Online Learning: Beyond Regret

Alexander Rakhlin
Department of Statistics
University of Pennsylvania

Karthik Sridharan
TTIC
Chicago, IL

Ambuj Tewari
Computer Science Department
University of Texas at Austin

November 18, 2010

Abstract

We study online learnability of a wide class of problems, extending the results of [26] 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, adaptive regret, 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 [26]. 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 [26], 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 [26] 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, adaptive regret, 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 [26], plays a key role in these 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 [26], the question of learnability can now be addressed in a variety of situations in a unified manner. In fact, [26] 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 [26] 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 [9]; 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 5.2, we discuss the relation of our results to approachability in greater detail.

Our contributions can be broken down into three parts.

- The first contribution lies in the formulation of the online learning problem, with a performance measure (a form of *regret*), defined in terms of certain payoff transformation mappings. While this formulation might appear unusual, we show that it is general enough to encompass many seemingly different frameworks (games), yet specific enough that we can provide generic upper bounds.

- The second contribution is in developing 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 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.

- We apply our machinery to a number of well-known problems. (a) First, for the usual notion of external regret, the results boil down to those of [26]. (b) For the more general Φ-regret (see e.g. [27, 16, 17]), we recover and improve several known results. In particular, for convergence to Φ-correlated equilibria, we improve upon the results of Stoltz and Lugosi [27]. (c) We study the game of Blackwell's approachability [4] in (possibly infinite-dimensional) separable Banach spaces. Specifically, we show that martingale convergence in these spaces (along with Blackwell's one-shot approachability condition) is both necessary and sufficient for Blackwell's approachability to hold. (d) We also consider the game of calibrated forecasting. We improve upon the results of Mannor and Stoltz [23] and prove (to the best of our knowledge) the first known $O(T^{-1/3})$ 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 [11]. (f) We provide techniques for bounding notions of regret where algorithm's performance is measured against a time-varying comparator (see e.g. [19, 7, 28]). Such notions of regret are better suited for reactive environments. Using the general tools we developed, we not only recover the results in [19, 7] but also extend them to prove learnability and obtain rates for much more general settings. Our last example shows that adaptive regret notion of Hazan and Seshadhri [18] can be defined in greater generality while still preserving learnability.

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 [14]. In [26, 1], 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.

Let us spend a minute describing the organization of this paper. Since our results are meant to serve as a unifying framework, we faced the question of whether to build up the level of generality as we progress through the paper, or whether to start with the most general results and then make them more specific. We decided to do the latter. While we find this flow of general-to-specific more natural, we risk losing potential readers on the first few pages. In hopes of avoiding this, after defining the online learning problem in full generality in Section 2, we briefly state how various well-known frameworks appear as particular instances. Then, in Section 3, learnability is established under various very general assumptions. Next, in Section 4, techniques for proving lower bounds are shown. Finally, various examples and frameworks are considered in more detail in Section 5. For an overview of the results without the painful details, one may read Section 2 and then skip to Section 5. For the sake of readability, most of the proofs are deferred to the appendix. Let us remark that [26] is not required for reading this paper. In a few places, however, if a proof is basically the same as in [26] except for notation, we will omit the proof.

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.

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 [26], 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 $\ell(f_t, x_t) \in \mathcal{H}$
End

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 of performance measure:

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

where

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

---

[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.

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

- 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: $\boldsymbol{B}, \ell, \mathcal{F}, \mathcal{X}$, and Φ_T. In [26], 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 $\boldsymbol{B}, \ell, \mathcal{F}, \mathcal{X}$, 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 $\boldsymbol{B}$. We emphasize that most flexibility comes from the payoff mapping ℓ and from the transformations Φ_T of the payoffs.

In particular, important classes of payoff transformation mappings are the *departure mappings* that transform the payoff function ℓ by acting only on the first argument of ℓ, i.e. only modifying the row (player's action) choice.

Definition 1. 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 the property that, for all $t \in [T]$, $f \in \mathcal{F}$ and $x \in \mathcal{X}$, the payoff transformations can be written as $\ell_{\phi_t}(f, x) := \ell(\phi_t'(f), x)$.

For payoff transformation classes that are departure mapping classes, the transformations Φ_T can be identified in terms of a corresponding class of departure mapping from $\mathcal{F}$ to itself, and we shall 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.

Definition 2. 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:

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

(2)

where q_t and p_t range over $\mathcal{Q}$ and $\mathcal{P}$, respectively. With this definition of a value, the strategy of the adversary is a sequence of mappings $(\mathcal{Q} \times \mathcal{F} \times \mathcal{X})^{t-1} \times \mathcal{Q} \mapsto \mathcal{P}$ for each $t \in [T]$.

Definition 3. 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. For "almost sure" convergence, however, more work is required in many situations. In this paper, we will be content with "in expectation" or "in probability" notions of convergence.

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)**.** *Let* $\mathcal{H} = \mathbb{R}$ *and*

- $\boldsymbol{B}(z_1, \ldots, z_T) = \frac{1}{T} \sum_{t=1}^{T} z_t$
- $\Phi_T = \{(\phi_f, \ldots, \phi_f) : f \in \mathcal{F} \quad and \quad \phi_f : \mathcal{F} \mapsto \mathcal{F} \quad is\ a\ constant\ mapping\ \phi_f(g) = f\ \forall g \in \mathcal{F}\}$

It is easy to see that Eq. (1) becomes

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

External regret is discussed in Section 5.1.1.

Example 2 (Φ-Regret)**.** *Let* $\mathcal{H} = \mathbb{R}$ *and*

- $\boldsymbol{B}(z_1, \ldots, z_T) = \frac{1}{T} \sum_{t=1}^{T} z_t$
- $\Phi_T = \{(\phi, \ldots, \phi) : \phi \in \Phi\}$ *for some fixed family* Φ *of* $\mathcal{F} \mapsto \mathcal{F}$ *mappings.*

It is easy to see that Eq. (1) becomes

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

This example covers a variety of notions such as external, internal, and swap regrets (see Section 5.1).

Example 3 (Blackwell's Approachability)**.** *Let* $\mathcal{H}$ *a subset of a Banach space* $\mathcal{B}$, $S \subset \mathcal{B}$ *be a closed convex set, and*

- $\boldsymbol{B}(z_1, \ldots, z_T) = \inf_{c \in S} \left\| \frac{1}{T} \sum_{t=1}^{T} z_t - c \right\|$
- 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* $1 \le t \le T$.

It is easy to see that Eq. (1) becomes

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

the distance to the set S. *Indeed, our definition of* Φ_T *ensures that the comparator term is zero. Blackwell's approachability is discussed in Section 5.2.*

Example 4 (Calibration of Forecasters). *Let $\mathcal{H} = \mathbb{R}^k$, $\mathcal{F} = \Delta(k)$ (the k-dimensional probability simplex) and $\mathcal{X}$ the set of standard unit vectors in $\mathbb{R}^k$ (vertices of $\Delta(k)$). Define $\ell(f,x) = 0$. Further,*

- $\boldsymbol{B}(z_1, \ldots, z_T) = -\left\| \frac{1}{T} \sum_{t=1}^{T} z_t \right\|$ *for some norm $\|\cdot\|$ on $\mathbb{R}^k$*

- $\Phi_T = \{(\phi_{p,\lambda}, \ldots, \phi_{p,\lambda}) : p \in \Delta(k), \lambda > 0\}$ *contains time-invariant mappings defined by*

$$\ell_{\phi_{p,\lambda}}(f,x) = \mathbf{1}\left\{\|f - p\| \le \lambda\right\} \cdot (f - x).$$

It is easy to see that Eq. (1) becomes

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

Calibration is discussed in more detail in Section 5.3.

Example 5 (Global Cost Online Learning Game [11]). *Let $\mathcal{H} = \mathbb{R}^k$, $\mathcal{X} = [0,1]^k$, $\mathcal{F} = \Delta(k)$, $\ell(f,x) = f \odot x = (f^1 \cdot x^1, \ldots, f^k \cdot x^k)$.*

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

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

It is easy to see that Eq. (1) becomes

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

A generalization of this scenario is considered in Section 5.4.

2.2 Notation

Let $\mathbb{E}_{x \sim p}$ denote expectation with respect to a random variable x with a distribution p. Note that we do not use capital letters for random variables in order to ease reading of already cumbersome equations. For a collection of 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 with respect to 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 Y is uniformly distributed on $\{\pm 1\}$. 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 $\mathcal{X} \times \ldots \times \mathcal{X}$ is denoted by $\mathcal{X}^t$. Whenever a supremum (infimum) is written in the form $\sup_a$ without a being quantified, it is assumed that a ranges over the set of all possible values which will be understood from the context. Convex hulls will be denoted by $\mathrm{conv}(\cdot)$.

Following [26], we define binary trees as follows.

Definition 4. *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 $\mathbf{z}_t(\epsilon_{1:t-1})$.

