Competing With Strategies

Wei Han
Univ. of Pennsylvania

Alexander Rakhlin
Univ. of Pennsylvania

Karthik Sridharan
Univ. of Pennsylvania

February 13, 2013

Abstract

We study the problem of online learning with a notion of regret defined with respect to a set of strategies. We develop tools for analyzing the minimax rates and for deriving regret-minimization algorithms in this scenario. While the standard methods for minimizing the usual notion of regret fail, through our analysis we demonstrate existence of regret-minimization methods that compete with such sets of strategies as: autoregressive algorithms, strategies based on statistical models, regularized least squares, and follow the regularized leader strategies. In several cases we also derive efficient learning algorithms.

1 Introduction

The common criterion for evaluating an online learning algorithm is *regret*, that is the difference between the cumulative loss of the algorithm and the cumulative loss of the best fixed decision, chosen in hindsight. While much work has been done on understanding no-regret algorithms, such a definition of regret against a fixed decision often draws criticism: even if regret is small, the cumulative loss of a best *fixed* action can be large, thus rendering the result uninteresting. To address this problem, various generalizations of the regret notion have been proposed, including regret with respect to the cost of a "slowly changing" compound decision. While being a step in the right direction, such definitions are still "static" in the sense that the decision of each compound comparator per step does not depend on the sequence of realized outcomes.

Arguably, a more interesting (and more difficult to deal with) notion is that of performing as well as a set of *strategies* (or, *algorithms*). A strategy π is a sequence of functions π_t, for each time period t, mapping the observed outcomes to the next action. Of course, if the collection of such strategies is finite, we may disregard their dependence on the actual sequence and treat each strategy as a black box expert. This is precisely the reason the Multiplicative Weights and other expert algorithms gained such popularity. However, this "black box" approach is not always desirable since some measure of the "effective number" of experts must play a role in the complexity of the problem: experts that predict similarly should not count as two independent ones. But what is a notion of closeness of two strategies? Imagine that we would like to develop an algorithm that incurs loss comparable to that of the best of an infinite family of strategies. To obtain such a statement, one may try to discretize the space of strategies and invoke the black-box experts method. As we show in this paper, such an approach will not always work. Instead, we present a theoretical framework

1

for the analysis of "competing against strategies" and for algorithmic development, based on the ideas in [11, 9].

The strategies considered in this paper are termed "simulatable experts" in [3]. The authors also distinguish *static* and *non-static* experts. In particular, for static experts and absolute loss, [2] were able to show that problem complexity is governed by the geometry of the class of static experts as captured by its i.i.d. Rademacher averages. For nonstatic experts, however, the authors note that "unfortunately we do not have a characterization of the minimax regret by an empirical process", due to the fact that the sequential nature of the online problems is at odds with the i.i.d.-based notions of classical empirical process theory. In recent years, however, a martingale generalization of empirical process theory has emerged, and these tools were shown to characterize learnability of online supervised learning, online convex optimization, and other scenarios [11, 1]. Yet, the machinery developed so far is not directly applicable to the case of general simulatable experts which can be viewed as mappings from an ever-growing set of histories to the space of actions. The goal of this paper is precisely this: to extend the non-constructive as well as constructive techniques of [11, 9] to simulatable experts. We analyze a number of examples with the developed techniques, but we must admit that our work only scratches the surface. We can imagine further research developing methods that compete with interesting gradient descent methods (parametrized by step size choices), with Bayesian procedures (parametrized by choices of priors), and so on. We also note the connection to online algorithms, where one typically aims to prove a bound on the competitive ratio. Our results can be seen in that light as implying a competitive ratio of one.

We close the introduction with a high-level outlook, which builds on the ideas of [8]. Imagine we are faced with a sequence of data from a probabilistic source, such as a k-Markov model with unknown transition probabilities. A well developed statistical theory tells us how to estimate the parameter *under the assumption that the model is correct*. We may view an estimator as a *strategy* for predicting the next outcome. Suppose we have a set of possible models, with a good prediction strategy for each model. Now, let us lift the assumption that the sequence is generated by one of these models, and set the goal as that of performing as well as the best prediction strategy. In this case, if the observed sequence is indeed given by one of the models, our loss will be small because one of the strategies will perform well. If not, we still have a valid statement that does not rely on the fact that the model is "well specified". To illustrate the point, we will exhibit an example where we can compete with the set of all Bayesian strategies (parametrized by priors). We then obtain a statement that we perform as well as the best of them without assuming that the model is correct.

The paper is organized as follows. In Section 2, we extend the minimax analysis of online learning problems to the case of competing with a set of strategies. In Section 3, we show that it is possible to compete with a set of autoregressive strategies, and that the usual online linear optimization algorithms do not attain the optimal bounds. We then derive an optimal and computationally efficient algorithm for one of the proposed regimes. In Section 4 we describe the general idea of competing with statistical models that use sufficient statistics, and demonstrate an example of competing with a set of strategies parametrized by priors. For this example, we derive an optimal and efficient randomized algorithm. In Section 5, we turn to the question of competing with regularized least squares algorithms indexed by the choice of a shift and a regularization parameter. In Section 6, we consider online linear optimization and show that it is possible to compete with Follow the Regularized Leader methods parametrized by a shift and by a step size

schedule.

2 Minimax Regret and Sequential Rademacher Complexity

We consider the problem of online learning, or sequential prediction, that consists of T rounds. At each time $t = \{1, \ldots, T\} \triangleq [T]$, the learner makes a prediction $f_t \in \mathcal{F}$ and observes an outcome $z_t \in \mathcal{Z}$, where $\mathcal{F}$ and $\mathcal{Z}$ are abstract sets of decisions and outcomes. Let us fix a loss function $\ell : \mathcal{F} \times \mathcal{Z} \mapsto \mathbb{R}$ that measures the quality of prediction. A *strategy* $\pi = (\pi_t)_{t=1}^T$ is a sequence of functions $\pi_t : \mathcal{Z}^{t-1} \mapsto \mathcal{F}$ mapping history of outcomes to a decision. Let Π denote a set of strategies. The regret with respect to Π is the difference between the cumulative loss of the player and the cumulative loss of the best strategy

$$\mathbf{Reg}_T = \sum_{t=1}^T \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^T \ell(\pi_t(z_{1:t-1}), z_t).$$

where we use the notation $z_{1:k} \triangleq \{z_1, \ldots, z_k\}$. We now define the value of the game against a set Π of strategies as

$$\mathcal{V}_T(\Pi) \triangleq \inf_{q_1 \in \mathcal{Q}} \sup_{z_1 \in \mathcal{Z}} \mathbb{E}_{f_1 \sim q_1} \ldots \inf_{q_T \in \mathcal{Q}} \sup_{z_T \in \mathcal{Z}} \mathbb{E}_{f_T \sim q_T} \left[\mathbf{Reg}_T \right]$$

where $\mathcal{Q}$ and $\mathcal{P}$ are the sets of probability distributions on $\mathcal{F}$ and $\mathcal{Z}$, correspondingly. It was shown in [11] that one can derive non-constructive upper bounds on the value through a process of sequential symmetrization, and in [9] it was shown that these non-constructive bounds can be used as relaxations to derive an algorithm. This is the path we take in this paper.

Let us describe an important variant of the above problem – that of *supervised learning*. Here, before making a real-valued prediction $\hat{y}_t$ on round t, the learner observes side information $x_t \in \mathcal{X}$. Simultaneously, the actual outcome $y_t \in \mathcal{Y}$ is chosen by Nature. A strategy can therefore depend on the history $x_{1:t-1}, y_{t-1}$ *and* the current x_t, and we write such strategies as $\pi_t(x_{1:t}, y_{1:t-1})$, with $\pi_t : \mathcal{X}^t \times \mathcal{Y}^{t-1} \mapsto \mathcal{Y}$. Fix some loss function $\ell(\hat{y}, y)$. The value $\mathcal{V}_T^S(\Pi)$ is then defined as

$$\sup_{x_1} \inf_{q_1 \in \Delta(\mathcal{Y})} \sup_{y_1 \in \mathcal{Y}} \mathbb{E}_{\hat{y}_1 \sim q_1} \ldots \sup_{x_T} \inf_{q_T \in \Delta(\mathcal{Y})} \sup_{y_T \in \mathcal{Y}} \mathbb{E}_{\hat{y}_T \sim q_T} \left[\sum_{t=1}^T \ell(\hat{y}_t, y_t) - \inf_{\pi \in \Pi} \sum_{t=1}^T \ell(\pi_t(x_{1:t}, y_{1:t-1}), y_t) \right]$$

To proceed, we need to define a notion of a tree. A $\mathcal{Z}$-valued tree $\mathbf{z}$ is a sequence of mappings $\{\mathbf{z}_1, \ldots, \mathbf{z}_T\}$ with $\mathbf{z}_t : \{\pm 1\}^{t-1} \mapsto \mathcal{Z}$. Throughout the paper, $\epsilon_t \in \{\pm 1\}$ are i.i.d. Rademacher variables, and a realization of $\epsilon = (\epsilon_1, \ldots, \epsilon_T)$ defines a *path* on the tree, given by $\mathbf{z}_{1:t}(\epsilon) \triangleq (\mathbf{z}_1(\epsilon), \ldots, \mathbf{z}_t(\epsilon))$ for any $t \in [T]$. We write $\mathbf{z}_t(\epsilon)$ for $\mathbf{z}_t(\epsilon_{1:t-1})$. By convention, a sum $\sum_a^b = 0$ for $a > b$ and for simplicity assume that no loss is suffered on the first round.

Definition 1. Sequential Rademacher complexity of the set Π of strategies is defined as

$$\mathfrak{R}(\ell, \Pi) \triangleq \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \ell(\pi_t(\mathbf{w}_1(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon)), \mathbf{z}_t(\epsilon)) \right] \tag{1}$$

where the supremum is over two $\mathcal{Z}$-valued trees $\mathbf{z}$ and $\mathbf{w}$ of depth T.

3

The **w** tree can be thought of as providing "history" while **z** providing "outcomes". We shall use these names throughout the paper. The reader might notice that in the above definition, the outcomes and history are decoupled. We now state the main result:

Theorem 1. *The value of prediction problem with a set Π of strategies is upper bounded as*

$$\mathcal{V}_T(\Pi) \le 2\mathfrak{R}(\ell, \Pi)$$

While the statement is visually similar to those in [11, 12], it does not follow from these works. Indeed, the proof (which appears in Appendix) needs to deal with the additional complications stemming from the dependence of strategies on the history. Further, we provide the proof for a more general case when sequences $z_1, \ldots, z_T$ are not arbitrary but need to satisfy constraints.

As we show below, the sequential Rademacher complexity on the right-hand side allows us to analyze general non-static experts, thus addressing the question raised in [2]. As the first step, we can "erase" a Lipschitz loss function (see [10] for more details), leading to the sequential Rademacher complexity of Π without the loss and without the **z** tree:

$$\mathfrak{R}(\Pi) \triangleq \sup_{\mathbf{w}} \mathfrak{R}(\Pi, \mathbf{w}) \triangleq \sup_{\mathbf{w}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{w}_{1:t-1}(\epsilon)) \right]$$

For example, suppose $\mathcal{Z} = \{0, 1\}$, the loss function is the indicator loss, and strategies have potentially dependence on the full history. Then one can verify that

$$\sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi^k} \left[\sum_{t=1}^T \epsilon_t \mathbf{1} \left\{ \pi_t(\mathbf{w}_{1:t-1}(\epsilon)) \ne \mathbf{z}_t(\epsilon) \right\} \right]$$

$$= \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi^k} \left[\sum_{t=1}^T \epsilon_t \left(\pi_t(\mathbf{w}_{1:t-1}(\epsilon))(1 - 2\mathbf{z}_t(\epsilon)) + \mathbf{z}_t(\epsilon) \right) \right] = \mathfrak{R}(\Pi) \tag{2}$$

The same result holds when $\mathcal{F} = [0, 1]$ and ℓ is the absolute loss. The process of "erasing the loss" (or, contraction) extends quite nicely to problems of supervised learning. Let us state the second main result:

Theorem 2. *Suppose the loss function $\ell : \mathcal{Y} \times \mathcal{Y} \mapsto \mathbb{R}$ is convex and L-Lipschitz in the first argument, and let $\mathcal{Y} = [-1, 1]$. Then*

$$\mathcal{V}_T^S(\Pi) \le 2L \sup_{\mathbf{x}, \mathbf{y}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon)) \right]$$

where $(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon))$ naturally takes place of $\mathbf{w}_{1:t-1}(\epsilon)$ in Theorem 1. Further, if $\mathcal{Y} = [-1, 1]$ and $\ell(\hat{y}, y) = |\hat{y} - y|$, $\mathcal{V}_T^S(\Pi) \ge \sup_{\mathbf{x}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \epsilon_{1:t-1}) \right]$.

