JMLR: Workshop and Conference Proceedings 19 (2011) 559–594 24th Annual Conference on Learning Theory

Online Learning: Beyond Regret

Alexander Rakhlin
Department of Statistics
University of Pennsylvania

Karthik Sridharan
TTI-Chicago

Ambuj Tewari
Department of Computer Science
University of Texas at Austin

Editor: Sham Kakade, Ulrike von Luxburg

Abstract

We study online learnability of a wide class of problems, extending the results of Rakhlin et al. (2010a) to general notions of performance measure well beyond external regret. Our framework simultaneously captures such well-known notions as internal and general Φ-regret, learning with non-additive global cost functions, Blackwell's approachability, calibration of forecasters, and more. We show that learnability in all these situations is due to control of the same three quantities: a martingale convergence term, a term describing the ability to perform well if future is known, and a generalization of sequential Rademacher complexity, studied in Rakhlin et al. (2010a). Since we directly study complexity of the problem instead of focusing on efficient algorithms, we are able to improve and extend many known results which have been previously derived via an algorithmic construction.

1. Introduction

In the companion paper Rakhlin et al. (2010a) (hereafter referred to as *RST*), we analyzed learnability in the `Online Learning Model` when the value of the game is defined through minimax *regret*. However, regret (also known as *external regret*) is not the only way to measure performance of an online learning procedure. In the present paper, we extend the results of *RST* to other performance measures, encompassing a wide spectrum of notions which appear in the literature. Our framework gives the same footing to external regret, internal and general Φ-regret, learning with non-additive global cost functions, Blackwell's approachability, calibration of forecasters, and more. We recover, extend, and improve some existing results, and (what is more important) show that they all follow from control of the same quantities. In particular, sequential Rademacher complexity, introduced in *RST*, plays a key role in our derivations.

A reflection on the past two decades of research in learning theory reveals (in our somewhat biased view) an interesting difference between Statistical Learning Theory and Online Learning. In the former, the focus has been primarily on understanding *complexity measures* rather than *algorithms*. There are good reasons for this: if a supervised problem with i.i.d. data is learnable, Empirical Risk Minimization is the algorithm that will perform well if one disregards computational aspects. In contrast, Online Learning has been mainly centered around algorithms. Given an algorithm, a non-trivial bound serves as a certificate that the problem is learnable. This algorithm-focused approach has dominated research in Online Learning for several decades. Many important tools (such as optimization-based algorithms for online convex optimization) have emerged, yet the results lacked a unified approach for determining learnability.

With the tools developed in *RST*, the question of learnability can now be addressed in a variety of situations in a unified manner. In fact, *RST* presents a number of examples of provably learnable problems for which computationally feasible online learning methods have not yet been developed. In the present paper, we show that the scope of problems whose learnability and precise rates can be characterized is much larger than those defined in *RST* through external regret. Within this circle of problems are such well-known results as Blackwell's approachability and calibration of forecasters. For instance, our complexity-based (rather than algorithm-based) approach yields a proof of Blackwell's approachability in Banach spaces without ever mentioning an algorithm. Let us remark that Blackwell's approachability has been a key tool for showing learnability (Cesa-Bianchi and Lugosi, 2006); as our results imply approachability, they can be utilized whenever Blackwell's approachability has been successful. The results can also be used in situations where phrasing a problem as an approachability question is not necessarily natural. In Section 4.2, we discuss the relation of our results to approachability in greater detail. Our contributions can be broken down into three parts:

1. We formulate the online learning problem, with a performance measure (a form of *regret*), defined in terms of certain payoff transformations. While this formulation might appear unusual, we show that it is general enough to encompass many seemingly different frameworks, yet specific enough that we can provide generic upper bounds.

2. We develop upper and lower bounds on the value of the game under various natural assumptions. These tools allow us to deal with performance measures well beyond the standard additive notion of external regret. Such performance measures include smooth non-additive functions of payoffs, generalizing the "cumulative payoff" notion often considered in the literature. The abstract definition in terms of payoff transformations lets us consider rich classes of mappings whose complexity can be studied through random averages, covering numbers, and combinatorial parameters.

3. We apply our machinery to a number of well-known problems. Unfortunately, in this extended abstract we are not able to fit all the details. We refer the reader to Rakhlin et al. (2010b).

   (a) For the usual notion of external regret, the results boil down to those of Rakhlin et al. (2010a).

   (b) For the more general Φ-regret (see e.g. Stoltz and Lugosi (2007); Gordon et al. (2008); Hazan and Kale (2007)), we recover and improve several known results. In particular, for convergence to Φ-correlated equilibria, we improve upon the results of Stoltz and Lugosi (Stoltz and Lugosi, 2007).

(c) We study the game of Blackwell's approachability (Blackwell, 1956) in (possibly infinite-dimensional) separable Banach spaces. Specifically, we show that variation of the worst-case martingale upper and lower-bounds (to within a constant) the rate of convergence to the set.

(d) We also consider the game of calibrated forecasting. We improve upon the results of Mannor and Stoltz (Mannor and Stoltz, 2010) and prove (to the best of our knowledge) the first known $O(\sqrt{T})$ rates for calibration with more than 2 outcomes. Our approach is markedly different from those found in the literature.

(e) We use our framework to study games with global cost functions and as an example we extend the bounds recently obtained by Even-Dar et al. (2009).

(f) We provide techniques for bounding notions of regret where algorithm's performance is measured against a time-varying comparator (see e.g. Herbster and Warmuth (1998); Bousquet and Warmuth (2002); Zinkevich (2003)).

The intent of this paper is to provide a framework and tools for studying problems that can be phrased as repeated games. However, unlike much of existing research in online learning, we are not solving the general problem by exhibiting an algorithm and studying its performance. Rather, we proceed by directly attacking the value of the game. Alas, the value is a complicated object, and the non-invitingly long sequence of infima and suprema can single-handedly extinguish any desire to study it. Our results attest to the power of *symmetrization*, which emerges as a key tool for studying the value of the game. In the literature, symmetrization has been used for i.i.d. data (Giné and Zinn, 1984). In *RST* (see also Abernethy et al. (2009)), it was shown that symmetrization can also be used in situations beyond the traditional setting. What is even more surprising, we are able to employ symmetrization ideas even when the objective function is not a summation of terms but rather a global function of many variables. We hope that these tools can have an impact not only on online learning but also on game theory.

We believe that there are many more examples falling under the present framework. We only chose a few to demonstrate how upper and lower bounds arise from the complexity of the problem. Along with an upper bound, a (computationally inefficient) algorithm can always be recovered from the minimax analysis. Finding efficient algorithms is often a difficult enterprise, and it is important to be able to understand the inherent complexity even before focusing on computation.

2. The Setting

At a very abstract level, the problem of online learning can be phrased as that of optimization of a given function $\mathbf{R}_T(f_1, x_1, \ldots, f_T, x_T)$ with coordinates being chosen *sequentially* by the player and the adversary. Of course, at this level of generality not much can be said. Hence, we make some minimal assumptions on the function $\mathbf{R}_T$ which lead to meaningful guarantees on the online optimization process.[1] These assumptions are satisfied by a number of natural performance measures, as illustrated by the examples below.

1. The question of general conditions on the function under which such sequential minimization is possible was put forth by Peter Bartlett a few years ago in a coffee conversation. This paper paves way towards addressing this question.

Let $\mathcal{F}$ and $\mathcal{X}$ be the sets of moves of the learner (player) and the adversary, respectively. Generalizing the `Online Learning Model` considered in *RST*, we study the following T-round interaction between the learner and the adversary: On round $t = 1, \ldots, T$, the learner chooses a mixed strategy q_t (distribution on $\mathcal{F}$), the adversary picks $x_t \in \mathcal{X}$, the learner draws $f_t \in \mathcal{F}$ from q_t and receives payoff (loss) signal $\ell(f_t, x_t) \in \mathcal{H}$.

We would like to specify that we are in the full information setting and that at the end of each round both the player and the adversary observe each other's moves f_t, x_t. The payoff space $\mathcal{H}$ is a (not necessarily convex) subset of a separable Banach space $\mathcal{B}$. Both the player and the adversary can be randomized and adaptive. The goal of the learner is to minimize the following general form

$$\mathbf{R}_T = \mathbf{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \inf_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \tag{1}$$

of performance measure, where

(1) The function $\ell : \mathcal{F} \times \mathcal{X} \mapsto \mathcal{H}$ is an $\mathcal{H}$-valued payoff (or loss) function.

(2) The function $\mathbf{B} : \mathcal{H}^T \mapsto \mathbb{R}$ is a (not necessarily additive or convex) form of cumulative payoff.

(3) The set Φ_T consists of sequences $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_T)$ of measurable payoff transformation mappings $\phi_t : \mathcal{H}^{\mathcal{F} \times \mathcal{X}} \mapsto \mathcal{H}^{\mathcal{F} \times \mathcal{X}}$ that transform the payoff function ℓ into a payoff function ℓ_{ϕ_t}.

The goal of the adversary is to maximize the same quantity (1), making it a zero-sum game.

This paper is concerned with learnability and with identifying *complexity measures* that govern learnability. But complexity of what should we focus on? After all, the general online learning problem is defined by the choice of five components: $\mathbf{B}, \ell, \mathcal{F}, \mathcal{X}$, and Φ_T. In *RST*, the choice was easy: it should be the complexity of the function class $\mathcal{F}$ that plays the key role. That was natural because the payoff was written as $\ell(f, x) = f(x)$, which suggested that the function class $\mathcal{F}$ is the object of study. The present formulation, however, is much more general. When this work commenced, it seemed likely that complexity of the problem will be some interaction between the complexity of Φ_T and complexity of $\mathcal{F}$. As we show below, one may just focus on the complexity of Φ_T, while $\mathcal{F}$ and $\mathcal{X}$ are now on the same footing. For instance, even if it might seem unusual at first, we will introduce a notion of a cover of the set of sequences of payoff transformations Φ_T. In summary, while all five components $\mathbf{B}, \ell, \mathcal{F}, \mathcal{X}$, and Φ_T play a role in determining learnability, we will mainly refer to the complexity of the payoff mapping ℓ and the payoff transformation Φ_T without an explicit reference to $\mathcal{F}$, $\mathcal{X}$, and $\mathbf{B}$. We emphasize that most flexibility comes from the payoff mapping ℓ and from the transformations Φ_T of the payoffs.

Important classes of payoff transformation mappings are those that transform the payoff function ℓ by acting only on the first argument of ℓ, i.e. only modifying the player's action. Formally, a class of sequences of payoff transformations Φ_T is said to be a *departure mapping class* if there exists a class Φ'_T of sequences $\boldsymbol{\phi}' = (\phi'_1, \ldots, \phi'_T)$ with $\phi'_i : \mathcal{F} \mapsto \mathcal{F}$ such that for each $\boldsymbol{\phi} \in \Phi_T$ there exists a $\boldsymbol{\phi}' \in \Phi'_T$ with $\ell_{\phi_t}(f, x) := \ell(\phi'_t(f), x)$ that for all $t \in [T]$, $f \in \mathcal{F}$ and $x \in \mathcal{X}$. We shall slightly abuse notation and use Φ_T to represent both the class of payoff transformation and the class of departure mappings from $\mathcal{F}$ to itself. Another class of interest are payoff transformations that do not vary with time. We say that Φ_T is *time-invariant* if all sequences of payoff transformation are constant in time: $\Phi_T = \{(\phi, \ldots, \phi) : \phi \in \Phi\}$, where Φ is a "basis" class of mappings $\mathcal{H}^{\mathcal{F}\times\mathcal{X}} \mapsto \mathcal{H}^{\mathcal{F}\times\mathcal{X}}$.

In the following, we assume that $\mathcal{F}$ and $\mathcal{X}$ are subsets of a separable metric space. Let $\mathcal{Q}$ and $\mathcal{P}$ be the sets of probability distributions on $\mathcal{F}$ and $\mathcal{X}$, respectively. Assume that $\mathcal{Q}$ and $\mathcal{P}$ are weakly compact. From the outset, we assume that the adversary is non-oblivious (that is, adaptive). Formally, define a learner's strategy π as a sequence of mappings $\pi_t : (\mathcal{P} \times \mathcal{F} \times \mathcal{X})^{t-1} \mapsto \mathcal{Q}$ for each $t \in [T]$. The form (1) of the performance measure gives rise to the value of the game:

$$\mathcal{V}_T(\ell, \Phi_T) = \inf_{q_1} \sup_{x_1} \mathop{\mathbb{E}}_{f_1 \sim q_1} \ldots \inf_{q_T} \sup_{x_T} \mathop{\mathbb{E}}_{f_T \sim q_T} \sup_{\phi \in \Phi_T} \{\mathbf{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) \\ -\mathbf{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T))\} \tag{2}$$

where q_t and x_t range over $\mathcal{Q}$ and $\mathcal{X}$, respectively. With this definition of a value, the (deterministic) strategy of the adversary is a sequence of mappings $(\mathcal{Q}\times\mathcal{F}\times\mathcal{X})^{t-1}\times\mathcal{Q} \mapsto \mathcal{X}$ for each $t \in [T]$. The problem is said to be *online learnable* if $\limsup_{T\to\infty} \mathcal{V}_T(\ell, \Phi_T) = 0$.

The value of the game is defined as an *expected* performance measure. As such, it yields "in probability" statements. While beyond the scope of this paper, we can also define the value of the game using a *high probability* performance measure, leading to "almost sure" convergence (Rakhlin et al., 2010b).

2.1. Examples

A reader might wonder why we have defined the game in terms of abstract payoff transformation mappings. It turns out that with this definition, various seemingly different frameworks become nothing but special cases, as illustrated by the following examples.

Example 1 (External Regret Game, Section 4.1.1) *Let* $\mathcal{H} = \mathbb{R}$, *let* $\mathbf{B}(z_1, \ldots, z_T) = \frac{1}{T}\sum_{t=1}^T z_t$, *and*

$$\Phi_T = \{(\phi_f, \ldots, \phi_f) : f \in \mathcal{F} \quad \textit{and} \quad \phi_f : \mathcal{F} \mapsto \mathcal{F} \quad \textit{is a constant mapping } \phi_f(g) = f \ \forall g \in \mathcal{F}\}.$$

It is easy to see that (1) *becomes external regret:*

$$\mathbf{R}_T = \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) - \inf_{f\in\mathcal{F}} \frac{1}{T}\sum_{t=1}^T \ell(f, x_t).$$

Example 2 (Φ-Regret, Section 4.1) *Let* $\mathcal{H} = \mathbb{R}$, *let* $\mathbf{B}(z_1, \ldots, z_T) = \frac{1}{T}\sum_{t=1}^T z_t$, *and* $\Phi_T = \{(\phi, \ldots, \phi) : \phi \in \Phi\}$ *for a fixed family* Φ *of* $\mathcal{F} \mapsto \mathcal{F}$ *mappings. Performance measure in* (1) *becomes*

$$\mathbf{R}_T = \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) - \inf_{\phi\in\Phi} \frac{1}{T}\sum_{t=1}^T \ell(\phi(f_t), x_t). \tag{3}$$

This example covers a variety of notions such as external, internal, and swap regrets.