Let ϕ_{id} denote the identity payoff transformation $\ell_{\phi_{\mathrm{id}}}(f, x) = \ell(f, x)$ for all $f \in \mathcal{F}$, $x \in \mathcal{X}$. Let $\mathcal{I} = \{(\phi_{\mathrm{id}}, \ldots, \phi_{\mathrm{id}})\}$ be the singleton set containing the time-invariant sequence of identity transformations.

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 $\mathcal{B}$. A Hilbert space is dual to itself.

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 $\boldsymbol{B}$, 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 [26]. 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 [26], to "global" functions $\boldsymbol{B}$ of the payoff sequence.

Definition 5. The *sequential complexity* with respect to the payoff function ℓ and payoff transformation mappings Φ_T is defined as

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

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 $\boldsymbol{B}$ is clear from the context, it will be omitted 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 [26] (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 5.1.1, we show that the notion studied in [26] is indeed recovered for the case of external regret.

An equivalent way to write sequential complexity is through the expanded version

$$\mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}) = \sup_{f_1, x_1} \mathbb{E}_{\epsilon_1} \sup_{f_2, x_2} \mathbb{E}_{\epsilon_2} \ldots \sup_{f_T, x_T} \mathbb{E}_{\epsilon_T} \sup_{\phi \in \Phi_T} \boldsymbol{B}\Big(\epsilon_1 \ell_{\phi_1}(f_1, x_1), \ldots, \epsilon_T \ell_{\phi_T}(f_T, x_T)\Big) \tag{3}$$

where the supremum on t-th step is over $f_t \in \mathcal{F}$, $x_t \in \mathcal{X}$. We shall use Eq. (3) and the more succinct Definition 5 interchangeably.

3.1 Triplex Inequality

The following theorem is the main starting point for all further analysis. Because of its importance, we shall refer to it as the *Triplex Inequality*. The three terms in the upper bound of the theorem can be thought of as 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.

7

Theorem 1 (**Triplex Inequality**)**.** *The following* 3*-term upper bound on the value of the game holds:*

$$\mathcal{V}_T(\ell, \Phi_T)$$

$$\leq \sup_{p_1, q_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T, q_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \right\} \qquad (4)$$

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

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

First, we remark that convexity of $\boldsymbol{B}$ is *not required* for the Triplex Inequality to hold. Under a weak subadditivity condition, the following Theorem gives upper bounds on the first and the third term.

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

Discussion of Theorem 1 and Theorem 2

- First, let us mention that Triplex Inequality is not the only way to decompose the value of the game into useful and interpretable terms. In fact, slightly different decompositions yield better constants for some of the examples in this paper. Nonetheless, the Triplex Inequality seems to capture the essence of all the problems we considered and allows us to give a unified treatment to all of them.

- We note that the first and the third terms are similar in their form. In fact, the first term can be equivalently written as

$$\sup_{p_1, q_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T, q_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \sup_{\phi \in \mathcal{I}} \left\{ \boldsymbol{B}\left(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T) \right) - \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}\left(\ell_{\phi_1}(f'_1, x'_1), \ldots, \ell_{\phi_T}(f'_T, x'_T) \right) \right\}$$

  where $\mathcal{I}$ only contains the identity mapping. If $\mathcal{I} \subseteq \Phi_T$, then, trivially, $\mathfrak{R}_T(\ell, \mathcal{I}, \boldsymbol{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B})$ 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 2? Let us compare the third term in the Triplex Inequality to its sequential complexity upper bound given by Eq. (3). 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. (3), 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 5.2 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 $\boldsymbol{B}$, Theorem 2 upper bounds the first and third terms by the sequential complexity.

- We remark that the first and third terms in Triplex Inequality contain *suprema* over the player's strategies q_t instead of *infima* as in the definition of the value of the game. The proof of Theorem 1 points out the step where this over-bounding is done. While this might appear as a loose step, in all the examples we considered, this still yields the needed results. Nevertheless, as mentioned in the proof, one can substitute a particular strategy q_t^* for the first and third terms instead of passing to the supremum. For instance, q_t^* can be the strategy which makes the second term in the Triplex Inequality small. To simplify the presentation, we decided not to include such analysis.

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

$$\mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}') \ . \tag{5}$$

- Finally, let us mention that we could have defined the performance measure in (1) as

$$\mathbf{R}_T = \sup_{(\boldsymbol{\phi}', \boldsymbol{\phi}) \in (\Phi_T' \times \Phi_T)} \boldsymbol{B}(\ell_{\phi_1'}(f_1, x_1), \ldots, \ell_{\phi_T'}(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \ . \tag{6}$$

Clearly, (1) can be expressed as an instance of (6) by setting $\Phi_T' = \mathcal{I}$. Conversely, if $\boldsymbol{B}$ is, for instance, an average of its coordinates, we can view definition (6) as a particular case of (1). Indeed, given a payoff ℓ and sets Φ_T', Φ_T of transformations, define a new payoff $\bar{\ell}(f, x) = 0$ and $\bar{\ell}_{(\phi_t', \phi_t)}(f, x) = -(\ell_{\phi_t'}(f, x) - \ell_{\phi_t}(f, x))$. Then (1) becomes exactly (6). While the analysis presented in this paper can be extended for (6), in the examples we consider, the definition (1) of performance measure is expressive enough.

We now detail upper bounds on this complexity under the smoothness assumption on $\boldsymbol{B}$. The smoothness assumption covers many important cases, such as norms.

3.2   General Bounds for Smooth $\boldsymbol{B}$

As shown by Pisier [25] and Pinelis [24], existence of a smooth norm in a Banach spaces is crucial in the study of exponential inequalities for martingales. Using similar techniques, we show that a smooth function $\boldsymbol{B}$ will admit upper bounds in terms of certain increments. This will yield general tools for studying sequential complexity for smooth functions $\boldsymbol{B}$. Informally, the smoothness assumption provides a link from a "global" function 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.

Consider the following definition of smoothness.

Definition 6. 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,
$$G(z) \leq G(z') + \langle \nabla G(z'), z - z' \rangle + \frac{\sigma}{p} \|z - z'\|^p$$

We say that G is *uniformly smooth* if there exist finite σ and p such that G is (σ, p)-uniformly smooth.

A function which is smooth in its arguments can be "sequentially linearized", with additional second-order terms as norms of the increments. We establish the following upper bound on the first term of the Triplex Inequality.

Lemma 3. *Suppose $\boldsymbol{B}$ is subadditive and for some $q \geq 1$, $\boldsymbol{B}^q$ is (σ, p)-uniformly smooth in each of its arguments. Suppose $\boldsymbol{B}(0, \ldots, 0) = 0$ and that for any $x \in \mathcal{X}$ and $f \in \mathcal{F}$ it is true that $\|\ell(f, x)\| \leq \eta$. Then the first term in the Triplex Inequality is bounded by $((2\eta)^p \sigma T)^{1/q}$.*

Under the assumptions of Lemma 3, we can also provide an upper bound on the third term. Lemma 4 below says that the sequential complexity defined through a smooth function $\boldsymbol{B}$ can be upper bounded by the sequential complexity involving a sum of first-order expansions of $\boldsymbol{B}$.

Lemma 4. *Assume that for some $q \geq 1$, $\boldsymbol{B}^q$ is (σ, p)-uniformly smooth in each of its arguments, $\boldsymbol{B}(0, \ldots, 0) = 0$ and that for any $x \in \mathcal{X}$, $f \in \mathcal{F}$, $\phi \in \Phi_T$ and $t \in [T]$, it is true that $\|\ell_{\phi_t}(f, x)\| \leq \eta$, then we have that*

$$\mathfrak{R}_T(\ell, \Phi_T) \leq \left(\sup_{\mathbf{f}, \mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi_T} \sum_{t=1}^{T} \epsilon_t g_t \big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \big) \right)^{1/q} + (\sigma \eta^p)^{1/q} T^{1/q}$$

where

$$g_t \big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \big)$$
$$= \big\langle \nabla_t \boldsymbol{B}^q \big(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_{t-1} \ell_{\phi_{t-1}}(\mathbf{f}_{t-1}(\epsilon), \mathbf{x}_{t-1}(\epsilon)), 0, \ldots, 0 \big), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \big\rangle \ .$$

By taking gradients at successive time steps, we reduced the study of a global function $\boldsymbol{B}$ to the study of its gradients. A reader familiar with [26] will notice that the first term of Lemma 4 (under the power of $1/q$) resembles sequential Rademacher complexity. The first step in studying this term is to ask what can be done with a finite class Φ_T. To approach this question, we state a lemma from [26].