Let us present a few simple examples as a warm-up.

Example 1 (History-independent strategies)**.** *Let $\pi^f \in \Pi$ be constant history-independent strategies $\pi_1^f = \ldots = \pi_T^f = f \in \mathcal{F}$. Then (1) recovers the definition of sequential Rademacher complexity in [11].*

Example 2 (Static experts). *For static experts, each strategy π is a predetermined sequence of outcomes, and we may therefore associate each π with a vector in $\mathcal{Z}^T$. A direct consequence of Theorem 2 for any convex L-Lipschitz loss is that*

$$\mathcal{V}(\Pi) \le 2L\mathbb{E}_\epsilon \sup_{\pi\in\Pi}\left[\sum_{t=1}^T \epsilon_t\pi_t\right]$$

which is simply the classical i.i.d. Rademacher averages. For the case of $\mathcal{F} = [0,1]$, $\mathcal{Z} = \{0,1\}$, and the absolute loss, this is the result of [2].

Example 3 (Finite-order Markov strategies). *Let Π^k be a set of strategies that only depend on the k most recent outcomes to determine the next move. Theorem 1 implies that the value of the game is upper bounded as*

$$\mathcal{V}(\Pi^k) \le 2\sup_{\mathbf{w},\mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi^k}\left[\sum_{t=1}^T \epsilon_t\ell(\pi_t(\mathbf{w}_{t-k}(\epsilon),\dots,\mathbf{w}_{t-1}(\epsilon)),\mathbf{z}_t(\epsilon))\right]$$

Now, suppose that $\mathcal{Z}$ is a finite set, of cardinality s. Then there are effectively s^{s^k} strategies π. The bound on the sequential Rademacher complexity then scales as $\sqrt{2s^k\log(s)T}$, recovering the result of [5] (see [3, Cor. 8.2]).

In addition to providing an understanding of minimax regret against a set of strategies, sequential Rademacher complexity can serve as a starting point for algorithmic development. As shown in [9], any *admissible relaxation* can be used to define a succinct algorithm with a regret guarantee. For the setting of this paper, this means the following. Let $\mathbf{Rel}:\mathcal{Z}^t\mapsto\mathbb{R}$, for each t, be a collection of functions satisfying two conditions:

$$\forall t, \ \inf_{q_t\in\mathcal{Q}}\sup_{z_t\in\mathcal{Z}}\left\{\mathop{\mathbb{E}}_{f_t\sim q_t}\ell(f_t,z_t)+\mathbf{Rel}(z_{1:t})\right\} \le \mathbf{Rel}(z_{1:t-1}), \ \text{and} \ -\inf_{\pi\in\Pi}\sum_{t=1}^T\ell(\pi_t(z_{1:t-1}),z_t) \le \mathbf{Rel}(z_{1:T}) \ .$$

Then we say that *the relaxation is admissible*. It is then easy to show that regret of any algorithm that ensures above inequalities is bounded by $\mathbf{Rel}(\{\})$.

Theorem 3. *The conditional sequential Rademacher complexity with respect to Π*

$$\mathfrak{R}(\ell,\Pi|z_1,\dots,z_t) \triangleq \sup_{\mathbf{z},\mathbf{w}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{\pi\in\Pi}\left[2\sum_{s=t+1}^T \epsilon_s\ell(\pi_s((z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon))-\sum_{s=1}^t\ell(\pi_s(z_{1:s-1}),z_s)\right]$$

is admissible.

Conditional sequential Rademacher complexity can therefore be used as a starting point for possibly deriving computationally attractive algorithms, as shown throughout the paper.

We may now define covering numbers for the set Π of strategies over the history trees. The development is a straightforward modification of the notions we developed in [11], where we replace "any tree $\mathbf{x}$" with a tree of histories $\mathbf{w}_{1:t-1}$.

Definition 2. A set V of $\mathbb{R}$-valued trees is an α-*cover* (with respect to ℓ_p) of a set of strategies Π on an $\mathcal{Z}^*$-valued history tree $\mathbf{w}$ if

$$\forall\pi\in\Pi, \ \forall\epsilon\in\{\pm1\}^T, \ \exists\mathbf{v}\in V \quad \text{s.t.} \quad \left(\tfrac{1}{T}\sum_{t=1}^T|\pi_t(\mathbf{w}_{1:t-1}(\epsilon))-\mathbf{v}_t(\epsilon)|^p\right)^{1/p}\le\alpha \ . \tag{3}$$

An α-*covering number* $\mathcal{N}_p(\Pi,\mathbf{w},\alpha)$ is the size of the smallest α-cover.

For supervised learning, $(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon))$ takes place of $\mathbf{w}_{1:t-1}(\epsilon)$. Now, for any history tree $\mathbf{w}$, sequential Rademacher averages of a class of $[-1,1]$-valued strategies Π satisfy

$$\mathfrak{R}(\Pi, \mathbf{w}) \leq \inf_{\alpha \geq 0} \left\{ \alpha T + \sqrt{2 \log \mathcal{N}_1(\Pi, \mathbf{w}, \alpha) T} \right\}$$

and the Dudley entropy integral type bound also holds:

$$\mathfrak{R}(\Pi, \mathbf{w}) \leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_{\alpha}^{1} \sqrt{\log \mathcal{N}_2(\Pi, \mathbf{w}, \delta)} \, d\delta \right\} \tag{4}$$

In particular, this bound should be compared with Theorem 7 in [2], which employs a covering number in terms of a pointwise metric between strategies that requires closeness for *all histories and all time steps*. Second, the results of [2] for real-valued prediction require strategies to be bounded away from 0 and 1 by $\delta > 0$ and this restriction spoils the rates.

In the rest of the paper, we show how the results of this section (a) yield proofs of existence of regret-minimization strategies with certain rates and (b) guide in the development of algorithms. For some of these examples, standard methods (such as Exponential Weights) come close to providing an optimal rate, while for others – fail miserably.

3 Competing with Autoregressive Strategies

In this section, we consider strategies that depend linearly on the past outcomes. To this end, we fix a set $\Theta \subset \mathbb{R}^k$, for some $k > 0$, and parametrize the set of strategies as

$$\Pi_\Theta = \left\{ \pi^\theta : \pi_t^\theta(z_1, \ldots, z_{t-1}) = \sum_{i=0}^{k-1} \theta_{i+1} z_{t-k+i}, \quad \theta = (\theta_1, \ldots, \theta_k) \in \Theta \right\}$$

For consistency of notation, we assume that the sequence of outcomes is padded with zeros for $t \leq 0$. First, as an example where known methods *can* recover the correct rate, we consider the case of a constant look-back of size k. We then extend the study to cases where neither the regret behavior nor the algorithm is known in the literature, to the best of our knowledge.

3.1 Finite Look-Back

Suppose $\mathcal{Z} = \mathcal{F} \subset \mathbb{R}^d$ are ℓ_2 unit balls, the loss is $\ell(f, z) = \langle f, z \rangle$, and $\Theta \subset \mathbb{R}^k$ is also a unit ℓ_2 ball. Denoting by $\mathbf{W}_{(t-k:t-1)} = [\mathbf{w}_{t-k}(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon)]$ a matrix with columns in $\mathcal{Z}$,

$$\mathfrak{R}(\ell, \Pi_\Theta) = \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\theta \in \Theta} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \pi^\theta(\mathbf{w}_{t-k:t-1}(\epsilon)), \mathbf{z}_t(\epsilon) \right\rangle \right] = \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\theta \in \Theta} \left[\sum_{t=1}^{T} \epsilon_t \mathbf{z}_t(\epsilon)^\top \mathbf{W}_{(t-k:t-1)} \cdot \theta \right]$$

$$= \sup_{\mathbf{w}, \mathbf{z}} \mathbb{E}_\epsilon \left\| \sum_{t=1}^{T} \epsilon_t \mathbf{z}_t(\epsilon)^\top \mathbf{W}_{(t-k:t-1)} \right\| \leq \sqrt{kT} \tag{5}$$

In fact, this bound against all strategies parametrized by Θ is achieved by the gradient descent (GD) method with the simple update $\theta_{t+1} = \mathrm{Proj}_\Theta(\theta_t - \eta [z_{t-k}, \ldots, z_{t-1}]^\top z_t)$ where Proj_Θ is the Euclidean projection onto the set Θ. This can be seen by writing the loss as

$$\langle [z_{t-k}, \ldots, z_{t-1}] \cdot \theta_t, z_t \rangle = \langle \theta_t, [z_{t-k}, \ldots, z_{t-1}]^\top z_t \rangle.$$

The regret of GD, $\sum_{t=1}^{T} \langle \theta_t, [z_{t-k}, \ldots, z_{t-1}]^\top z_t \rangle - \inf_{\theta \in \Theta} \sum_{t=1}^{T} \langle \theta, [z_{t-k}, \ldots, z_{t-1}]^\top z_t \rangle$, is precisely regret against strategies in Θ, and analysis of GD yields the rate in (5).

3.2 Full Dependence on History

The situation becomes less obvious when $k = T$ and strategies depend on the full history. The regret bound in (5) is vacuous, and the question is whether a better bound can be proved, under some additional assumptions on Θ. Can such a bound be achieved by GD?

For simplicity, consider the case of $\mathcal{F} = \mathcal{Z} = [-1, 1]$, and assume that $\Theta = B_p(1) \subset \mathbb{R}^T$ is a unit ℓ_p ball, for some $p \geq 1$. Since $k = T$, it is easier to re-index the coordinates so that

$$\pi_t^\theta(z_{1:t-1}) = \sum_{i=1}^{t-1} \theta_i z_i.$$

The sequential Rademacher complexity of the strategy class is

$$\mathfrak{R}(\ell, \Pi_\Theta) = \sup_{\mathbf{w},\mathbf{z}} \mathbb{E} \sup_{\theta \in \Theta} \left[\sum_{t=1}^{T} \epsilon_t \pi^\theta(\mathbf{w}_{1:t-1}(\epsilon)) \cdot \mathbf{z}_t(\epsilon) \right] = \sup_{\mathbf{w},\mathbf{z}} \mathbb{E} \sup_{\theta \in \Theta} \left[\sum_{t=1}^{T} \left(\sum_{i=1}^{t-1} \theta_i \mathbf{w}_i(\epsilon) \right) \epsilon_t \mathbf{z}_t(\epsilon) \right].$$

Rearranging the terms, the last expression is equal to

$$\sup_{\mathbf{w},\mathbf{z}} \mathbb{E} \sup_{\theta \in \Theta} \left[\sum_{t=1}^{T-1} \theta_t \mathbf{w}_t(\epsilon) \cdot \left(\sum_{i=t+1}^{T} \epsilon_i \mathbf{z}_i(\epsilon) \right) \right] \leq \sup_{\mathbf{w},\mathbf{z}} \mathbb{E} \left[\|\mathbf{w}_{1:T-1}(\epsilon)\|_q \cdot \max_{1 \leq t \leq T} \left| \sum_{i=t+1}^{T} \epsilon_i \mathbf{z}_i(\epsilon) \right| \right]$$

where q is the Hölder conjugate of p. Observe that

$$\sup_{\mathbf{z}} \mathbb{E} \sup_{1 \leq t \leq T} \left| \sum_{i=t}^{T} \epsilon_i \mathbf{z}_i(\epsilon) \right| \leq \sup_{\mathbf{z}} \mathbb{E} \left[\left| \sum_{i=1}^{T} \epsilon_i \mathbf{z}_i(\epsilon) \right| + \sup_{1 \leq t \leq T} \left| \sum_{i=1}^{t-1} \epsilon_i \mathbf{z}_i(\epsilon) \right| \right] \leq 2 \sup_{\mathbf{z}} \mathbb{E} \sup_{1 \leq t \leq T} \left| \sum_{i=1}^{t} \epsilon_i \mathbf{z}_i(\epsilon) \right|$$

Since $\{\epsilon_t \mathbf{z}_t(\epsilon) : t = 1, \ldots, T\}$ is a bounded martingale difference sequence, the last term is of the order of $\mathcal{O}(\sqrt{T})$. Now, suppose there is some $\beta > 0$ such that $\|\mathbf{w}_{1:T-1}(\epsilon)\|_q \leq T^\beta$ for all ϵ. This assumption can be implemented if we consider constrained adversaries, where such ℓ_q-bound is required to hold for any prefix $\mathbf{w}_{1:t}(\epsilon)$ of history (In Appendix, we prove Theorem 1 for the case of constrained sequences). Then $\mathfrak{R}(\ell, \Pi_\Theta) \leq C \cdot T^{\beta+1/2}$ for some constant C. We now compare the rate of convergence of sequential Rademacher and the rate of the mirror descent algorithm for different settings of q in Table 3.2. If $\|\theta\|_p \leq 1$ and $\|\mathbf{w}\|_q \leq T^\beta$ for $q \geq 2$, the convergence rate of mirror descent with Legendre function $F(\theta) = \frac{1}{2}\|\theta\|_p^2$ is $\sqrt{q-1}T^{\beta+1/2}$ (see [13]).

