) - \ell(f_{t+1}(x_t), y_t)\right] \le \mathbb{E}\left[\ell(f_{t+1}(x'_t), (y'_t)^\varepsilon) - \ell(f_{t+1}(x_t), y_t^\varepsilon)\right] + 2L\varepsilon \tag{33}$$

Now, note that

$$
\begin{aligned}
L_{t-1}(f_{t+1}) + \ell(f_{t+1}(x_t), y_t^\varepsilon) + \eta\hat{\omega}_{t+1,n}(f_{t+1}) &\le L_{t-1}(f_{t+1}) + \ell(f_{t+1}(x_t), y_t) + \eta\hat{\omega}_{t+1,n}(f_{t+1}) + L\varepsilon \\
&\le \inf_{f \in \mathcal{F}} L_{t-1}(f) + \ell(f(x_t), y_t) + \eta\hat{\omega}_{t+1,n}(f) + \zeta + L\varepsilon \\
&\le \inf_{f \in \mathcal{F}} L_{t-1}(f) + \ell(f(x_t), y_t^\varepsilon) + \eta\hat{\omega}_{t+1,n}(f) + \zeta + 2L\varepsilon
\end{aligned}
$$

Noting again that (x_t, y_t^ε) is $(\varepsilon\sigma/2)$-smooth with respect to $\mu \otimes \mathrm{Unif}(S^\varepsilon)$, we apply Lemma 38, adjusting ζ to $\zeta + 2L\varepsilon$, to get

$$
\mathbb{E}\left[\ell(f_{t+1}(x_t'), (y_t')^\varepsilon) - \ell(f_{t+1}(x_t), y_t^\varepsilon)\right] \le 8\frac{\log T}{c_0\sigma\varepsilon n}\mathcal{R}_{c_0\sigma\varepsilon n/(4\log T)}(\mathcal{F}) + \frac{n\sigma\varepsilon}{T^2} + 2\zeta + 2L\varepsilon
$$

after noting that $|S^\varepsilon| \le \frac{2}{\varepsilon}$. Combining this with (33) gives

$$
\mathbb{E}\left[\ell(f_{t+1}(x_t'), y_t') - \ell(f_{t+1}(x_t), y_t)\right] \le 4\frac{\log T}{c_0\sigma\varepsilon n}\mathcal{R}_{c_0\sigma\varepsilon n/(2\log T)}(\mathcal{F}) + \frac{2n\sigma\varepsilon}{T^2} + 2\zeta + 4L\varepsilon
$$

Plugging back in to Lemma 32 concludes the proof. ■

As a corollary, we have the following bounds.

Corollary 43 *Suppose we are in the setting of Proposition 42 and, furthermore,* $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ *for some* $p < 2$*. Then if* $\eta = T^{2/3}\sigma^{-1/3}$*,* $n = \sqrt{T/\sigma}$*, and* $\varepsilon = T^{-1/3}$*, we have*

$$
\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim T^{\frac{2}{3}}\sigma^{-\frac{1}{3}} + \zeta T
$$

If $p \ge 2$*, we may choose* $n = T$*,* $\varepsilon = (\sigma T)^{-\frac{1}{p+1}}$*, and* $\eta = T^{\frac{2}{p}}$ *to yield* $\mathbb{E}\left[\mathrm{Reg}_T(f_T)\right] = o(T)$*.*

Proof The first statement follows immediately from Proposition 42 and (32).

The second statement holds by direct computation and the fact that for all k,

$$
\max\left(\mathcal{G}_k(\mathcal{F}), \mathcal{R}_k(\mathcal{F})\right) \lesssim k^{1-\frac{1}{p}}
$$

■

We see that Corollary 43 contains Theorem 10.

Finally, at the cost of a slightly worse regret bound, we may simplify the algorithm by considering a single perturbation. Note that in Corollary 43, we may tune η and n independently because we have two distinct perturbations. In the case where the perturbations are the same, Lemma 38 tells us that $\eta \ge \sqrt{n}$. We thus have the following regret bound for the simpler algorithm:

Corollary 44 *Suppose we are in the situation of Proposition 42 and* $\mathsf{vc}(\mathcal{F}, \delta) \lesssim \delta^{-p}$ *for some* $p < 2$*. Suppose that* f_t *is chosen such that*

$$L_{t-1}(f_t) + \frac{\eta}{\sqrt{n}}\hat{\omega}'_{t,n}(f_t) \leq \inf_{f \in \mathcal{F}} L_{t-1}(f) + \frac{\eta}{\sqrt{n}}\hat{\omega}_{t,n}(f) + \zeta$$

Then if we set $\eta = T^{5/12}\sigma^{-1/4}$*,* $n = \eta^2$*, and* $\varepsilon = T^{-3/4}\sigma^{-1/4}$*, we have*

$$\mathbb{E}\left[\mathrm{Reg}_T(f_t)\right] \lesssim T^{\frac{3}{4}}\sigma^{-\frac{1}{4}} \log\left(\frac{T}{\sigma}\right) + \zeta T$$

Proof Note that Proposition 41 may be proved with a single perturbation in much the same way, with the caveat that $\eta \geq \sqrt{n}$, and achieve the regret bound given in (31) with $m = n$. This proof may then be extended by discretization in the same way as Proposition 42, again with the caveat that $\eta \geq \sqrt{n}$. We may use (32) to control the size of the Gaussian and Rademacher complexities as before and then tune the parameters such that $\eta \geq \sqrt{n}$. Plugging in the assumed parameters yields the desired result. ■

Note that Corollary 44 suffices to prove the more general case of Theorem 9.

Appendix F. Proofs from Section 6

In this section we prove the lower bounds on oracle-efficient algorithms from Section 6. The proof structure closely follows that from Hazan and Koren (2016), but some additional work is required since our setup allows more powerful algorithms than Hazan and Koren (2016): in particular, the ERM oracle allows (possibly negative) real-valued weights to be attached to each pair (x_i, y_i). In Section F.1 we recall the definition of Aldous' problem and introduce a slight variant; a known oracle lower bound for Aldous' problem forms the basis for our hardness results. In Section F.2 we introduce an intermediate problem, namely that of approximating the Nash value in a two-player zero-sum game given a value oracle and best response oracles; we then show an oracle lower bound for this problem by reducing from Aldous' problem. Using this result, in Section F.3, we prove Theorem 11 and Corollary 12 using a standard reduction from finding Nash equilibria to no-regret learning (Freund and Schapire, 1999).