Lemma 5. *[26] For any finite set V of $\mathbb{R}$-valued trees of depth T we have that*

$$\mathbb{E}_\epsilon \left[\max_{\mathbf{v} \in V} \sum_{t=1}^{T} \epsilon_t \mathbf{v}_t(\epsilon) \right] \leq \sqrt{2 \log(|V|) \max_{\mathbf{v} \in V} \max_{\epsilon \in \{\pm 1\}^T} \sum_{t=1}^{T} \mathbf{v}_t(\epsilon)^2} \ .$$

The above Lemma can be used to show the following result for any finite set of transformations Φ_T.

Proposition 6. *For any finite set of payoff transformations Φ_T, under the conditions of Lemma 4 and assuming*

$$\big\| \nabla_t \boldsymbol{B}^q \big(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_{t-1} \ell_{\phi_{t-1}}(\mathbf{f}_{t-1}(\epsilon), \mathbf{x}_{t-1}(\epsilon)), 0, \ldots, 0 \big) \big\| \leq R$$

then

$$\mathfrak{R}_T(\ell, \Phi_T) \leq \big(2\eta^2 R^2 \log(|\Phi_T|) T \big)^{1/2q} + (\sigma \eta^p)^{1/q} T^{1/q} \ .$$

Hence, if Φ_T is finite, sequential complexity is bounded whenever $\boldsymbol{B}$ is smooth and the gradients of $\boldsymbol{B}$ are bounded by R. Typically, R is of the order $O(1/T)$ if $\boldsymbol{B}$ is appropriately normalized to account for T (for instance, if $\boldsymbol{B}$ is an average of its coordinates). Similarly, σ is either zero or $o(1)$ for the examples considered in this paper. With the appropriate behavior of the online covering number, the bound yields learnability according to Definition 3.

For the rest of this sub-section we consider $\boldsymbol{B}$ of a particular form. We assume that for some $q > 0$, $\boldsymbol{B}^q(z_1, \ldots, z_T) = G\left(\frac{1}{T} \sum_{t=1}^{T} z_t \right)$, where G is (γ, p)-smooth function on the convex set $\text{conv}(\mathcal{H})$ for some $1 < p \leq 2$. This form of $\boldsymbol{B}$ occurs naturally in many games including Blackwell's approachability and calibration. Among the most basic smooth functions are powers of norms, as the next example shows.

Example 6. *Consider $\boldsymbol{B}$ of the form*

$$\boldsymbol{B}(z_1, \ldots, z_T) = \left\| \frac{1}{T} \sum_{t=1}^{T} z_t \right\|_q .$$

The three cases $q \in (1, \infty)$, $q = 1$, and $q = \infty$ are considered separately. In each case, we are interested in checking if $\boldsymbol{B}^s$ is uniformly smooth for some power s.

- ▶ **$q \in (\mathbf{1}, \infty)$**   *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.*

- ▶ **$q = \infty$**   *Unfortunately, for no finite power s is $\boldsymbol{B}^s$ uniformly smooth. However, for any $z \in \mathcal{H}$ and any $q' \in (1, \infty)$, $\|z\|_\infty \leq \|z\|_{q'}$. Hence we can use (5) and upper bound the sequential complexity*

$$\mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}')$$

  where $\boldsymbol{B}'(z_1, \ldots, z_T) = \left\| \frac{1}{T} \sum_{t=1}^{T} z_t \right\|_{q'}$. By choosing q' appropriately and using the smoothness of the $L_{q'}$ norm (previous case) we can provide upper bounds for the value of the game.

- ▶ **$q = \mathbf{1}$**   *As in the previous example, for no finite power s is $\boldsymbol{B}^s$ uniformly smooth. However if $\mathcal{H} \subseteq \mathbb{R}^d$, then for any $z \in \mathcal{H}$ and any $q' \in (1, \infty)$, $\|z\|_1 \leq C_{q',d} \|z\|_{q'}$ where $C_{q',d}$ is a constant dependent on q' and dimension of the space d. Again we can use (5) and upper bound*

$$\mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}) \leq \mathfrak{R}_T(\ell, \Phi_T, \boldsymbol{B}')$$

  where $\boldsymbol{B}'(z_1, \ldots, z_T) = \left\| \frac{1}{T} \sum_{t=1}^{T} z_t \right\|_{q'}$. Choosing q' appropriately and using the smoothness of the $L_{q'}$ norm we can provide upper bounds for the value of the game.

For a concrete example of a smooth norm, we refer to the calibration example of Section 5.3. We now specialize the statement of Proposition 6 to the specific assumption on $\boldsymbol{B}$.

Corollary 7. *Let Φ_T be a finite set of payoff transformations. Assume that for some $q \geq 1$, $\boldsymbol{B}^q(z_1, \ldots, z_T) = G\left(\frac{1}{T} \sum_{t=1}^{T} z_t \right)$ where G is (γ, p)-smooth function on the convex set $\mathrm{conv}(\mathcal{H})$ for some $1 < p \leq 2$. Also assume that $\|\nabla G(z)\|_* \leq \rho$ for any $z \in \mathrm{conv}(\mathcal{H})$. Further, suppose that for any $x \in \mathcal{X}$, $f \in \mathcal{F}$, $\phi \in \Phi_T$ and $t \in [T]$, it is true that $\|\ell_{\phi_t}(f, x)\| \leq \eta$. Then it holds that*

$$\mathfrak{R}_T(\ell, \Phi_T) \leq \left(\frac{2\eta^2 \log(|\Phi_T|)}{T} \right)^{1/2q} + (\gamma \eta^p)^{1/q} T^{(1-p)/q} .$$

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. Before proceeding, however, we need to define an appropriate notion of a cover. The following definition can be seen as a generalization of the corresponding notion introduced in [26]. We remark that the object, for which we would like to provide a cover, is the set Φ_T of payoff transformations. 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 $\mathcal{F}$ (see Section 5.1.1 for more details). In general, however, the set of payoff transformations can be much more complex than (or not even comparable to) $\mathcal{F}$.

Definition 7. 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

$$\forall \phi \in \Phi_T, \ \forall \epsilon \in \{\pm 1\}^T \ \exists \mathbf{v} \in V \text{ s.t.} \quad \left(\frac{1}{T} \sum_{t=1}^{T} \|\mathbf{v}_t(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\|^p \right)^{1/p} \leq \alpha \qquad (7)$$

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

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

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 over depth T trees.

This definition of the cover is indeed the most general for the setting we consider in this paper. In sections that follow, we specialize this definition to fit particular assumptions on Φ_T.

Theorem 8 shows that sequential complexity can be upper bounded in term of the covering number defined above, integrated over all the scales. This is a generalization of the analogous result in [26].

Theorem 8. *Assume that for some $q \geq 1$, $\boldsymbol{B}^q(z_1, \ldots, z_T) = G\left(\frac{1}{T}\sum_{t=1}^T z_t\right)$ where G is (γ, p)-smooth function on the convex set $\mathrm{conv}(\mathcal{H})$ for some $1 < p \leq 2$. Also assume that $\|\nabla G(z)\|_* \leq \rho$ for any $z \in \mathrm{conv}(\mathcal{H})$. Further, suppose that for any $x \in \mathcal{X}$, $f \in \mathcal{F}$, $\phi \in \Phi_T$ and $t \in [T]$, it is true that $\|\ell_{\phi_t}(f, x)\| \leq \eta$. Then it holds that*

$$\mathfrak{R}_T(\ell, \Phi_T)$$

$$\leq \inf_{\alpha > 0}\left(c_0\alpha + c_1\alpha^{p-1} + c_2\int_\alpha^\eta \sqrt{\frac{\log \mathcal{N}_2(\delta, \Phi_T, T)}{T}}\,d\delta + c_3\int_{\alpha^{p-1}}^\eta \sqrt{\frac{\log \mathcal{N}_2\left(\delta^{\frac{1}{p-1}}, \Phi_T, T\right)}{T}}\,d\delta\right)^{1/q} + c_4 T^{\frac{1-p}{q}}$$

where $c_0 = 4\rho$, $c_1 = 32\gamma\eta$, $c_2 = 12\sqrt{3}\rho$, $c_3 = (24\sqrt{3}\gamma\eta)/\sqrt{(2^{p-1} - 1)}$, $c_4 = (\gamma\eta^p)^{1/q}$.

3.3 General Bounds Under Linearity Assumptions on B

The general results of the previous section can be restated in simpler terms once more assumptions are made. In particular, some of the terms in the three-term decomposition in Theorem 1 can be dropped as soon as $\boldsymbol{B}$ is linear. While some of the results below can be repeated for a more general form $\boldsymbol{B}(z_1, \ldots, z_T) = \sum_{t=1}^T \langle c_t, z_t \rangle$ (for some $c_1, \ldots, c_T \in \mathcal{B}^*$ and $\mathcal{H} \subseteq \mathcal{B}$), for simplicity we assume that $\boldsymbol{B}$ is an average of its arguments and that $\mathcal{H} \subseteq \mathbb{R}$:

$$\boldsymbol{B}(z_1, \ldots, z_T) = \frac{1}{T}\sum_{t=1}^T z_t .$$

Of course, such $\boldsymbol{B}$ is trivially smooth (with $\sigma = 0$), so all the results of the previous section apply.