	Θ	$\mathbf{w}_{1:T}$	sequential Radem. rate	Mirror descent rate
	$B_1(1)$	$\|\mathbf{w}_{1:T-1}\|_\infty \leq 1$	$\sqrt{T}$	$\sqrt{T \log T}$
$q \geq 2$	$B_p(1)$	$\|\mathbf{w}_{1:T-1}\|_q \leq T^\beta$	$T^{\beta+1/2}$	$\sqrt{q-1}T^{\beta+1/2}$
	$B_2(1)$	$\|\mathbf{w}_{1:T-1}\|_2 \leq T^\beta$	$T^{\beta+1/2}$	$T^{\beta+1/2}$
$1 \leq q \leq 2$	$B_p(1)$	$\|\mathbf{w}_{1:T-1}\|_q \leq T^\beta$	$T^{\beta+1/2}$	$T^{\beta+1/q}$
	$B_\infty(1)$	$\|\mathbf{w}_{1:T-1}\|_1 \leq T^\beta$	$T^{\beta+1/2}$	T

Table 1: Comparison of the rates of convergence (up to constant factors)

We observe that mirror descent, which is known to be optimal for online linear optimization, and which gives the correct rate for the case of *bounded* look-back strategies, in several regimes fails to yield the correct rate for more general linearly parametrized strategies. Even in the most basic regime where Θ is a unit ℓ_1 ball and the sequence of data is not constrained (other than $\mathcal{Z} = [-1, 1]$), there is a gap of $\sqrt{\log T}$ between the Rademacher bound and the guarantee of mirror descent. Is there an algorithm that removes this factor?

7

3.2.1 Algorithms for $\Theta = B_1(1)$

For the example considered in the previous section, with $\mathcal{F} = \mathcal{Z} = [-1,1]$ and $\Theta = B_1(1)$, the conditional sequential Rademacher complexity of Theorem 3 becomes

$$\mathfrak{R}_T(\Pi|z_1,\ldots,z_t) = \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi\in\Pi} \left[2\sum_{s=t+1}^T \epsilon_s \pi_s(z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon))\cdot\mathbf{z}_s(\epsilon) - \sum_{s=1}^t \pi_s(z_{1:s-1})\cdot z_s \right]$$

$$\leq \sup_{\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi\in\Pi} \left[2\sum_{s=t+1}^T \epsilon_s \pi_s(z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)) - \sum_{s=1}^t z_s \pi_s(z_{1:s-1}) \right]$$

where the $\mathbf{z}$ tree is "erased", as at the end of the proof of Theorem 2. Define $a_s(\epsilon) = 2\epsilon_s$ for $s > t$ and $-z_s$ otherwise; $b_i(\epsilon) = \mathbf{w}_i(\epsilon)$ for $i > t$ and z_i otherwise. We can then simply write

$$\sup_{\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\theta\in\Theta} \left[\sum_{s=1}^T a_s(\epsilon) \sum_{i=1}^{s-1} \theta_i b_i(\epsilon) \right] = \sup_{\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\theta\in\Theta} \left[\sum_{s=1}^{T-1} \theta_s b_s(\epsilon) \sum_{i=s+1}^T a_i(\epsilon) \right] \leq \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\leq s\leq T} \left| \sum_{i=s}^T a_i(\epsilon) \right|$$

which we may use as a relaxation:

Lemma 4. *Define* $a_s^t(\epsilon) = 2\epsilon_s$ *for* $s > t$, *and* $-z_s$ *otherwise. Then,*

$$\mathbf{Rel}(z_{1:t}) = \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\leq s\leq T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right|$$

is an admissible relaxation.

With this relaxation, the following method attains $\mathcal{O}(\sqrt{T})$ regret: prediction at step t is

$$q_t = \operatorname*{argmin}_{q\in[-1,1]} \sup_{z_t\in\{\pm 1\}} \left\{ \mathbb{E}_{f_t\sim q} f_t\cdot z_t + \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\leq s\leq T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\}$$

where the sup over $z_t \in [-1,1]$ is achieved at $\{\pm 1\}$ due to convexity. Following [9], we can also derive randomized algorithms, which can be viewed as "randomized playout" generalizations of the Follow the Perturbed Leader algorithm.

Lemma 5. *Consider the randomized strategy where at round t we first draw $\epsilon_{t+1},\ldots,\epsilon_T$ uniformly at random and then further draw our move f_t according to the distribution*

$$q_t(\epsilon) = \operatorname*{argmin}_{q\in[-1,1]} \sup_{z_t\in\{-1,1\}} \left\{ \mathbb{E}_{f_t\sim q} f_t\cdot z_t + \max_{1\leq s\leq T} \left| \sum_{i=s}^T a_i^t(\epsilon) \right| \right\}$$

$$= \frac{1}{2}\left(\max\left\{ \max_{s=1,\ldots,t} \left| -\sum_{i=s}^{t-1} z_i + 1 + 2\sum_{i=t+1}^T \epsilon_i \right|, \quad \max_{s=t+1,\ldots,T} \left| 2\sum_{i=s}^T \epsilon_i \right| \right\} \right.$$
$$\left. - \max\left\{ \max_{s=1,\ldots,t} \left| -\sum_{i=s}^{t-1} z_i - 1 + 2\sum_{i=t+1}^T \epsilon_i \right|, \quad \max_{s=t+1,\ldots,T} \left| 2\sum_{i=s}^T \epsilon_i \right| \right\} \right)$$

The expected regret of this randomized strategy is upper bounded by sequential Rademacher complexity: $\mathbb{E}[\mathbf{Reg}_T] \leq 2\mathfrak{R}_T(\Pi)$, *which was shown to be* $\mathcal{O}(\sqrt{T})$ *(see Table 3.2).*

The time consuming parts of the above randomized method are to draw $T-t$ random bits at round t and to calculate the partial sums. However, we may replace Rademacher random variables by Gaussian $\mathcal{N}(0,1)$ random variables and use known results on the distributions of extrema of a

Brownian motion. To this end, define a Gaussian analogue of conditional sequential Rademacher complexity

$$\mathcal{G}_T(\Pi|z_1,\ldots,z_t) = \sup_{\mathbf{z},\mathbf{w}} \mathbb{E} \sup_{\sigma_{t+1:T}} \sup_{\pi\in\Pi}\left[\sqrt{2\pi}\sum_{s=t+1}^{T}\sigma_s\ell(\pi_s((z_{1:t},\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - \sum_{s=1}^{t}\ell(\pi_s(z_{1:s-1}),z_s)\right]$$

where $\sigma_t \sim \mathcal{N}(0,1)$, and $\epsilon = (\text{sign}(\sigma_1),\ldots,\text{sign}(\sigma_T))$. For our example the $\mathcal{O}(\sqrt{T})$ bound can be shown for $\mathcal{G}_T(\Pi)$ by calculating the expectation of the maximum of Brownian motion. Proofs similar to Theorem 1 and Theorem 3 show that the conditional Gaussian complexity $\mathcal{G}_T(\Pi|z_1,\ldots,z_t)$ is an upper bound on $\mathfrak{R}_T(\Pi|z_1,\ldots,z_t)$ and is admissible (see Theorem 9 in Appendix). Furthermore, the proof of Lemma 5 holds for Gaussian random variables, and gives the randomized algorithm as in Lemma 5 with ϵ_t replaced by σ_t. It is not difficult to see that we can keep track of the maximum and minimum of $\left\{-\sum_{i=s}^{t-1}z_i\right\}$ between rounds in $\mathcal{O}(1)$ time. We can then draw three random variables from the joint distribution of the maximum, the minimum and the endpoint of a Brownian Motion and calculate the prediction in $\mathcal{O}(1)$ time per round of the game (the joint distribution can be found in [7]). In conclusion, we have derived an algorithm that for the case of $\Theta = B_1(1)$, with time complexity of $\mathcal{O}(1)$ per round and the optimal regret bound of $\mathcal{O}(\sqrt{T})$. We leave it as an open question to develop efficient and optimal algorithms for the other settings in Table 3.2.

4   Competing with Statistical Models

In this section we consider competing with a set of strategies that arise from statistical models. For example, for the case of Bayesian models, strategies are parametrized by the choice of a prior. Regret bounds with respect to a set of such methods can be thought of as a robustness statement: we are aiming to perform as well as the strategy with the best choice of a prior. We start this section with a general setup that needs further investigation.

4.1   Compression and Sufficient Statistics

Assume that strategies in Π have a particular form: they all work with a "sufficient statistic", or, more loosely, *compression* of the past data. Suppose "sufficient statistics" can take values in some set Γ. Fix a set $\bar{\Pi}$ of mappings $\bar{\pi} : \Gamma \mapsto \mathcal{F}$. We assume that all the strategies in Π are of the form $\pi_t(z_1,\ldots,z_{t-1}) = \bar{\pi}(\gamma(z_1,\ldots,z_{t-1}))$ for some $\bar{\pi} \in \bar{\Pi}$ and $\gamma : \mathcal{Z}^* \mapsto \Gamma$. Such a bottleneck Γ can arise due to a finite memory or finite precision, but can also arise if the strategies in Π are actually solutions to a statistical problem. If we assume a certain stochastic source for the data, we may estimate the parameters of the model, and there is often a natural set of sufficient statistics associated with it. If we collect all such solutions to stochastic models in a set Π, we may compete with all these strategies as long as Γ is not too large and the dependence of estimators on these sufficient statistics is smooth. With the notation introduced in this paper, we need to study the sequential Rademacher complexity for strategies Π, which can be upper bounded by the complexity of $\bar{\Pi}$ on Γ-valued trees:

$$\mathfrak{R}(\Pi) \le \sup_{\mathbf{g},\mathbf{z}}\mathbb{E}_\epsilon \sup_{\bar{\pi}\in\bar{\Pi}}\left[\sum_{t=1}^{T}\epsilon_t\ell(\bar{\pi}(\mathbf{g}_t(\epsilon)),\mathbf{z}_t(\epsilon)\right]$$

This complexity corresponds to our intuition that with sufficient statistics the dependence on the ever-growing history can be replaced with the dependence on a *summary* of the data. Next, we consider one particular case of this general idea, and refer to [6] for more details on these types of bounds.