Example 3 (Blackwell Approachability, Section 4.2) *Let $\mathcal{H}$ a subset of a Banach space $\mathcal{B}$, $S \subset \mathcal{B}$ be a closed convex set, and $\mathbf{B}(z_1, \ldots, z_T) = \inf_{c\in S} \left\| \frac{1}{T}\sum_{t=1}^T z_t - c \right\|$. The set Φ_T contains sequences $(\phi_1, \ldots, \phi_T)$ such that $\ell_{\phi_t}(f,x) = c_t \in S$ for all $f \in \mathcal{F}$, $x \in \mathcal{X}$, and $1 \le t \le T$. Eq.* (1) *becomes the distance to the set S:*

$$\mathbf{R}_T = \inf_{c\in S} \left\| \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) - c \right\| \tag{4}$$

Example 4 (Calibration of Forecasters, Section 4.3) *Let $\mathcal{H} = \mathbb{R}^k$, $\mathcal{F}$ the probability simplex in $\mathbb{R}^k$, and $\mathcal{X}$ the vertices of $\mathcal{F}$. Define $\ell(f,x) = 0$. Further, $\mathbf{B}(z_1, \ldots, z_T) = -\left\| \frac{1}{T}\sum_{t=1}^T z_t \right\|$ for some norm $\|\cdot\|$ on $\mathbb{R}^k$, and $\Phi_T = \{(\phi_{p,\lambda}, \ldots, \phi_{p,\lambda}) : p \in \Delta(k), \lambda > 0\}$ contains time-invariant mappings defined by $\ell_{\phi_{p,\lambda}}(f,x) = \mathbf{1}\{\|f-p\| \le \lambda\} \cdot (f-x)$. Performance measure in* (1) *then becomes*

$$\mathbf{R}_T = \sup_{\lambda>0} \sup_{p\in\Delta(k)} \left\| \frac{1}{T}\sum_{t=1}^T \mathbf{1}\{\|f_t - p\| \le \lambda\} \cdot (f_t - x_t) \right\|. \tag{5}$$

Example 5 (Global Cost Online Learning Game, Section 4.4) *(see also the original paper Even-Dar et al. (2009)) Let $\mathcal{H} = \mathbb{R}^k$, $\mathcal{X} = [0,1]^k$, $\mathcal{F} = \Delta(k)$, $\ell(f,x) = f \odot x = (f^1 \cdot x^1, \ldots, f^k \cdot x^k)$. Let $\mathbf{B}(z_1, \ldots, z_T) = \left\| \frac{1}{T}\sum_{t=1}^T z_t \right\|$ and*

$$\Phi_T = \{(\phi_f, \ldots, \phi_f) : f \in \mathcal{F} \quad \textit{and} \quad \phi_f : \mathcal{F} \mapsto \mathcal{F} \quad \textit{is a constant mapping } \phi_f(g) = f \; \forall g \in \mathcal{F}\}.$$

Then

$$\mathbf{R}_T = \left\| \frac{1}{T}\sum_{t=1}^T f_t \odot x_t \right\| - \inf_{f\in\mathcal{F}} \left\| \frac{1}{T}\sum_{t=1}^T f \odot x_t \right\|. \tag{6}$$

2.2. Notation

We let $\mathbb{E}_{x\sim p}$ denote expectation w.r.t. a random variable x with a distribution p. For random variables $x_1, \ldots, x_T$ with distributions $p_1, \ldots, p_T$, we will use the shorthand $\mathbb{E}_{x_{1:T}\sim p_{1:T}}$ to denote expectation w.r.t. all these variables. Let q and p be distributions on $\mathcal{F}$ and $\mathcal{X}$, respectively. We define a shorthand $\ell(q,p) = \mathbb{E}_{f\sim q, x\sim p}\ell(f,x)$ and $\ell_\phi(q,p) = \mathbb{E}_{f\sim q,x\sim p}\ell_\phi(f,x)$. The Dirac delta distribution is denoted by δ_x. A Rademacher random variable is symmetric $\{\pm 1\}$. The notation $x_{a:b}$ denotes the sequence $x_a, \ldots, x_b$. The indicator of an event A is denoted by $\mathbf{1}\{A\}$. The set $\{1, \ldots, T\}$ is denoted by $[T]$, while the k-dimensional probability simplex is denoted by $\Delta(k)$. The set of all functions from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $\mathcal{Y}^{\mathcal{X}}$, and the t-fold product is denoted by $\mathcal{X}^t$. Whenever a supremum (infimum) is written as $\sup_a$ without a being quantified, it is assumed that a ranges over the set of all possible values which will be understood from the context. For a separable Banach space $\mathcal{B}$ equipped with a norm $\|\cdot\|$, let $B_{\|\cdot\|}$ be the unit ball. Let $\mathcal{B}^*$ denote the dual space and $B_{\|\cdot\|_*}$ the corresponding dual ball. For $a \in \mathcal{B}^*$, $\|a\|_* = \sup_{b\in B_{\|\cdot\|}} |\langle a, b\rangle|$. For $b \in \mathcal{B}$, we write $\langle a, b\rangle = a(b)$ for the continuous linear functional $a \in \mathcal{B}^*$ on $\mathcal{B}$. Let ϕ_{id} be identity payoff transformation $\ell_{\phi_{\mathrm{id}}}(f,x) = \ell(f,x)$ for all $f \in \mathcal{F}$, $x \in \mathcal{X}$. The singleton set containing the time-invariant sequence of identity transformations is denoted by $\mathcal{I} = \{(\phi_{\text{id}}, \ldots, \phi_{\text{id}})\}$. Following *RST*, we define binary trees as follows. Given some set $\mathcal{Z}$, a *$\mathcal{Z}$-valued tree of depth T* is a sequence $(\mathbf{z}_1, \ldots, \mathbf{z}_T)$ of T mappings $\mathbf{z}_i : \{\pm 1\}^{i-1} \mapsto \mathcal{Z}$. The *root* of the tree $\mathbf{z}$ is the constant function $\mathbf{z}_1 \in \mathcal{Z}$. Unless specified otherwise, $\epsilon = (\epsilon_1, \ldots, \epsilon_T) \in \{\pm 1\}^T$ will define a path. Slightly abusing the notation, we will write $\mathbf{z}_t(\epsilon)$ instead of $\mathbf{z}_t(\epsilon_{1:t-1})$.

3. General Upper Bounds

This section is devoted to upper bounds on the value of the game. We start by introducing the Triplex Inequality, which requires no assumptions beyond those described in Section 2. Under the additional weak assumption of subadditivity of $\mathbf{B}$, we can perform symmetrization and further upper bound two of the three terms in Triplex Inequality by a non-additive version of sequential Rademacher complexity. As we progress through the section, we make additional assumptions and specialize and refine the upper bounds. The following definition generalizes the notion of sequential Rademacher complexity, introduced in *RST*, to "global" functions $\mathbf{B}$ of the payoff sequence.

Definition 1 *The* sequential complexity *with respect to the payoff function ℓ and payoff transformation mappings Φ_T is defined as*

$$\mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}) = \sup_{\mathbf{f},\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi_T} \mathbf{B}\Big(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_T \ell_{\phi_T}(\mathbf{f}_T(\epsilon), \mathbf{x}_T(\epsilon))\Big)$$

where the outer supremum is taken over all $(\mathcal{F} \times \mathcal{X})$-valued trees of depth T and $\epsilon = (\epsilon_1, \ldots, \epsilon_T)$ is a sequence of i.i.d. Rademacher random variables.

Whenever $\mathbf{B}$ is clear from the context, we will omit it from the notation: $\mathfrak{R}_T(\ell, \Phi_T)$. If Φ_T is a set of sequences of time-invariant transformations obtained from the base class Φ, we will simply write $\mathfrak{R}_T(\ell, \Phi)$. Let us remark that the moves of the player and the adversary appear "on the same footing" in $\mathbf{R}_T$ and in the above definition of sequential complexity. The "asymmetry" of sequential Rademacher complexity as studied in *RST* (where the supremum is taken over the *player's* best choice) arises precisely from the asymmetry of the notion of external regret, which, in turn, is due to Φ_T acting on the player choice only. In Section 4.1.1, we show that the notion studied in *RST* is indeed recovered for the case of external regret. An equivalent way to write sequential complexity is

$$\mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}) = \sup_{f_1, x_1} \mathbb{E}_{\epsilon_1} \sup_{f_2, x_2} \mathbb{E}_{\epsilon_2} \ldots \sup_{f_T, x_T} \mathbb{E}_{\epsilon_T} \sup_{\phi \in \Phi_T} \mathbf{B}\Big(\epsilon_1 \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_T \ell_{\phi_T}(f_T, x_T)\Big) \tag{7}$$

where the supremum on t-th step is over $f_t \in \mathcal{F}$, $x_t \in \mathcal{X}$.

3.1. Triplex Inequality

The following theorem is the starting point for all further analysis. Because of its importance, we shall call it the *Triplex Inequality.* The three terms in the upper bound of the theorem are the three key players in the process of online learning: martingale convergence, the ability to perform well if the future is known, and complexity of the class in terms of sequential complexity.

Theorem 2 (Triplex Inequality) *The following* 3*-term upper bound on the value of the game holds:*

$$
\begin{aligned}
&\mathcal{V}_T(\ell, \Phi_T) \qquad (8)\\
&\quad \le \sup_{p_1,q_1} \mathop{\mathbb{E}}_{x_1,f_1} \dots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{x_T,f_T} \Big\{ \mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T)) - \mathop{\mathbb{E}}_{x'_{1:T},f'_{1:T}} \mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T)) \Big\} \\
&\quad + \sup_{p_1} \inf_{q_1} \dots \sup_{p_T} \inf_{q_T} \sup_{\phi\in\Phi_T} \mathop{\mathbb{E}}_{x_{1:T},f_{1:T}} \Big\{ \mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T)) - \mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T)) \Big\} \\
&\quad + \sup_{p_1,q_1} \mathop{\mathbb{E}}_{x_1,f_1} \dots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{x_T,f_T} \\
&\qquad\qquad \sup_{\phi\in\Phi_T} \left\{ \mathop{\mathbb{E}}_{x'_{1:T},f'_{1:T}} \mathbf{B}\Big(\ell_{\phi_1}(f'_1,x'_1),\dots,\ell_{\phi_T}(f'_T,x'_T)\Big) - \mathbf{B}\Big(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T)\Big) \right\}
\end{aligned}
$$

In the statement of the theorem, the random variables f_t, f'_t have distribution q_t while x_t, x'_t have distribution p_t. We remark that convexity of $\mathbf{B}$ is *not required* for the Triplex Inequality to hold. Under a subadditivity condition, the following result gives upper bounds on the first and third terms.

Theorem 3 *If* $\mathbf{B}$ *is subadditive, then the last term in the Triplex Inequality is upper bounded by twice the sequential complexity,* $2\mathfrak{R}_T(\ell, \Phi_T, \mathbf{B})$*, and the first term is bounded by* $2\mathfrak{R}_T(\ell, \mathcal{I}, \mathbf{B})$ *where* $\mathcal{I}$ *is the singleton set consisting of the identity mapping. Similarly, if* $-\mathbf{B}$ *is subadditive, then the last term is upper bounded by* $2\mathfrak{R}_T(\ell, \Phi_T, -\mathbf{B})$ *and the first term is bounded by* $2\mathfrak{R}_T(\ell, \mathcal{I}, -\mathbf{B})$*.*

Discussion of Theorem 2 and Theorem 3 We note that the first and the third terms are similar in their form. In fact, the first term can be equivalently written in a form similar to the third term, with only one difference that $\boldsymbol{\phi}$ belongs to a singleton set $\mathcal{I}$ containing the identity mapping. If $\mathcal{I} \subseteq \Phi_T$, then, trivially, $\mathfrak{R}_T(\ell, \mathcal{I}, \mathbf{B}) \le \mathfrak{R}_T(\ell, \Phi_T, \mathbf{B})$ and, therefore, an upper bound on the third term yields and upper bound on the first. However, in some situations Φ_T is "simpler" or incomparable to $\mathcal{I}$ and, hence, the first and the third term in the Triplex Inequality are distinct.

What exactly is achieved by Theorem 3? Let us compare the third term in the Triplex Inequality to its sequential complexity upper bound given by Eq. (7). Both quantities involve interleaved suprema and expected values. However, in the former, the suprema are over the choice of distributions p_t, q_t and the expected values are draws of x_t, f_t from these mixed strategies. In contrast, sequential complexity, as written in Eq. (7), contains suprema over the choices x_t, f_t followed by a random draw of the next sign ϵ_t. Crucially, it is easier to work with the sequential complexity as opposed to the third term in the Triplex Inequality since in the former the only randomness comes from the random signs. In mathematical terms, the σ-algebra is generated by $\{\epsilon_t\}$ rather than a complicated stochastic process arising from the Triplex Inequality. This is one of the key observations of the paper.

Depending on a particular problem, some of the terms in the Triplex Inequality might be easier to control than others. However, it is often the case that the first term is the easiest, as it naturally leads to the question of martingale convergence. The second term is typically bounded by providing a specific response strategy for the player if the mixed strategy of the adversary is known. This response strategy is similar to the so-called Blackwell's condition for approachability (see Section 4.2 for further comparison). The third term is arguably the most difficult as it captures complexity of the set of payoff transformations Φ_T. Under the subadditivity assumption on $\mathbf{B}$, Theorem 3 upper bounds the first and third terms by the sequential complexity.

The following observation gives us a simple condition under which we can replace $\mathbf{B}$ with some other $\mathbf{B}'$, and we shall find it useful in scenarios when it is difficult to directly deal with $\mathbf{B}$. If $\mathbf{B} : \mathcal{H}^T \mapsto \mathbb{R}$ and $\mathbf{B}' : \mathcal{H}^T \mapsto \mathbb{R}$ are such that $\forall z_1, \ldots, z_T \in \mathcal{H}$, $\mathbf{B}(z_1, \ldots, z_T) \leq \mathbf{B}'(z_1, \ldots, z_T)$ then we have that for any class of transformations Φ_T, $\mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}')$.

This completes our discussion of the main theorems. We now turn to the question of upper bounding the terms in the Triplex Inequality. To this end, we need to define the notion of a smooth function. A function $g : \mathcal{H} \mapsto \mathbb{R}$ is said to be (σ, p)-uniformly smooth for some $p \in (1, 2]$ and $\sigma \geq 0$ if for all $z, z' \in \mathcal{H}$ we have,

$$g(z) \leq g(z') + \left\langle \nabla g(z'), z - z' \right\rangle + \tfrac{\sigma}{p} \|z - z'\|^p \ .$$

We say that g is *uniformly smooth* if there exist finite σ and p such that g is (σ, p)-uniformly smooth. We say that a norm $\|\cdot\|$ is (σ, p)-smooth if $\|\cdot\|^p/p$ is a (σ, p)-smooth function.

A function $\mathbf{B}$ which is smooth in its arguments can be "sequentially linearized", with additional second-order terms appearing as norms of the increments. Informally, the smoothness assumption provides a link from a "global" function $\mathbf{B}$ of coordinates to a sum of its parts. From the point of view of online learning, this is very promising, as it appears to be difficult to sequentially optimize a "global" function of many decisions. Due to limited space in this extended abstract, we will not present the most general bounds based solely on smoothness of $\mathbf{B}$ (we refer the reader to Rakhlin et al. (2010b) for these results). However, we will state bounds for a smooth function of the average of coordinates.

3.2. When B is a Function of the Average

For the rest of this sub-section we assume that $\mathbf{B} = G\left(\frac{1}{T}\sum_{t=1}^T z_t\right)$, where (some power of) G is an appropriately smooth function on the convex set $\mathrm{conv}(\mathcal{H})$. This form of $\mathbf{B}$ occurs naturally in many games including Blackwell's approachability and calibration. Among the most basic smooth functions are powers of norms. For the ℓ_q norms, the following smoothness results are known. For any $q \in (1, 2]$, $G(z) = \|z\|_q^q$ is (q, q)-uniformly smooth and for any $q \in [2, \infty)$ the function $G(z) = \|z\|_q^2$ is $(2(q-1), 2)$-uniformly smooth. The ℓ_∞ cannot be made smooth by raising it to any finite power s. However, for any $z \in \mathcal{H}$ and any $q' \in (1, \infty)$, $\|z\|_\infty \leq \|z\|_{q'}$. Hence as discussed above, $\mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \mathbf{B}')$ where $\mathbf{B}'(z_1, \ldots, z_T) = \left\|\frac{1}{T}\sum_{t=1}^T z_t\right\|_{q'}$. By choosing q' appropriately and using the smoothness of the $\ell_{q'}$ norm we can provide upper bounds for the value of the game. Similarly to ℓ_∞, no finite power of the ℓ_1 norm is smooth. However if $\mathcal{H} \subseteq \mathbb{R}^d$, we can upper bound the ℓ_1 norm by, say, ℓ_2 norm multiplied by a factor $\sqrt{d}$. Smoothness of this latter norm can then be used. This is indeed the approach that is employed for proving rates for calibration.