F.1. Modified Aldous' Problem

We begin by recalling the definition of Aldous' problem and a slight variation we will use. Consider a function $\phi : \{0,1\}^d \to \mathbb{Z}$; for all such functions in this section, we assume that $|\phi(x)| \leq 2^{O(d)}$ for all $x \in \{0,1\}^d$. A point $x \in \{0,1\}^d$ is a *local maximum* if $\phi(x) \geq \phi(x')$ for all x' of Hamming distance at most 1 from x. The function ϕ is *globally consistent* if it has a single local maximum (i.e., the only local maximum is also a global maximum). *Aldous' problem* is the following problem: suppose we are given a globally consistent function $\phi : \{0,1\}^d \to \mathbb{N}$ with black-box oracle access in the sense that we can query a value $x \in \{0,1\}^d$ and the oracle will respond with the value $\phi(x)$. The objective is to determine whether the maximum value of ϕ is even or odd (with a minimum number of oracle queries). The following lower bound on the number of oracle calls needed to solve Aldous' problem is known:

Theorem 45 (Aaronson (2006); Aldous (1983); Hazan and Koren (2016)) *There is a constant $c > 0$ so that the following holds. Fix any $d \in \mathbb{N}$, and consider any randomized algorithm for Aldous' problem that makes at most $c \cdot 2^{d/2}/d^2$ oracle queries in the worst case. Then there is a globally consistent function $\phi : \{0,1\}^d \to \mathbb{N}$ so that the algorithm cannot determine with probability higher than $2/3$ whether the maximum value of ϕ over $\{0,1\}^d$ is even or odd.*

For our purposes we require a lower bound applying to a slightly more restricted class of functions than Theorem 45, specified in Definintion 46 below.

Definition 46 *We say that a function $\phi : \{0,1\}^d \to \mathbb{Z}$ is* min-max consistent *if it has both a single local maximum and a single local minimum.*

As an immediate corollary of Theorem 45 we get an exponential lower bound for local search with min-max consistent functions:

Corollary 47 *There is a constant $c' > 0$ so that the following holds. Consider any randomized algorithm for Aldous' problem that makes at most $c' \cdot 2^{d/2}/d^2$ oracle queries in the worst case. Then there is a min-max consistent function $\phi : \{0,1\}^d \to \mathbb{Z}$ so that the algorithm cannot determine with probability higher than $2/3$ whether the maximum value of ϕ over $\{0,1\}^d$ is even or odd.*

Proof Suppose to the contrary that $\mathcal{A}$ is a (randomized) algorithm that makes at most $c' \cdot 2^{d/2}/d^2$ oracle queries in the worst case and determines, for any min-max consistent function $\phi : \{0,1\}^d \to \mathbb{N}$, the parity of its maximum value with probability at least 2/3.

Consider a globally consistent function $\phi : \{0,1\}^d \to \mathbb{N}$. We define a min-max consistent function $\phi' : \{0,1\}^{d+1} \to \mathbb{Z}$ as follows: for $x \in \{0,1\}^{d+1}$,

$$\phi'(x) = \begin{cases} \phi(x_1, \ldots, x_d) & x_{d+1} = 0 \\ -\phi(x_1, \ldots, x_d) & x_{d+1} = 1. \end{cases} \tag{34}$$

To see that ϕ' is min-max consistent, note that any local maximum $x^\star = (x_1^\star, \ldots, x_{d+1}^\star)$ of ϕ' must satisfy $x_{d+1}^\star = 0$, and furthermore, the point $(x_1^\star, \ldots, x_d^\star) \in \{0,1\}^d$ must be a local maximum of ϕ. Thus ϕ' has a single local maximum. Similarly, for any local minimum $x_\star$ of ϕ', we must have $x_{\star,d+1} = 1$ and $(x_{\star,1}, \ldots, x_{\star,d})$ is a local maximum of ϕ; clearly there is a unique such point $x_\star \in \{0,1\}^{d+1}$.

We use $\mathcal{A}$ to determine the parity of the maximum value of ϕ using in the worst case no more than $c' 2^{(d+1)/2}/(d+1)^2$ oracle queries (to ϕ): we run the algorithm $\mathcal{A}$ with the function ϕ', and for each oracle query $x \in \{0,1\}^{d+1}$, we can return the value of $\phi'(x)$ per (34) using a single oracle query to ϕ. By assumption $\mathcal{A}$ determines the parity of the maximum value of ϕ', which is the same as the parity of the maximum value of ϕ, with probability at least 2/3.

Letting c be the constant of Theorem 45, as long as c' is chosen so that $c' \cdot 2^{(d+1)/2}/(d+1)^2 < c \cdot 2^{d/2}/d^2$ for all d, we get a contradiction to Theorem 45, as desired. ■

F.2. Hardness of Computing Nash Equilibria with Best-Response Oracles

Fix $N \in \mathbb{N}$ which is a power of 2, and define $d = \log_2 N$. Throughout this section, we identify each vertex v of the d-dimensional hypercube $\{0,1\}^d$ with the integer in $[N]$ whose binary representation corresponds to v. Let $\phi : [N] \to \mathbb{Z}$ be a min-max consistent input (Definition 46) to Aldous' problem, with maximum value $\phi^\star = \max_{i\in[N]}\{\phi(i)\}$. We construct a 0-sum game with value $\lambda = \lambda(\phi^\star)$, with

$$\lambda(k) = \begin{cases} -1 & \text{if } k \text{ is even} \\ 1 & \text{if } k \text{ is odd.} \end{cases}$$

Further, for a subset $V \subset [N]$ (identified with the corresponding subset of the hypercube), let $\Gamma(V) \subset [N]$ denote the set of neighbors of V in the hypercube (including the elements of V).

Given the function ϕ, we construct the following game matrix $G^\phi \in \{-1, 1\}^{N\times N}$:

$$\forall i, j \in [N], \qquad G^\phi_{ij} = \begin{cases} \lambda(\phi(i)) & \text{if } i, j \text{ are local maxima of } \phi \\ -1 & \text{if } \phi(i) \geq \phi(j) \text{ (and the first case does not apply)} \\ 1 & \text{otherwise.} \end{cases} \tag{35}$$

We let $k^\star := \max_{i\in[N]}\{\phi(i)\}$ denote the (unique) global maximum of ϕ. As an intermediate problem between Aldous' problem and the problem of oracle-efficient online (smoothed) learning, we consider the problem of approximating the Nash equilibrium value in the two-player zero-sum game induced by the matrix G, given access to the following 3 oracles:

- The value oracle $\mathsf{Val}(i, j)$ returns G^ϕ_{ij} for $i, j \in [N]$.
- The best response oracle $\mathsf{BR}^1(q)$, for $q \in \mathbb{R}^N$, returns

$$\mathsf{BR}^1(q) = \begin{cases} \operatorname{argmin}_{i\in\Gamma(\operatorname{supp}(q))}\{e_i^\top G^\phi q\} & k^\star \notin \operatorname{supp}(q) \\ \operatorname{argmin}_{i\in[N]}\{e_i^\top G^\phi q\} & \text{otherwise.} \end{cases} \tag{36}$$

- The best response oracle $\mathsf{BR}^2(p)$, for $p \in \mathbb{R}^n$, returns

$$\mathsf{BR}^2(p) = \begin{cases} \arg\max_{j\in\Gamma(\operatorname{supp}(p))}\{p^\top G^\phi e_j\} & k^\star \notin \operatorname{supp}(p) \\ \arg\max_{j\in[N]}\{p^\top G^\phi e_j\} & \text{otherwise.} \end{cases}$$

We define computation given access to the above oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ in an analogous way as to how computation was defined with respect to the ERM oracle in Section 2.3: p is represented as a list of atoms $\{(i, p_i) : p_i > 0\}$ and q is represented as a list of atoms $\{(j, q_j) : q_j > 0\}$, and changing a single atom on either list takes unit time. Further, calling any of the oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ takes unit time. Next we show that given the oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$, computing the approximate Nash equilibrium value of an $N \times N$ game G cannot be done in $o(\sqrt{N})$ time (up to logarithmic factors).