4.2 Bernoulli Model with a Beta Prior

Suppose the data $z_t \in \{0, 1\}$ is generated according to Bernoulli distribution with parameter p, and the prior on $p \in [0, 1]$ is $p \sim Beta(\alpha, \beta)$. Given the data $\{z_1, \ldots, z_{t-1}\}$, the maximum a posteriori (MAP) estimator of p is $\hat{p} = (\sum_{i=1}^{t-1} z_i + \alpha - 1)/(t - 1 + \alpha + \beta - 2)$. We now consider the problem of competing with $\Pi = \{\pi^{\alpha,\beta} : \alpha > 1, \beta \in (1, C_\beta]\}$ for some C_β, where each $\pi^{\alpha,\beta}$ predicts the corresponding MAP value for the next round:

$$\pi_t^{\alpha,\beta}(z_1, \ldots, z_{t-1}) = (\textstyle\sum_{i=1}^{t-1} z_i + \alpha - 1)/(t - 1 + \alpha + \beta - 2) .$$

Let us consider the absolute loss, which is equivalent to probability of a mistake of the randomized prediction[1] with bias $\pi_t^{\alpha,\beta}$. Thus, the loss of a strategy $\pi^{\alpha,\beta}$ on round t is $\left|\pi_t^{\alpha,\beta}(z_{1:t-1}) - z_t\right|$. Using Theorem 1 and the argument in (2) to erase the outcome tree, we conclude that there exists a regret minimization algorithm against the set Π which attains regret of at most

$$2 \sup_{\mathbf{w}} \mathbb{E}_\epsilon \sup_{\alpha,\beta} \left[\textstyle\sum_{t=1}^T \epsilon_t \frac{\sum_{i=1}^{t-1} \mathbf{w}_i(\epsilon) + \alpha - 1}{t - 1 + \alpha + \beta - 2} \right] .$$

To analyze the rate exhibited by this upper bound, construct a new tree with $\mathbf{g}_1(\epsilon) = 1$ and $\mathbf{g}_t(\epsilon) = \frac{\sum_{i=1}^{t-1} \mathbf{w}_i(\epsilon) + \alpha - 1}{t + \alpha - 2} \in [0, 1]$ for $t \geq 2$. With this notation, we can simply re-write the last expression as twice

$$\sup_{\mathbf{g}} \mathbb{E}_\epsilon \sup_{\alpha,\beta} \left[\textstyle\sum_{t=1}^T \epsilon_t \mathbf{g}_t(\epsilon) \frac{t + \alpha - 2}{t + \alpha + \beta - 3} \right]$$

The supremum ranges over all $[0, 1]$-valued trees $\mathbf{g}$, but we can pass to the supremum over all $[-1, 1]$-valued trees (thus making the value larger). We then observe that the supremum is achieved at a $\{\pm 1\}$-valued tree $\mathbf{g}$, which can then be erased as in the end of the proof of Theorem 2 (roughly speaking, it amounts to renaming ϵ_t into $\epsilon_t \mathbf{g}_t(\epsilon_{1:t-1})$). We obtain an upper bound

$$\mathfrak{R}(\Pi) \leq \mathbb{E}_\epsilon \sup_{\alpha,\beta} \sum_{t=1}^T \frac{\epsilon_t(t + \alpha - 2)}{t + \alpha + \beta - 3} \leq \mathbb{E}_\epsilon \left| \sum_{t=1}^T \epsilon_t \right| + \mathbb{E}_\epsilon \sup_{\alpha,\beta} \left| \sum_{t=1}^T \frac{\epsilon_t(\beta - 1)}{t + \alpha + \beta - 3} \right| = (\sqrt{C_\beta} + 1)\sqrt{T} \qquad (6)$$

where we used Cauchy-Schwartz inequality for the second term. We note that an experts algorithm would require a discretization that depends on T and will yield a regret bound of order $O(\sqrt{T \log T})$. It is therefore interesting to find an algorithm that avoids the discretization and obtains this regret. To this end, we take the derived upper bound on the sequential Rademacher complexity and prove that it is an admissible relaxation.

Lemma 6. *The relaxation*

$$\mathbf{Rel}(z_{1:t}) = \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^T \epsilon_s \cdot \frac{s + \alpha - 2}{s + \alpha + \beta - 3} - \sum_{s=1}^t \left| \frac{\sum_{i=1}^{s-1} z_i}{s + \alpha + \beta - 3} - z_s \right| \right]$$

is admissible.

---

[1]Alternatively, we can consider strategies that predict according to $\mathbf{1}\{\hat{p} \geq 1/2\}$, which better matches the choice of an absolute loss. However, in this situation, an experts algorithm on an appropriate discretization attains the bound.

Given that this relaxation is admissible, we have a guarantee that the following algorithm attains the rate $(\sqrt{C_\beta}+1)\sqrt{T}$ given in (6):

$$q_t = \arg\min_{q\in[0,1]}\max_{z_t\in\{0,1\}}\left\{\mathbb{E}_{f\sim q}|f-z_t| + \mathbb{E}_{\epsilon_{t+1:T}}\sup_{\alpha,\beta}\left[2\sum_{s=t+1}^{T}\epsilon_s\cdot\frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t}\left|\frac{\sum_{i=1}^{s-1}z_i}{s+\alpha+\beta-3} - z_s\right|\right]\right\}$$

In fact, q_t can be written as

$$q_t = \frac{1}{2}\left\{\mathbb{E}_{\epsilon_{t+1:T}}\sup_{\alpha,\beta}\left[2\sum_{s=t+1}^{T}\epsilon_s\cdot\frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t-1}(1-2z_s)\cdot\frac{\sum_{i=1}^{s-1}z_i}{s+\alpha+\beta-3} + \frac{\sum_{i=1}^{t-1}z_i}{t+\alpha+\beta-3}\right]\right.$$
$$\left. - \mathbb{E}_{\epsilon_{t+1:T}}\sup_{\alpha,\beta}\left[2\sum_{s=t+1}^{T}\epsilon_s\cdot\frac{s+\alpha-2}{s+\alpha+\beta-3} - \sum_{s=1}^{t-1}(1-2z_s)\cdot\frac{\sum_{i=1}^{s-1}z_i}{s+\alpha+\beta-3} - \frac{\sum_{i=1}^{t-1}z_i}{t+\alpha+\beta-3}\right]\right\}$$

For a given realization of random signs, the supremum is an optimization of a sum of linear fractional functions of two variables. Such an optimization can be carried out in time $O(T\log T)$ (see [4]). To deal with the expectation over random signs, one may either average over many realizations or use the random playout idea and only draw one sequence. Such an algorithm is admissible for the above relaxation, obtains the $O(\sqrt{T})$ bound, and runs in $O(T\log T)$ time per step. We leave it as an open problem whether a more efficient algorithm with $O(\sqrt{T})$ regret exists.

5 Competing with Regularized Least Squares

Consider the supervised learning problem with $\mathcal{Y} = [-1,1]$ and some set $\mathcal{X}$. Consider the Regularized Least Squares (RLS) strategies, parametrized by a regularization parameter λ and a shift w_0. That is, given data $(x_1,y_1),\ldots,(x_t,y_t)$, the strategy solves

$$\arg\min_w \sum_{i=1}^{t}(y_i - \langle x_i, w\rangle)^2 + \lambda\|w - w_0\|^2 \ .$$

For a given pair λ and w_0, the solution is

$$w_{t+1}^{\lambda,w_0} = w_0 + (X^\top X + \lambda I)^{-1}X^\top Y,$$

where $X \in \mathbb{R}^{t\times d}$ and $Y \in \mathbb{R}^{t\times 1}$ are the usual matrix representations of the data $x_{1:t}, y_{1:t}$. We would like to compete against a set of such RLS strategies which make prediction $\langle w_{t-1}^{\lambda,w_0}, x_t\rangle$, given side information x_t. Since the outcomes are in $[-1,1]$, without loss of generality we clip the predictions of strategies to this interval, thus making our regret minimization goal only harder. To this end, let $c(a) = a$ if $a \in [-1,1]$ and $c(a) = \text{sign}(a)$ for $|a| > 1$. Thus, given side-information $x_t \in \mathcal{X}$, the prediction of strategies in $\Pi = \left\{\pi^{\lambda,w_0} : \lambda \geq \lambda_{\min} > 0, \|w_0\|_2 \leq 1\right\}$ is simply the clipped product

$$\pi_t^{\lambda,w_0}(x_{1:t}, y_{1:t-1}) = c\left(\left\langle w_{t-1}^{\lambda,w_0}, x_t\right\rangle\right) \ .$$

Let us take the squared loss function $\ell(\hat{y}, y) = (\hat{y} - y)^2$.

Lemma 7. *For the set Π of strategies defined above, the minimax regret of competing against Regularized Least Squares strategies is*

$$\mathcal{V}_T(\Pi) \leq c\sqrt{T\log(T\lambda_{\min}^{-1})}$$

for an absolute constant c.

Observe that $\lambda_{\min}^{-1}$ enters only logarithmically, which allows us to set, for instance, $\lambda_{\min} = 1/T$. Finally, we mention that the set of strategies includes $\lambda = \infty$. This setting corresponds to a static strategy $\pi_t^{\lambda,w_0}(x_{1:t}, y_{1:t-1}) = \langle w_0, x_t \rangle$ and regret against such a static family parametrized by $w_0 \in B_2(1)$ is exactly the objective of online linear regression [14]. Lemma 7 thus shows that it is possible to have vanishing regret with respect to a much larger set of strategies. It is an interesting open question of whether one can develop an efficient algorithm with the above regret guarantee.

6 Competing with Follow the Regularized Leader Strategies

Consider the problem of online linear optimization with the loss function $\ell(f_t, x_t) = \langle f_t, z_t \rangle$ for $f_t \in \mathcal{F}$, $z_t \in \mathcal{Z}$. For simplicity, assume that $\mathcal{F} = \mathcal{Z} = B_2(1)$. An algorithm commonly used for online linear and online convex optimization problems is the Follow the Regularized Leader (FTRL) algorithm. We now consider competing with a family of FTRL algorithms $\pi^{w_0,\lambda}$ indexed by $w_0 \in \{w : \|w\| \leq 1\}$ and $\lambda \in \Lambda$ where Λ is a family of functions $\lambda : \mathbb{R}^+ \times [T] \mapsto \mathbb{R}^+$ specifying a schedule for the choice of regularization parameters. Specifically we consider strategies $\pi^{w_0,\lambda}$ such that $\pi_t^{w_0,\lambda}(z_1, \ldots, z_{t-1}) = w_t$ where

$$w_t = w_0 + \underset{w:\|w\|\leq 1}{\operatorname{argmin}} \left\{ \sum_{i=1}^{t-1} \langle w, z_i \rangle + \tfrac{1}{2}\lambda\left(\left\|\sum_{i=1}^{t-1} z_i\right\|, t\right)\|w\|^2 \right\} \tag{7}$$

This can be written in closed form as $w_t = w_0 - \left(\sum_{i=1}^{t-1} z_i\right)/\max\left\{\lambda\left(\left\|\sum_{i=1}^{t-1} z_i\right\|, t\right), \left\|\sum_{i=1}^{t-1} z_i\right\|\right\}$.

Lemma 8. *For a given class Λ of functions indicating choices of the regularization parameters, define a class Γ of functions on $[0,1] \times [1/T, 1]$ specified by*

$$\Gamma = \left\{ \gamma : \forall b \in [1/T, 1], a \in [0,1], \gamma(a,b) = \min\left\{ \frac{a/(b-1)}{\lambda(a/(b-1), 1/b)}, 1 \right\}, \lambda \in \Lambda \right\}$$

Then the value of the online learning game competing against FTRL strategies given by Equation 7 is bounded as

$$\mathcal{V}_T(\Pi_\Lambda) \leq 4\sqrt{T} + 2\,\mathcal{R}_T(\Gamma)$$

where $\mathcal{R}_T(\Gamma)$ is the sequential Rademacher complexity [11] of Γ.

Notice that if $|\Lambda| < \infty$ then the second term is bounded as $\mathcal{R}_T(\Gamma) \leq \sqrt{T \log|\Lambda|}$. However, we may compete with an infinite set of step-size rules. Indeed, each $\gamma \in \Gamma$ is a function $[0,1]^2 \mapsto [0,1]$. Hence, even if one considers Γ to be the set of all 1-Lipschitz functions (Lipschitz w.r.t., say, ℓ_∞ norm), it holds that $\mathcal{R}_T(\Gamma) \leq 2\sqrt{T \log T}$. We conclude that it is possible to compete with set of FTRL strategies that pick any w_0 in unit ball as starting point and further use for regularization parameter schedule any $\lambda : \mathbb{R}^2 \mapsto \mathbb{R}$ that is such that $\frac{a/(b-1)}{\lambda(a/(b-1),1/b)}$ is a 1-Lipchitz function for every $a, b \in [1/T, 1]$.

Beyond the finite and Lipschitz cases shown above, it would be interesting to analyze richer families of step size schedules, and possibly derive efficient algorithms.

References

[1] S. Ben-David, D. Pal, and S. Shalev-Shwartz. Agnostic online learning. In *Proceedings of the 22th Annual Conference on Learning Theory*, 2009.

[2] N. Cesa-Bianchi and G. Lugosi. On prediction of individual sequences. *The Annals of Statistics*, 27(6):pp. 1865–1895, 1999.