The following result shows that if G is 1-Lipschitz and G^2 is 2-smooth, we obtain a $O(1/\sqrt{T})$ convergence rate whenever Φ_T is a finite set. We refer to Rakhlin et al. (2010b) for the case when G is not Lipschitz, as well as for the case of a (γ, p)-smooth function G for $1 < p \leq 2$.

Lemma 4 *Let Φ_T be a finite set of payoff transformations. Let*

$$\mathbf{B}(z_1, \ldots, z_T) = G\left(\frac{1}{T}\sum_{t=1}^{T} z_t\right)$$

where G is 1-Lipschitz with respect to a norm $\|\cdot\|$ and $G(0) = 0$. Suppose that G^2 is $(\gamma, 2)$-smooth function on the convex set $\mathrm{conv}(\mathcal{H})$. *Further, suppose that for any $x \in \mathcal{X}$, $f \in \mathcal{F}$, $\boldsymbol{\phi} \in \Phi_T$ and $t \in [T]$, it is true that $\|\ell_{\phi_t}(f, x)\| \leq \eta$. Then it holds that*

$$\mathfrak{R}_T(\ell, \Phi_T) \leq \sqrt{\frac{\gamma\eta^2 \log(2|\Phi_T|)}{T}} \ .$$

Having a bound on the complexity of a finite set of payoff transformations, we seek to extend the results to infinite sets. A natural approach is to pass to a finite cover of the set at an expense of losing an amount proportional to the resolution of the cover. The following definition can be seen as a generalization of the corresponding notion introduced in *RST*. We remark that the object, for which we would like to provide a cover, is the set Φ_T. Whenever payoff transformations are simply constant time-invariant departure mappings, complexity of Φ_T identical to that of $\mathcal{F}$, yielding the online cover of class $\mathcal{F}$. In general, however, the set of payoff transformations can be much more complex than (or not even comparable to) $\mathcal{F}$.

Definition 5 *A set V of $\mathcal{H}$-valued trees of depth T is* an α-cover *(with respect to ℓ_p-norm) of Φ_T on an $(\mathcal{F} \times \mathcal{X})$-valued tree $(\mathbf{f}, \mathbf{x})$ of depth T if*

$$\forall \boldsymbol{\phi} \in \Phi_T,\ \forall \epsilon \in \{\pm 1\}^T\ \exists \mathbf{v} \in V \text{ s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} \|\mathbf{v}_t(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\|^p \leq \alpha^p \ . \qquad (9)$$

The covering number *of the set of payoff transformations Φ_T on a given tree $(\mathbf{f}, \mathbf{x})$ is defined as*

$$\mathcal{N}_p(\alpha, \Phi_T, (\mathbf{f}, \mathbf{x})) = \min\{|V| : V \textit{ is an } \alpha\textit{-cover w.r.t. } \ell_p\textit{-norm of } \Phi_T \textit{ on } (\mathbf{f}, \mathbf{x}) \textit{ tree}\} \ .$$

Further define $\mathcal{N}_p(\alpha, \Phi_T, T) = \sup_{(\mathbf{f},\mathbf{x})} \mathcal{N}_p(\alpha, \Phi_T, (\mathbf{f}, \mathbf{x}))$, the maximal ℓ_p covering number of Φ_T.

In sections that follow, we specialize this definition to fit particular assumptions on Φ_T. The next theorem shows that sequential complexity can be bounded above in terms of the covering number, integrated over all the scales. This is a generalization of the analogous result in *RST*.

Theorem 6 *Assume that $\mathbf{B}(z_1, \ldots, z_T) = G\left(\frac{1}{T}\sum_{t=1}^{T} z_t\right)$ where G is 1-Lipschitz with respect to a norm $\|\cdot\|$ and $G(0) = 0$. Suppose that G^2 is $(\gamma, 2)$-smooth function on the convex set* $\mathrm{conv}(\mathcal{H})$. *Further, suppose that for any $x \in \mathcal{X}$, $f \in \mathcal{F}$, $\boldsymbol{\phi} \in \Phi_T$ and $t \in [T]$, it is true that $\|\ell_{\phi_t}(f, x)\| \leq \eta$. Then it holds that*

$$\mathfrak{R}_T(\ell, \Phi_T) \leq 4 \inf_{\alpha > 0} \left\{ \alpha + 3\sqrt{\frac{\gamma}{T}} \int_{\alpha}^{\eta} \sqrt{\log \mathcal{N}_\infty(\beta, \Phi_T, T)}\, d\beta \right\} \ .$$

3.3. General Bounds Under Linearity Assumptions on B

The general results of the previous section can be restated in simpler terms if more assumptions are made. In particular, some of the terms in the Triplex Inequality can be dropped as soon as $\mathbf{B}$ is linear. While some of the results below can be repeated for a more general form of $\mathbf{B}$, for simplicity we assume that $\mathbf{B}$ is an average of its arguments and that $\mathcal{H} \subseteq \mathbb{R}$: $\mathbf{B}(z_1, \ldots, z_T) = \frac{1}{T}\sum_{t=1}^{T} z_t$.

Corollary 7 *Under the assumption on* $\mathbf{B}$ *made above, the following statements hold:*

(a) The first term in the Triplex Inequality is zero.

(b) If Φ_T *is a class of departure mappings, then the second term in the Triplex Inequality is non-positive. Hence,*

$$\mathcal{V}_T(\ell, \Phi_T) \leq 2\mathfrak{R}_T(\ell, \Phi_T) \ .$$

(c) For $\mathcal{H} \subseteq [-1, 1]$*, and assuming* $|\ell(f, x)| \leq \eta$ *for any* $x \in \mathcal{X}$ *and* $f \in \mathcal{F}$*,*

$$\mathfrak{R}_T(\ell, \Phi_T) \leq 4 \inf_{\alpha \geq 0} \left\{ \alpha + 3\sqrt{2} \int_{\alpha}^{\eta} \sqrt{\frac{\log \mathcal{N}_\infty(\delta, \Phi_T, T)}{T}} d\delta \right\} \ .$$

Experts in the area will notice the use of ℓ_∞ (as opposed to ℓ_2 in the classical Dudley integral bound) covering numbers in the two results above. This can certainly be done (Rakhlin et al., 2010b) for Corollary 7 and most probably even for the more general Theorem 6. However, in applications, one seldom loses more than a mild logarithmic factor (in T) due to the use of ℓ_∞ covering numbers.

When $\mathbf{B}$ is the average of its coordinates, the sequential complexity takes on a familiar form:

$$\mathfrak{R}_T(\ell, \Phi_T) = \sup_{\mathbf{f}, \mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \left\{ \sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\}.$$

Further, for $\mathcal{H} \subseteq \mathbb{R}$, Eq. (9) in definition of the cover becomes

$$\forall \boldsymbol{\phi} \in \Phi_T, \ \forall \epsilon \in \{\pm 1\}^T \ \exists \mathbf{v} \in V \text{ s.t.} \quad \frac{1}{T} \sum_{t=1}^{T} |\mathbf{v}_t(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))|^p \leq \alpha^p$$

where V is now a set of $\mathbb{R}$-valued trees. A further simplification of various notions is obtained for time-invariant payoff transformations. Moreover, for time-invariant payoff transformations we can define combinatorial parameters, generalizing the Littlestone (Littlestone, 1988; Ben-David et al., 2009) and (online) fat-shattering dimensions (Rakhlin et al., 2010a). This is the subject of the next section.

3.3.1. Combinatorial Parameters for Time-Invariant Payoff Transformations

Assume $\mathcal{H} \subseteq \mathbb{R}$. Consider time-invariant payoff transformations generated from some base class of payoff transformations Φ. That is, $\Phi_T = \{(\phi, \ldots, \phi) : \phi \in \Phi\}$. We have the following definition of a generalized shattering dimension.

Definition 8 *Let $\mathcal{H} = \{\pm 1\}$. An $(\mathcal{F}\times\mathcal{X})$-valued tree $(\mathbf{f},\mathbf{x})$ of depth d is* shattered[2] *by a payoff transformation class Φ if for all $\epsilon \in \{\pm1\}^d$, there exists $\phi \in \Phi$ such that $\ell_\phi(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) = \epsilon_t$ for all $t \in [d]$. The* shattering dimension $\mathrm{Sdim}(\Phi)$ *is the largest d such that Φ shatters an $(\mathcal{F}\times\mathcal{X})$-valued tree of depth d.*

We can also define the scale-sensitive version of the shattering dimension, generalizing the fat-shattering dimension of *RST*.

Definition 9 *An $(\mathcal{F}\times\mathcal{X})$-valued tree $(\mathbf{f},\mathbf{x})$ of depth d is* α-shattered *by a payoff transformation class Φ, if there exists an $\mathbb{R}$-valued tree $\mathbf{s}$ of depth d such that*

$$\forall \epsilon \in \{\pm1\}^d,\ \exists \phi \in \Phi \quad s.t.\ \forall t \in [d],\ \epsilon_t\Big(\ell_\phi(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) - \mathbf{s}_t(\epsilon)\Big) \geq \alpha/2$$

The tree $\mathbf{s}$ is called the witness to shattering. *The* fat-shattering dimension $\mathrm{fat}_\alpha(\Phi)$ *at scale α is the largest d such that Φ α-shatters an $(\mathcal{F}\times\mathcal{X})$-valued tree of depth d.*

Slightly abusing notation, we write $\mathcal{N}_p(\alpha, \Phi, (\mathbf{f},\mathbf{x}))$ instead of $\mathcal{N}_p(\alpha, \Phi_T, (\mathbf{f},\mathbf{x}))$ whenever Φ_T consists of sequences of time-invariant payoff transformations with a base class Φ.

The combinatorial parameters are useful if they can be shown to control problem complexity through, for instance, covering numbers. We state the following two results without proofs, as the arguments are identical to the ones given in the companion paper *RST*. To be precise, the $(\mathbf{f},\mathbf{x})$ tree here plays the role of the $\mathbf{x}$ tree in *RST*, ℓ_ϕ for $\phi \in \Phi$ plays the role of $f \in \mathcal{F}$ in *RST*.

Theorem 10 *Let $\mathcal{H} \subseteq \{0, \ldots, k\}$ and $\mathrm{fat}_2(\Phi) = d_2$, $\mathrm{fat}_1(\Phi) = d_1$. Then*

$$\mathcal{N}_\infty(1/2, \Phi, T) \leq \sum_{i=0}^{d_2} \binom{T}{i} k^i \leq (ekT)^{d_2} \quad , \qquad \mathcal{N}(0, \Phi, T) \leq \sum_{i=0}^{d_1} \binom{T}{i} k^i \leq (ekT)^{d_1} .$$

In particular, the result holds for binary-valued payoffs ($k = 1$), in which case $\mathrm{fat}_1(\Phi) = \mathrm{Sdim}(\Phi)$. We now show that the covering numbers are bounded in terms of the fat-shattering dimension.

Corollary 11 *Suppose $\mathcal{H} \subseteq [-1, 1]$. Then for any $\alpha > 0$, any $T > 0$, and any $(\mathcal{F}\times\mathcal{X})$-valued tree $(\mathbf{f},\mathbf{x})$ of depth T,*

$$\mathcal{N}_1(\alpha, \Phi, (\mathbf{f},\mathbf{x})) \leq \mathcal{N}_2(\alpha, \Phi, (\mathbf{f},\mathbf{x})) \leq \mathcal{N}_\infty(\alpha, \Phi, (\mathbf{f},\mathbf{x})) \leq \left(\frac{2eT}{\alpha}\right)^{\mathrm{fat}_\alpha(\Phi)}$$

The generality of these results is evident, as both the combinatorial parameters and covering numbers are defined for any performance measure (1) with time-invariant payoff transformations. In particular, this includes Φ-regret (see Section 4.1).

2. As an aside, the term "shattered set" was introduced by J. Michael Steele in his Ph.D. thesis in 1975.

3.4. Covering Number Bounds for Slowly-Varying Payoff Transformations

In this subsection, we lift the assumption of time-invariance and observe that size of Φ_T or an appropriately behaving covering number $\mathcal{N}(\alpha, \Phi_T, T)$ is key for bounding the sequential complexity. If payoff transformations change wildly in time, there is little hope of getting non-trivial bounds. Under some assumptions on the variability of the sequences in Φ_T, we can get a bound on the covering number of Φ_T. It has been shown in Herbster and Warmuth (1998); Bousquet and Warmuth (2002) that it is possible to have small external regret against comparators that change a limited number of times. In Zinkevich (2003), *dynamic regret* is defined with respect to a comparator whose path length is bounded. In general, one can consider situations where we would like to compete with a budgeted comparator. We now show that the assumptions of slowly-varying or budgeted comparators are naturally captured by our framework through the notion of slowly-changing payoff transformations Φ_T. Furthermore, the control of covering numbers of Φ_T becomes transparent under such assumptions. Our goal here is not to provide a comprehensive list of possible results, but rather to show versatility of our framework.

Suppose Φ_T consists of payoff transformations $(\phi_1, \dots, \phi_T)$ which are "almost" time-invariant within each of $k+1$ intervals. Consider the following definition:

$$\Phi_T^{k,\alpha} = \Big\{(\phi_1, \dots, \phi_T) : 1 = i_0 \le \dots \le i_k \le T, \\ \sup_{f,x} \|\ell_{\phi_t}(f,x) - \ell_{\phi_{t'}}(f,x)\| \le \alpha \text{ if } i_s \le t \le t' < i_{s+1} \text{ for } s \ge 0\Big\}.$$

One can think of the time-invariant segments as "accumulation points" where the payoff transformations do not vary much. This, of course, includes the case when Φ_T is constant over the $k+1$ intervals by setting $\alpha = 0$. The following result controls the covering number of $\Phi_T^{k,\alpha}$.

Lemma 12 *If $\mathcal{N}_\infty(\alpha, \Phi, T)$ is finite, $\mathcal{N}_\infty(2\alpha, \Phi_T^{k,\alpha}, T) \le \binom{T}{k} \cdot \mathcal{N}_\infty(\alpha, \Phi, T)^{k+1}$.*

Further extending the above results, we will now study the size of an online cover if Φ_T consists of payoff transformations of bounded length. In general, "length" can be defined as some budget given by the setting at hand. Here, we present a straightforward approach without an attempt to give very general and tight bounds. The length of a sequence $(\phi_1, \dots, \phi_T)$ of payoff transformations (with respect to L_∞ distance) is defined as $\mathrm{len}(\phi_1, \dots, \phi_T) := \sum_{t=1}^{T-1} \sup_{f,x} \left\|\ell_{\phi_t}(f,x) - \ell_{\phi_{t+1}}(f,x)\right\|$.