To begin, we establish some basic properties of the game G^ϕ constructed in (35).

Lemma 48 *For any globally consistent function ϕ, the minimax value of G^ϕ is λ.*

Proof Let $k^\star = \arg\max_{i\in[N]}\{\phi(i)\}$ denote the global maximum of ϕ. We show that the pure strategy profile $(k^\star, k^\star)$ is a Nash equilibrium of the game G^ϕ. The payoff with this profile is $\lambda(\phi(k^\star)) = \lambda$. For any $i \in [N]$, the strategy profile $(i, k^\star)$ generates a payoff of either λ (in the case $i = k^\star$) or $1 \geq \lambda$ since for all $i \neq k^\star$, $\phi(i) < \phi(k^\star)$. Thus there is no useful deviation for player 1. Similar, for any $j \in [N]$, the strategy profile $(k^\star, j)$ generates a payoff of either λ (in the case that $j = k^\star$) or of $-1 \leq \lambda$ since for all $j \neq k^\star$, $\phi(j) < \phi(k^\star)$. Thus $(k^\star, k^\star)$ is a Nash equilibrium, meaning that its value is the value of the game. ■

Lemma 49 *Fix any min-max consistent function $\phi : [N] \to \mathbb{Z}$. The oracles Val and $\mathsf{BR}^1, \mathsf{BR}^2$ are correct value and best-response oracles for the game G^ϕ, in that:*

$$\mathsf{Val}(i, j) = G^\phi_{ij}, \quad \mathsf{BR}^1(q) = \operatorname*{argmin}_{i \in [N]} \{e_i^\top G^\phi q\}, \quad \mathsf{BR}^2(p) = \operatorname*{arg\,max}_{j \in [N]} \{p^\top G^\phi e_j\}.$$

Proof The oracle Val is clearly valid as a value oracle for the game G^ϕ since $\mathsf{Val}(i, j) = G^\phi_{ij}$ for all $i, j \in [N]$ by definition. Furthermore, the best response oracles $\mathsf{BR}^1(q)$, $\mathsf{BR}^2(p)$ are clearly valid for G^ϕ in the case that $k^\star \in \mathrm{supp}(q)$ or $k^\star \in \mathrm{supp}(p)$, respectively. We next verify that they are valid in the remaining case.

We begin by considering the best response oracle BR^2: fix some input $p \in \mathbb{R}^N$ with $k^\star \not\in \mathrm{supp}(p)$, let $j = \mathsf{BR}^2(p)$ denote the output of the oracle defined above, and set $v = \max_{j \in [N]} p^\top G^\phi e_j$ to be the value of player 2's best response to p. Choose some $j^\star \in [N]$ so that $p^\top G^\phi e_{j^\star} = v$. Since $k^\star \not\in \mathrm{supp}(p)$, we have, for all $j \in [N]$,

$$p^\top G^\phi e_j = \sum_{i \in \mathrm{supp}(p) : \phi(i) < \phi(j)} p_i - \sum_{i \in \mathrm{supp}(p) : \phi(i) \geq \phi(j)} p_i. \tag{37}$$

We consider the following cases regarding the value of $\phi(j^\star)$:

1. $\phi(j^\star) = \phi(i)$ for some $i \in \mathrm{supp}(p)$. Then since $p^\top G^\phi e_j$ only depends on j through $\phi(j)$ (as is evident from (37), it follows that $v = p^\top G^\phi e_{j^\star} = p^\top G^\phi e_i \leq p^\top G^\phi e_j$, as desired.

2. $\phi(j^\star) > \max_{i \in \mathrm{supp}(p)} \{\phi(i)\}$. Since $k^\star \not\in \mathrm{supp}(p)$, and ϕ is min-max consistent, there is some $j' \in \Gamma(\mathrm{supp}(p))$ so that $\phi(j') > \max_{i \in \mathrm{supp}(p)}\{\phi(i)\}$. It is evident from (37) that $p^\top G^\phi e_{j'} = p^\top G^\phi e_{j^\star} = v$, which implies, by definition of j and since $j' \in \Gamma(\mathrm{supp}(p))$, that $p^\top G^\phi e_j \geq p^\top G^\phi e_{j^\star}$, as desired.

3. Suppose the previous two cases do not hold. Choose $j' \in \mathrm{supp}(p)$ with $\phi(j')$ as small as possible so that $\phi(j') \geq \phi(j^\star)$. It is again evident from (37) that $p^\top G^\phi e_{j^\star} = p^\top G^\phi e_{j'} \leq p^\top G^\phi e_j$, as desired.

We next consider the best response oracle BR^1: fix some input $q \in \mathbb{R}^N$ with $k^\star \not\in \mathrm{supp}(q)$, let $i = \mathsf{BR}^2(q)$ denote the output of the oracle defined above, and set $v = \min_{i \in [N]} e_i^\top G^\phi q$ to be the value of player 1's best response to q. Choose some $i^\star \in [N]$ so that $e_{i^\star}^\top G^\phi q = v$. Since $k^\star \not\in \mathrm{supp}(q)$, we have, for all $i \in [N]$,

$$e_i^\top G^\phi q = \sum_{j \in \mathrm{supp}(q) : \phi(j) > \phi(i)} q_j - \sum_{j \in \mathrm{supp}(q) : \phi(j) \leq \phi(i)} q_j. \tag{38}$$

We consider the following cases regarding the value of $\phi(i^\star)$:

1. $\phi(i^\star) = \phi(j)$ for some $j \in \mathrm{supp}(q)$. Then since $e_i^\top G^\phi q$ only depends on i through $\phi(i)$ (as is evident from (38)), it follows that $v = e_{i^\star}^\top G^\phi q = e_j^\top G^\phi q \geq e_i^\top G^\phi q$, as desired.