[3] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games.* Cambridge University Press, 2006.

[4] D.Z. Chen, O. Daescu, Y. Dai, N. Katoh, X. Wu, and J. Xu. Efficient algorithms and implementations for optimizing the sum of linear fractional functions, with applications. *Journal of Combinatorial Optimization*, 9(1):69–90, 2005.

[5] M. Feder, N. Merhav, and M. Gutman. Universal prediction of individual sequences. *Information Theory, IEEE Transactions on*, 38(4):1258–1270, 1992.

[6] D. P. Foster, A. Rakhlin, K. Sridharan, and A. Tewari. Complexity-based approach to calibration with checking rules. *Journal of Machine Learning Research - Proceedings Track*, 19:293–314, 2011.

[7] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, Berlin, 2nd edition, 1991.

[8] N. Merhav and M. Feder. Universal prediction. *IEEE Transactions on Information Theory*, 44:2124–2147, 1998.

[9] A. Rakhlin, O. Shamir, and K. Sridharan. Relax and randomize : From value to algorithms. In *Advances in Neural Information Processing Systems*, 2012.

[10] A. Rakhlin and K. Sridharan. http://www-stat.wharton.upenn.edu/~rakhlin/courses/stat928/stat928_notes.pdf, 2012.

[11] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Random averages, combinatorial parameters, and learnability. In *Advances in Neural Information Processing Systems*, 2010.

[12] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. In *NIPS*, pages 1764–1772, 2011.

[13] N. Srebro, K. Sridharan, and A. Tewari. On the universality of online mirror descent. In *NIPS*, pages 2645–2653, 2011.

[14] V. Vovk. Competitive on-line linear regression. In *NIPS '97: Proceedings of the 1997 conference on Advances in neural information processing systems 10*, pages 364–370, Cambridge, MA, USA, 1998. MIT Press.

A Proofs

Proof of Theorem 1. Let us prove a more general version of Theorem 1, which we do not state in the main text due to lack of space. The extra twist is that we allow constraints on the sequences $z_1, \ldots, z_T$ played by the adversary. Specifically, the adversary at round t can only play x_t that satisfy constraint $C_t(z_1, \ldots, z_t) = 1$ where $(C_1, \ldots, C_T)$ is a predetermined sequence of constraints with $C_t : \mathcal{Z}^t \mapsto \{0, 1\}$. When each C_t is the function that is always 1 then we are in the setting of the theorem statement where we play an unconstrained/worst case adversary. However the proof here allows us to even analyze constrained adversaries which come in handy in many cases. Following [12], a *restriction* $\mathcal{P}_{1:T}$ on the adversary is a sequence $\mathcal{P}_1, \ldots, \mathcal{P}_T$ of mappings $\mathcal{P}_t : \mathcal{Z}^{t-1} \mapsto 2^{\mathcal{P}}$ such that $\mathcal{P}_t(z_{1:t-1})$ is a *convex* subset of $\mathcal{P}$ for any $z_{1:t-1} \in \mathcal{Z}^{t-1}$. In the present proof we will only consider *constrained adversaries*, where $\mathcal{P}_t = \Delta(\mathcal{C}_t(z_{1:t-1}))$ the set of all distributions on the constrained subset

$$\mathcal{C}_t(z_{1:t-1}) \triangleq \{z \in \mathcal{Z} : C_t(z_1, \ldots, z_{t-1}, z) = 1\}.$$

defined at time t via a binary constraint $C_t : \mathcal{Z}^t \mapsto \{0, 1\}$. Notice that the set $\mathcal{C}_t(z_{1:t-1})$ is the subset of $\mathcal{Z}$ from which the adversary is allowed to pick instance z_t from given the history so far. It was shown in [12] that such constraints can model sequences with certain properties, such as slowly changing sequences, low-variance sequences, and so on. Let $\mathbf{C}$ be the set of $\mathcal{Z}$-valued trees $\mathbf{z}$ such that for every $\epsilon \in \{\pm 1\}^T$ and $t \in [T]$,

$$C_t(\mathbf{z}_1(\epsilon), \ldots, \mathbf{z}_t(\epsilon)) = 1,$$

that is, the set of trees such that the constraint is satisfied along any path. The statement we now prove is that the value of the prediction problem with respect to a set Π of strategies and against constrained adversaries (denoted by $\mathcal{V}_T(\Pi, \mathcal{C}_{1:T})$) is upper bounded by twice the sequential complexity

$$\sup_{\mathbf{w} \in \mathbf{C}, \mathbf{z}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \sum_{t=1}^T \epsilon_t \ell(\pi_t(\mathbf{w}_1(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon))), \mathbf{z}_t(\epsilon)) \tag{8}$$

where it is crucial that the $\mathbf{w}$ tree ranges over trees that respect the constraints along all paths, while $\mathbf{z}$ is allowed to be an arbitrary $\mathcal{Z}$-valued tree. This fact that $\mathbf{w}$ respects the constraints is the only difference with the original statement of Theorem 1 in the main body of the paper.

For ease of notation we use $\left\langle\!\left\langle \ \right\rangle\!\right\rangle_{t=1}^T$ to denote repeated application of operators such has sup or inf. For instance, $\left\langle\!\left\langle \sup_{a_t \in A} \inf_{b_t \in B} \mathbb{E}_{r_t \sim P} \right\rangle\!\right\rangle_{t=1}^T [F(a_1, b_1, r_1, \ldots, a_T, b_T, r_T)]$ denotes

$$\sup_{a_1 \in A} \inf_{b_1 \in B} \mathbb{E}_{r_1 \sim P} \ldots \sup_{a_T \in A} \inf_{b_T \in B} \mathbb{E}_{r_T \sim P} [F(a_1, b_1, r_1, \ldots, a_T, b_T, r_T)]$$

The value of a prediction problem with respect to a set of strategies and against constrained

adversaries can be written as :

$$\mathcal{V}_T(\Pi, \mathcal{C}_{1:T}) = \left\langle\!\!\!\left\langle \inf_{q_t \in \mathcal{Q}} \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathbb{E}_{f_t \sim q_t, z_t \sim p_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \left[\sum_{t=1}^{T} \ell(f_t, z_t) - \inf_{\pi \in \Pi} \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$= \left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathbb{E}_{z_t \sim p_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \inf_{f_t \in \mathcal{F}} \mathbb{E}_{z_t'} \ell(f_t, z_t') - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathbb{E}_{z_t \sim p_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \mathbb{E}_{z_t'} \ell(\pi_t(z_{1:t-1}), z_t') - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(z_{1:t-1})} \mathbb{E}_{z_t, z_t'} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \ell(\pi_t(z_{1:t-1}), z_t') - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

Let us now define the "selector function" $\chi : \mathcal{Z} \times \mathcal{Z} \times \{\pm 1\} \mapsto \mathcal{Z}$ by

$$\chi(z, z', \epsilon) = \left\{ \begin{array}{ll} z' & \text{if } \epsilon = -1 \\ z & \text{if } \epsilon = 1 \end{array} \right.$$

In other words, χ_t selects between z_t and z_t' depending on the sign of ϵ. We will use the shorthand $\chi_t(\epsilon_t) \triangleq \chi(z_t, z_t', \epsilon_t)$ and $\chi_{1:t}(\epsilon_{1:t}) \triangleq (\chi(z_1, z_1', \epsilon_1), \ldots, \chi(z_t, z_t', \epsilon_t))$. We can then re-write the last statement as

$$\left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))} \mathbb{E}_{z_t, z_t'} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), \chi_t(-\epsilon_t)) - \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), \chi_t(\epsilon_t)) \right) \right]$$

One can indeed verify that we simply used χ_t to switch between z_t and z_t' according to ϵ_t. Now, we can replace the second argument of the loss in both terms by a larger value to obtain an upper bound

$$\left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))} \mathbb{E}_{z_t, z_t'} \sup_{z_t'', z_t'''} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \left(\ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t'') - \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t''') \right) \right]$$

$$\leq 2 \left\langle\!\!\!\left\langle \sup_{p_t \in \mathcal{P}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))} \mathbb{E}_{z_t, z_t'} \sup_{z_t''} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t'') \right]$$

since the two terms obtained by splitting the suprema are the same. We now pass to the suprema over z_t, z_t', noting that the constraints need to hold:

$$2 \left\langle\!\!\!\left\langle \sup_{z_t, z_t' \in \mathcal{C}_t(\chi_{1:t-1}(\epsilon_{1:t-1}))} \sup_{z_t''} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\!\right\rangle_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\chi_{1:t-1}(\epsilon_{1:t-1})), z_t'') \right]$$

$$= 2 \sup_{(\mathbf{z}, \mathbf{z}') \in \mathcal{C}'} \sup_{\mathbf{z}''} \mathbb{E}_{\epsilon} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \pi_t(\chi(\mathbf{z}_1, \mathbf{z}_1', \epsilon_1), \ldots, \chi(\mathbf{z}_{t-1}(\epsilon), \mathbf{z}_{t-1}'(\epsilon), \epsilon_{t-1})), \mathbf{z}''(\epsilon) \right] = (*)$$

where in the last step we passed to the tree notation. Importantly, the pair $(\mathbf{z}, \mathbf{z}')$ of trees does not range over all pairs, but only over those which satisfy the constraints:

$$\mathcal{C}' = \left\{ (\mathbf{z}, \mathbf{z}') : \forall \epsilon \in \{\pm 1\}^T, \forall t \in [T], \quad \mathbf{z}_t(\epsilon), \mathbf{z}_t'(\epsilon) \in \mathcal{C}_t(\chi(\mathbf{z}_1, \mathbf{z}_1', \epsilon_1), \ldots, \chi(\mathbf{z}_{t-1}(\epsilon), \mathbf{z}_{t-1}'(\epsilon), \epsilon_{t-1})) \right\}$$

15

Now, given the pair $(\mathbf{z}, \mathbf{z}') \in \mathcal{C}'$, define a $\mathcal{Z}$-valued tree of depth T as:

$$\tilde{\mathbf{w}}_1 = \varnothing, \qquad \tilde{\mathbf{w}}_t(\epsilon) = \chi(\mathbf{z}_{t-1}(\epsilon), \mathbf{z}'_{t-1}(\epsilon), \epsilon_{t-1}) \ \text{ for all } \ t > 1$$

Clearly, this is a well-defined tree, and we now claim that it satisfies the constraints along every path. Indeed, we need to check that for any ϵ and t, both $\tilde{\mathbf{w}}_t(\epsilon_{1:t-2}, +1), \tilde{\mathbf{w}}_t(\epsilon_{1:t-2}, -1) \in \mathcal{C}_t(\tilde{\mathbf{w}}_1, \ldots, \tilde{\mathbf{w}}_{t-1}(\epsilon_{1:t-2}))$. This amounts to checking, by definition of $\tilde{\mathbf{w}}$ and the selector χ, that

$$\mathbf{z}_{t-1}(\epsilon_{1:t-2}), \mathbf{z}'_{t-1}(\epsilon_{1:t-2}) \in \mathcal{C}_{t-1}(\chi(\mathbf{z}_1, \mathbf{z}'_1, \epsilon_1), \ldots, \chi(\mathbf{z}_{t-2}(\epsilon), \mathbf{z}'_{t-2}(\epsilon), \epsilon_{t-2})) \ .$$

But this is true because $(\mathbf{z}, \mathbf{z}') \in \mathcal{C}'$. Hence, $\tilde{\mathbf{w}}$ constructed from $\mathbf{z}, \mathbf{z}'$ satisfies the constraints along every path.