Corollary 9. *The following statements hold:*

- *The first term in the Triplex Inequality is zero.*

- *If Φ_T is a class of departure mappings, then the second term in the Triplex Inequality is non-positive. In this case,*
$$\mathcal{V}_T(\ell, \Phi_T) \leq 2\mathfrak{R}_T(\ell, \Phi_T).$$

- *Let $\mathcal{H} \subseteq [-1, 1]$. Assuming $|\ell(f, x)| \leq 1$ for any $x \in \mathcal{X}$ and $f \in \mathcal{F}$,*
$$\mathfrak{R}_T(\ell, \Phi_T) \leq 4\inf_{\alpha \geq 0}\left\{\alpha + 3\sqrt{3}\int_\alpha^1 \sqrt{\frac{\log \mathcal{N}_2(\delta, \Phi_T, T)}{T}}\,d\delta\right\} .$$

When B is the average of its coordinates, the sequential complexity takes on a familiar form:

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

Further, for $\mathcal{H} \subseteq \mathbb{R}$, Eq. (7) in definition of the cover becomes

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

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's [22, 3] and fat-shattering dimensions [26]. 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 (see Definition 2). That is, $\Phi_T = \{(\phi, \dots, \phi) : \phi \in \Phi\}$. 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 $\epsilon \in \{\pm 1\}^d$, 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 [26].

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

$$\forall \epsilon \in \{\pm 1\}^d, \ \exists \phi \in \Phi \quad \text{s.t. } \forall t \in [d], \ \epsilon_t \left(\ell_{\phi}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) - \mathbf{s}_t(\epsilon) \right) \geq \alpha/2$$

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 three results without proofs, as the arguments are identical to the ones given in [26]. To be precise, the $(\mathbf{f}, \mathbf{x})$ tree here plays the role of the $\mathbf{x}$ tree in [26], ℓ_{ϕ} for $\phi \in \Phi$ plays the role of $f \in \mathcal{F}$ in [26].

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

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

Furthermore, for $T \geq d$

$$\sum_{i=0}^{d} \binom{T}{i} k^i \leq \left(\frac{ekT}{d} \right)^d.$$

---

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

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,*

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

Theorem 12. *Let $\mathcal{H} \subseteq \{0, \ldots, k\}$ and $\mathrm{fat}_1(\Phi) = d$. Then*

$$\mathcal{N}(0, \Phi, T) \leq \sum_{i=0}^{d} \binom{T}{i} k^i \leq (ekT)^d.$$

Furthermore, for $T \geq d$

$$\sum_{i=0}^{d} \binom{T}{i} k^i \leq \left(\frac{ekT}{d} \right)^d.$$

In particular, the result holds for binary-valued function classes ($k = 1$), in which case $\mathrm{fat}_1(\Phi) = \mathrm{Sdim}(\Phi)$.

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 5.1).

3.4 General Bounds for Slowly-Varying Payoff Transformations

In Section 3.3.1, we assumed that the set Φ_T of sequences of payoff transformations is time-invariant. This assumption naturally leads to a control on the complexity of Φ_T. Lifting the assumption of time-invariance, we now go back to the level of generality of Proposition 6 and Theorem 8. We observe that size of Φ_T or an appropriately behaving covering number $\mathcal{N}_2(\alpha, \Phi_T, T)$ is key for bounding the sequential complexity. If payoff transformations change wildly in time, there is little hope of getting non-trivial bounds. The good news is that, 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 [19, 7] that it is possible to have small external regret against comparators that change a limited number of times. This alleviates an obvious limitation of the classical notion of external regret, viz., comparison to the fixed best decision. Another result of this flavor appears in [28], where *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.

3.4.1 Tracking the Best Transformation

Suppose Φ is a finite set of payoff transformations. Let Φ_T^k be obtained by considering all piecewise constant sequences with k changes:

$$\Phi_T^k = \{(\phi_1, \ldots, \phi_T) : 1 = i_0 \leq i_1 \leq \ldots \leq i_k \leq T \text{ and } \phi_t = \phi_{t'} \text{ if } i_s \leq t \leq t' < i_{s+1} \text{ for some } s \geq 0\}.$$

If cardinality $|\Phi| = N$, it is easy to check that $|\Phi_T^k| \leq \binom{T}{k} \cdot N^{k+1}$. Under the assumptions of Proposition 6, this immediately implies a bound of the order

$$\left(R^2 (k \log N + k \log T) T \right)^{1/2q} + \sigma^{1/q} T^{1/q}.$$

It is natural to extend the above results by lifting the assumption that Φ is a finite set of payoff transformations. This can be done by considering an online cover $\mathcal{N}_p(\ell, \Phi, \alpha)$ of Φ in some ℓ_p norm along with the same definition of Φ_T^k. Next we do this in an even more general setting.

3.4.2 Slowly Changing Transformations

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

$$\Phi_T^{k,\alpha} = \Big\{ (\phi_1, \ldots, \phi_T) : 1 = i_0 \le i_1 \le \ldots \le i_k \le T$$
$$\text{and} \sup_{f,x} \|\ell_{\phi_t}(f,x) - \ell_{\phi_{t'}}(f,x)\| \le \alpha \text{ if } i_s \le t \le t' < i_{s+1} \text{ for some } s \ge 0 \Big\}.$$

One can think of the time-invariant segments as "accumulation points" where the payoff transformations do not vary much.

Suppose that we have a finite cover V of Φ at scale α, of cardinality $|V| = \mathcal{N}_\infty(\alpha, \Phi, T)$. The L_∞ covering is chosen for the purposes of simplicity, though tighter (and more difficult) results are expected from directly studying L_2 covering numbers.

Lemma 13. *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.

Suppose that Φ_T is a set of sequences $(\phi_1, \ldots, \phi_T)$ of payoff transformations which do not "vary much", according to the following definition. The length of a sequence $(\phi_1, \ldots, \phi_T)$ of payoff transformations (with respect to L_∞ distance) is defined as

$$\text{len}(\phi_1, \ldots, \phi_T) := \sum_{t=1}^{T-1} \sup_{f,x} \left\| \ell_{\phi_t}(f,x) - \ell_{\phi_{t+1}}(f,x) \right\| .$$

Again, we consider the L_∞ distance between payoffs (as functions over $\mathcal{F} \times \mathcal{X}$) for simplicity, though one can also define weaker notions. Assume that for all sequences in Φ_T, their length is bounded by some $L > 0$. We will now claim that by choosing k large enough, the set of covering trees V^k defined in the proof of Lemma 13 provides a cover for Φ_T at a given scale $\alpha > 0$. Consider any $(\phi_1, \ldots, \phi_T) \in \Phi_T$. We construct the nondecreasing sequence $i_1, \ldots, i_j, \ldots \in \{1, \ldots, T\}$ 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 the previous section: 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:

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

and
$$\log \mathcal{N}_\infty(2\alpha, \Phi_T, T) \leq O\left(\frac{L}{\alpha}\log T + \frac{L}{\alpha}\log \mathcal{N}_\infty(\alpha, \Phi, T)\right).$$

The covering number can be now used in Theorem 8 to control sequential complexity.

4 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 10. 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 $\{\overline{q_t^*}\}$ 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 14. *For any Equalizer strategy $\{p_t^*\}$ we have that for any $f \in \mathcal{F}$,*

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

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

Remark 15. *For many interesting games we consider it is often the case that for any $x_1, \ldots, x_T$ and any $f_1, \ldots, f_T, f_1', \ldots, f_T'$,*

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

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

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

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. [9, p. 252]). In particular, in [1], 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 [1]. Furthermore, [20] 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.

5 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 and ways of making more comprehensive statements.

5.1 Φ-Regret

In this section, we consider a particular notion of performance measure, known as Φ-regret [27, 16, 17]. 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 (see Definitions 1 and 2). The particular choices of Φ lead to various notions, such as external, internal, swap regret, and more.

To define Φ-regret (Example 2), 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:

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

where $\mathcal{H} \subseteq \mathbb{R}$. Since $\boldsymbol{B}$ is the average of its arguments, Corollary 9 implies

Corollary 16. *In the setting of Φ-regret,*

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

Specializing the definition of sequential complexity to Φ-regret, we obtain the following definition.

Definition 11. The sequential complexity for Φ-regret is defined as

$$\mathfrak{R}_T(\ell, \Phi) = \sup_{(\mathbf{f}, \mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell(\phi \circ \mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \tag{8}$$

where, as before, the first supremum is over $\mathcal{F} \times \mathcal{X}$-valued trees $(\mathbf{f}, \mathbf{x})$ of depth T.

The following property allows us to immediately obtain bounds for convex hulls of finite sets Φ.