Lemma 13 *Assume that for all $(\phi_1, \dots, \phi_T) \in \Phi_T$, we have $\mathrm{len}(\phi_1, \dots, \phi_T) \le L$. Then, we have,*

$$\mathcal{N}_\infty(2\alpha, \Phi_T, T) \le \binom{T}{L/\alpha} \cdot \mathcal{N}_\infty(\alpha, \Phi, T)^{L/\alpha+1} .$$

4. Examples and Comparison to Known Results

We now turn to several specific settings studied in the literature and look at them through the prism of our general results. While we believe that online learnability in many different scenarios can be established through our framework, we decided to focus on several major problems. On the surface, these problems are quite different; yet, through our unified approach we show that learnability can be seamlessly established for all of them. The unification not only leads to simpler proofs and sharper results, but also yields insight into the inherent complexity of problems.

4.1. Φ-Regret

In this section, we consider a particular notion of performance measure, known as Φ-regret (Stoltz and Lugosi, 2007; Gordon et al., 2008; Hazan and Kale, 2007). In our framework, this means that we restrict ourselves to only *time-invariant departure mapping classes* Φ_T specified by a base class Φ of mappings from $\mathcal{F}$ to itself. The particular choices of Φ lead to various notions, such as external, internal, swap regret, and more. To define Φ-regret, we fix a set Φ of departure mappings which map $\mathcal{F}$ to $\mathcal{F}$ and define the set of time-invariant departure mappings $\Phi_T := \{(\phi, \ldots, \phi) : \phi \in \Phi\}$. Then the measure of performance becomes Φ-regret (Eq. (3)). Since $\mathbf{B}$ is the average of its arguments, Corollary 7 implies that in the setting of Φ-regret, $\mathcal{V}_T(\ell, \Phi) \le 2\mathfrak{R}(\ell, \Phi)$. Specializing the definition of sequential complexity to Φ-regret, we obtain:

Definition 14 *The sequential complexity for Φ-regret is defined as*

$$\mathfrak{R}_T(\ell, \Phi) = \sup_{(\mathbf{f},\mathbf{x})} \; \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell(\phi \circ \mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \; . \tag{10}$$

Sequential complexity for Φ-regret enjoys some of the nice properties of the sequential Rademacher complexity for external regret. Suppose ℓ is convex in the first argument and $\mathrm{conv}(\Phi)$ maps $\mathcal{F}$ into $\mathcal{F}$. Then $\mathfrak{R}_T(\ell, \mathrm{conv}(\Phi)) = \mathfrak{R}_T(\ell, \Phi)$. This allows us to obtain bounds for convex hulls of finite sets Φ.

To capture complexity via covering numbers, Definition 5 can be specialized to the case of Φ-regret:

Definition 15 *A set V of $\mathbb{R}$-valued trees of depth T is* an α-cover *(with respect to ℓ_p-norm) of Φ_T on the $\mathcal{F} \times \mathcal{X}$-valued tree $(\mathbf{f}, \mathbf{x})$ of depth T if*

$$\forall \phi \in \Phi, \; \forall \epsilon \in \{\pm 1\}^T \; \exists \mathbf{v} \in V \text{ s.t.} \quad \frac{1}{T} \sum_{t=1}^{T} |\mathbf{v}_t(\epsilon) - \ell(\phi \circ \mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))|^p \le \alpha^p$$

4.1.1. External Regret

External regret is the simplest example of Φ-regret. We separate it from the general discussion in order to show that for external regret the various notions introduced in this paper reduce to the ones proposed in *RST*. Considering the definitions in Example 1, notice that the time-invariant departure mappings class Φ_T is chosen to be the class of sequences of *constant* mappings $\{(\phi_f, \ldots, \phi_f) : f \in \mathcal{F} \text{ and } \phi_f(g) = f \;\forall g \in \mathcal{F}\}$. It is precisely because of this constancy of $\boldsymbol{\phi}$ that the dependence on the $\mathcal{F}$-valued tree $\mathbf{f}$ disappears from all the definitions and results. Further, because of the obvious bijection between elements of Φ_T and $\mathcal{F}$, minimization (maximization) over Φ_T can be written as minimization (maximization) over $\mathcal{F}$. Notice that the action of ϕ_f on the payoff is $\ell_{\phi_f}(f_t, x_t) = \ell(f, x_t)$. Let us turn to Definition 14 of the sequential complexity for Φ-regret. Because each $\phi_f \in \Phi$ is a constant mapping, we have

$$\mathfrak{R}_T(\ell, \Phi) = \sup_{\mathbf{f},\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{f\in\mathcal{F}} \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)) = \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{f\in\mathcal{F}} \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)). \tag{11}$$

If payoff is written as $\ell(f, x) = f(x)$, this is precisely the sequential Rademacher complexity defined in *RST*. Next, we show that Definition 15 reduces to the definition of online covering given in *RST*. Indeed, $\ell_{\phi_f}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) = \ell(f, \mathbf{x}_t(\epsilon))$ for the constant mappings $\boldsymbol{\phi} = (\phi_f, \ldots, \phi_f)$. Further, the payoff space $\mathcal{H} \subseteq \mathbb{R}$. With these simplifications, the closeness to a covering element in Definition 15 becomes

$$\forall f \in \mathcal{F},\ \forall \epsilon \in \{\pm 1\}^T\ \exists \mathbf{v} \in V \text{ s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} |\mathbf{v}_t(\epsilon) - \ell(f, \mathbf{x}_t(\epsilon))|^p \le \alpha^p$$

where V is a set of $\mathbb{R}$-valued trees. It is then immediate that Corollary 7 recovers the corresponding result of *RST*. For a detailed study of external regret, we refer to the companion paper *RST*.

4.1.2. Internal and Swap Regret

Assume the cardinality $N = |\mathcal{F}|$ is finite. For internal regret, Φ is the set of mappings $\{\phi_{f\to g} : \phi_{f\to g}(f) = g \quad \text{and} \quad \phi_{f\to g}(h) = h \quad \forall h \neq f, h \in \mathcal{F}\}$. For swap regret (Blum and Mansour, 2005; Cesa-Bianchi and Lugosi, 2006), Φ contains all N^N functions from $\mathcal{F}$ to itself. It is easy to see that applying Corollary 7 in the finite class case ($|\Phi_T| < \infty$) immediately recovers the $O(\sqrt{T \log N})$ bound for internal and external regret and the $O(\sqrt{TN \log N})$ bound for the swap regret (Cesa-Bianchi and Lugosi, 2006). Our general tools, however, allow us to go well beyond finite sets of departure mappings. In the following sections, we consider several examples of infinite classes of departure mappings which have been considered in the literature. In some of these cases, an explicit strategy requires computation of a fixed-point (Foster and Vohra, 1997; Hazan and Kale, 2007; Gordon et al., 2008). Since we are not providing efficient algorithms in order to obtain bounds, we are able to get sharp results by directly focusing on the complexity of these infinite classes of departure mappings.

4.1.3. Convergence to Φ-correlated Equilibria

A beautiful result of Foster and Vohra (1997) shows that convergence to the set of correlated equilibria can be achieved if players follow *internal* regret minimization strategies. What is surprising, no coordination is required to achieve this goal. Stoltz and Lugosi (2007) extended this result to compact and convex sets of strategies in normed spaces. In this section we show that their results can be improved in certain situations. Let us consider their setting in a bit more detail. Suppose there are N players each playing in a strategy set $\mathcal{F}$. We could make the strategy set player dependent but it only complicates notation. There are N loss functions mapping a strategy profile $(f_1, \ldots, f_N)$ to $\{\ell_k(f_1, \ldots, f_N)\}_{k=1}^N$, the losses for each of the N players. Consider a set of departure mappings $\Phi \subseteq \{\phi : \mathcal{F} \to \mathcal{F}\}$. A Φ-correlated equilibrium is a distribution π over strategy profiles such that if the player jointly play according to it, no player has an incentive to unilaterally transform its action using a mapping from Φ. That is, $\mathbb{E}_{(f_1,\ldots,f_N)\sim\pi}[\ell_k(f_k, f_{-k})] \le \mathbb{E}_{(f_1,\ldots,f_N)\sim\pi}[\ell_k(\phi(f_k), f_{-k})]$ for all $k \in [N], \phi \in \Phi$. Theorem 18 in Stoltz and Lugosi (2007) shows the following. If $\mathcal{F}$ is convex compact subset of a normed vector space, ℓ_k's are continuous and Φ is a separable subset of $\mathcal{C}(\mathcal{F})$[3], then there exist regret minimizing algorithms such that, if every player follows the algorithm then the sequence of empirical plays jointly converges to the set of Φ-correlated equilibria.

Consider the case where $\mathcal{F}$ is some compact subset of the unit ball in some normed space with a norm $\|\cdot\|$, the loss function ℓ_k is a 1-Lipschitz convex function, and the class Φ of departure functions has finite metric entropy $\mathcal{N}_{\text{metric}}(\Phi, \alpha)$ for all $\alpha > 0$. Metric entropy is simply the log covering number where covers of Φ are built for the supremum norm $\|\phi\|_\infty = \sup_{f\in\mathcal{F}} \|\phi(f)\|$. Let us consider a typical situation where $\mathcal{N}_{\text{metric}}(\Phi, \alpha) = \Theta(1/\alpha^p)$. The adversary's set $\mathcal{X}$ here is simply $\{f \mapsto \ell_k(f, g) \ : \ g \in \mathcal{F}^{k-1}\}$, where g is a strategy profile over the remaining $k-1$ players. To upper bound the Φ-regret we can always make the set of adversary's moves larger. In fact, we may set $\mathcal{X} = \mathcal{C}_\mathcal{F}$, where $\mathcal{C}_\mathcal{F} = \{x : \mathcal{F} \to \mathbb{R} \ : \ x \text{ convex and 1-Lipschitz}\}$. Moreover, the value of the convex-Lipschitz game is equal to the value of the linear game (see Rakhlin et al. (2010b)): $\mathcal{V}_T(\mathcal{C}_\mathcal{F}, \mathcal{F}, \Phi) = \mathcal{V}_T(\mathcal{L}_\mathcal{F}, \mathcal{F}, \Phi)$ where $\mathcal{L}_\mathcal{F} = \{x : \mathcal{F} \to \mathbb{R} \text{ is linear and 1-Lipschitz}\}$. Then the sequential complexity bound is

$$\sup_{(\mathbf{f},\mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi\in\Phi} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle \phi(\mathbf{f}_t(\epsilon)), \mathbf{x}_t(\epsilon) \right\rangle \ . \tag{12}$$

Note that the set $\mathcal{X}$ is now just the set of 1-Lipschitz linear functions. Since $\|\phi_1 - \phi_2\|_\infty \le \alpha$ implies $|\langle \phi_1(f), x\rangle - \langle \phi_2(f), x\rangle| \le \alpha$ for any $x \in \mathcal{X}$, we can use metric entropy inside Dudley's integral to upper bound the sequential complexity by $c \inf_\alpha\{\alpha T + \sqrt{T}\int_{\alpha'=\alpha}^{1} \sqrt{1/\alpha'^p} d\alpha'\}$. This behaves as $O(\sqrt{T})$, if $p < 2$, as $O(\sqrt{T\log(T)})$ if $p = 2$, and as $O(T^{(p-1)/p})$ if $p > 2$. These are better than the $O(T^{(p+1)/(p+2)})$ bound (derived using an explicit learning algorithm) given in Example 23 of Stoltz and Lugosi (2007).

4.2. Blackwell's Approachability

Blackwell's Approachability Theorem (Blackwell, 1956; Mertens et al., 1994; Lehrer, 2003; Cesa-Bianchi and Lugosi, 2006) is a fundamental result for repeated two-player zero-sum games. By means of this theorem, learnability (Hannan consistency) can be established for a wide array of problems, as illustrated in Cesa-Bianchi and Lugosi (2006). For instance, existence of calibrated forecasters can be deduced from Blackwell's Approachability Theorem (Mannor and Stoltz, 2010; Foster and Vohra, 1997). Let us first discuss the relation of our results to Blackwell's Theorem. A proof of Blackwell's Theorem (e.g. Cesa-Bianchi and Lugosi (2006)) reveals that (a) martingale convergence has to take place in the payoff space, and (b) the so-called Blackwell's one-shot approachability condition has to be satisfied. The former is closely related to the first term in our Triplex Inequality, while the latter is related to the second term (ability to play well if the next move is known). What is interesting, in the literature, Blackwell's Theorem is applied by embedding the problem at hand into an often high-dimensional space. The dimensionality represents the complexity of the problem, but this embedding is often artificial. In contrast, the problem complexity is captured by

3. The set of continuous functions on $\mathcal{F}$ equipped with the supremum norm.

the third term of our decomposition, the *sequential complexity*, and it is explicitly written as a complexity measure rather than an embedding into some other space. The ability to upper bound problem complexity with tools similar to those developed in RST (e.g. covering numbers) means that learnability can be established for a wide class of problems. In this section, we show that Blackwell's approachability can be viewed as an online game with a particular performance measure (distance to the set). Using the techniques developed in this paper, we prove Blackwell's approachability in Banach spaces for which martingale convergence holds (Theorem 16). We also show that martingale convergence is necessary for the result to hold (Theorem 17). To the best of our knowledge, both of these results are novel. To define the problem precisely, suppose $\mathcal{H}$ a subset of a Banach space $\mathcal{B}$ and $S \subset \mathcal{B}$ is a closed convex set. For the moves $f \in \mathcal{F}$ of the player and $x \in \mathcal{X}$ of the adversary, $\ell(f, x) \in \mathcal{H}$ is a Banach space valued signal. The goal of the player is to keep the average of the signals $\frac{1}{T}\sum_{t=1}^{T}\ell(f_t, x_t)$ close to the set S. By defining $\mathbf{B}$ as in Example 3, $\mathbf{R}_T$ becomes distance of the average payoff to the set (see Eq. (4)). The comparator term is zero by our assumption that Φ_T contains sequences $(\phi_1, \ldots, \phi_T)$ of constant mappings which transform our actions to a point inside S: $\ell_{\phi_t}(f, x) = c_t \in S$ for all $f \in \mathcal{F}$, $x \in \mathcal{X}$, and $1 \leq t \leq T$.

The Blackwell's approachability game is said to be *one shot approachable* if for every mixed strategy p of the adversary, there exists a mixed strategy q for a player such that $\ell(q, p) \in S$. This condition says that the player should be able to choose a "good" mixed strategy q in response to a given adversarial strategy p. Recall that $\ell(q, p)$ is simply a shorthand for the expected payoff $\mathbb{E}_{f\sim q, x\sim p}\ell(f, x)$ (we make no assumptions about linearity of ℓ). Blackwell's one-shot approachability condition is akin the second term in the Triplex Inequality, where the order of who plays first is switched. If the one-shot condition is satisfied, it remains to check martingale convergence. We now show that, under the one-shot approachability condition, a variation of the worst-case martingale in the subset of the Banach space provides an upper bound on the distance to the set.

Theorem 16 *For any game that is one shot approachable, we have that*

$$\mathcal{V}_T(\ell, \Phi_T) \leq 4 \sup_{\mathbf{M}} \mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^{T} d_t\right\|\right]$$

where sup *is over distributions* $\mathbf{M}$ *of* $\mathrm{conv}(\mathcal{H} \bigcup -\mathcal{H})$*-valued martingale difference sequences* $\{d_t\}_{t\in\mathbb{N}}$.

The notion of approachability considered in this paper corresponds to weak approachability. Extending the techniques of this work to a slightly different notion of a value (see Rakhlin et al. (2010b)), we can guarantee almost sure convergence and, hence, strong approachability.

It is straightforward that for any Blackwell's approachability game to have vanishing regret, one shot approachability for the game is a necessary condition. We now show that martingale convergence in the space of payoffs is necessary for Blackwell's approachability. To the best of our knowledge, this result has not appeared in the literature.