We can therefore upper bound the expression in $(*)$ by twice

$$\sup_{\tilde{\mathbf{w}} \in \mathbf{C}} \sup_{\mathbf{z}''} \mathbb{E} \sup_{\epsilon} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\tilde{\mathbf{w}}_1(\epsilon), \ldots, \tilde{\mathbf{w}}_{t-1}(\epsilon)), \mathbf{z}''(\epsilon)) \right].$$

Define $\mathbf{w}^* = \tilde{\mathbf{w}}(-1)$ and $\mathbf{w}^{**} = \tilde{\mathbf{w}}(+1)$, we can expend the expectation with respect to ϵ_1 of the above expression by

$$\frac{1}{2} \sup_{\mathbf{w}^* \in \mathbf{C}} \sup_{\mathbf{z}''} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi \in \Pi} \left[-\ell(\pi_1(\cdot), \mathbf{z}''_1(\cdot)) + \sum_{t=2}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}^*(\epsilon)), \mathbf{z}''(\epsilon)) \right]$$

$$+ \frac{1}{2} \sup_{\mathbf{w}^{**} \in \mathbf{C}} \sup_{\mathbf{z}''} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi \in \Pi} \left[\ell(\pi_1(\cdot), \mathbf{z}''_1(\cdot)) + \sum_{t=2}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}^{**}(\epsilon)), \mathbf{z}''(\epsilon)) \right].$$

With the assumption that we do not suffer lose at the first round, which means $\ell(\pi_1(\cdot), \mathbf{z}''_1(\cdot)) = 0$, we can see that both terms achieve the suprema with the same $\mathbf{w}^* = \mathbf{w}^{**}$. Therefore, the above expression can be rewrite as

$$\sup_{\mathbf{w} \in \mathbf{C}} \sup_{\mathbf{z}''} \mathbb{E}_{\epsilon_{2:T}} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}(\epsilon)), \mathbf{z}''(\epsilon)) \right]$$

which is precisely (8). This concludes the proof of Theorem 1.

$\square$

Proof of Theorem 2. By convexity of the loss,

$$\left\| \left\langle \sup_{x_t \in \mathcal{X}} \inf_{q_t \in \Delta(\mathcal{Y})} \sup_{y_t \in \mathcal{Y}} \mathbb{E}_{\hat{y}_t \sim q_t} \right\rangle \right\|_{t=1}^{T} \left[\sum_{t=1}^{T} \ell(\hat{y}_t, y_t) - \inf_{\pi \in \Pi} \sum_{t=1}^{T} \ell(\pi_t(x_{1:t}, y_{1:t-1}), y_t) \right]$$

$$\leq \left\| \left\langle \sup_{x_t \in \mathcal{X}} \inf_{q_t \in \Delta(\mathcal{Y})} \sup_{y_t \in \mathcal{Y}} \mathbb{E}_{\hat{y}_t \sim q_t} \right\rangle \right\|_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} \ell'(\hat{y}_t, y_t)(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$\leq \left\| \left\langle \sup_{x_t \in \mathcal{X}} \inf_{q_t \in \Delta(\mathcal{Y})} \sup_{y_t \in \mathcal{Y}} \mathbb{E}_{\hat{y}_t \sim q_t} \sup_{s_t \in [-L, L]} \right\rangle \right\|_{t=1}^{T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^{T} s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

where in the last step we passed to an upper bound by allowing for the worst-case choice s_t of the derivative. We will often omit the range of the variables in our notation, and it is understood that

16

s_t's range over $[-L, L]$, while $y_t, \hat{y}_t$ over $\mathcal{Y}$ and x_t's over $\mathcal{X}$. Now, by Jensen's inequality, we pass to an upper bound by exchanging $\mathbb{E}_{\hat{y}_t}$ and $\sup_{y_t \in \mathcal{Y}}$:

$$\left\langle\!\!\left\langle \sup_{x_t} \inf_{q_t \in \Delta(\mathcal{Y})} \mathbb{E}_{\hat{y}_t \sim q_t} \sup_{y_t} \sup_{s_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \left\langle\!\!\left\langle \sup_{x_t} \inf_{\hat{y}_t \in \mathcal{Y}} \sup_{y_t, s_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

Consider the last step, assuming all the other variables fixed:

$$\sup_{x_T} \inf_{\hat{y}_T} \sup_{y_T, s_T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^T s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \sup_{x_T} \inf_{\hat{y}_T} \sup_{p_T \in \Delta(\mathcal{Y} \times [-L,L])} \mathbb{E}_{(y_T, s_T) \sim p_T} \sup_{\pi \in \Pi} \left[\sum_{t=1}^T s_t(\hat{y}_t - \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

where the distribution p_T ranges over all distributions on $\mathcal{Y} \times [-L, L]$. Now observe that the function inside the infimum is convex in $\hat{y}_T$, and the function inside $\sup_{p_T}$ is linear in the distribution p_T. Hence, we can appeal to the minimax theorem, obtaining equality of the last expression to

$$\sup_{x_T} \sup_{p_T \in \Delta(\mathcal{Y} \times [-L,L])} \inf_{\hat{y}_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[\sum_{t=1}^T s_t \hat{y}_t - \inf_{\pi \in \Pi} \sum_{t=1}^T s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \inf_{\hat{y}_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[s_T \hat{y}_T - \inf_{\pi \in \Pi} \sum_{t=1}^T s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \mathbb{E}_{(y_T, s_T) \sim p_T} \inf_{\pi \in \Pi} \sum_{t=1}^T s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T} \sup_{p_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \inf_{\pi \in \Pi} \sum_{t=1}^T s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

We can now upper bound the choice of $\hat{y}_T$ by that given by π_T, yielding an upper bound

$$\sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T, p_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \sup_{\pi \in \Pi} \left[\inf_{\hat{y}_T} \left(\mathbb{E}_{(y_T, s_T) \sim p_T} s_T \right) \hat{y}_T - \sum_{t=1}^T s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

$$= \sum_{t=1}^{T-1} s_t \hat{y}_t + \sup_{x_T, p_T} \mathbb{E}_{(y_T, s_T) \sim p_T} \sup_{\pi \in \Pi} \left[\left(\mathbb{E}_{(y_T', s_T') \sim p_T} s_T' - s_T \right) \pi_T(x_{1:T}, y_{1:T-1}) - \sum_{t=1}^{T-1} s_t \pi_t(x_{1:t}, y_{1:t-1})) \right]$$

It is not difficult to verify that this process can be repeated for $T - 1$ and so on. The resulting

upper bound is therefore

$$\mathcal{V}_T^S(\Pi) \leq \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathbb{E}_{(y_t, s_t) \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \left(\mathbb{E}_{(y_t', s_t') \sim p_t} s_t' - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

$$\leq \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathbb{E}_{\substack{(y_t, s_t) \sim p_t \\ (y_t', s_t') \sim p_t}} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \left(s_t' - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

$$= \left\langle\!\!\left\langle \sup_{x_t, p_t} \mathbb{E}_{\substack{(y_t, s_t) \sim p_t \\ (y_t', s_t') \sim p_t}} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(s_t' - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

$$\leq \left\langle\!\!\left\langle \sup_{x_t} \sup_{\substack{(y_t, s_t) \\ (y_t', s_t')}} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(s_t' - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

$$\leq \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s_t', s_t} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \left(s_t' - s_t \right) \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

$$\leq 2 \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s_t \in [-L, L]} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t s_t \pi_t(x_{1:t}, y_{1:t-1}) \right]$$

Since the expression is convex in each s_t, we can replace the range of s_t by $\{-L, L\}$, or, equivalently,

$$\mathcal{V}_T^S(\Pi) \leq 2L \left\langle\!\!\left\langle \sup_{x_t, y_t} \sup_{s_t \in \{-1,1\}} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t s_t \pi_t(x_{1:t}, y_{1:t-1}) \right] \tag{9}$$

Now consider any arbitrary function $\psi : \{\pm 1\} \mapsto \mathbb{R}$, we have that

$$\sup_{s \in \{\pm 1\}} \mathbb{E}_\epsilon \left[\psi(s \cdot \epsilon) \right] = \sup_{s \in \{\pm 1\}} \frac{1}{2} \left(\psi(+s) + \psi(-s) \right) = \frac{1}{2} \left(\psi(+1) + \psi(-1) \right) = \mathbb{E}_\epsilon \left[\psi(\epsilon) \right]$$

Since in Equation (9), for each t, s_t and ϵ_t appear together as $\epsilon_t \cdot s_t$ using the above equation repeatedly, we conclude that

$$\mathcal{V}_T^S(\Pi) \leq 2L \left\langle\!\!\left\langle \sup_{x_t, y_t} \mathbb{E}_{\epsilon_t} \right\rangle\!\!\right\rangle_{t=1}^T \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(x_{1:t}, y_{1:t-1}) \right] = \sup_{\mathbf{x}, \mathbf{y}} \mathbb{E}_\epsilon \sup_{\pi \in \Pi} \left[\sum_{t=1}^T \epsilon_t \pi_t(\mathbf{x}_{1:t}(\epsilon), \mathbf{y}_{1:t-1}(\epsilon)) \right]$$

The lower bound is obtained by the same argument as in [11].

$\square$

Proof of Theorem 3. Denote $L_t(\pi) = \sum_{s=1}^t \ell(\pi_s(z_{1:s-1}), z_s)$. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \left\{ \mathbb{E}_{f_t \sim q_t} \left[\ell(f_t, z_t) \right] + \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^T \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \right\}$$

$$= \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \left\{ \mathbb{E}_{z_t \sim p_t} \left[\ell(f_t, z_t) \right] + \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^T \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi) \right] \right\}$$

18

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$\mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$
$$\left. + \inf_{f_t \in \mathcal{F}} \mathbb{E}_{z_t \sim p_t} [\ell(f_t, z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$
$$\left. + \mathbb{E}_{z_t \sim p_t} [\ell(\pi_t(z_{1:t-1}), z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \mathbb{E}_{z_t, z_t' \sim p_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$
$$\left. + \ell(\pi_t(z_{1:t-1}), z_t') - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

We now argue that the independent z_t and z_t' have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$\mathbb{E}_{z_t, z_t' \sim p_t} \mathbb{E}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$
$$\left. + \epsilon_t(\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \right]$$

$$\leq \mathbb{E}_{z_t, z_t' \sim p_t} \sup_{z'', z'''} \mathbb{E}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$
$$\left. + \epsilon_t(\ell(\pi_t(z_{1:t-1}), z_t'') - \ell(\pi_t(z_{1:t-1}), z_t''')) \right]$$

Splitting the resulting expression into two parts, we arrive at the upper bound of

$$2 \mathbb{E}_{z_t, z_t' \sim p_t} \sup_{z''} \mathbb{E}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[\sum_{s=t+1}^{T} \epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - \frac{1}{2} L_{t-1}(\pi) + \epsilon_t \ell(\pi_t(z_{1:t-1}), z_t'') \right]$$

$$\leq \sup_{z, z', z''} \mathbb{E}_{\epsilon_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\pi \in \Pi} \left[\sum_{s=t+1}^{T} 2\epsilon_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) + \epsilon_t \ell(\pi_t(z_{1:t-1}), z_t'') \right]$$

$$\leq \mathfrak{R}_T(\Pi | z_1, \ldots, z_{t-1}).$$

The first inequality is true as we upper bounded the expectation by the supremum. The last inequality is easy to verify, as we are effectively filling in a root z_t and z_t' for the two subtrees, for $\epsilon_t = +1$ and $\epsilon_t = -1$, respectively, and jointing the two trees with a $\varnothing$ root.