Proposition 17. *Suppose ℓ is convex in the first argument and $\mathrm{conv}(\Phi)$ maps $\mathcal{F}$ into $\mathcal{F}$. Then*

$$\mathfrak{R}_T(\ell, \mathrm{conv}(\Phi)) = \mathfrak{R}_T(\ell, \Phi) \ .$$

We also have the following version of the contraction lemma, whose proof is identical to that given in [26].

Lemma 18. *Fix a function $\psi : \mathbb{R} \times \mathcal{F} \times \mathcal{X} \mapsto \mathbb{R}$ such that for any $f \in \mathcal{F}, x \in \mathcal{X}$, $\psi(\cdot, f, x)$ is a Lipschitz function with a constant L. Then*

$$\mathfrak{R}(\psi \circ \ell, \Phi) \leq L \cdot \mathfrak{R}(\ell, \Phi)$$

where $\psi \circ \ell$ is defined by the mapping $(f, x) \mapsto \psi(\ell(f, x), f, x)$ for all $f \in \mathcal{F}, x \in \mathcal{X}$.

Next, we specialize Definition 7 to the particular case of Φ-regret.

Definition 12. 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

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

The *covering number* of Φ_T on a given tree $(\mathbf{f}, \mathbf{x})$ is defined as the size of the minimum cover, as in Definition 7.

We now turn to particular examples to utilize the results and definitions stated above.

5.1.1 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 [26].

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 ϕ 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 11 of the sequential complexity for Φ-regret. Because each $\phi_f \in \Phi$ is a constant mapping, we have

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

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

If payoff is written as $\ell(f, x) = f(x)$, this is precisely the sequential Rademacher complexity defined in [26].

Next, we show that Definition 12 reduces to the definition of online covering given in [26]. Indeed, $\ell_{\phi_f}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) = \ell(f, \mathbf{x}_t(\epsilon))$ for the constant mappings $\phi = (\phi_f, \ldots, \phi_f)$. Further, the payoff space $\mathcal{H} \subseteq \mathbb{R}$. With these simplifications, the closeness to a covering element in Definition 12 becomes

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

where V is a set of $\mathbb{R}$-valued trees. It is then immediate that Corollary 9 recovers the corresponding result of [26]. For a detailed study of external regret, we refer the reader to the companion paper [26].

Lower Bounds in the Supervised Setting We provide a lower bound for external regret in the supervised learning setting using the notion of an equalizer (see Section 4). To this end, we assume that $\mathcal{X} = \mathcal{Z} \times \mathcal{Y}$ where $\mathcal{Z}$ is the space of predictors and $\mathcal{Y}$ is the space of responses (outcomes). The setting is called *supervised* because, in the machine learning terminology, the observed data is thought of as examples together with labels. Assume $\mathcal{F}$ is a class of bounded real-valued functions and the space of outcomes is a bounded interval; for simplicity let $\mathcal{F} \subseteq [-1, 1]^{\mathcal{Z}}$ and $\mathcal{Y} = [-1, 1]$. Suppose the loss is of the form $\ell(f, (z, y)) = |f(z) - y|$.

Proposition 19. *The value of the supervised game defined above is lower bounded by sequential Rademacher complexity:*

$$\mathcal{V}_T^S(\ell, \Phi_T) \geq \mathfrak{R}_T(\ell, \Phi)$$

Proof. Recall that we have a fixed set Φ of *constant* departure mappings. We will now exhibit an equalizer strategy. Following Remark 15, observe that for any $(z_1, y_1), \ldots, (z_T, y_T)$ and any $f_1, \ldots, f_T, f_1', \ldots, f_T'$,

$$\inf_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} |(\phi \circ f_t)(z_t) - y_t| = \inf_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} |(\phi \circ f_t')(z_t) - y_t|$$

because any $\phi \in \Phi$ is a constant mapping. Thus, for a strategy to be an equalizer, it only needs to "equalize" the cumulative loss of the player. Here is how we construct such a strategy. Let p^y be defined as the distribution of a Rademacher ± 1 random variable Y; this will define the labels y_t as independent coin flips. Now, fix any $\mathcal{Z}$-valued tree $\mathbf{z}$ of depth T. Let $\{p_t^*\}$ be a strategy defined by $p_t^*(y_{1:t-1}) = \delta_{\mathbf{z}_t(y_{1:t-1})} \times p^y$, a delta distribution on $\mathbf{z}_t(y_{1:t-1})$ defined by the tree $\mathbf{z}$ and p^y on $\mathcal{Y}$. In plain words, the strategy of the adversary for each t is to choose a particular $z_t \in \mathcal{Z}$ given the labels $y_1, \ldots, y_{t-1}$, and let the label be an independent Rademacher random variable.

By Remark 15, it is enough to check

$$\underset{\substack{(z_1,y_1) \sim p_1^* \\ f_1 \sim q_1^*}}{\mathbb{E}} \cdots \underset{\substack{(z_T,y_T) \sim p_T^* \\ f_T \sim q_T^*}}{\mathbb{E}} \frac{1}{T} \sum_{t=1}^{T} |f_t(z_t) - y_t| = \underset{\substack{(z_1,y_1) \sim p_1^* \\ f_1 \sim \overline{q_1^*}}}{\mathbb{E}} \cdots \underset{\substack{(z_T,y_T) \sim p_T^* \\ f_T \sim \overline{q_T^*}}}{\mathbb{E}} \frac{1}{T} \sum_{t=1}^{T} |f_t(z_t) - y_t|$$

for all strategies $\{q_t^*\}$ and $\{\overline{q_t^*}\}$ of the player. This equality is indeed true because $\underset{y_t \sim p^y}{\mathbb{E}} |a - y_t| = 1$ independently of the constant $a \in [-1, 1]$. By Proposition 14, for any $g \in \mathcal{F}$

$$\mathcal{V}_T^S(\ell, \Phi_T) \geq \underset{(z_1,y_1) \sim p_1^*}{\mathbb{E}} \cdots \underset{(z_T,y_T) \sim p_T^*}{\mathbb{E}} \left[\frac{1}{T} \sum_{t=1}^{T} |g(z_t) - y_t| - \inf_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} |f(z_t) - y_t| \right]$$

$$= \underset{y_1, \ldots, y_T}{\mathbb{E}} \left[1 - \inf_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} |f(\mathbf{z}_t(y_{1:t-1})) - y_t| \right]$$

$$= \underset{y_1, \ldots, y_T}{\mathbb{E}} \left[\sup_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} y_t f(\mathbf{z}_t(y_{1:t-1})) \right]$$

where $y_1, \ldots, y_T$ are i.i.d. Rademacher random variables. Since the lower bound holds for any $\mathcal{Z}$-valued tree $\mathbf{z}$ of depth T, it also holds for the supremum:

$$\mathcal{V}_T^S(\ell, \Phi_T) \geq \sup_{\mathbf{z}} \underset{y_1, \ldots, y_T}{\mathbb{E}} \left[\sup_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} y_t f(\mathbf{z}_t(y_{1:t-1})) \right] = \mathfrak{R}_T(\ell, \Phi) \ .$$

Hence, the lower bound on the value of the supervised game is the sequential Rademacher complexity of $\mathcal{F}$. $\qquad \square$

Lower Bounds for Online Convex Optimization We first provide a lower bound for a linear game. By Lemma 30, this lower bound will also serve as a lower bound for a convex Lipschitz game. We remark that these lower bounds are not entirely new (see e.g. [1, 2]), and we derive them here for the purposes of completeness, as well as to stress that they arise from an equalizing strategy.

Suppose $\mathcal{F}$ is a unit ball in some norm $\| \cdot \|$ and $\mathcal{X}$ is a unit ball in the dual norm $\| \cdot \|_*$. The loss $\ell(f, x) = x(f) = \langle f, x \rangle$ and the set Φ is, again, a set of constant departure mappings.

Proposition 20. *The value of the linear game defined above is lower bounded by sequential Rademacher complexity:*

$$\mathcal{V}_T(\ell, \Phi_T) \geq \mathfrak{R}_T(\ell, \Phi).$$

Hence, the value of the convex Lipschitz game (where $\mathcal{X}$ is the set of all 1-Lipschitz convex functions on $\mathcal{F}$) is also lower bounded by the same quantity.

Proof. Similarly to the proof for the supervised game, observe that for any $x_1, \ldots, x_T$ and any $f_1, \ldots, f_T, f_1', \ldots, f_T'$,

$$\inf_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \langle \phi(f_t), x_t \rangle = \inf_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \langle \phi(f_t'), x_t \rangle$$

because any $\phi \in \Phi$ is a constant mapping. Following Remark 15, we only need to exhibit a strategy that equalizes the player's loss. To this end, fix an $\mathcal{X}$-valued tree $\mathbf{x}$ of depth T. Consider the adversary's strategy where at each step an ϵ_t is chosen uniformly at random from $\{\pm 1\}$ and $x_t = \epsilon_t \cdot \mathbf{x}(\epsilon_{1:t-1}) \in \mathcal{X}$.