Theorem 17 *For every symmetric convex set* $\mathcal{H}$ *there exists a one shot approachable game with payoff's mapping to* $\mathcal{H}$ *such that*

$$\mathcal{V}_T(\ell, \Phi_T) \geq \frac{1}{2} \sup_{\mathbf{M}} \mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^{T} d_t\right\|\right]$$

where sup *is over distributions* $\mathbf{M}$ *of* $\mathcal{H}$*-valued martingale difference sequences* $\{d_t\}_{t\in\mathbb{N}}$.

4.3. Calibration

Calibration is an important notion for forecasting binary sequences (Dawid, 1982). Example 4 corresponds to the notion of λ-calibration for $\{1,\ldots,k\}$-valued sequences (Cesa-Bianchi and Lugosi, 2006) and defines the measure of performance $\mathbf{R}$. We are interested in sharp rates on the value of the calibration game and compare our results with the recent work of Mannor and Stoltz (2010). Note that the definition of value allows the worst scale λ to be chosen at the end of the game, making it a stronger requirement than what is required for building a well calibrated forecaster. Using our techniques, for the ℓ_1-calibration game with k outcomes, for $T \geq 3$ and some absolute constant c, we show that $\mathcal{V}_T(\ell, \Phi_T) \leq ck^2\left((\log T)/T\right)^{1/2}$. That is, the rate of calibration is $\tilde{O}(T^{-1/2})$. For $k>2$, the best rates known to us (due to Mannor and Stoltz (2010)) deteriorate with k because the authors in fact calibrate with respect to all Borel sets.

4.4. External Regret with Global Costs

Let us first state a more general setting where the (vector) loss is $\ell(f,x)$ rather than the specific choice $f\odot x$ in Example 5. To state the result we need the following Assumption.

Assumption 1 *For any* p_1, p_2, $\inf_f \|\ell(f,p_1)+\ell(f,p_2)\| \geq \inf_f \|\ell(f,p_1)\| + \inf_f \|\ell(f,p_2)\|$.

Theorem 18 *For the setting of Example 5 with vector valued loss* $\ell(f,x)$*, under Assumption 1 :*

$$\mathcal{V}_T \;\leq\; 4\sup_{\mathbf{M}} \mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^T d_t\right\|\right] \;+2\;\sup_{\mathbf{x}}\;\mathbb{E}_{\epsilon_{1:T}}\sup_{f\in\mathcal{F}}\left\|\frac{1}{T}\sum_{t=1}^T \epsilon_t \ell(f,\mathbf{x}_t(\epsilon))\right\|\;, \tag{13}$$

where sup *is over distributions* $\mathbf{M}$ *of* $\mathrm{conv}(\mathcal{H}\bigcup -\mathcal{H})$*-valued martingale difference sequences* $\{d_t\}_{t\in\mathbb{N}}$.

Let us see what this implies in a specific case of Example 5, the setting studied in Even-Dar et al. (2009), i.e. $\ell(f,x) = f\odot x$. Let us first verify if Assumption 1 holds here. By linearity of the vector loss, we just have to verify whether, for arbitrary p_1, p_2, we have

$$\inf_{q\in\Delta(k)}\left\|q\odot \underline{p_1} + q\odot\underline{p_2}\right\| \geq \inf_{q\in\Delta(k)}\left\|q\odot\underline{p_1}\right\| + \inf_{q\in\Delta(k)}\left\|q\odot\underline{p_2}\right\| \;.$$

where the notation $\underline{p_i}$ stands for the mean of the distribution p_i. This is equivalent to asking whether the function $x \mapsto \inf_{f\in\mathcal{F}}\|f\odot x\|$ is *concave*. Lemma 22 in the appendix proves that it is. Note that in Even-Dar et al. (2009), it is shown that the above function is concave for the ℓ_p norms (including $p=\infty$). It turns out that it remains concave no matter what norm is chosen. Thus, the general upper bound (13) holds. In the case we are considering, we can further massage the second term in that upper bound. Note that for any f and y, $\|f\odot y\| \leq \|f\|_\infty\|y\| \leq \|y\|$. Using this in (13) we see that

$$\mathcal{V}_T \;\leq\; 4\sup_{\mathbf{M}}\mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^T d_t\right\|\right] + 2\;\sup_{\mathbf{x}}\;\mathbb{E}_{\epsilon_{1:T}}\left\|\frac{1}{T}\sum_{t=1}^T \epsilon_t\mathbf{x}_t(\epsilon)\right\| \leq\; 6\;\sup_{\mathbf{M}}\mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^T d_t\right\|\right]$$

where the last inequality is because $(\epsilon_t \mathbf{x}_t(\epsilon))_{t=1}^T$ is a martingale difference sequence. In the last inequality the supremum is over distributions $\mathbf{M}$ of $[-1,1]^k$-valued martingale difference sequences $\{d_t\}_{t\in\mathbb{N}}$. For ℓ_p norms we recover the rates in Even-Dar et al. (2009), specifically for ℓ_∞ norm the bound is $6\sqrt{\log(k)/T}$

Acknowledgments

We thank Dean Foster for many insightful discussions about calibration and Blackwell's approachability. A. Rakhlin gratefully acknowledges the support of NSF under grant CAREER DMS-0954737 and Dean's Research Fund.

References

J. Abernethy, A. Agarwal, P. Bartlett, and A. Rakhlin. A stochastic view of optimal regret through minimax duality. In *Proceedings of the 22nd Annual Conference on Learning Theory*, 2009.

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

D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.

A. Blum and Y. Mansour. From external to internal regret. In *Proceedings of the 18th Annual Conference on Learing Theory*, pages 621–636. Springer, 2005.

O. Bousquet and M. K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3:363–396, 2002.

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

A.P. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379):605–610, 1982.

E. Even-Dar, R. Kleinberg, S. Mannor, and Y. Mansour. Online learning for global cost functions. In *Proceedings of the 22nd Annual Conference on Learning Theory*, 2009.

D. P. Foster and R. V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2):40–55, October 1997.

E. Giné and J. Zinn. Some limit theorems for empirical processes. *Ann. Prob.*, 12(4): 929–989, 1984.

P.W. Goldberg and M.R. Jerrum. Bounding the Vapnik-Chervonenkis dimension of concept classes parameterized by real numbers. *Machine Learning*, 18(2):131–148, 1995.

G.J. Gordon, A. Greenwald, and C. Marks. No-regret learning in convex games. In *Proceedings of the 25th International Conference on Machine learning*, pages 360–367. ACM, 2008.

E. Hazan and S. Kale. Computational equivalence of fixed points and no regret algorithms, and convergence to equilibria. In *NIPS*, 2007.

M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine Learning*, 32(2): 151–178, 1998.

E. Lehrer. Approachability in infinite dimensional spaces. *International Journal of Game Theory*, 31(2):253–268, 2003.

N. Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2(4):285–318, 04 1988.

S. Mannor and G. Stoltz. A geometric proof of calibration. *Mathematics of Operations Research*, 35(4):721–727, 2010.

J.-F. Mertens, S. Sorin, and S. Zamir. Repeated games: Part A Background material. Technical Report 9420, CORE, Universite Catholique de Louvain, 1994.

I. Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Prob.*, 22(4):1679–1706, 1994.

G. Pisier. Martingales with values in uniformly convex spaces. *Israel Journal of Mathematics*, 20(3):326–350, 1975.

A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Random averages, combinatorial parameters, and learnability. In *NIPS*, 2010a. Full version available at arXiv:1006.1138.

A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Beyond regret. *ArXiv preprint arXiv:1011.3168v2*, 2010b.

G. Stoltz and G. Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*, 2003.

Appendix

Appendix A. Proofs for General Upper Bounds (Section 3)

Proof [**of Theorem 2**] The value of the game, defined in (2), is

$$\begin{aligned}
\mathcal{V}_T(\ell,\Phi_T) &= \inf_{q_1}\sup_{p_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\inf_{q_T}\sup_{p_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\sup_{\phi\in\Phi_T}\left\{\mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T))-\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T))\right\}\\
&= \sup_{p_1}\inf_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\sup_{\phi\in\Phi_T}\left\{\mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T))-\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T))\right\}
\end{aligned}$$

via an application of the minimax theorem. Adding and subtracting terms to the expression above leads to

$$\begin{aligned}
\mathcal{V}_T(\ell,\Phi_T) &= \sup_{p_1}\inf_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\left[\mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T))-\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))\right.\\
&\qquad\left.+\sup_{\phi\in\Phi_T}\left\{\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))-\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T))\right\}\right]\\
&\le \sup_{p_1}\inf_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\left[\mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T))-\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))\right.\\
&\qquad+\sup_{\phi\in\Phi_T}\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\left\{\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))-\mathbf{B}(\ell_{\phi_1}(f'_1,x'_1),\dots,\ell_{\phi_T}(f'_T,x'_T))\right\}\\
&\qquad\left.+\sup_{\phi\in\Phi_T}\left\{\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell_{\phi_1}(f'_1,x'_1),\dots,\ell_{\phi_T}(f'_T,x'_T))-\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T))\right\}\right]
\end{aligned}$$

At this point, we would like to break up the expression into three terms. To do so, notice that expectation is linear and sup is a convex function, while for the infimum,

$$\inf_a\left[C_1(a)+C_2(a)+C_3(a)\right]\le\left[\sup_a C_1(a)\right]+\left[\inf_a C_2(a)\right]+\left[\sup_a C_3(a)\right]$$

for functions C_1, C_2, C_3. We use these properties of inf, sup, and expectation, starting from the inside of the nested expression and splitting the expression in three parts. We arrive at

$$\begin{aligned}
&\mathcal{V}_T(\ell,\Phi_T)\\
&\le \sup_{p_1}\sup_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\sup_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\left[\mathbf{B}(\ell(f_1,x_1),\dots,\ell(f_T,x_T))-\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))\right]\\
&+\sup_{p_1}\inf_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\left[\sup_{\phi\in\Phi_T}\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\left\{\mathbf{B}(\ell(f'_1,x'_1),\dots,\ell(f'_T,x'_T))-\mathbf{B}(\ell_{\phi_1}(f'_1,x'_1),\dots,\ell_{\phi_T}(f'_T,x'_T))\right\}\right]\\
&+\sup_{p_1}\sup_{q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\x_1\sim p_1}}\dots\sup_{p_T}\sup_{q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\x_T\sim p_T}}\left[\sup_{\phi\in\Phi_T}\left\{\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\x'_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell_{\phi_1}(f'_1,x'_1),\dots,\ell_{\phi_T}(f'_T,x'_T))-\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\dots,\ell_{\phi_T}(f_T,x_T))\right\}\right]
\end{aligned}$$

The replacement of infima by suprema in the first and third terms appears to be a loose step and, indeed, one can pick a particular response strategy $\{q_t^*\}$ instead of passing to the supremum. For instance, this can be the best-response strategy for the second term. However, in the examples we have considered so far, passing to the supremum still yields the results we need. This is due to the fact that the online learning setting is worst-case.

Consider the second term in the above decomposition. We claim that

$$\sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \dots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\sup_{\phi \in \Phi_T} \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \left[\mathbf{B}(\ell(f'_1, x'_1), \dots, \ell(f'_T, x'_T)) - \mathbf{B}(\ell_{\phi_1}(f'_1, x'_1), \dots, \ell_{\phi_T}(f'_T, x'_T)) \right] \right]$$

$$= \sup_{p_1} \inf_{q_1} \dots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{1:T} \\ x_{1:T} \sim p_{1:T}}} \left[\mathbf{B}(\ell(f_1, x_1), \dots, \ell(f_T, x_T)) - \mathbf{B}(\ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f_T, x_T)) \right]$$

because the objective

$$\mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \left[\mathbf{B}(\ell(f'_1, x'_1), \dots, \ell(f'_T, x'_T)) - \mathbf{B}(\ell_{\phi_1}(f'_1, x'_1), \dots, \ell_{\phi_T}(f'_T, x'_T)) \right]$$

does not depend on the random draws $f_1, x_1, \dots, f_T, x_T$. We then rename f'_t, x'_t into f_t, x_t. This concludes the proof of the Triplex Inequality. ■

Proof [**of Theorem 3**] We turn to the third term in the Triplex Inequality. If $\mathbf{B}$ is subadditive,

$$\begin{aligned} &\mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \mathbf{B}(\ell_{\phi_1}(f'_1, x'_1), \dots, \ell_{\phi_T}(f'_T, x'_T)) - \mathbf{B}(\ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f_T, x_T)) \\ &\quad \le \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \mathbf{B}(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T)). \end{aligned}$$

If, on the other hand, $-\mathbf{B}$ is subadditive,

$$\begin{aligned} &\mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \mathbf{B}(\ell_{\phi_1}(f'_1, x'_1), \dots, \ell_{\phi_T}(f'_T, x'_T)) - \mathbf{B}(\ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f_T, x_T)) \\ &\quad \le - \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \mathbf{B}(\ell_{\phi_1}(f_1, x_1) - \ell_{\phi_1}(f'_1, x'_1), \dots, \ell_{\phi_T}(f_T, x_T) - \ell_{\phi_T}(f'_T, x'_T)). \end{aligned} \tag{14}$$

Below assume that $\mathbf{B}$ is subadditive, and the proof of the other case is identical.

To prove the bound on the third term in terms of twice sequential complexity, we proceed as in Rakhlin et al. (2010a), applying the symmetrization technique from inside out. To this end, first note that,

$$\sup_{p_1, q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \dots \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \sup_{\phi \in \Phi_T} \mathop{\mathbb{E}}_{\substack{f'_1 \sim q_1, \dots, f'_T \sim q_T \\ x'_1 \sim p_1, \dots x'_T \sim p_T}} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T) \Big)$$

$$\le \sup_{p_1, q_1} \mathop{\mathbb{E}}_{\substack{f_1, f'_1 \sim q_1 \\ x_1, x'_1 \sim p_1}} \dots \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\phi \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \dots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T) \Big) .$$

The above is true because the expectations are pulled outside the suprema, thus resulting in an upper bound. Now notice that conditioned on history f_T, f'_T are distributed identically and independently drawn from q_T. Similarly x_T, x'_T are also identically distributed conditioned on history. Hence renaming them we see that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T)\Big) \\
&\quad = \mathop{\mathbb{E}}_{\substack{f'_T, f_T \sim q_T \\ x'_T, x_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T) - \ell_{\phi_T}(f'_T, x'_T)\Big) \\
&\quad = \mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, -(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big)
\end{aligned}$$

where only the last argument of $\mathbf{B}$ is changing sign. Thus,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T)\Big) \\
&\quad = \mathbb{E}_{\epsilon_T} \mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big)
\end{aligned}$$

where ϵ_T is a Rademacher random variable. Furthermore,

$$\begin{aligned}
&\sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T)\Big) \\
&\quad = \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f'_T, f_T \sim q_T \\ x'_T, x_T \sim p_T}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big) \\
&\quad \leq \sup_{\substack{x_T, x'_T \in \mathcal{X} \\ f_T, f'_T \in \mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi} \in \Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big)
\end{aligned}$$