2. $\phi(i^\star) < \min_{j \in \mathrm{supp}(q)}\{\phi(j)\}$. It cannot be the case that $k_\star \in \mathrm{supp}(q)$ since then we would have $\phi(i^\star) < \phi(k_\star)$. Therefore, since ϕ is min-max consistent, there is some $i' \in \Gamma(\mathrm{supp}(q))$ so that $\phi(i') < \min_{j \in \mathrm{supp}(q)}\{\phi(j)\}$. It follows that $e_{i^\star}^\top G^\phi q \leq e_{i'}^\top G^\phi q = e_{i^\star}^\top G^\phi q = v$, as desired.

3. Suppose the previous two cases do not hold. Choose $i' \in \operatorname{supp}(q)$ with $\phi(i')$ as large as possible so that $\phi(i') \leq \phi(i^\star)$. It is again evident from (38) that $e_{i^\star}^\top G^\phi q = e_{i'}^\top G^\phi q \geq e_i^\top G^\phi q$, as desired.

■

Lemma 50 *There is a constant $c_0 > 0$ so that the following holds. Fix any $N \in \mathbb{N}$. Any randomized algorithm $\mathcal{A}$ for approximating the equilibrium of $N \times N$ $\{-1, 1\}$-valued zero-sum games with the oracles BR^1, BR^2, Val cannot guarantee with probability greater than $2/3$ that the algorithm $\mathcal{A}$'s output value is at most $1/4$ from the game's true value in time $c_0 \cdot \sqrt{N}/\log^3 N$.*

Proof Fix any $N \in \mathbb{N}$. At the cost of a constant factor (and by a standard padding argument) we may assume that N is a power of 2. Let $\mathcal{A}$ be an algorithm as in the theorem statement, and suppose for the purpose of contradiction that with probability greater than $2/3$, for any $N \times N$, $\{-1, 1\}$-valued zero-sum game, $\mathcal{A}$'s output value is at most $1/4$ away from the game's value and $\mathcal{A}$ runs in time $c_0 \cdot \sqrt{N}/\log^3 N$ for some constant $c_0 > 0$.

We use the algorithm $\mathcal{A}$ to derive a contradiction to Corollary 47. Accordingly, let $\phi : [N] \to \mathbb{Z}$ be a min-max consistent function to which we can make black-box value queries. Consider the $N \times N$ $\{-1, 1\}$-valued game G^ϕ defined in (35). We run algorithm $\mathcal{A}$ on the game G^ϕ, simulating the oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ as follows:

- The value oracle $\mathsf{Val}(i, j)$ can be simulated using at most $\log(N) + 1$ queries to ϕ (namely, to $\phi(i)$ and $\phi(j)$, as well as, in the case that $i = j$, to all neighbors of i to check whether it is a local maximum).
- Fix some $q \in \mathbb{R}^N$, and set $m_q := |\operatorname{supp}(q)|$; the best response oracle $\mathsf{BR}^1(q)$ may be simulated as follows:
  1. Query the value of $\phi(j)$ for all $j \in \Gamma(\operatorname{supp}(q))$; this requires $m_q \cdot (\log(N) + 1)$ oracle queries to ϕ.
  2. By comparing, for each $j \in \operatorname{supp}(q)$, the value of $\phi(j)$ to the value of $\phi(j')$ for each neighbor j' of j (all of which were queried in the previous step), we may check if $k^\star \in \operatorname{supp}(q)$.
  3. If $k^\star \in \operatorname{supp}(q)$, then output the parity of $\phi(k^\star)$ and terminate the algorithm early.
  4. Otherwise, if $k^\star \notin \operatorname{supp}(q)$, then using the queried values of $\phi(j)$, $j \in \Gamma(\operatorname{supp}(q))$, we may compute $\mathsf{BR}^1(q)$ per (36) – here we use that $\operatorname{argmin}_{i \in \Gamma(\operatorname{supp}(q))}\{e_i^\top G^\phi q\}$ may be computed entirely from the values of $\phi(j)$ for $j \in \Gamma(\operatorname{supp}(q))$.
- For $p \in \mathbb{R}^N$ and $m_p := |\operatorname{supp}(p)|$, the best response oracle $\mathsf{BR}^2(p)$ may be simualted analogously to above, using at most $m_p \cdot (\log(N) + 1)$ oracle queries to ϕ.

If none of the calls to $\mathsf{BR}^1, \mathsf{BR}^2$ terminates early, then given the output $\hat{v} \in \mathbb{R}$ of the algorithm $\mathcal{A}$, we simply output the sign of $\hat{v}$.

Write $\phi^\star = \max_{j \in [N]}\{\phi(j)\}$. We claim that the resulting algorithm described above outputs with probability at least 2/3, -1 if $\phi^\star$ is even and 1 if $\phi^\star$ is odd. To see this, we argue as follows: with probability 2/3 over the randomness of the algorithm $\mathcal{A}$, one of the following must occur:

- Some call to either $\mathsf{BR}^1, \mathsf{BR}^2$ causes the algorithm to terminate early, in which case it is clear that the algorithm correctly outputs the parity of the maximum value of ϕ.
- The output of the algorithm $\mathcal{A}$ is within $1/4$ of the value of the game G^ϕ, which we denote by $\lambda \in \{-1, 1\}$. By Lemma 48, λ is equal to -1 if $\phi^\star$ is even and 1 if $\phi^\star$ is odd. Thus, the output of the algorithm $\mathcal{A}$ is -1 if $\phi^\star$ is even and 1 if $\phi^\star$ is odd.

Having verified correctness (with probability at least 2/3) of the algorithm above to find the parity of $\phi^\star$, we proceed to analyze its oracle cost. The algorithm $\mathcal{A}$ is assumed to take time $c_0 \cdot \sqrt{N}/\log^3(N)$, for some sufficiently small constant c_0. Let us denote the number of oracle calls $\mathcal{A}$ makes to Val by ω_{Val}; further, denote the total time consumed by all oracle calls $\mathcal{A}$ makes to $\mathsf{BR}^2(p)$ (including the oracle calls themselves and the time spent writing the input atoms (i, p_i)) by ω_{BR^2}; define ω_{BR^1} similarly for the oracle BR^1. By the definition of our oracle model above, it holds that $\omega_{\mathsf{Val}} + \omega_{\mathsf{BR}^1} + \omega_{\mathsf{BR}^2} \leq c_0 \cdot \sqrt{N}/\log^3(N)$.