By Remark 15, it is enough to check

$$\underset{f_1 \sim q_1^* \epsilon_1}{\mathbb{E}} \underset{f_T \sim q_T^* \epsilon_T}{\mathbb{E}} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \langle f_t, \mathbf{x}(\epsilon_{1:t-1}) \rangle = \underset{f_1 \sim \overline{q_1^*} \epsilon_1}{\mathbb{E}} \underset{f_T \sim \overline{q_T^*} \epsilon_T}{\mathbb{E}} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \langle f_t, \mathbf{x}(\epsilon_{1:t-1}) \rangle$$

for all strategies $\{q_t^*\}$ and $\{\overline{q_t^*}\}$ of the player. This equality is indeed true because both terms are identically zero. By Proposition 14, for any $g \in \mathcal{F}$

$$\mathcal{V}_T(\ell, \Phi_T) \geq \underset{\epsilon_1, \ldots, \epsilon_T}{\mathbb{E}} \left[\frac{1}{T} \sum_{t=1}^{T} \epsilon_t \langle g, \mathbf{x}(\epsilon_{1:t-1}) \rangle - \inf_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \langle f, \mathbf{x}(\epsilon_{1:t-1}) \rangle \right]$$

$$= \underset{\epsilon_1, \ldots, \epsilon_T}{\mathbb{E}} \sup_{f \in \mathcal{F}} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \langle f, \mathbf{x}(\epsilon_{1:t-1}) \rangle .$$

Since this holds for any $\mathcal{X}$-valued tree $\mathbf{x}$, we have proven the statement. $\qquad\square$

5.1.2   Internal and Swap Regret

Assume the cardinality $N = |\mathcal{F}|$ is finite. For internal regret, Φ is the set of mappings $\{\phi_{f \to g} : \phi_{f \to g}(f) = g$ and $\phi_{f \to g}(h) = h \ \forall h \neq f, h \in \mathcal{F}\}$. For swap regret [5, 9], Φ contains all N^N functions from $\mathcal{F}$ to itself. It is easy to see that the finite class lemma (Lemma 5) 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 [9].

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 [17, 16]. 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.

5.1.3   Convergence to Φ-correlated Equilibria

A beautiful result of Foster and Vohra [12] 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 [27] 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 is 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 equilibrim 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,

$$\forall k \in [N], \forall \phi \in \Phi, \qquad \mathbb{E}_{(f_1,\ldots,f_N) \sim \pi} \left[\ell_k(f_k, f_{-k}) \right] \leq \mathbb{E}_{(f_1,\ldots,f_N) \sim \pi} \left[\ell_k(\phi(f_k), f_{-k}) \right] \; .$$

Theorem 18 in [27] 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 a particular player k. The regret minimizing algorithm for it is simply a $\tilde{\Phi}$-regret minimizing algorithm with $\ell(f, x) = x(f)$ where we have identified the adversary set $\mathcal{X}$ with the class of functions $\{f \mapsto \ell_k(f, g) \; : \; g \in \mathcal{F}^{k-1}\}$, where g is a strategy profile over the remaining $k-1$ players. Examining Stoltz and Lugosi's proof reveals that $\tilde{\Phi}$ is taken to be a dense countable subset of Φ and an explicit regret minimizing algorithm for countably infinite classes of departure mappings is used. The regret w.r.t. each $\phi \in \Phi$ does go to zero but the rate is not uniform in ϕ. In particular, it depends on the order in which the class $\tilde{\Phi}$ is enumerated. Later, they also consider examples of uncountable classes Φ of departure mapping where non-asymptotic rates of convergence for Φ-regret can be obtained. Specifically, they use the metric entropy of Φ. We show how to improve their bounds using sequential complexity.

As an example, 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 $\mathcal{N}_{\mathrm{metric}}(\Phi, \alpha)$ 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 $\mathcal{N}_{\mathrm{metric}}(\Phi, \alpha) = \Theta(1/\alpha^p)$. To upper bound the Φ-regret we can always make the set of adversary's moves larger. In fact, we make 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, by Lemma 30, we have $\mathcal{V}_T(\mathcal{C}_\mathcal{F}, \mathcal{F}, \Phi) = \mathcal{V}_T(\mathcal{L}_\mathcal{F}, \mathcal{F}, \Phi)$ where

$$\mathcal{L}_\mathcal{F} = \{x : \mathcal{F} \to \mathbb{R} \; : \; x \text{ linear and 1-Lipschitz}\} \; .$$

Then the sequential complexity bound is

$$\sup_{(\mathbf{f}, \mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle \phi(\mathbf{f}_t(\epsilon)), \mathbf{x}_t(\epsilon) \right\rangle \; . \tag{9}$$

Note that the set $\mathcal{X}$ is now just the set of 1-Lipschitz linear functions, i.e. elements in the unit ball of the dual space. Since $\|\phi_1 - \phi_2\|_\infty \leq \alpha$ implies

$$|\langle \phi_1(f), x \rangle - \langle \phi_2(f), x \rangle| \leq \alpha$$

for any $x \in \mathcal{X}$, we can use metric entropy inside Dudley's integral to upper bound the sequential complexity by

$$c \inf_\alpha \left(\alpha T + \sqrt{T} \int_{\alpha'=\alpha}^{1} \sqrt{\frac{1}{\alpha'^p}} d\alpha' \right) \; .$$

This bound 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 general bound of $O(T^{(p+1)/(p+2)})$ given in Example 23 of [27].

---

[3]The set of continuous function on $\mathcal{F}$ equipped with the supremum norm

5.1.4 Linear Transformations

In this section we consider the following scenario, discussed in [16]. Suppose $\mathcal{F}$ is a subset of a Hilbert space $\mathcal{M}$. Let Φ be the set of Lipschitz linear transformations on $\mathcal{F}$, i.e. $\Phi = \{M \in \mathcal{F} \to \mathcal{F} : \|M\| \le R\}$ for some operator norm $\|\cdot\|$. Let $\|\cdot\|_*$ be dual to $\|\cdot\|$. We are assuming the Online Convex Optimization scenario, i.e. $\mathcal{X}$ is a set of L-Lipschitz real-valued convex functions on $\mathcal{F}$ and the loss is defined as $\ell(f,x) = x(f)$. Furthermore,

$$\ell_{\phi_M}(f,x) = x(Mf).$$

Therefore, we are in the setting of the well-studied online convex optimization (possibly in an infinite-dimensional Hilbert space), yet instead of being compared to the value of the best fixed point f^* in hindsight, the player is being evaluated according to the best linear transformation of his trajectory $f_1, \ldots, f_T$. Is this problem learnable?

By Lemma 30, the value of the convex game is equal to the value of the associated linear game. Suppose functions $x \in \mathcal{X}$ have gradients bounded by L in the ℓ_2 norm. The value of the convex game is upper bounded by the sequential complexity of the class of linear payoffs $\ell^{\mathrm{lin}}(f, \tilde{x}) = \langle f, \tilde{x} \rangle$. Then the sequential complexity bound is

$$\sup_{(\mathbf{f},\mathbf{x})} \mathbb{E}_{\epsilon_{1:T}} \sup_{M \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle M\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon) \right\rangle , \tag{10}$$

which can be upper bounded by $R \cdot L \cdot \mathrm{diam}_2(\mathcal{F})$. Note that these results hold in infinite-dimensional Hilbert spaces, where a metric entropy-type cover of $\mathcal{F}$ would not even be finite.

5.2 Blackwell's Approachability

Blackwell's Approachability Theorem [4, 21, 9] 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 [9]. For instance, existence of calibrated forecasters can be deduced from Blackwell's Approachability Theorem [23, 12].

Let us first discuss the relation of our results to Blackwell's Theorem. A proof of Blackwell's Theorem (see for instance [9]) 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 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 [26] (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 21). We also show that martingale convergence is necessary for the result to hold (Theorem 23). 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. To view this

problem as an instance of our general framework, define

$$\boldsymbol{B}(z_1, \ldots, z_T) = \inf_{c \in S} \left\| \frac{1}{T} \sum_{t=1}^{T} z_t - c \right\|.$$

The comparator term is zero by our assumption that Φ_T contain 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 $1 \le t \le T$. Thus, indeed, the performance measure is

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

the distance to the set S. The next condition on the payoff ℓ says that it must that the player can choose a "good" mixed strategy q in response to a given mixed strategy p of the adversary. This strategy q should, on average, put the payoff inside the set S. Recall that $\ell(q, p)$ is simply a short-hand for the expected payoff $\mathbb{E}_{f \sim q, x \sim p} \ell(f, x)$ (that is, we do not make any assumptions about linearity of ℓ).

Definition 13. Given a set S, 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 $\ell(q, p) \in S$.

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.