Proceeding similarly notice that since given history x_{T-1}, x'_{T-1} and f_{T-1}, f'_{T-1} are distributed independently and identically we have,

$$\begin{aligned}
&\sup_{p_{T-1}, q_{T-1}} \mathop{\mathbb{E}}_{\substack{f_{T-1}, f'_{T-1} \sim q_{T-1} \\ x_{T-1}, x'_{T-1} \sim p_{T-1}}} \sup_{\substack{x_T, x'_T \in \mathcal{X} \\ f_T, f'_T \in \mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi} \in \Phi_T} \\
&\qquad \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_{T-1}}(f'_{T-1}, x'_{T-1}) - \ell_{\phi_{T-1}}(f_{T-1}, x_{T-1}), \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big) \\
&= \sup_{p_{T-1}, q_{T-1}} \mathop{\mathbb{E}}_{\substack{f_{T-1}, f'_{T-1} \sim q_{T-1} \\ x_{T-1}, x'_{T-1} \sim p_{T-1}}} \mathbb{E}_{\epsilon_{T-1}} \sup_{\substack{x_T, x'_T \in \mathcal{X} \\ f_T, f'_T \in \mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi} \in \Phi_T} \\
&\qquad \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_{T-1}(\ell_{\phi_T}(f'_{T-1}, x'_{T-1}) - \ell_{\phi_{T-1}}(f_{T-1}, x_{T-1})), \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big) \\
&\leq \sup_{\substack{x_{T-1}, x'_{T-1} \in \mathcal{X} \\ f_{T-1}, f'_{T-1} \in \mathcal{F}}} \mathbb{E}_{\epsilon_{T-1}} \sup_{\substack{x_T, x'_T \in \mathcal{X} \\ f_T, f'_T \in \mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi} \in \Phi_T} \\
&\qquad \mathbf{B}\Big(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_{T-1}(\ell_{\phi_{T-1}}(f'_{T-1}, x'_{T-1}) - \ell_{\phi_{T-1}}(f_{T-1}, x_{T-1})), \epsilon_T(\ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T))\Big)
\end{aligned}$$

Proceeding in similar fashion introducing Rademacher random variables all the way to ϵ_1 we arrive at

$$\sup_{p_1,q_1} \mathop{\mathbb{E}}_{\substack{f_1,f_1'\sim q_1\\ x_1,x_1'\sim p_1}} \ldots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{\substack{f_T,f_T'\sim q_T\\ x_T,x_T'\sim p_T}} \sup_{\boldsymbol{\phi}\in\Phi_T} \mathbf{B}\Big(\ell_{\phi_1}(f_1',x_1') - \ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_T}(f_T',x_T') - \ell_{\phi_T}(f_T,x_T)\Big)$$

$$\le \sup_{\substack{x_1,x_1'\in\mathcal{X}\\ f_1,f_1'\in\mathcal{F}}} \mathbb{E}_{\epsilon_1} \ldots \sup_{\substack{x_T,x_T'\in\mathcal{X}\\ f_T,f_T'\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi}\in\Phi_T} \mathbf{B}\Big(\epsilon_1(\ell_{\phi_1}(f_1',x_1') - \ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f_T',x_T') - \ell_{\phi_T}(f_T,x_T))\Big)$$

Subadditivity of $\mathbf{B}$ implies $\mathbf{B}(a-b) \le \mathbf{B}(a) + \mathbf{B}(-b)$, and thus

$$\begin{aligned}
&\mathbf{B}\Big(\epsilon_1(\ell_{\phi_1}(f_1',x_1') - \ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f_T',x_T') - \ell_{\phi_T}(f_T,x_T))\Big)\\
&\quad\le \mathbf{B}\Big(\epsilon_1\ell_{\phi_1}(f_1',x_1'),\ldots,\epsilon_T\ell_{\phi_T}(f_T',x_T')\Big) + \mathbf{B}\Big(-\epsilon_1\ell_{\phi_1}(f_1,x_1),\ldots,-\epsilon_T\ell_{\phi_T}(f_T,x_T)\Big)
\end{aligned}$$

We, therefore, arrive at

$$\begin{aligned}
&\sup_{\substack{x_1,x_1'\in\mathcal{X}\\ f_1,f_1'\in\mathcal{F}}} \mathbb{E}_{\epsilon_1} \ldots \sup_{\substack{x_T,x_T'\in\mathcal{X}\\ f_T,f_T'\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi}\in\Phi_T} \mathbf{B}\Big(\epsilon_1(\ell_{\phi_1}(f_1',x_1') - \ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f_T',x_T') - \ell_{\phi_T}(f_T,x_T))\Big)\\
&\quad\le 2 \sup_{f_1\in\mathcal{F},x_1\in\mathcal{X}} \mathbb{E}_{\epsilon_1} \ldots \sup_{f_T\in\mathcal{F},x_T\in\mathcal{X}} \mathbb{E}_{\epsilon_T} \sup_{\boldsymbol{\phi}\in\Phi_T} \mathbf{B}\Big(\epsilon_1\ell_{\phi_1}(f_1,x_1),\ldots,\epsilon_T\ell_{\phi_T}(f_T,x_T)\Big)\\
&\quad= 2\sup_{(\mathbf{f},\mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{\boldsymbol{\phi}\in\Phi_T} \mathbf{B}\Big(\epsilon_1\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\epsilon_T\ell_{\phi_T}(\mathbf{f}_T(\epsilon),\mathbf{x}_T(\epsilon))\Big)
\end{aligned}$$

where in the last step we passed to the supremum over $(\mathcal{F}\times\mathcal{X})$-valued trees. This concludes the proof for the case of $\mathbf{B}$ being subadditive. Starting from Eq. (14), the proof for the case of $-\mathbf{B}$ being subadditive and convex in each of its coordinates leads to the bound of

$$2\sup_{(\mathbf{f},\mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{\boldsymbol{\phi}\in\Phi_T} -\mathbf{B}\Big(\epsilon_1\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\epsilon_T\ell_{\phi_T}(\mathbf{f}_T(\epsilon),\mathbf{x}_T(\epsilon))\Big).$$

The complete proof can be repeated for the first term in the Triplex Inequality in order to bound it by $2\mathfrak{R}_T(\ell,\mathcal{I},\mathbf{B})$ (or respectively $2\mathfrak{R}_T(\ell,\mathcal{I},-\mathbf{B})$). ■

Proof [**of Lemma 4**] The lemma follows directly from a result on concentration of 2-smooth functions of martingales, due to Pinelis Pinelis (1994). A detailed proof appears in Rakhlin et al. (2010b). ■

Proof [**of Theorem 6**] Define $\beta_0 = \eta$ and $\beta_j = 2^{-j}$. For a fixed tree $(\mathbf{f},\mathbf{x})$ of depth T, let V_j be an ℓ_∞-cover at scale β_j. For any path $\epsilon\in\{\pm1\}^T$ and any $\boldsymbol{\phi}\in\Phi_T$, let $\mathbf{v}[\boldsymbol{\phi},\epsilon]^j \in V_j$ a

β_j-close element of the cover in the ℓ_∞ sense. Now, for any $\boldsymbol{\phi} \in \Phi_T$,

$$
\begin{aligned}
& G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right) \\
& \leq G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t (\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \mathbf{v}[\phi,\epsilon]_t^N)\right) + \sum_{j=1}^{N} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1}\right)\right) \\
& \leq \left\|\frac{1}{T}\sum_{t=1}^{T}\epsilon_t (\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \mathbf{v}[\phi,\epsilon]_t^N)\right\| + \sum_{j=1}^{N} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1}\right)\right) \\
& \leq \max_{t=1}^{T}\left\|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \mathbf{v}[\phi,\epsilon]_t^N\right\| + \sum_{j=1}^{N} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t(\mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1})\right)
\end{aligned}
$$

Thus,

$$
\sup_{\boldsymbol{\phi}\in\Phi_T} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right) \leq \beta_N + \sup_{\phi\in\Phi_T}\left\{\sum_{j=1}^{N} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t(\mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1})\right)\right\}
$$

We now proceed to upper bound the second term. Consider all possible pairs of $\mathbf{v}^s \in V_j$ and $\mathbf{v}^r \in V_{j-1}$, for $1 \leq s \leq |V_j|$, $1 \leq r \leq |V_{j-1}|$, where we assumed an arbitrary enumeration of elements. For each pair $(\mathbf{v}^s, \mathbf{v}^r)$, define a real-valued tree $\mathbf{w}^{(s,r)}$ by

$$
\mathbf{w}_t^{(s,r)}(\epsilon) = \begin{cases} \mathbf{v}_t^s(\epsilon) - \mathbf{v}_t^r(\epsilon) & \text{if there exists } \boldsymbol{\phi} \in \Phi_T \text{ s.t. } \mathbf{v}^s = \mathbf{v}[\boldsymbol{\phi},\epsilon]^j, \mathbf{v}^r = \mathbf{v}[\boldsymbol{\phi},\epsilon]^{j-1} \\ 0 & \text{otherwise.} \end{cases}
$$

for all $t \in [T]$ and $\epsilon \in \{\pm 1\}^T$. It is crucial that $\mathbf{w}^{(s,r)}$ can be non-zero only on those paths ϵ for which $\mathbf{v}^s$ and $\mathbf{v}^r$ are indeed the members of the covers (at successive resolutions) close in the ℓ_2 sense *to some* $\boldsymbol{\phi} \in \Phi_T$. It is easy to see that $\mathbf{w}^{(s,r)}$ is well-defined. Let the set of trees W_j be defined as

$$
W_j = \left\{\mathbf{w}^{(s,r)} : 1 \leq s \leq |V_j|, 1 \leq r \leq |V_{j-1}|\right\}
$$

Using the above notations we see that

$$
\begin{aligned}
& \mathbb{E}_\epsilon\left[\sup_{\phi\in\Phi_T} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right)\right] \\
& \leq \beta_N + \mathbb{E}_\epsilon\left[\sup_{\phi\in\Phi_T}\left\{\sum_{j=1}^{N} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t(\mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1})\right)\right\}\right] \\
& \leq \beta_N + \mathbb{E}_\epsilon\left[\sum_{j=1}^{N}\sup_{\mathbf{w}^j\in W_j} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t \mathbf{w}_t^j(\epsilon)\right)\right] \qquad (15)
\end{aligned}
$$

Similarly to the corresponding proof in Rakhlin et al. (2010a), we can show that $\max_{t=1}^{T} \|\mathbf{w}_t^j(\epsilon)\| \leq 3\beta_j$ for any $\mathbf{w}^j \in \mathcal{W}^j$ and any path ϵ. Using concentration inequalities for 2-smooth functions in Banach spaces (see Pinelis (1994) or the full version Rakhlin et al. (2010b) of this extended abstract), we get

$$\begin{aligned}
\mathbb{E}_\epsilon\left[\sup_{\phi\in\Phi_T} G\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right)\right] &\leq \beta_N + \sum_{j=1}^{N}\beta_j\sqrt{\frac{\gamma\log(2|W_j|}{T}} \\
&\leq \beta_N + \sum_{j=1}^{N}\beta_j\sqrt{\frac{\gamma\log(2|V_j|\cdot|V_{j-1}|}{T}} \\
&\leq \beta_N + \frac{6\sqrt{\gamma}}{\sqrt{T}}\sum_{j=1}^{N}\beta_j\sqrt{\log(|V_j|)} \\
&\leq \beta_N + \frac{12\sqrt{\gamma}}{\sqrt{T}}\sum_{j=1}^{N}(\beta_j-\beta_{j+1})\sqrt{\log\mathcal{N}_\infty(\beta_j,\Phi_T,T)}\ .
\end{aligned}$$

Using standard arguments, this gives the bound,

$$\inf_{\alpha}\ 4\alpha + \frac{12\sqrt{\gamma}}{\sqrt{T}}\int_{\alpha}^{\eta}\sqrt{\log\mathcal{N}_\infty(\beta,\Phi_T,T)}d\beta\ .$$

■

Proof [**of Corollary 7**] The first statement is trivially verified. In fact, for this to hold we only require that $\mathbf{B}$ is subadditive, affine in its arguments, and $\mathbf{B}(0,\ldots,0)=0$. Indeed, the expectations can be sequentially moved inside of $\mathbf{B}$, making the coordinates of $\mathbf{B}$ zero, and making the suprema over the distributions irrelevant.

For the second claim, consider the second term in (8), specialized to the case of departure mappings:

$$\sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\sup_{\phi\in\Phi_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\left\{\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)-\ell(\phi_t(f_t),x_t)\right\} \tag{16}$$

Pick a particular (sub)optimal response q_t which puts all mass on $f_t^* = \arg\min_{f\in\mathcal{F}}\mathbb{E}_{x\sim p_t}\ell(f,x)$. It follows that $\ell(f_t,x_t)-\ell(\phi_t(f_t),x_t)\leq 0$, ensuring that the quantity in (16) is non-positive.

The third claim is a straightforward consequence of Theorem 6. Indeed, $\mathcal{H}\subset[-\eta,\eta]$ and G is the identity mapping, hence G^2 is $(2,2)$-smooth. ■

Proof [**of Lemma 12**] Fix an $(\mathcal{F}\times\mathcal{X})$-valued tree $(\mathbf{f},\mathbf{x})$ of depth T. Let $(i_0,\ldots,i_k)$ be the sequence which defines intervals of time-invariant mappings for the sequence $(\phi_1,\ldots,\phi_T)$. Fix $\epsilon\in\{\pm1\}^T$. Let $\mathbf{v}^{i_0},\ldots,\mathbf{v}^{i_k}\in V$ be the elements of the L_∞ cover closest to $\phi_{i_0},\ldots,\phi_{i_k}$, respectively, on the path ϵ. That is, for any $a\in\{i_0,\ldots,i_k\}$,

$$\max_t\|\ell_{\phi_a}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))-\mathbf{v}_t^a(\epsilon)\|\leq\alpha.$$

By our assumption, on any interval I, defined by the endpoints $a = i_j$ and $b = i_{j+1}$,

$$\max_{t\in\{a,\dots,b-1\}} \|\ell_{\phi_a}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\| \le \alpha,$$

Hence,

$$\max_{t\in\{a,\dots,b-1\}} \|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \mathbf{v}_t^a(\epsilon)\| \le 2\alpha$$

Denoting by $a(t) \in \{i_0,\dots,i_k\}$ the left endpoint of an interval to which t belongs,

$$\max_{t\in\{1,\dots,T\}} \|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon)) - \mathbf{v}_t^{a(t)}(\epsilon)\| \le 2\alpha$$

It is then clear that to construct a 2α-cover for $\Phi_T^{k,\alpha}$ in L_∞ norm, it is enough to concatenate trees in V. More precisely, this is done as follows. Construct a set V^k of $\mathcal{H}$-valued trees as

$$V^k = \{\mathbf{v}' = \mathbf{v}'\left(\mathbf{v}^0,\dots,\mathbf{v}^k, i_0,\dots,i_k\right) : 1 = i_0 \le i_1 \le \dots \le i_k \le T, \quad \mathbf{v}^0,\dots,\mathbf{v}^k \in V\}$$

and $\mathbf{v}' = \mathbf{v}'\left(\mathbf{v}^0,\dots,\mathbf{v}^k, i_0,\dots,i_k\right)$ is defined as a sequence of T mappings

$$\mathbf{v}'_t(\epsilon) = \mathbf{v}_t^{a(t)}(\epsilon) \qquad t \in I_{a(t)}$$

for any $\epsilon \in \{\pm 1\}^T$. Here $I_a = \{i_j,\dots,i_{j+1}-1\}$ and $a(t)$ is the index of the interval to which t belongs. In plain words, we consider all ways of partitioning $\{1,\dots,T\}$ into $k+1$ intervals and defining a new set of trees out of V in such a way that within the interval, the values are given by a fixed tree from V. As before, it is clear that

$$\mathcal{N}_\infty(2\alpha, \Phi_T^{k,\alpha}, T) = |V^k| \le \binom{T}{k} \cdot \mathcal{N}_\infty(\alpha,\Phi,T)^{k+1},$$

providing a control on the complexity of $\Phi_T^{k,\alpha}$. ∎

Proof [**of Lemma 13**] We claim that by choosing k large enough, the set of covering trees V^k defined in the proof of Lemma 12 provides a cover for Φ_T at a given scale $\alpha > 0$. Consider any $(\phi_1,\dots,\phi_T) \in \Phi_T$. We construct the nondecreasing sequence $i_1,\dots,i_j,\dots \in \{1,\dots,T\}$ of "change-points" as follows: increase t until the next payoff transformation is farther than α from the payoff transformation at i_j:

$$i_{j+1} = \inf_{t > i_j} \left\{ \sup_{f,x} \left\| \ell_{\phi_{i_j}}(f,x) - \ell_{\phi_t}(f,x) \right\| \ge \alpha \right\}$$

Let k be the length of the largest such sequence for all elements of Φ_T. We have simply reduced the problem to the one studied in Lemma 12: within each block, all the payoff transformations are close. Clearly, $k = k(\alpha) \le L/\alpha$, but can potentially be smaller under additional assumptions on Φ_T. We then have a bound on the size of a 2α-cover of Φ_T:

$$\mathcal{N}_\infty(2\alpha,\Phi_T,T) \le \binom{T}{k(\alpha)} \cdot \mathcal{N}_\infty(\alpha,\Phi,T)^{k(\alpha)+1} \le \binom{T}{L/\alpha} \cdot \mathcal{N}_\infty(\alpha,\Phi,T)^{L/\alpha+1} \ .$$

∎

Appendix B. Techniques for Lower Bounds

It is well-known that an *equalizing strategy* (i.e. a strategy that makes the move of the other player "irrelevant") can often be shown to be minimax optimal. In this section, we define a notion of an equalizer for our repeated game and show that it can be used to prove *lower bounds* on the value of the game. While existence of an equalizer has to be established for particular problems at hand, the lower bounds below hold whenever such an equalizer exists.