Since each call by $\mathcal{A}$ to Val makes at most $\log(N) + 1$ oracle queries to ϕ, the total number of oracle calls to ϕ as a result of calls to the Val oracle by $\mathcal{A}$ is bounded above by $(\log(N) + 1) \cdot \omega_{\mathsf{Val}}$. Similarly, since we can store the result of oracle calls to ϕ for previously used atoms (i, p_i) or (j, q_j) (in step 1 above), the total number of oracle calls to ϕ as a result of calls to the BR^2 oracle by $\mathcal{A}$ is bounded above by $(\log(N) + 1) \cdot \omega_{\mathsf{BR}^2}$. Using similar reasoning for calls to BR^1, we get that the total number of oracle calls to ϕ in our algorithm above is at most

$$(\log(N) + 1) \cdot \left(\omega_{\mathsf{Val}} + \omega_{\mathsf{BR}^1} + \omega_{\mathsf{BR}^2}\right) \leq (\log(N) + 1) \cdot c_0\sqrt{N}/\log^3(N) < c' \cdot \sqrt{N}/\log^2(N),$$

where c' is the constant of Corollary 47 (as long as the constant c_0 is chosen sufficiently small). This is a contradiction to the conclusion of Corollary 47, thus completing the proof of Lemma 50. ■

F.3. Hardness of oracle-efficient proper no-regret learning

In this section we use the oracle lower bounds for finding Nash equilibria in two-player zero-sum games to derive oracle lower bounds for no-regret online learning against a σ-smooth adversary.

In particular, we first prove Theorem 11, stated below with precise logarithmic factors. The proof is a standard reduction from finding Nash equilibria in two-player zero-sum games to no-regret learning (Freund and Schapire, 1999), but we provide the details for completeness:

Theorem 11 (Restated, precise) *For some constant $c > 0$, we have the following: fix any $T \in \mathbb{N}$ and $\sigma \in (0, 1]$. In the ERM oracle model, any randomized algorithm cannot guarantee expected regret smaller than $\frac{T}{200}$ against a σ-smooth online adversary and any $\mathcal{F}$ with $|\mathcal{F}| \leq 1/\sigma$ over T rounds in total time smaller than $c \cdot \frac{1/\sqrt{\sigma}}{\log^3 1/\sigma}$; further, this result holds even for binary-valued classes.*

Proof [Theorem 11] Fix T, σ as in the theorem statement; at the cost of a constant factor we may assume that $1/\sigma$ is an integer. Suppose $\mathcal{A}$ is an algorithm which guarantees expected regret smaller than $\frac{T}{200}$ against all σ-smooth adversaries in time $\leq c \cdot \frac{1/\sqrt{\sigma}}{\log^3 1/\sigma}$. By Markov's inequality, for any σ-smooth adversary, the regret of $\mathcal{A}$ is bounded above by $\frac{T}{20}$ with probability at least $9/10$.

Set $N := 1/\sigma$, and consider any $N \times N$ $\{-1, 1\}$-valued zero-sum game, represented by a game matrix $G \in \{-1, 1\}^{N \times N}$, with entries G_{f_1, f_2}, $f_1, f_2 \in [N]$; as a matter of convention we suppose that the min-player chooses the first coordinate f_1 and the max-player chooses the second coordinate f_2. Now consider the following procedure for approximating the Nash equilibrium value of G (we will show below how to implement the below using the oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ introduced in the previous section):

1. Initialize instances $\mathcal{A}_1, \mathcal{A}_2$ of the algorithm $\mathcal{A}$ given the time horizon T; for $\mathcal{A}_1$ the function class is $\{f_2 \mapsto G_{f_1,f_2} : f_1 \in [N]\}$, and for $\mathcal{A}_2$ the function class is $\{f_1 \mapsto G_{f_1,f_2} : f_2 \in [N]\}$. The loss functions of the algorithms are given as follows:
   - The loss function of $\mathcal{A}_1$ is $\ell(\hat{y}, y) = \hat{y}$; thus $\mathcal{A}_1$ incurs loss of G_{f_1,f_2} for predicting f_1 when it observes f_2.
   - The loss function of $\mathcal{A}_2$ is $\ell(\hat{y}, y) = -\hat{y}$; thus $\mathcal{A}_2$ incurs loss of $-G_{f_1,f_2}$ for predicting f_2 when it observes f_1.
2. For $t = 1, 2, \ldots, T$:
   (a) Let the algorithms $\mathcal{A}_1, \mathcal{A}_2$ produce (random) decisions $f_{1,t}, f_{2,t} \in [N]$, respectively.
   (b) Update $\mathcal{A}_1$ with the context $f_{2,t}$.
   (c) Update $\mathcal{A}_2$ with the context $f_{1,t}$.
3. Define mixed strategies $\bar{f}_{i,T} := \frac{1}{T}\sum_{t=1}^{T} f_{i,t}$ for $i = 1, 2$.
4. Output the value $\hat{v} := \bar{f}_{1,T}^{\top} G \bar{f}_{2,T} = \frac{1}{T^2} \cdot \sum_{t,s=1}^{T} G_{f_{1,t},f_{2,s}}$.

Note that we do not need to specify the labels y_t for either algorithm $\mathcal{A}_1, \mathcal{A}_2$ above, since their loss functions do not depend on the true labels y_t. By the union bound, with probability at least $4/5$, the regret of both $\mathcal{A}_1, \mathcal{A}_2$ is bounded above by $T/20$; in particular, with probability at least $4/5$ we have:

$$\sum_{t=1}^{T} G_{f_{1,t},f_{2,t}} - \min_{f_1 \in [N]} \sum_{t=1}^{T} G_{f_1,f_{2,t}} \le \frac{T}{20}, \qquad \max_{f_2 \in [N]} \sum_{t=1}^{T} G_{f_{1,t},f_2} - \sum_{t=1}^{T} G_{f_{1,t},f_{2,t}} \le \frac{T}{20}.$$

Adding the two preceding equations, we obtain

$$\max_{f_2 \in [N]} \bar{f}_{1,T}^{\top} G e_{f_2} - \min_{f_1 \in [N]} e_{f_1}^{\top} G \bar{f}_{2,T} \le \frac{1}{10}, \tag{39}$$

where $e_f,\ f \in [N]$ denotes the unit vector corresponding to f. Set $\varepsilon := 1/10$. Letting $(f_1^\star, f_2^\star)$ denote a Nash equilibrium of G and $v^\star := (f_1^\star)^\top G f_2^\star$ denotes the value of G, we have

$$v^\star - \varepsilon \le \bar{f}_T^{\top} G f_2^\star - \varepsilon \le \max_{f_2 \in [N]} \bar{f}_{1,T}^{\top} G e_{f_2} - \varepsilon \overset{(39)}{\le} \min_{f_1 \in [N]} e_{f_1}^{\top} G \bar{f}_{2,T} \le \bar{f}_{1,T}^{\top} G \bar{f}_{2,T}$$
$$\le \max_{f_2 \in [N]} \bar{f}_{1,T}^{\top} G e_{f_2} \overset{(39)}{\le} \min_{f_1 \in [N]} e_{f_1}^{\top} G \bar{f}_{2,T} + \varepsilon \le (f_1^\star)^\top G \bar{f}_{2,T} + \varepsilon \le v^\star + \varepsilon.$$

Thus we have $|\hat{v} - v^\star| \le \varepsilon = 1/10$, meaning that the above procedure determines the game G's value up to error $1/10$.