Definition 14. We will say that *martingale convergence holds* if

$$\lim_{T \to \infty} \sup_{\mathbf{M}} \mathbb{E} \left[\left\| \frac{1}{T} \sum_{t=1}^{T} d_t \right\| \right] = 0,$$

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

We now show that, under the one-shot approachability condition, the set is approachable whenever martingale convergence holds in the subset of the Banach space.

Theorem 21. *For any game that is one shot approachable, we have that*

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

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

Proof. Now we apply Theorem 1 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, $\boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) = 0$ for any $f_1, \ldots, f_T \in \mathcal{F}$, $x_1, \ldots, x_T \in \mathcal{X}$. We then conclude that

$$\mathcal{V}_T(\ell, \Phi_T) \le \sup_{p_1, q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T, q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \right\}$$

(11)

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

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

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

$$\leq \sup_{\substack{p_1,q_1 \\ }} \underset{\substack{f_1\sim q_1 \\ x_1\sim p_1}}{\mathbb{E}} \ldots \sup_{\substack{p_T,q_T \\ }} \underset{\substack{f_T\sim q_T \\ x_T\sim p_T}}{\mathbb{E}} \underset{\substack{f'_{1:T}\sim q_{1:T} \\ x'_{1:T}\sim p_{1:T}}}{\mathbb{E}} \left\{ \left\|\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f'_t,x'_t)\right\| \right\}$$

$$\leq \sup_{\substack{p_1,q_1 \\ }} \underset{\substack{f_1,f'_1\sim q_1 \\ x_1,x'_1\sim p_1}}{\mathbb{E}} \ldots \sup_{\substack{p_T,q_T \\ }} \underset{\substack{f_T,f'_T\sim q_T \\ x_T,x'_T\sim p_T}}{\mathbb{E}} \left\{ \left\|\frac{1}{T}\sum_{t=1}^{T}\ell(f_t,x_t)-\frac{1}{T}\sum_{t=1}^{T}\ell(f'_t,x'_t)\right\| \right\}$$

where in the first inequality we used $\inf_a[C_1(a)] - \inf_a[C_2(a)] \leq \sup_a[C_1(a) - C_2(a)]$ 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. (11) is

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

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

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

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

$$\leq \sup_{p_1}\inf_{q_1}\ldots\sup_{p_T}\inf_{q_T} \left\{\inf_{c\in S}\left\|c-\frac{1}{T}\sum_{t=1}^{T}\ell(q_t,p_t)\right\|\right\} \qquad (12)$$

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

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. (12)) 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. (12) 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. □

We now discuss lower bounds on the value of Blackwell's approachability game. The first lower bound is straightforward.

Proposition 22. *Suppose martingale convergence holds. 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 23. *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 the supremum is over distributions $\mathbf{M}$ of martingale difference sequences $\{d_t\}_{t \in \mathbb{N}}$ such that each $d_t \in \mathcal{H}$.

Proof. 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 $\ell(f, x) = f \cdot x$. 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 $x_t = \epsilon_t \mathbf{x}_t(f_1 \cdot \epsilon_1, \ldots, f_{t-1} \cdot \epsilon_{t-1})$ that is a random sign multiplied with the instance given by the path on the tree specified by $f_1 \cdot \epsilon_1, \ldots, f_{t-1} \cdot \epsilon_{t-1}$. 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

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

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

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

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

The first equality above is due to the fact that $f_T \cdot \epsilon_T$ is a Rademacher random variable conditioned on $f_1, \ldots, f_{T-1}$ and $\epsilon_1, \ldots, \epsilon_{T-1}$ 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 $f_1, \ldots, f_{t-1}$ and $\epsilon_1, \ldots, \epsilon_{t-1}$; and replacing each q_t^* by $\overline{q_t^*}$. Hence we see that the adversary strategy is an equalizer strategy. Hence using Proposition 14 and picking the fixed $f = 1$ we see that

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

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 [25]. □

5.3 Calibration

Calibration, introduced by Brier [8] and Dawid [10], is an important notion for forecasting binary sequences. In the context of weather forecasting, calibration means that, for the days the forecaster announced "30% chance of rain", the empirical frequency of rain should indeed be close to 30% [9, p. 85]; moreover, this has to hold for any forecasted value. The existence of calibrated forecasters, a fact which is not obvious a priori, was shown by Foster and Vohra [13]. Following [9], we consider the notion of λ-calibration. If a forecaster is λ-calibrated for all $\lambda > 0$, we say that the forecaster is well calibrated.

In what follows, we formulate the calibration problem of forecasting $\{1, \ldots, k\}$-valued sequences in our general framework. In particular, we are interested in sharp rates on the resulting value of the calibration game, and we will compare our results with the recent work of Mannor and Stoltz [23].

Fix a norm $\|\cdot\|$ on $\mathbb{R}^k$. Let $\mathcal{H} = \mathbb{R}^k$, $\mathcal{F} = \Delta(k)$, and $\mathcal{X}$ the set of standard unit vectors in $\mathbb{R}^k$ (vertices of $\Delta(k)$). Define $\ell(f, x) = 0$; that is, the forecaster is penalized only through the comparator term. We define $\boldsymbol{B}(z_1, \ldots, z_T) = -\left\| \frac{1}{T} \sum_{t=1}^T z_t \right\|$. Define $\Phi_T = \{ (\phi_{p,\lambda}, \ldots, \phi_{p,\lambda}) : p \in \Delta(k), \lambda > 0 \}$ to contain time-invariant mappings defined by

$$\ell_{\phi_{p,\lambda}}(f, x) = \mathbf{1}\{\|f - p\| \leq \lambda\} \cdot (f - x) .$$

This definition of the loss is indeed natural for the λ-calibration problem. It says that, for any p chosen after the game, if we consider a round when the player predicted $f \in \Delta(k)$ close to p, the loss should be the difference between the actual outcome x and f. Indeed, when we put all the definitions together, we obtain

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

Note that this notion of regret allows the worst scale λ to be chosen at the end of the game. This makes it a stronger requirement than what is required for building a well calibrated forcaster. Nevertheless, we can bound the value of this game, improving on the results of Mannor and Stoltz [23]. Theorem 24 shows that the rate of calibration is $\tilde{O}(T^{-1/3})$ no matter what k is. The rate of $\tilde{O}(T^{-1/3})$ has been established for $k = 2$ previously. For $k > 2$, however, the best rates known to us (due to [23]) deteriorate with k. Let us remark that some looseness of the approach of [23] comes from discretization in order to phrase the problem as Blackwell's approachability. A reader will note that we also pass to a discretization in the proof below. However, this is done late in the analysis in order to upper bound the sequential complexity. This seems to speak in favor of our approach, aimed at directly looking at the complexity of the problem through the notion of sequential complexity.

Theorem 24. *For the calibration game with k outcomes and with ℓ_1 norm, we have that for $T \geq 3$ and some absolute constant c*

$$\mathcal{V}_T(\ell, \Phi_T) \leq ck \left(\frac{\log T}{T} \right)^{1/3}$$

whenever $T > k^3$. If, on the other hand, $T < k^3$, the statement holds with $k^{4/3}$ instead of k.

Proof. 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 1 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

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

$$= \sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\lambda > 0} \sup_{p \in \Delta(k)} \mathop{\mathbb{E}}_{\substack{x_{1:T} \sim p_{1:T} \\ f_{1:T} \sim q_{1:T}}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell_{\phi_{p,\lambda}}(f_t, x_t) \right\|$$

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

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

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

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

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

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

$$\sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \left[- \mathop{\mathbb{E}}_{\substack{x_{1:T} \sim p_{1:T} \\ f_{1:T} \sim q_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \right] \leq 2\delta + 2\sqrt{\frac{k}{T}} . \tag{13}$$

We now proceed to upper bounded the third term in the Triplex Inequality. Since $-\boldsymbol{B}$ is a subadditive, by Theorem 2, we have that the third term is bounded by twice the sequential complexity

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

$$= 2 \sup_{\mathbf{f}, \mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{\lambda > 0} \sup_{p \in \Delta(k)} \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \mathbf{1} \left\{ \|\mathbf{f}_t(\epsilon) - p\| \leq \lambda \right\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon)) \right\|$$

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 Proposition 6.