Definition 19 *A strategy $\{p_t^*\}$ for the adversary is said to be an* equalizer strategy *if*

$$\mathop{\mathbb{E}}_{\substack{x_1\sim p_1^*\\ f_1\sim q_1^*}}\cdots\mathop{\mathbb{E}}_{\substack{x_T\sim p_T^*\\ f_T\sim q_T^*}}\mathbf{R}_T\left((f_1,x_1),\ldots,(f_T,x_T)\right)=\mathop{\mathbb{E}}_{\substack{x_1\sim p_1^*\\ f_1\sim \overline{q_1^*}}}\cdots\mathop{\mathbb{E}}_{\substack{x_T\sim p_T^*\\ f_T\sim \overline{q_T^*}}}\mathbf{R}_T\left((f_1,x_1),\ldots,(f_T,x_T)\right)$$

for all strategies $\{q_t^*\}$ and $\left\{\overline{q_t^*}\right\}$ of the player. Here $\mathbf{R}_T$ is defined as in (1).

Using the above definition of an equalizer we have the following proposition as an immediate consequence.

Proposition 20 *For any Equalizer strategy $\{p_t^*\}$ we have that for any $f\in\mathcal{F}$,*

$$\mathcal{V}_T(\ell,\Phi_T)\geq\mathop{\mathbb{E}}_{x_1\sim p_1}\cdots\mathop{\mathbb{E}}_{x_T\sim p_T}\left[\mathbf{B}\left(\ell(f,x_1),\ldots,\ell(f,x_T)\right)-\inf_{\phi\in\Phi_T}\mathbf{B}\left(\ell_{\phi_1}(f,x_1),\ldots,\ell_{\phi_T}(f,x_T)\right)\right]$$

where $p_t=p_t^*\left(\{f_s=f,x_s\}_{s=1}^{t-1}\right)$

Remark 21 *For many interesting games we consider it is often the case that for any $x_1,\ldots,x_T$ and any $f_1,\ldots,f_T,f_1',\ldots,f_T'$,*

$$\inf_{\phi\in\Phi_T}\mathbf{B}\left(\ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_T}(f_T,x_T)\right)=\inf_{\phi\in\Phi_T}\mathbf{B}\left(\ell_{\phi_1}(f_1',x_1),\ldots,\ell_{\phi_T}(f_T',x_T)\right)$$

In these cases since the player's actions do not even affect the second term of the regret, to check if a strategy $\{p_t^*\}$ is an equalizer or not we only need to check if

$$\mathop{\mathbb{E}}_{\substack{x_1\sim p_1^*\\ f_1\sim q_1^*}}\cdots\mathop{\mathbb{E}}_{\substack{x_T\sim p_T^*\\ f_T\sim q_T^*}}\mathbf{B}\left(\ell(f_1,x_1),\ldots,\ell(f_T,x_T)\right)=\mathop{\mathbb{E}}_{\substack{x_1\sim p_1^*\\ f_1\sim \overline{q_1^*}}}\cdots\mathop{\mathbb{E}}_{\substack{x_T\sim p_T^*\\ f_T\sim \overline{q_T^*}}}\mathbf{B}\left(\ell(f_1,x_1),\ldots,\ell(f_T,x_T)\right)$$

for all strategies $\{q_t^*\}$ and $\{\overline{q_t^*}\}$ of the player.

Interestingly enough, many of the existing lower bounds in online learning literature are, in fact, equalizers (see e.g. (Cesa-Bianchi and Lugosi, 2006, p. 252)). In particular, in Abernethy et al. (2009), a lower bound on the value of the game was derived by looking at a certain *face* of a convex hull of loss vectors. The face, supported by a probability distribution p, corresponds to the set of functions with the same expected loss under the distribution p. Hence, p is an equalizing strategy for those functions. Since these functions are the "best" with respect to this distribution, a lower bound in terms of complexity of this set was derived in Abernethy et al. (2009). Furthermore, (Lee, Bartlett and Williamson, 1998) shows that a lower bound on the rate of convergence in the i.i.d. setting is achieved when there are two distinct minimizers of expected error for a given distribution. Again, this distribution can be viewed as an equalizer for the non-singleton set of minimizers of expected error.

Appendix C. Proofs for Blackwell Approachability (Section 4.2)

Proof [**of Theorem 16**] Now we apply Theorem 2 to the Blackwell Approachability game. Note that for any sequence $(\phi_1, \ldots, \phi_T)$, ϕ_t maps the payoff to some element of S. Hence,

$$\mathbf{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) = 0$$

for any $f_1, \ldots, f_T \in \mathcal{F}$, $x_1, \ldots, x_T \in \mathcal{X}$. We then conclude that

$$\begin{aligned}\mathcal{V}_T(\ell, \Phi_T) \le & \sup_{p_1,q_1} \mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\ x_1\sim p_1}} \ldots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\ x_T\sim p_T}} \Big\{ \mathbf{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\ x'_{1:T}\sim p_{1:T}}} \mathbf{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \Big\} \\ & + \sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}} \mathbf{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) \ .\end{aligned} \tag{17}$$

We remark for the upper bound to hold it is enough to assume that Φ_T contains *some* sequence that maps the payoffs to some element of S.

Consider the two terms in the above bound separately. The first term can be written as

$$\begin{aligned} & \sup_{p_1,q_1} \mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\ x_1\sim p_1}} \ldots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\ x_T\sim p_T}} \mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\ x'_{1:T}\sim p_{1:T}}} \left\{ \inf_{c\in S} \left\| c - \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) \right\| - \inf_{c'\in S} \left\| c' - \frac{1}{T}\sum_{t=1}^T \ell(f'_t, x'_t) \right\| \right\} \\ \le & \sup_{p_1,q_1} \mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\ x_1\sim p_1}} \ldots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\ x_T\sim p_T}} \mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\ x'_{1:T}\sim p_{1:T}}} \left\{ \left\| \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) - \frac{1}{T}\sum_{t=1}^T \ell(f'_t, x'_t) \right\| \right\} \\ \le & \sup_{p_1,q_1} \mathop{\mathbb{E}}_{\substack{f_1, f'_1\sim q_1\\ x_1, x'_1\sim p_1}} \ldots \sup_{p_T,q_T} \mathop{\mathbb{E}}_{\substack{f_T, f'_T\sim q_T\\ x_T, x'_T\sim p_T}} \left\{ \left\| \frac{1}{T}\sum_{t=1}^T \ell(f_t, x_t) - \frac{1}{T}\sum_{t=1}^T \ell(f'_t, x'_t) \right\| \right\} \end{aligned}$$

where in the first inequality we used $\inf_a[C_1(a)] - \inf_a[C_2(a)] \le \sup_a[C_1(a) - C_2(a)]$ along with a triangle inequality. This is now bounded by

$$2 \sup_{\mathbf{M}} \mathbb{E} \left[\left\| \frac{1}{T} \sum_{t=1}^T d_t \right\| \right]$$

where the supremum is over distributions $\mathbf{M}$ of martingale difference sequences $\{d_t\}_{t\in\mathbb{N}}$ such that each $d_t \in \mathrm{conv}(\mathcal{H} \bigcup -\mathcal{H})$.

The second term in Eq. (17) is

$$\sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\mathbf{B}(\ell(f_1,x_1),\ldots,\ell(f_T,x_T))$$

$$= \sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\inf_{c\in S}\left\|c-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|$$

$$\leq \sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\inf_{c\in S}\left\{\left\|c-\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)\right\|+\left\|\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|\right\}$$

$$\leq \sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\left\{\inf_{c\in S}\left\|c-\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)\right\|+\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\left\|\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|\right\}$$

$$\leq \sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\left\{\inf_{c\in S}\left\|c-\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)\right\|\right\} \tag{18}$$

$$+\sup_{p_1,q_1}\ldots\sup_{p_T,q_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\left\|\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|$$

where the last inequality uses the fact that supremum is convex and infimum satisfies the following property: $\inf_a [C_1(a)+C_2(a)] \leq [\inf_a C_1(a)] + [\sup_a C_2(a)]$. By one shot approachability assumption, we can choose a particular response q_t (in the first term of Eq. (18)) for a given p_t to be the mixed strategy that satisfies $\ell(q_t,p_t)\in S$. Since S is a convex set, we conclude that

$$\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)\in S$$

and the first term in Eq. (18) is zero. The second term is trivially upper bounded as

$$\sup_{p_1,q_1}\ldots\sup_{p_T,q_T}\mathop{\mathbb{E}}_{\substack{f_{1:T}\sim q_{1:T}\\ x_{1:T}\sim p_{1:T}}}\left\|\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|$$

$$\leq \sup_{p_1,q_1}\mathop{\mathbb{E}}_{\substack{f_1\sim q_1\\ x_1\sim p_1}}\ldots\sup_{p_T,q_T}\mathop{\mathbb{E}}_{\substack{f_T\sim q_T\\ x_T\sim p_T}}\left\|\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)\right\|$$

$$\leq 2\sup_{\mathbf{M}}\mathbb{E}\left[\left\|\frac{1}{T}\sum_{t=1}^{T}d_t\right\|\right] .$$

Combining the two upper bounds yields the desired result. ■

Proof [**of Theorem 17**] Consider the game where adversary plays from set $\mathcal{X}=\mathcal{H}$, the player plays from set $\mathcal{F}=\{\pm 1\}$, and $S=\{0\}$. Suppose the payoff is given by $\ell(f,x)=f\cdot x$. This game is clearly one-shor approachable since the player can always play ± 1 with equal probability to ensure that $\ell(p,q)=0$ irrespective of q.

Now consider the adversary strategy where adversary fixes a $\mathcal{H}$ valued tree $\mathbf{x}$ and at each time t picks a random $\epsilon_t \in \{\pm 1\}$ and plays $x_t = \epsilon_t \mathbf{x}_t(f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1})$ that is a random sign multiplied with the instance given by the path on the tree specified by $f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1}$. Further note that since $\epsilon_t \in \{\pm 1\}$ are Rademacher random variables, we see that irrespective of choice of distribution from which f_t is drawn, $f_t \cdot \epsilon_t$ is a Rademacher random variable conditioned on history. This shows that for the above prescribed adversary strategy, we have that for any $\mathcal{X}$ valued tree $\mathbf{x}$ and any two player strategies $\{q_t^*\}$ and $\{\overline{q_t^*}\}$ we have

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{f_1 \sim q_1^* \\ \epsilon_1 \sim \mathrm{Unif}\{\pm 1\}}} \cdots \mathop{\mathbb{E}}_{\substack{f_T \sim q_T^* \\ \epsilon_T \sim \mathrm{Unif}\{\pm 1\}}} \left\| \frac{1}{T} \sum_{t=1}^{T} (f_t \cdot \epsilon_t) \mathbf{x}(f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1}) \right\| \\
&\quad = \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1^* \\ \epsilon_1 \sim \mathrm{Unif}\{\pm 1\}}} \cdots \mathop{\mathbb{E}}_{\substack{f_T \sim \overline{q_T^*} \\ \epsilon_T \sim \mathrm{Unif}\{\pm 1\}}} \left\| \frac{1}{T} \sum_{t=1}^{T} (f_t \cdot \epsilon_t) \mathbf{x}(f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1}) \right\| \\
&\quad = \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1^* \\ \epsilon_1 \sim \mathrm{Unif}\{\pm 1\}}} \cdots \mathop{\mathbb{E}}_{\substack{f_{T-1} \sim \overline{q_{T-1}^*} \\ \epsilon_{T-1} \sim \mathrm{Unif}\{\pm 1\}}} \mathop{\mathbb{E}}_{\substack{f_T \sim \overline{q_T^*} \\ \epsilon_T \sim \mathrm{Unif}\{\pm 1\}}} \left\| \frac{1}{T} \sum_{t=1}^{T} (f_t \cdot \epsilon_t) \mathbf{x}(f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1}) \right\| \\
&\quad \dots = \mathop{\mathbb{E}}_{\substack{f_1 \sim \overline{q_1^*} \\ \epsilon_1 \sim \mathrm{Unif}\{\pm 1\}}} \cdots \mathop{\mathbb{E}}_{\substack{f_T \sim \overline{q_T^*} \\ \epsilon_T \sim \mathrm{Unif}\{\pm 1\}}} \left\| \frac{1}{T} \sum_{t=1}^{T} (f_t \cdot \epsilon_t) \mathbf{x}(f_1 \cdot \epsilon_1, \dots, f_{t-1} \cdot \epsilon_{t-1}) \right\|
\end{aligned}
$$

The first equality above is due to the fact that $f_T \cdot \epsilon_T$ is a Rademacher random variable conditioned on $f_1, \dots, f_{T-1}$ and $\epsilon_1, \dots, \epsilon_{T-1}$ which means we can replace q_T^* with $\overline{q_T^*}$. The subsequent equalities are got similarly by replacing each q_t^* by $\overline{q_t^*}$ one by one inside out by conditioning on $f_1, \dots, f_{t-1}$ and $\epsilon_1, \dots, \epsilon_{t-1}$; and replacing each q_t^* by $\overline{q_t^*}$. Hence we see that the adversary strategy is an equalizer strategy. Hence using Proposition 20 and picking the fixed $f = 1$ we see that

$$
\mathcal{V}_T \geq \sup_{\mathbf{x}} \mathbb{E}_{\epsilon \sim \mathrm{Unif}\{\pm 1\}^T} \left[\left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \mathbf{x}(\epsilon) \right\| \right] \geq \frac{1}{2} \sup_{\mathbf{M}} \mathbb{E} \left[\left\| \frac{1}{T} \sum_{t=1}^{T} d_t \right\| \right]
$$

where the last inequality is because the worst-case martingale difference sequence generated by random signs (Walsh Paley martingales) are lower bounded by the worst case martingale difference sequences within a factor of at most two Pisier (1975). ■

Appendix D. Proofs for Calibration (Section 4.3)

Let $\delta > 0$ to be determined later. Let $\|\cdot\|$ denote the ℓ_1 norm. Let C_δ be the maximal 2δ-packing of $\Delta(\mathcal{X})$ in this norm. Consider the calibration game defined in Example 4, augmented with the restriction that the player's choice belongs to C_δ instead of $\Delta(k)$. The corresponding minimax expression with this restriction is clearly an upper bound on the value of the game defined in Example 4.