We next analyze the time complexity of the above procedure, which involves showing how to implement it efficiently using the oracles $\mathsf{BR}^1, \mathsf{BR}^2, \mathsf{Val}$:

- Each time $\mathcal{A}_1$ makes an ERM oracle call of the form $\operatorname{argmin}_{f_1 \in [N]} \sum_{i=1}^m w_i \cdot \ell_i(G_{f_1, f_{2,i}}, y_i)$, we do the following: we may assume without loss of generality that all $f_{2,i}$ are distinct. Now write $\ell_{i,1} := \ell_i(1, y_i)$, $\ell_{i,-1} := \ell_i(-1, y_i)$. This ERM call may be simulated by the oracle call $\mathsf{BR}^1(q)$, where $q_{f_{2,i}} = w_i \cdot \frac{\ell_{i,1} - \ell_{i,-1}}{2}$ for all $i \in [m]$, and $q_{f_2} = 0$ for all other f_2.
- Each time $\mathcal{A}_2$ makes an ERM oracle call of the form $\operatorname{argmin}_{f_2 \in [N]} \sum_{i=1}^m w_i \cdot \ell_i(G_{f_{1,i}, f_2})$, we define $\ell_{i,1}, \ell_{i,-1}$ as above and simulate it using the oracle call $\mathsf{BR}^2(p)$ where $p_{f_{1,i}} = w_i \cdot \frac{\ell_{i,-1} - \ell_{i,1}}{2}$ for all $i \in [m]$ and $p_{f_1} = 0$ for all other f_2.
- It only remains to show how the estimation of $\hat{v}$ in step 4 can be implemented efficiently: for any $f_1, f_2 \in [N]$, the value G_{f_1, f_2} can be queried with a single oracle call as $\mathsf{Val}(f_1, f_2)$, so $\hat{v}$ may trivially be computed in time $O(T^2)$. We may in fact obtain a stronger bound as follows: fix $\delta > 0$ and a sufficiently large constant $C > 0$, and for $1 \leq j \leq C \log(1/\delta)$ sample i.i.d. pairs (i_j^1, i_j^2) uniformly from $[T] \times [T]$, and output $\hat{v}' := \frac{1}{C \log 1/\delta} \sum_{j=1}^{C \log 1/\delta} G_{f_{1, i_j^1}, f_{2, i_j^2}}$ tuples, for a total of $O(C \log 1/\delta)$ time (including the oracle calls to Val). By the Chernoff bound, we have that $|\hat{v}' - \hat{v}| \leq 1/100$ with probability $1 - \delta$, as long as C is sufficiently large.

As long as δ in the third bullet above satisfies $\delta \leq 4/5 - 2/3$, we have established that there is an algorithm that with probability $2/3$ estimates the value $v^\star$ of G up to accuracy of $1/9$. Furthermore, it is straightforward to see that implementing the oracle calls to $\mathsf{BR}^1, \mathsf{BR}^2, \mathsf{Val}$ as described above only lead to a constant factor blowup in the total time. It is also evident that since the space of contexts for both $\mathcal{A}_1, \mathcal{A}_2$ is $[N]$, arbitrary adaptive adversaries (in particular, the adversaries faced by $\mathcal{A}_1, \mathcal{A}_2$ above) are $1/N$-smooth with respect to the uniform distribution on $[N]$. Thus, by the assumed time complexity upper bound of $\mathcal{A}$, we have that the algorithm to estimate $v^\star$ runs in time $c' \cdot \frac{\sqrt{N}}{\log^3(N)}$ for some constant c', which can be made arbitrarily small by choosing c to be arbitrarily small. This contradicts Lemma 50.

■

Now we prove Corollary 12 (restated below with precise logarithmic factors), which is a straightforward consequence of Theorem 11:

Corollary 12 (Restated, precise) *Fix any $\alpha \geq 1$, $\varepsilon < 1/200, \sigma \in (0, 1]$, and $d \geq \log 1/\sigma$. Any algorithm whose total time in the ERM oracle model over T rounds is bounded as T^α requires that $T \geq \Omega\left(\max\left\{\frac{d}{\varepsilon^2}, \frac{\sigma^{-1/(2\alpha)}}{\log^3 1/\sigma}\right\}\right)$ to achieve regret εT for classes $\mathcal{F}$ of VC dimension at most d against a σ-smooth adversary.*

Furthermore, any algorithm which achieves regret εT for classes of VC dimension at most d against a σ-smooth adversary must have computation time at least $\Omega\left(\max\left\{\frac{d}{\varepsilon^2}, \frac{\sigma^{-1/2}}{\log^3 1/\sigma}\right\}\right)$.

Proof [Corollary 12] Fix any $\varepsilon < 1/200, \sigma \in (0, 1]$, and $d \geq \log 1/\sigma$, as in the statement of the corollary. We begin by proving the first statement of the lemma. We consider the following cases:

Case 1. $d/\varepsilon^2 > \sigma^{-\frac{1}{2\alpha}} / \log^3 1/\sigma$. For any fixed distribution Q on $\mathcal{X} \times \{-1, 1\}$, consider the i.i.d. adversary which chooses (x_t, y_t) according to Q for each t. An online-to-batch reduction (Cesa-Bianchi et al., 2004; Shalev-Shwartz et al., 2011) establishes that if an online algorithm can achieve expected regret at most εT, then there is an offline algorithm that achieves expected error at most ε given T samples from Q. But Vapnik and Chervonenkis (1974) shows that for any binary function class $\mathcal{F}$ with $\mathsf{vc}(\mathcal{F}) = d$, no algorithm using only $c \cdot d/\varepsilon^2$ samples (for a sufficiently small constant c) can achieve expected error at most ε for all distributions Q whose $\mathcal{X}$-marginal is uniform on a shattered set of $\mathcal{F}$ of size d. Taking μ to be such a uniform marginal, we see that there is no online algorithm (regardless of oracle efficiency) that achieves regret $\leq \varepsilon T$ against any 1-smooth adversary with respect to μ if $T < c \cdot d/\varepsilon^2$.

Case 2. $d/\varepsilon^2 \leq \sigma^{-\frac{1}{2\alpha}} / \log^3 1/\sigma$. Consider any algorithm in the ERM oracle model, $\mathcal{A}$, whose total computation time over T rounds is bounded above by T^α, and suppose that for some value of T, $\mathcal{A}$ achieves regret at most εT against a σ-smooth adversary for any class of VC dimension at most d. Since any class $\mathcal{F}$ with $|\mathcal{F}| \leq 1/\sigma$ must have $\mathsf{vc}(\mathcal{F}) \leq \log 1/\sigma \leq d$, and since $\varepsilon < 1/200$, by Theorem 11, we must have that $T^\alpha \geq \Omega\left(\frac{1/\sqrt{\sigma}}{\log^3 1/\sigma}\right)$. Thus $T \geq \Omega\left(\frac{\sigma^{-1/(2\alpha)}}{\log^3 1/\sigma}\right)$, as desired.