One can see that the proof of admissibility corresponds to one step minimax swap and symmetrization in the proof of [11]. In contrast, in the latter paper, all T minimax swaps are performed at once, followed by T symmetrization steps. $\qquad\square$

Proof of Lemma 4. The first step of the proof is an application of the minimax theorem (we

assume the necessary conditions hold):

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \left\{ \underset{f_t \sim q_t}{\mathbb{E}} f_t \cdot z_t + \underset{\epsilon_{t+1:T}}{\mathbb{E}} \max_{1 \leq s \leq T} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right| \right\} = \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \left\{ f_t \cdot \underset{z_t \sim p_t}{\mathbb{E}} z_t + \underset{z_t \sim p_t}{\mathbb{E}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \max_{1 \leq s \leq T} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right| \right\}$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$- \left| \underset{z_t \sim p_t}{\mathbb{E}} z_t \right| + \underset{z_t \sim p_t}{\mathbb{E}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right| \right\}$$

$$\leq \underset{z_t \sim p_t}{\mathbb{E}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) + \underset{z_t' \sim p_t}{\mathbb{E}} z_t' \right| \right\}$$

$$\leq \underset{z_t, z_t' \sim p_t}{\mathbb{E}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i \geq s, i \neq t} a_i^t(\epsilon) + (z_t' - z_t) \right| \right\}$$

We now argue that the independent z_t and z_t' have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$\underset{z_t, z_t' \sim p_t}{\mathbb{E}} \underset{\epsilon_{t:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i \geq s, i \neq t} a_i^t(\epsilon) + \epsilon_t(z_t' - z_t) \right| \right\}$$

$$\leq \underset{z_t \sim p_t}{\mathbb{E}} \underset{\epsilon_{t:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i \geq s, i \neq t} a_i^t(\epsilon) + 2\epsilon_t z_t \right| \right\}$$

Now, the supremum over p_t is achieved at a delta distribution, yielding an upper bound

$$\sup_{z_t \in [-1,1]} \underset{\epsilon_{t:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i \geq s, i \neq t} a_i^t(\epsilon) + 2\epsilon_t z_t \right| \right\}$$

$$\leq \underset{\epsilon_{t:T}}{\mathbb{E}} \max \left\{ \max_{s > t} \left| \sum_{i=s}^{T} a_i^t(\epsilon) \right|, \max_{s \leq t} \left| \sum_{i \geq s, i \neq t} a_i^t(\epsilon) + 2\epsilon_t \right| \right\}$$

$$= \underset{\epsilon_{t:T}}{\mathbb{E}} \max_{1 \leq s \leq T} \left| \sum_{i=s}^{T} a_i^{t-1}(\epsilon) \right|$$

$\square$

Proof of Lemma 6. Denote

$$L_t(\alpha, \beta) = \sum_{s=1}^{t} \left| \frac{\frac{\sum_{i=1}^{s-1} z_i}{s + \alpha - 2}}{1 + \frac{\beta - 1}{s + \alpha - 2}} - z_s \right|.$$

The first step of the proof is an application of the minimax theorem:

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \left\{ \underset{f_t \sim q_t}{\mathbb{E}} |f_t - z_t| + \underset{\epsilon_{t+1:T}}{\mathbb{E}} \sup_{\alpha, \beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_t(\alpha, \beta) \right] \right\}$$

$$= \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \left\{ \underset{z_t \sim p_t}{\mathbb{E}} |f_t - z_t| + \underset{z_t \sim p_t}{\mathbb{E}} \underset{\epsilon_{t+1:T}}{\mathbb{E}} \sup_{\alpha, \beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_t(\alpha, \beta) \right] \right\}$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$\mathop{\mathbb{E}}_{z_t \sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \inf_{f_t \in \mathcal{F}} \mathop{\mathbb{E}}_{z_t \sim p_t} |f_t - z_t| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right]$$

$$\leq \mathop{\mathbb{E}}_{z_t \sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \mathop{\mathbb{E}}_{z'_t \sim p_t} \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right]$$

$$\leq \mathop{\mathbb{E}}_{z_t, z'_t \sim p_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right]$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$\mathop{\mathbb{E}}_{z_t, z'_t \sim p_t} \mathbb{E}_{\epsilon_t} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \epsilon_t \left(\left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right) \right]$$

$$\leq \sup_{z_t, z'_t \in \mathcal{Z}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[2 \sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - L_{t-1}(\alpha,\beta) + \epsilon_t \left(\left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z'_t \right| - \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right) \right]$$

where we upper bounded the expectation by the supremum. Splitting the resulting expression into two parts, we arrive at the upper bound of

$$2 \sup_{z_t \in \mathcal{Z}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \left| \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} - z_t \right| \right]$$

$$= 2 \sup_{z_t \in \mathcal{Z}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} (1 - 2z_t) - \epsilon_t z_t \right]$$

$$= 2 \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{\frac{\sum_{i=1}^{t-1} z_i}{t+\alpha-2}}{1 + \frac{\beta-1}{t+\alpha-2}} \right]$$

where the last step is due to the fact that for any $z_t \in \{0,1\}$, $\epsilon_t(1 - 2z_t)$ has the same distribution as ϵ_t. We then proceed to upper bound

$$2 \sup_p \mathbb{E}_{a \sim p} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{a}{1 + \frac{\beta-1}{t+\alpha-2}} \right]$$

$$\leq 2 \sup_{a \in \{\pm 1\}} \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t+1}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) + \epsilon_t \cdot \frac{a}{1 + \frac{\beta-1}{t+\alpha-2}} \right]$$

$$\leq 2 \mathbb{E}_{\epsilon_{t:T}} \sup_{\alpha,\beta} \left[\sum_{s=t}^{T} \epsilon_s \cdot \frac{1}{1 + \frac{\beta-1}{s+\alpha-2}} - \frac{1}{2} L_{t-1}(\alpha,\beta) \right]$$

The initial condition is trivially satisfied as

$$\mathbf{Rel}(z_{1:T}) = -\inf_{\alpha,\beta} \sum_{s=1}^{T} \left| \frac{\frac{\sum_{i=1}^{s-1} z_i}{s+\alpha-2}}{1 + \frac{\beta-1}{s+\alpha-2}} - z_s \right|$$

21

$\square$

Theorem 9. *The conditional sequential Rademacher complexity with respect to* Π

$$\mathcal{G}_T(\ell, \Pi | z_1, \ldots, z_t) \triangleq \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\sqrt{2\pi} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - \sum_{s=1}^{t} \ell(\pi_s(z_{1:s-1}), z_s) \right]$$

is admissible.

Proof of Theorem 9. Denote $L_t(\pi) = \sum_{s=1}^{t} \ell(\pi_s(z_{1:s-1}), z_s)$. Let $c = \mathbb{E}_\sigma |\sigma| = \sqrt{2/\pi}$. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{z_t \in \mathcal{Z}} \left\{ \mathbb{E}_{f_t \sim q_t} [\ell(f_t, z_t)] + \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi)) \right] \right\}$$

$$= \sup_{p_t \in \Delta(\mathcal{Z})} \inf_{f_t \in \mathcal{F}} \left\{ \mathbb{E}_{z_t \sim p_t} [\ell(f_t, z_t)] + \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_t(\pi)) \right] \right\}$$

For any $p_t \in \Delta(\mathcal{Z})$, the infimum over f_t of the above expression is equal to

$$\mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$

$$\left. + \inf_{f_t \in \mathcal{F}} \mathbb{E}_{z_t \sim p_t} [\ell(f_t, z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \mathbb{E}_{z_t \sim p_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$

$$\left. + \mathbb{E}_{z_t \sim p_t} [\ell(\pi_t(z_{1:t-1}), z_t)] - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

$$\leq \mathbb{E}_{z_t, z'_t \sim p_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t}, \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$

$$\left. + \ell(\pi_t(z_{1:t-1}), z'_t) - \ell(\pi_t(z_{1:t-1}), z_t) \right]$$

We now argue that the independent z_t and z'_t have the same distribution p_t, and thus we can introduce a gaussian random variable σ_t and a random sign $\epsilon_t = \text{sign}(\sigma_t)$. The above expression then equals to

$$\mathbb{E}_{z_t, z'_t \sim p_t} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$

$$\left. + \epsilon_t (\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \right]$$

$$\leq \mathbb{E}_{z_t, z'_t \sim p_t} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z}, \mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi \in \Pi} \left[\frac{2}{c} \sum_{s=t+1}^{T} \sigma_s \ell(\pi_s((z_{1:t-1}, \chi_t(\epsilon_t), \mathbf{w}_{1:s-t-1}(\epsilon)), \mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.$$

$$\left. + \epsilon_t \mathbb{E}_{\sigma_t} \left| \frac{\sigma_t}{c} \right| (\ell(\pi_t(z_{1:t-1}), \chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}), \chi_t(\epsilon_t))) \right]$$

22

Put the expectation outside and use the fact $\epsilon_t|\sigma_t| = \sigma_t$, we get

$$
\mathbb{E}_{z_t,z_t'\sim p_t} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi\in\Pi} \left[\frac{2}{c} \sum_{s=t+1}^T \sigma_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.
$$

$$
\left. + \frac{\sigma_t}{c}(\ell(\pi_t(z_{1:t-1}),\chi_t(-\epsilon_t))) - \ell(\pi_t(z_{1:t-1}),\chi_t(\epsilon_t))) \right]
$$

$$
\leq \mathbb{E}_{z_t,z_t'\sim p_t} \sup_{z'',z'''} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi\in\Pi} \left[\frac{2}{c} \sum_{s=t+1}^T \epsilon_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.
$$

$$
\left. + \frac{\sigma_t}{c}(\ell(\pi_t(z_{1:t-1}),z_t'') - \ell(\pi_t(z_{1:t-1}),z_t''')) \right]
$$

Splitting the resulting expression into two parts, we arrive at the upper bound of

$$
2 \mathbb{E}_{z_t,z_t'\sim p_t} \sup_{z''} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi\in\Pi} \left[\frac{1}{c} \sum_{s=t+1}^T \sigma_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - \frac{1}{2}L_{t-1}(\pi) \right.
$$

$$
\left. + \frac{\sigma_t}{c}\ell(\pi_t(z_{1:t-1}),z_t'') \right]
$$

$$
\leq \sup_{z,z',z''} \mathbb{E}_{\sigma_t} \sup_{\mathbf{z},\mathbf{w}} \mathbb{E}_{\sigma_{t+1:T}} \sup_{\pi\in\Pi} \left[\frac{2}{c} \sum_{s=t+1}^T \sigma_s \ell(\pi_s((z_{1:t-1},\chi_t(\epsilon_t),\mathbf{w}_{1:s-t-1}(\epsilon)),\mathbf{z}_{s-t}(\epsilon)) - L_{t-1}(\pi) \right.
$$

$$
\left. + \frac{2\sigma_t}{c}\ell(\pi_t(z_{1:t-1}),z_t'') \right]
$$

$$
\leq \mathcal{G}_T(\ell,\Pi|z_1,\dots,z_{t-1}).
$$

$\square$

Proof of Lemma 5. Let q_t be the randomized strategy where we draw $\epsilon_{t+1},\dots,\epsilon_T$ uniformly at random and pick

$$
q_t(\epsilon) = \operatorname*{argmin}_{q\in[-1,1]} \sup_{z_t\in\{-1,1\}} \left\{ \mathbb{E}_{f_t\sim q} f_t\cdot z_t + \max_{1\leq s\leq T} \left|\sum_{i=s}^T a_i^t(\epsilon)\right| \right\} \tag{10}
$$

Then,

$$
\sup_{z_t\in\{-1,1\}} \left\{ \mathbb{E}_{f_t\sim q_t} f_t\cdot z_t + \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\leq s\leq T} \left|\sum_{i=s}^T a_i^t(\epsilon)\right| \right\}
$$

$$
= \sup_{z_t\in\{-1,1\}} \left\{ \mathbb{E}_{\epsilon_{t+1:T}} \mathbb{E}_{f_t\sim q_t(\epsilon)} f_t\cdot z_t + \mathbb{E}_{\epsilon_{t+1:T}} \max_{1\leq s\leq T} \left|\sum_{i=s}^T a_i^t(\epsilon)\right| \right\}
$$

$$
\leq \mathbb{E}_{\epsilon_{t+1:T}} \left[\sup_{z_t} \left\{ \mathbb{E}_{f_t\sim q_t(\epsilon)} f_t\cdot z_t + \max_{1\leq s\leq T} \left|\sum_{i=s}^T a_i^t(\epsilon)\right| \right\} \right]
$$

$$
= \mathbb{E}_{\epsilon_{t+1:T}} \left[\inf_{q_t\in\Delta(\mathcal{F})} \sup_{z_t} \left\{ \mathbb{E}_{f_t\sim q_t} f_t\cdot z_t + \max_{1\leq s\leq T} \left|\sum_{i=s}^T a_i^t(\epsilon)\right| \right\} \right]
$$

where the last step is due to the way we pick our predictor $f_t(\epsilon)$ given random draw of ϵ's in Equation (10). We now apply the minimax theorem, yielding the following upper bound on the term above:

$$\mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\inf_{f_t}\left\{\mathbb{E}_{z_t\sim p_t}f_t\cdot z_t + \mathbb{E}_{z_t\sim p_t}\max_{1\leq s\leq T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

This expression can be re-written as

$$\mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\inf_{f_t}\left\{\mathbb{E}_{z_t'\sim p_t}f_t\cdot z_t' + \max_{1\leq s\leq T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

$$\leq \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\left\{-\left|\mathbb{E}_{z_t'\sim p_t}z_t'\right| + \max_{1\leq s\leq T}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

$$\leq \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t\sim p_t}\max\left\{\max_{s\leq t}\left|\sum_{i=s}^T a_i^t(\epsilon) + \mathbb{E}_{z_t'\sim p_t}z_t'\right|, \max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

$$\leq \mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t,z_t'\sim p_t}\max\left\{\max_{s\leq t}\left|\sum_{i\geq s, i\neq t}^T a_i^t(\epsilon) + (z_t'-z_t)\right|, \max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