Observe that the first term in the Triplex Inequality of Theorem 2 is zero. The second term is upper bounded by a particular (sub)optimal response q_t being the point mass on p_t^δ, the element of C_δ closest to p_t. Note that any 2δ packing is also a 2δ cover. Thus, the second term becomes

$$\begin{aligned}
&\sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\sup_{\phi\in\Phi_T}\left[-\mathop{\mathbb{E}}_{\substack{x_{1:T}\sim p_{1:T}\\ f_{1:T}\sim q_{1:T}}}\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_T}(f_T,x_T))\right]\\
&=\sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\mathop{\mathbb{E}}_{\substack{x_{1:T}\sim p_{1:T}\\ f_{1:T}\sim q_{1:T}}}\left\|\frac{1}{T}\sum_{t=1}^{T}\ell_{\phi_{p,\lambda}}(f_t,x_t)\right\|\\
&\le\sup_{p_1}\ldots\sup_{p_T}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\mathop{\mathbb{E}}_{x_{1:T}\sim p_{1:T}}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\left\{\|p_t^\delta-p\|\le\lambda\right\}\cdot(p_t^\delta-x_t)\right\|
\end{aligned}$$

which, in turn, is upper bounded via triangle inequality by

$$\begin{aligned}
&\sup_{p_1}\ldots\sup_{p_T}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\mathop{\mathbb{E}}_{x_{1:T}\sim p_{1:T}}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\left\{\|p_t^\delta-p\|\le\lambda\right\}\cdot(p_t^\delta-p_t)\right\|\\
&\qquad+\sup_{p_1}\ldots\sup_{p_T}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\mathop{\mathbb{E}}_{x_{1:T}\sim p_{1:T}}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\left\{\|p_t^\delta-p\|\le\lambda\right\}\cdot(p_t-x_t)\right\|\\
&\le 2\delta+\sup_{p_1}\ldots\sup_{p_T}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\mathop{\mathbb{E}}_{x_{1:T}\sim p_{1:T}}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\left\{\|p_t^\delta-p\|\le\lambda\right\}\cdot(p_t-x_t)\right\|
\end{aligned}$$

Now note that for a given $\lambda>0$, $p_1,\ldots,p_T$ and $p\in\Delta(k)$, we have that $\{\mathbf{1}\left\{\|p_t^\delta-p\|\le\lambda\right\}\cdot(p_t-x_t)\}_{t\in\mathbb{N}}$ is a martingale difference sequence and so the second term in the triplex inequality is bounded as :

$$\sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T}\sup_{\phi\in\Phi_T}\left[-\mathop{\mathbb{E}}_{\substack{x_{1:T}\sim p_{1:T}\\ f_{1:T}\sim q_{1:T}}}\mathbf{B}(\ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_T}(f_T,x_T))\right]\le 2\delta+2\sqrt{\frac{k}{T}}\ . \tag{19}$$

We now proceed to upper bounded the third term in the Triplex Inequality. Since $-\mathbf{B}$ is a subadditive, by Theorem 3, we have that the third term is bounded by twice the sequential complexity

$$\begin{aligned}
2\mathfrak{R}_T(\ell,\Phi_T,-\mathbf{B})&=2\sup_{\mathbf{f},\mathbf{x}}\ \mathbb{E}_{\epsilon_{1:T}}\sup_{\phi\in\Phi_T}-\mathbf{B}\Big(\epsilon_1\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\epsilon_T\ell_{\phi_T}(\mathbf{f}_T(\epsilon),\mathbf{x}_T(\epsilon))\Big)\\
&=2\sup_{\mathbf{f},\mathbf{x}}\ \mathbb{E}_{\epsilon_{1:T}}\sup_{\lambda>0}\sup_{p\in\Delta(k)}\left\|\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\mathbf{1}\left\{\|\mathbf{f}_t(\epsilon)-p\|\le\lambda\right\}\cdot(\mathbf{f}_t(\epsilon)-\mathbf{x}_t(\epsilon))\right\|
\end{aligned}$$

where $\mathbf{f}$ is a C_δ-valued tree. Using the fact that $\mathbf{f}$ is a discrete-valued tree, not a $\Delta(k)$-valued tree, we would like to pass from the supremum over $\lambda>0$ and $p\in\Delta(k)$ to a supremum over finite discrete set in order to appeal to Lemma 4.

To this end, fix $\mathbf{f}, \mathbf{x}$ and $\epsilon_{1:T}$ and let us see how many genuinely different functions can we get by varying $\lambda > 0$ and $p \in \Delta(k)$. This question boils down to looking at the size of the class

$$\mathcal{G} := \{g_{p,\lambda}(f) = \mathbf{1}\{\|f - p\| \leq \lambda\} : p \in \Delta(k), \lambda > 0\}$$

over the possible values of $f \in C_\delta$. Indeed, if $g_{p,\lambda}(f) = g_{p',\lambda'}(f)$ for all $f \in C_\delta$, then

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\| \leq \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p'\| \leq \lambda'\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)).$$

We appeal to VC theory for bounding the size of $\mathcal{G}$ over C_δ. First, we claim that the VC dimension of $\mathcal{G}$ is $O(k^2)$. Note that $\mathcal{G}$ is the class of indicators over ℓ_1 balls of radius λ centered at p for various values of p, λ. A result of Goldberg and Jerrum Goldberg and Jerrum (1995) states that for a class $\mathcal{G}$ of functions parametrized by a vector of length d, if for $g \in \mathcal{G}$ and $f \in \mathcal{F}$, $\mathbf{1}\{g(f) = 1\}$ can be computed using m arithmetic operations, the VC dimension of $\mathcal{G}$ is $O(md)$. In our case, the functions in $\mathcal{G}$ are parametrized by k values and membership $\|f - p\|_1 \leq \lambda$ can be established in $O(k)$ operations. This yields $O(k^2)$ bound on the VC dimension of $\mathcal{G}$. By Sauer-Shelah Lemma, the number of different labelings of the set C_δ by $\mathcal{G}$ is bounded by $|C_\delta|^{c \cdot k^2}$ for some absolute constant c. We conclude that the effective number of different (p, λ) is finite. Let us remark that the VC upper bound is *not* used in place of the sequential Littlestone's dimension. It is only used to show that the set Φ_T is finite, and such technique can be useful when the set of player's actions is finite.

Hence, there exists a finite set S of pairs (λ, p) with cardinality $|S| \leq |C_\delta|^{c \cdot k^2}$ such that

$$\begin{aligned}
2\mathfrak{R}_T(\ell, \Phi_T, -\mathbf{B}) &\leq 2 \sup_{\mathbf{f},\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \sup_{\lambda > 0} \sup_{p \in \Delta(k)} \left\| \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\|_1 \leq \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)) \right\|_1 \\
&= 2 \sup_{\mathbf{f},\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \max_{(p,\lambda) \in S} \left\| \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\|_1 \leq \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)) \right\|_1 \\
&\leq 2 \ k^{1/2} \sup_{\mathbf{f},\mathbf{x}} \mathbb{E}_{\epsilon} \max_{(p,\lambda) \in S} \left\| \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\|_1 \leq \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)) \right\|_2
\end{aligned}$$

Now note that $\|\cdot\|_2^2$ is $(2,2)$-smooth and so applying Lemma 4 with $G = \|\cdot\|_2$, $\gamma = 2$, $\eta = 2$, we see that

$$\begin{aligned}
2\mathfrak{R}_T(\ell, \Phi_T, -\mathbf{B}) &\leq 2k^{1/2} \left(\frac{8 \log(2|S|)}{T} \right)^{1/2} \\
&\leq 2k^{1/2} \left(\frac{16ck^2 \log(|C_\delta|)}{T} \right)^{1/2} \\
&= c' k^{3/2} \left(\frac{\log(|C_\delta|)}{T} \right)^{1/2}
\end{aligned}$$

for some small absolute constant c'.

Now note that the size of set C_δ the 2δ packing of $\Delta(k)$ is upper bounded by the size of the minimal δ cover of $\Delta(k)$ which can be bounded as $|C_\delta| \le \left(\frac{1}{\delta}\right)^{k-1}$ and so we see that

$$2\mathfrak{R}_T(\ell, \Phi_T, -\mathbf{B}) \le c' k^2 \left(\frac{\log(1/\delta)}{T}\right)^{1/2} .$$

Combining the above upper bound on the third term of triplex inequality and Equation 19 that bounds the second term of the triplex inequality (and since first term is anyway 0) we see that,

$$\mathcal{V}_T \le 2\delta + 2\sqrt{\frac{k}{T}} + c' k^2 \left(\frac{\log(1/\delta)}{T}\right)^{1/2} .$$

Choosing $\delta = 1/T$ concludes the proof.

Appendix E. Proofs for Global Cost (Section 4.4)

Proof [**of Theorem 18**] The Triplex Inequality and Theorem 3 give

$$\begin{aligned}\mathcal{V}_T \le & \sup_{p_1, q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \dots \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \left\| \frac{1}{T} \sum_{t=1}^T (\ell(f_t, x_t) - \ell(f'_t, x'_t)) \right\| \\ & + \sup_{p_1} \inf_{q_1} \dots \sup_{p_T} \inf_{q_T} \sup_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{1:T} \\ x_{1:T} \sim p_{1:T}}} \left\{ \left\| \frac{1}{T} \sum_{t=1}^T \ell(f_t, x_t) \right\| - \left\| \frac{1}{T} \sum_{t=1}^T \ell(f, x_t) \right\| \right\} \\ & + 2 \sup_{\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \sup_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^T \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)) \right\| .\end{aligned}$$

Consider the first term in the Triplex Inequality. Observe that $(\ell(f_t, x_t) - \ell(f'_t, x'_t))_{t=1}^T$ is a (vector valued) martingale difference sequence and so

$$\sup_{p_1, q_1} \mathop{\mathbb{E}}_{\substack{f_1, f'_1 \sim q_1 \\ x_1, x'_1 \sim p_1}} \dots \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T, f'_T \sim q_T \\ x_T, x'_T \sim p_T}} \left\| \frac{1}{T} \sum_{t=1}^T (\ell(f_t, x_t) - \ell(f'_t, x'_t)) \right\| \le 2 \sup_{\mathbf{M}} \mathbb{E}\left[\left\| \frac{1}{T} \sum_{t=1}^T d_t \right\| \right] .$$

where the supremum is over distributions $\mathbf{M}$ of martingale difference sequences $\{d_t\}_{t \in \mathbb{N}}$ such that each $d_t \in \mathrm{conv}(\mathcal{H} \bigcup -\mathcal{H})$.

Now, consider the second summand above:

$$\begin{aligned} & \sup_{p_1} \inf_{q_1} \dots \sup_{p_T} \inf_{q_T} \sup_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{1:T} \\ x_{1:T} \sim p_{1:T}}} \left\{ \left\| \frac{1}{T} \sum_{t=1}^T \ell(f_t, x_t) \right\| - \left\| \frac{1}{T} \sum_{t=1}^T \ell(f, x_t) \right\| \right\} \\ & = \sup_{p_1} \inf_{q_1} \dots \sup_{p_T} \inf_{q_T} \left\{ \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{1:T} \\ x_{1:T} \sim p_{1:T}}} \left\| \frac{1}{T} \sum_{t=1}^T \ell(f_t, x_t) \right\| - \inf_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^T \ell(f, x_t) \right\| \right\} \\ & \le \sup_{p_1} \dots \sup_{p_T} \left\{ \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^T \ell(f_t, x_t) \right\| - \inf_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^T \ell(f, x_t) \right\| \right\}\end{aligned}$$

where in the last step a (sub)optimal choice was made for q_t: the distribution $q_t = \delta_{f_t}$ puts all the mass on f_t such that

$$\|\ell(f_t, p_t)\| = \inf_{f \in \mathcal{F}} \|\ell(f, p_t)\|.$$

Observe that by several applications of triangle and Jensen's inequalities,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, x_t) \right\| - \inf_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, x_t) \right\| \\
&\quad \le \left\{ \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, p_t) \right\| - \inf_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, p_t) \right\| \right\} + \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} (\ell(f_t, x_t) - \ell(f_t, p_t)) \right\|
\end{aligned} \tag{20}$$

Under Assumption 1, along with the way we chose f_t, the first term in (20) becomes

$$\left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, p_t) \right\| - \inf_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, p_t) \right\| \le \frac{1}{T} \sum_{t=1}^{T} \|\ell(f_t, p_t)\| - \frac{1}{T} \sum_{t=1}^{T} \inf_{f \in \mathcal{F}} \|\ell(f, p_t)\| = 0 \ .$$

We conclude that the second term in the Triplex Inequality can be upper bounded by

$$\sup_{p_1} \dots \sup_{p_T} \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} (\ell(f_t, x_t) - \ell(f_t, p_t)) \right\|,$$

which, in turn, is no worse than the supremum over distributions $\mathbf{M}$ of martingale difference sequences used to bound the first term.

This gives us the general upper bound on the value of the game:

$$\mathcal{V}_T \ \le \ 4 \sup_{\mathbf{M}} \mathbb{E} \left[\left\| \frac{1}{T} \sum_{t=1}^{T} d_t \right\| \right] \ + 2 \ \sup_{\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \sup_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)) \right\| \ . \tag{21}$$

■

Lemma 22 *Let $\mathcal{F}$ be the probability simplex in any dimension. Let $\|\cdot\|$ be any norm. The function*

$$x \mapsto \inf_{f \in \mathcal{F}} \|f \odot x\| \ ,$$

defined on the positive orthant, is concave.

Proof Since the function above is absolutely homogeneous and continuous, all we need to prove is

$$\inf_{f \in \mathcal{F}} \|f \odot (x + y)\| \ge \inf_{f \in \mathcal{F}} \|f \odot x\| + \inf_{f \in \mathcal{F}} \|f \odot y\| \ .$$

for arbitrary x, y. That is, for arbitrary f', x, y,

$$\|f' \odot (x + y)\| \ge \inf_{f \in \mathcal{F}} \|f \odot x\| + \inf_{f \in \mathcal{F}} \|f \odot y\| \ .$$

Define $h, g \in \mathcal{F}$ as follows:

$$g_i = \frac{f_i'(1 + y_i/x_i)}{Z_g} \qquad\qquad h_i = \frac{f_i'(1 + x_i/y_i)}{Z_h} \ ,$$

where

$$Z_g = \sum_i f_i'(1 + y_i/x_i) \qquad\qquad Z_h = \sum_i f_i'(1 + x_i/y_i) \ .$$

Now, as we show below, $1/Z_g + 1/Z_h \leq 1$. Thus,

$$\begin{aligned}
\|f' \odot (x+y)\| &\geq \frac{1}{Z_g}\|f' \odot (x+y)\| + \frac{1}{Z_h}\|f' \odot (x+y)\| \\
&= \|g \odot x\| + \|h \odot y\| \\
&\geq \inf_{f \in \mathcal{F}} \|f \odot x\| + \inf_{f \in \mathcal{F}} \|f \odot y\| \ .
\end{aligned}$$

To finish the proof, note that, by Cauchy-Schwarz,

$$\left(\sum_i f_i'(1 + y_i/x_i)\right) \cdot \left(\sum_i f_i' \frac{x_i}{x_i + y_i}\right) \geq \left(\sum_i f_i'\right)^2 = 1 \ .$$

This shows,

$$\frac{1}{Z_g} \leq \sum_i f_i' \frac{x_i}{x_i + y_i} \ .$$

Similarly, we get

$$\frac{1}{Z_h} \leq \sum_i f_i' \frac{y_i}{x_i + y_i} \ .$$

Adding them, we get

$$\frac{1}{Z_g} + \frac{1}{Z_h} \leq \sum_i f_i' = 1$$

as claimed. This completes the proof. ■