The second statement of the corollary follows from the above casework by noting that, in Case 1, the computation time is at least the number of rounds $T \geq \Omega(d/\varepsilon^2)$, and in Case 2, we get immediately from Theorem 11 that the computation time is $\Omega(\sigma^{-1/2} / \log^3 1/\sigma)$. ■

F.4. Lower bound on oracle calls for approximate ERM oracle

One limitation of the lower bounds of Theorem 11 and Corollary 12 is that they only lower bound the total computation time in the ERM oracle model and thus, for instance, do not rule out an algorithm which makes a single ERM oracle call with a large number of points (x_i, y_i). In this section we amend this issue, showing a lower bound on the number of ERM oracle calls that any proper online learning algorithm obtaining sublinear regret must make. To obtain this result, we have to slightly weaken the oracle, namely by working with the *approximate ERM oracle model* (i.e., where we have $\zeta > 0$ in Definition 2).

First, we need a slight variant of Lemma 50, which establishes a lower bound on the number of oracle calls (which in general is less than computation time), but under the additional assumption that all oracle calls to $\mathsf{BR}^1, \mathsf{BR}^2$ are made with small-support vectors.

Lemma 51 *There is a constant $c_0 \in (0,1)$ so that the following holds. Fix $N, S \in \mathbb{N}$. Any randomized algorithm $\mathcal{A}$ for approximating the equilibrium of $N \times N$ $\{-1,1\}$-valued zero-sum games with the oracles $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ cannot guarantee with probability greater than $2/3$ that $\mathcal{A}$'s output value is at most $1/4$ from the game's true value with fewer than $\frac{1}{S} \cdot c_0 \cdot \sqrt{N} / \log^3 N$ oracle calls, assuming that each oracle call to $\mathsf{BR}^1, \mathsf{BR}^2$ is made on a vector of support at most S.*

Proof The proof exactly mirrors that of Lemma 50, with the exception of the analysis of how the oracles $\mathsf{BR}^1, \mathsf{BR}^2$ are simulated using oracle calls to the min-max consistent function $\phi : [N] \to \mathbb{Z}$. In particular, for any $q \in \mathbb{R}^N$, $\mathsf{BR}^1(q)$ and $\mathsf{BR}^2(p)$ may each be simulated using at most $S \cdot (\log(N) + 1)$ oracle calls to ϕ assuming that q, p have at most S nonzero values.

Thus, if $\gamma \leq \frac{1}{S} c_0 \sqrt{N} / \log^3 N$ denotes the total number of oracle calls to $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$, then the total number of oracle calls to ϕ is at most

$$S \cdot (\log(N) + 1) \cdot \gamma \leq S \cdot (\log(N) + 1) \cdot \frac{1}{S} \cdot c_0 \sqrt{N} / \log^3 N < c' \cdot \sqrt{N} / \log^2(N),$$

where c' is the constant of Corollary 47 (as long as c_0 is chosen sufficiently small). This gives the desired contradiction to Corollary 47. ■

Given Lemma 51 we may prove in a manner analogously to Theorem 11 a lower bound on the number of oracle calls for any no-regret algorithm in the approximate ERM oracle model:

Theorem 52 *For some constant $c > 0$ we have the following: fix any $T \in \mathbb{N}$, $\sigma, \zeta \in (0,1]$. In the ζ-approximate ERM oracle model, any randomized algorithm cannot guarantee expected regret smaller than $\frac{T}{200}$ against a σ-smooth online adversary and any $\mathcal{F}$ with $|\mathcal{F}| \leq 1/\sigma$ over T rounds in using fewer than $c\zeta^2 \cdot \frac{1/\sqrt{\sigma}}{\log^4 1/\sigma}$ oracle calls; further, this result holds even for binary-valued classes.*

Proof We use the notation from the proof of Theorem 11. The proof exactly follows that of Theorem 11, except for how the ERM oracle calls are simulated. To describe this difference, recall the definition of $N = 1/\sigma$ to denote the size of the given game G, set $\delta = \frac{1}{100N^2}$, and write $S := \frac{C \log 1/\delta}{\zeta^2}$. Then the ζ-approximate ERM oracles are simulated as follows:

- To make an ERM oracle call of the form

  $$\operatorname*{argmin}_{f_1 \in [N]} \sum_{i=1}^{m} w_i \cdot \ell_i(G_{f_1, f_{2,i}}, y_i), \tag{40}$$

  we do the following:

  1. Draw S i.i.d. samples $i_1, \ldots, i_S$ from the distribution over $[m]$ whose mass at i is propositional to $|w_i|$.
  2. Use the procedure as in the proof of Theorem 11 to make the ERM oracle call

     $$\operatorname*{argmin}_{f_1 \in [N]} \sum_{j=1}^{S} \operatorname{sign}(w_{i_j}) \cdot \ell_{i_j}(G_{f_1, f_{2,i_j}}, y_{i_j}). \tag{41}$$

     Notice that this will lead to an oracle call $\mathsf{BR}^1(q)$ for some distribution q of support size at most S.

- We perform the same sampling procedure for an ERM oracle call of the form $\operatorname{argmin}_{f_2 \in [N]} \sum_{i=1}^{m} w_i \cdot \ell_i(G_{f_{1,i}, f_2}, y_i)$, which leads to an oracle call $\mathsf{BR}^2(p)$ for some distribution p of support size at most S.

We claim that each such oracle call of the above form, with probability at least $1 - N \cdot \delta$, satisfies the requirement of ζ-approximate ERM oracle. To establish this, we simply note that by the Chernoff bound and union bound, with probability $1 - N \cdot \delta$, we have, for each oracle call of the form (40), for $W := \sum_{i=1}^{m} |w_i|$,

$$\sup_{f_1 \in [N]} \left| \sum_{i=1}^{m} \frac{w_i}{W} \cdot \ell_i(G_{f_1, f_{2,i}}, y_i) - \frac{1}{S} \sum_{j=1}^{S} \operatorname{sign}(w_{i_j}) \cdot \ell_{i_j}(G_{f_1, f_{2,i_j}}, y_{i_j}) \right| \leq \zeta,$$

which implies that the result of (41) returns some $\hat{f}_1$ which is ζW within the minimum of (40). A similar argument applies to the ERM oracle calls taking a minimum over $f_2 \in [N]$.