We now argue that the independent z_t and z_t' have the same distribution p_t, and thus we can introduce a random sign ϵ_t. The above expression then equals to

$$\mathbb{E}_{\epsilon_{t+1:T}}\left[\sup_{p_t\in\Delta(\mathcal{Z})}\mathbb{E}_{z_t,z_t'\sim p_t}\mathbb{E}_{\epsilon_t}\max\left\{\max_{s\leq t}\left|\sum_{i\geq s, i\neq t}^T a_i^t(\epsilon) + \epsilon_t(z_t'-z_t)\right|, \max_{s>t}\left|\sum_{i=s}^T a_i^t(\epsilon)\right|\right\}\right]$$

$$\leq \mathbb{E}_{\epsilon_{t+1:T}}\sup_{z_t\in\{-1,1\}}\mathbb{E}_{\epsilon_t}\max_{1\leq s\leq T}\left|\sum_{i=s}^T a_i^{t-1}(\epsilon)\right| = \mathbb{E}_{\epsilon_{t:T}}\max_{1\leq s\leq T}\left|\sum_{i=s}^T a_i^{t-1}(\epsilon)\right|$$

$\square$

Proof of Lemma 7. Given an $\mathcal{X}$-valued tree $\mathbf{x}$ and a $\mathcal{Y}$-valued tree $\mathbf{y}$, let us write $\mathbf{X}_t(\epsilon)$ for the matrix consisting of $(\mathbf{x}_1(\epsilon),\ldots,\mathbf{x}_{t-1}(\epsilon))$ and $\mathbf{Y}_t$ for the vector $(\mathbf{y}_1(\epsilon),\ldots,\mathbf{y}_{t-1}(\epsilon))$. By Theorem 2, the minimax regret is bounded by

$$4\sup_{\mathbf{x},\mathbf{y}}\mathbb{E}_\epsilon\sup_{\pi^{\lambda,w_0}\in\Pi}\left[\sum_{t=1}^T \epsilon_t\pi_t^{\lambda,w_0}(\mathbf{x}_{1:t}(\epsilon),\mathbf{y}_{1:t-1}(\epsilon))\right]$$

$$= 4\sup_{\mathbf{x},\mathbf{y}}\mathbb{E}_\epsilon\sup_{\lambda,w_0}\left[\sum_{t=1}^T \epsilon_t c\left(\left\langle(\mathbf{X}_t(\epsilon)^\mathsf{T}\mathbf{X}_t(\epsilon)+\lambda I)^{-1}\mathbf{X}_t(\epsilon)^\mathsf{T}\mathbf{Y}_t(\epsilon),\mathbf{x}_t(\epsilon)\right\rangle + \left\langle w_0,\mathbf{x}_t(\epsilon)\right\rangle\right)\right]$$

Since the output of the clipped strategies in Π is between -1 and 1, the Dudley integral gives an upper bound

$$\mathfrak{R}(\Pi,(\mathbf{x},\mathbf{y})) \leq \inf_{\alpha\geq 0}\left\{4\alpha T + 12\sqrt{T}\int_\alpha^1\sqrt{\log\mathcal{N}_2(\Pi,(\mathbf{x},\mathbf{y}),\delta)}\,d\delta\right\}$$

Define the set of strategies before clipping:

$$\Pi' = \left\{ \pi' : \pi'_t(x_{1:t}, y_{1:t-1}) = \left\langle w_0 + (X^\top X + \lambda I)^{-1} X^\top Y, x_t \right\rangle, \|w_0\| \leq 1, \ \lambda > \lambda_{\min} \right\}$$

If V is a δ-cover of Π' on $(\mathbf{x}, \mathbf{y})$, then V is also an δ-cover of Π as $|c(x) - c(x')| \leq |x - y|$. Therefore, for any $(\mathbf{x}, \mathbf{y})$,

$$\mathcal{N}_2(\Pi, (\mathbf{x}, \mathbf{y}), \delta) \leq \mathcal{N}_2(\Pi', (\mathbf{x}, \mathbf{y}), \delta)$$

and

$$\mathfrak{R}(\Pi, (\mathbf{x}, \mathbf{y})) \leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_\alpha^1 \sqrt{\log \mathcal{N}_2(\Pi', (\mathbf{x}, \mathbf{y}), \delta)} \, d\delta \right\}.$$

If W is a $\delta/2$-cover of the set of strategies $\Pi_{w_0} = \{\langle w_0, \mathbf{x}_t(\epsilon)\rangle : \|w_0\| \leq 1\}$ on a tree $\mathbf{x}$, and Λ is a $\delta/2$-cover of the set of strategies

$$\Pi_\lambda = \left\{ \pi : \pi_t(x_{1:t}, y_{1:t-1}) = \left\langle (X^\top X + \lambda I)^{-1} X^\top Y, x_t \right\rangle : \lambda > \lambda_{\min} \right\}$$

then $W \times \Lambda$ is an δ-cover of Π'. Therefore,

$$\mathcal{N}_2(\Pi', (\mathbf{x}, \mathbf{y}), \delta) \leq \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x}, \mathbf{y}), \delta/2) \times \mathcal{N}_2(\Pi_\lambda, (\mathbf{x}, \mathbf{y}), \delta/2).$$

Hence,

$$\mathfrak{R}(\Pi, (\mathbf{x}, \mathbf{y})) \leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_\alpha^1 \sqrt{\log \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x}, \mathbf{y}), \delta/2) + \log \mathcal{N}_2(\Pi_\lambda, (\mathbf{x}, \mathbf{y}), \delta/2)} \, d\delta \right\}$$

$$\leq \inf_{\alpha \geq 0} \left\{ 4\alpha T + 12\sqrt{T} \int_\alpha^1 \sqrt{\log \mathcal{N}_2(\Pi_{w_0}, (\mathbf{x}, \mathbf{y}), \delta/2)} \, d\delta \right\}$$

$$+ 12\sqrt{T} \int_0^1 \sqrt{\log \mathcal{N}_2(\Pi_\lambda, (\mathbf{x}, \mathbf{y}), \delta/2)} \, d\delta$$

The first term is the Dudley integral of the set of static strategies Π_{w_0} given by $w_0 \in B_2(1)$, and it is exactly the complexity studied in [11] where it is shown to be $O(\sqrt{T \log(T)})$. We now provide a bound on the covering number for the second term. It is easy to verify that the following identity holds

$$(X^\top X + \lambda_2 I_d)^{-1} - (X^\top X + \lambda_1 I_d)^{-1} = (\lambda_1 - \lambda_2)(X^\top X + \lambda_1 I_d)^{-1}(X^\top X + \lambda_2 I_d)^{-1}$$

by right- and left-multiplying both sides by $(X^\top X + \lambda_2 I_d)$ and $(X^\top X + \lambda_1 I_d)$, respectively. Let $\lambda_1, \lambda_2 > 0$. Then, assuming that $\|x_t\|_2 \leq 1$ and $y_t \in [-1, 1]$ for all t,

$$\left\| (X_t X + \lambda_2 I_d)^{-1} X^\top Y - (X^\top X + \lambda_1 I_d)^{-1} X^\top Y \right\|_2$$

$$= |\lambda_2 - \lambda_1| \left\| (X^\top X + \lambda_1 I_d)^{-1}(X^\top X + \lambda_2 I_d)^{-1} X^\top Y \right\|_2 \leq |\lambda_2 - \lambda_1| \frac{1}{\lambda_1 \lambda_2} \left\| X^\top Y \right\|_2 \leq \left| \lambda_1^{-1} - \lambda_2^{-1} \right| t$$

Hence, for $\left| \lambda_1^{-1} - \lambda_2^{-1} \right| \leq \delta/T$, we have $\left\| (X^\top X + \lambda_2 I_d)^{-1} X^\top Y - (X^\top X + \lambda_1 I_d)^{-1} X^\top Y \right\|_2 \leq \delta$, and thus the discretization of λ^{-1} on $(0, \lambda_{\min}^{-1}]$ gives an ℓ_∞-cover, and the size of the cover at scale δ is $\lambda_{\min}^{-1} T \delta^{-1}$. The Dudley entropy integral yields the bound of,

$$\mathfrak{R}(\Pi, (\mathbf{x}, \mathbf{y})) \leq 12\sqrt{T} \int_0^1 \sqrt{\log(2T\lambda_{\min}^{-1}\delta^{-1})} d\delta \leq 12\sqrt{T} \left(1 + \sqrt{\log(2T\lambda_{\min}^{-1})} \right).$$

This concludes the proof. $\qquad \square$

Proof of Lemma 8. Using Theorem 1,

$$\mathcal{V}_T(\Pi_\Lambda) \le 2\Re(\ell, \Pi_\Lambda) = 2 \sup_{\mathbf{z}, \mathbf{z}'} \mathbb{E}_\epsilon \sup_{\mathbf{w}_0 : \|\mathbf{w}_0\| \le 1, \lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \mathbf{w}_0 - \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|, t \right), \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\| \right\}}, \mathbf{z}'_t(\epsilon) \right\rangle \right]$$

which we can upper bound by splitting the supremum into two:

$$2 \sup_{\mathbf{z}'} \mathbb{E}_\epsilon \sup_{\mathbf{w}_0 : \|\mathbf{w}_0\| \le 1} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \mathbf{w}_0, \mathbf{z}'_t(\epsilon) \right\rangle \right] + 2 \sup_{\mathbf{z}, \mathbf{z}'} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|, t \right), \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\| \right\}}, \mathbf{z}'_t(\epsilon) \right\rangle \right]$$

The first term is simply

$$2 \sup_{\mathbf{z}'} \mathbb{E}_\epsilon \left\| \sum_{t=1}^{T} \epsilon_t \mathbf{z}'_t(\epsilon) \right\| \le 2\sqrt{T}.$$

The second term can be written as

$$2 \sup_{\mathbf{z}, \mathbf{z}'} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \left\langle \frac{\sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon)}{\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|}, \mathbf{z}'_t(\epsilon) \right\rangle \frac{\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|}{\max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|, t \right), \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\| \right\}} \right]$$

$$\le 2 \sup_{\mathbf{z}} \sup_{\mathbf{s}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \mathbf{s}_t(\epsilon) \frac{\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|}{\max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|, t \right), \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\| \right\}} \right]$$

and the tree $\mathbf{s}$ can be erased (see end of the proof of Theorem 2), yielding an upper bound

$$2 \sup_{\mathbf{z}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|}{\max\left\{ \lambda\left(\left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\|, t \right), \left\| \sum_{i=1}^{t-1} \mathbf{z}_i(\epsilon) \right\| \right\}} \right]$$

$$\le 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t \mathbf{a}_t(\epsilon)}{\max\left\{ \lambda\left(\mathbf{a}_t(\epsilon), t \right), \mathbf{a}_t(\epsilon) \right\}} \right]$$

$$\le 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \frac{\epsilon_t}{\max\left\{ \frac{\lambda(\mathbf{a}_t(\epsilon), t)}{\mathbf{a}_t(\epsilon)}, 1 \right\}} \right]$$

$$= 2 \sup_{\mathbf{a}} \mathbb{E}_\epsilon \sup_{\lambda \in \Lambda} \left[\sum_{t=1}^{T} \epsilon_t \min\left\{ \frac{\mathbf{a}_t(\epsilon)}{\lambda(\mathbf{a}_t(\epsilon), t)}, 1 \right\} \right]$$

$$= 2 \sup_{\mathbf{b}} \mathbb{E}_\epsilon \sup_{\gamma \in \Gamma} \left[\sum_{t=1}^{T} \epsilon_t \gamma(\mathbf{b}_t(\epsilon), 1/t) \right]$$

$$\le 2 \, \mathcal{R}_T(\Gamma)$$

where in the above $\mathbf{a}$ is a $\mathbb{R}^+$-valued tree such that $\mathbf{a}_t : \{\pm 1\}^{t-1} \mapsto [0, t-1]$, $\mathbf{b}$ is a $[1/T, 1]$-value tree and $\Gamma = \left\{ \gamma : \forall b \in [1/T, 1], a \in [0, 1], \gamma(a, b) = \min\left\{ \frac{a/(b-1)}{\lambda(a/(b-1), 1/b)}, 1 \right\}, \lambda \in \Lambda \right\}$. $\qquad \square$