Since the total number of oracle calls of each of the algorithms $\mathcal{A}_1, \mathcal{A}_2$ in the proof of Theorem 11 is at most $c \cdot \frac{1/\sigma}{\log^4 1/\sigma} \leq N$, we have that with probability $1 - 2N^2\delta$, all oracle calls simulated as above are actually ζ-approximate ERM oracle calls. By the assumption of the theorem statement, it follows that with probability at least $2/3$ we can approximate the value of G up to accuracy of $1/4$. Further, the number of oracle calls made to $\mathsf{Val}, \mathsf{BR}^1, \mathsf{BR}^2$ is at most

$$c\zeta^2 \cdot \frac{1/\sqrt{\sigma}}{\log^4 1/\sigma} \leq \frac{1}{S} \cdot c_0 \frac{\sqrt{N}}{\log^3 N},$$

where c_0 is the constant of Lemma 51 (as long as c is sufficiently small), and each to $\mathsf{BR}^1, \mathsf{BR}^2$ is with a vector that has support size at most S. But this contradicts the statement of Lemma 51, completing the proof. ■

Finally, as a corollary of Theorem 52, we have the following analogue of Corollary 12, which shows a regret lower bound for any algorithm which makes polynomially many oracle queries to an oracle whose accuracy is an inverse polynomial. Notice that the upper bound of Theorem 10 obtains a regret bound under a ζ-approximate oracle that is the same as that under an exact oracle (up to a constant factor), as long as $\zeta < o\left(\frac{1}{T^2 \log T}\right)$; thus the assumption of $1/T^\alpha$-approximate oracle (for α constant) in the below corollary seems very reasonable.

Corollary 53 *Fix any $\alpha \geq 1, \varepsilon < 1/200, \sigma \in (0,1]$, and $d \geq \log 1/\sigma$. Any algorithm making at most T^α oracle calls over T rounds to a $1/T^\alpha$-approximate oracle requires that $T \geq \widetilde{\Omega}\left(\max\left\{\frac{d}{\varepsilon^2}, \sigma^{-\frac{1}{6\alpha}}\right\}\right)$ to achieve regret εT for classes $\mathcal{F}$ of VC dimension d against a σ-smooth adversary.*

Proof The proof is identical to that of Corollary 12 for $\frac{d}{\varepsilon^2} > \frac{\sigma^{-1/(6\alpha)}}{\log^4 1/\sigma}$.

For $\frac{d}{\varepsilon^2} \leq \frac{\sigma^{-1/(6\alpha)}}{\log^4 1/\sigma}$, we note that any algorithm making T^α oracle calls to a $1/T^\alpha$-approximate ERM oracle over T rounds, which achieves regret at most εT against a σ-smooth adversary must, by Theorem 52, have $T^\alpha \geq \Omega\left(T^{-2\alpha} \cdot \frac{1/\sqrt{\sigma}}{\log^4 1/\sigma}\right)$, i.e., $T \geq \Omega\left(\frac{\sigma^{-\frac{1}{6\alpha}}}{\log^4 1/\sigma}\right)$. ■

Appendix G. Proof of Theorem 13

We first prove a basic lemma about how smoothness behaves with product distributions.

Lemma 54 *Suppose that p is σ-smooth with respect to μ on $\mathcal{X}$ and for any $x \in \mathcal{X}$, $p'_x = p'(\cdot|x)$ is σ'-smooth with respect to μ' on $\mathcal{X}'$. Then $q(x,a) = p(x)p'(a|x)$ is $\sigma\sigma'$-smooth with respect to $\mu \otimes \mu'$.*

Proof Let $A \subset \mathcal{X}$ and $A' \subset \mathcal{X}'$ be measurable. Then

$$q(A \times A') = \mathbb{E}_{x\sim p}\left[p'(A'|x)\chi_{x\in A}\right] \leq \left(\frac{1}{\sigma'}\mu'(A')\right)p(A) \leq \frac{1}{\sigma\sigma'}\mu(A) \otimes \mu(A')$$

The result follows. ■

Note that any distribution on $[K]$ is $\frac{1}{K}$-smooth with respect to Unif($[K]$). Thus, by Lemma 54, independent of how a_t is chosen, we may assume that (x_t, a_t) is sampled from a distribution that is $\frac{\sigma}{K}$-smooth with respect to $\mu \otimes$ Unif($[K]$). Define the random quantity

$$\mathrm{Reg}_{Sq}(T) = \sum_{t=1}^{T} (\widehat{y}_t - \ell_t(a_t))^2 - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} (f(x_t, a_t) - \ell_t(a_t))^2$$

Then, by (Foster and Rakhlin, 2020, Theorem 1), with probability at least $1 - \delta$ over the randomization over actions, if we run SquareCB with parameter γ, we have

$$\mathrm{Reg}_{CB}(T) \le \frac{\gamma}{2}\mathrm{Reg}_{Sq}(T) + 4\gamma \log\left(\frac{2}{\delta}\right) + \frac{2KT}{\gamma} + \sqrt{2T \log\left(\frac{2}{\delta}\right)}$$

Setting $\delta = \frac{1}{T}$ and noting that the regret is always at most T, we have

$$\mathbb{E}\left[\mathrm{Reg}_{CB}(T)\right] \le \frac{\gamma}{2}\mathbb{E}\left[\mathrm{Reg}_{Sq}(T)\right] + 4\gamma\log(2T) + \frac{2KT}{\gamma} + \sqrt{2T\log\left(2T\right)} + 1 \tag{42}$$

If we set $\widehat{y}_t$ to be the prediction given by the relaxation-based algorithm from (4), setting $k = \frac{3K}{\sigma}\log T$ then we know that

$$\mathbb{E}\left[\mathrm{Reg}_{Sq}(T)\right] \le \frac{7LK\log T}{\sigma}\mathcal{R}_T(\mathcal{F})$$

can be achieved with $O\left(T^{\frac{3}{2}}\log T\right)$ calls to the ERM oracle. Letting

$$\gamma = 12\log(T)\sqrt{\frac{T\sigma}{L\mathcal{R}_T(\mathcal{F})}}$$

concludes the proof after noting that we may take $L = 2$ for the square loss in the range $[0, 1]$.

If we instead use the FTPL algorithm of Theorem 10, then $\mathsf{vc}(\mathcal{F}, \alpha) \lesssim \alpha^{-p}$ implies

$$\mathbb{E}\left[\mathrm{Reg}_{Sq}(T)\right] \le \widetilde{O}\left(\left(\frac{T\sqrt{K}}{\sqrt{\sigma}}\right)^{\max\left(1 - \frac{1}{3(p-1)}, \frac{2}{3}\right)}\right)$$

Plugging into (42) and minimizing over γ yields a regret of

$$\mathbb{E}\left[\mathrm{Reg}_{CB}(T)\right] \le \widetilde{O}\left(T^{\max\left(1 - \frac{1}{6(p-1)}, \frac{5}{6}\right)} K^{\max\left(\frac{3}{4} - \frac{1}{12(p-1)}, \frac{2}{3}\right)} \sigma^{-\frac{1}{4}}\right)$$

Note that for any $p < \infty$, this is $o(T)$ and so the result holds.