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

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

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

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

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 [15] 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 $\|f - p\|_1 \le \lambda$ 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 $|S| \le |C_\delta|^{c \cdot k^2}$ such that

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

$$= 2\sup_{\mathbf{f},\mathbf{x}}\ \mathbb{E}_{\epsilon_{1:T}}\max_{(p,\lambda)\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\|_1 \le \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon))\right\|_1$$

$$\le 2\,k^{\frac{q-1}{q}}\sup_{\mathbf{f},\mathbf{x}}\mathbb{E}_\epsilon\max_{(p,\lambda)\in S}\left\|\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\mathbf{1}\{\|\mathbf{f}_t(\epsilon) - p\|_1 \le \lambda\} \cdot (\mathbf{f}_t(\epsilon) - \mathbf{x}_t(\epsilon))\right\|_q$$

for some $q \in (1, 2]$ to be chosen later. Now note that $\|\cdot\|_q^q$ is (q, q)-smooth and so applying Corollary 7, we see that

$$2\mathfrak{R}_T(\ell, \Phi_T, -\boldsymbol{B}) \le 2k^{1-1/q}\left(\left(\frac{2\log(|S|)}{T}\right)^{1/2q} + 2\left(\frac{1}{T}\right)^{1-1/q}\right)$$

$$\le 2k^{1-1/q}\left(\frac{2ck^2\log(|C_\delta|)}{T}\right)^{1/2q} + 4k^{1-1/q}\left(\frac{1}{T}\right)^{1-1/q}$$

$$= c'k\left(\frac{\log(|C_\delta|)}{T}\right)^{1/2q} + 4k^{1-1/q}\left(\frac{1}{T}\right)^{1-1/q}$$

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

$$2\mathfrak{R}_T(\ell, \Phi_T, -\boldsymbol{B}) \le c'k^{1+1/2q}\left(\frac{\log(1/\delta)}{T}\right)^{1/2q} + 4k^{1-1/q}\left(\frac{1}{T}\right)^{1-1/q}.$$

Now we pick the $q \in (1, 2)$ that balances the the first and second term of the sum above (up to constant factors and the $\log 1/\delta$ term) as $q = \frac{3 \log T/k}{2 \log T}$ and in turn get :

$$2\mathfrak{R}_T(\ell, \Phi_T, -\boldsymbol{B}) \leq c'k \left(\frac{\log(1/\delta)}{T} \right)^{1/3} + 4k \left(\frac{1}{T} \right)^{1/3} .$$

Combining the above upper bound on the third term of triplex inequality and Equation 13 that bounds the second term of the triplex inequality (and since first term is anyway 0) we see that,

$$\mathcal{V}_T \leq 2\delta + 2\sqrt{\frac{k}{T}} + c'k \left(\frac{\log(1/\delta)}{T} \right)^{1/3} + 4k \left(\frac{1}{T} \right)^{1/3} .$$

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

5.4 Other Examples

5.4.1 External Regret with Global Costs

Let us consider a more general setting where the (vector) loss is $\ell(f, x)$ rather than the specific choice $f \odot x$ in Example 5. The Triplex Inequality and Theorem 2 then gives

$$\mathcal{V}_T \leq \sup_{\substack{p_1, q_1 \\ f_1 \sim q_1 \\ x_1 \sim p_1}} \mathbb{E} \ldots \sup_{\substack{p_T, q_T \\ f_T \sim q_T \\ x_T \sim p_T}} \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \left\| \frac{1}{T} \sum_{t=1}^{T} (\ell(f_t, x_t) - \ell(f'_t, x'_t)) \right\|$$

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

$$+ 2 \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)) \right\| .$$

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

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

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:

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

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

$$\leq \sup_{p_1} \ldots \sup_{p_T} \left\{ \mathbb{E}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, x_t) \right\| - \inf_{f \in \mathcal{F}} \mathbb{E}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, x_t) \right\| \right\}$$

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,

$$
\mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, x_t) \right\| - \inf_{f \in \mathcal{F}} \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, x_t) \right\|
$$
$$
\leq \left\{ \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, p_t) \right\| - \inf_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \ell(f, p_t) \right\| \right\} + \mathop{\mathbb{E}}_{x_{1:T} \sim p_{1:T}} \left\| \frac{1}{T} \sum_{t=1}^{T} (\ell(f_t, x_t) - \ell(f_t, p_t)) \right\| \quad (14)
$$

Now we make an important assumption.

Assumption 25. *Suppose that, for any* p_1, p_2,

$$\inf_f \|\ell(f, p_1) + \ell(f, p_2)\| \geq \inf_f \|\ell(f, p_1)\| + \inf_f \|\ell(f, p_2)\| \ .$$

Under Assumption 25, along with the way we chose f_t, the first term in (14) becomes

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

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:

$$\mathcal{V}_T \ \leq \ 4 \sup_{\mathbf{M}} \mathbb{E} \left[\left\| \frac{1}{T} \sum_{t=1}^{T} d_t \right\| \right] + 2 \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \ell(f, \mathbf{x}_t(\epsilon)) \right\| \ . \quad (15)$$

Let us see what this implies in a specific case of interest.

Global Cost Learning on the Simplex Here we consider Example 5, the setting studied in Even-Dar et al [11]. Let $\mathcal{F} = \Delta(k)$, $\mathcal{X} = [0,1]^k$ and $\ell(f, x) = f \odot x$. Let us first verify if Assumption 25 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 29 in the appendix proves that it is. Note that in [11], 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 (15) holds. In the case we are considering, we can further massage

30

the second term in that upper bound. Note that for any f and y, $\|f \odot y\| \leq \|f\|_\infty \|y\| \leq \|y\|$. Hence, we have

$$\sup_{f \in \mathcal{F}} \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t (f \odot \mathbf{x}_t(\epsilon)) \right\| = \sup_{f \in \mathcal{F}} \left\| f \odot \left(\frac{1}{T} \sum_{t=1}^{T} \epsilon_t \mathbf{x}_t(\epsilon) \right) \right\|$$

$$\leq \left\| \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \mathbf{x}_t(\epsilon) \right\|$$

Hence using the above in (15) 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 martingale difference sequences $\{d_t\}_{t \in \mathbb{N}}$ such that each $d_t \in [-1,1]^k$.

5.4.2 Adaptive Regret

To study online learning in changing environment Hazan and Seshadhri defined the notion of *adaptive regret* in [18]. The notion of adaptive regret introduced in [18] was mainly one where cumulative loss for any time interval is compared to the best predictor at hindsight for that particular interval. We first extend the notion of adaptive regret in [18] to include departure mappings as,

$$\mathbf{R}_T := \sup_{[r,s] \subseteq [T]} \left\{ \frac{1}{T} \sum_{t=r}^{s} \mathrm{loss}(f_t, x_t) - \inf_{\psi \in \Psi} \frac{1}{T} \sum_{t=r}^{s} \mathrm{loss}(\psi \circ f_t, x_t) \right\} \tag{16}$$

where $\mathrm{loss} : \mathcal{F} \times \mathcal{X} \mapsto \mathbb{R}$ is some arbitrary loss function and Ψ is some class of departure mappings. The key idea in the above definition of regret is that we consider the worst time interval and consider the regret for that time interval versus some fixed set of departure mappings.

We capture the above notion of regret in our framework by defining :

- $\ell(f, x) = 0$ for all $f \in \mathcal{F}$ and $x \in \mathcal{X}$

- Define the set of time-invariant payoff transformations $\Phi_T = \mathcal{I}_T \times \Psi_T$ where $\Psi_T = \{(\psi, \ldots, \psi) : \psi \in \Psi\}$ and Ψ is some class of departure mappings and $\mathcal{I}_T = \{([r,s], \ldots, [r,s]) : [r,s] \subseteq [T]\}$, the set of all intervals in $[T]$ repeated T times.

- For each $t \in [T]$ and $\phi_t = (I_t, \psi_t)$, define $\ell_{\phi_t}(f, x) = (-\mathrm{loss}(f,x) + \mathrm{loss}(\psi_t \circ f, x)) \mathbf{1}\{t \in I_t\}$

- $\boldsymbol{B}(z_1, \ldots, z_T) = \sum_{t=1}^{T} z_t / T$

Note that

$$\mathbf{R}_T = \sup_{[r,s] \subseteq [T]} \left\{ \frac{1}{T} \sum_{t=r}^{s} \mathrm{loss}(f_t, x_t) - \inf_{\psi \in \Psi} \frac{1}{T} \sum_{t=r}^{s} \mathrm{loss}(\psi \circ f_t, x_t) \right\}$$

$$= \sup_{I \in \mathcal{I}_T, \psi \in \Psi_T} \left\{ \frac{1}{T} \sum_{t=1}^{T} \mathrm{loss}(f_t, x_t) \mathbf{1}\{t \in I_t\} - \frac{1}{T} \sum_{t=1}^{T} \mathrm{loss}(\psi_t \circ f_t, x_t) \mathbf{1}\{t \in I_t\} \right\}$$

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

31

and thus we see that the adaptive regret defined in Equation (16) falls under our general framework. We would like to point out as an example that if we take $\Psi_T = \{(f, \ldots, f) : f \in \mathcal{F}\}$ the time invariant set of constant mappings then the regret defined in Equation (16) is identical to the one in [18]. Below we show a bound on the value of the game with adaptive regret in terms of covering number of the departure mapping class.

Theorem 26. *For the adaptive regret game we have that*

$$\mathcal{V}_T \leq 8 \inf_{\alpha > 0} \left\{ \alpha + 3 \int_{\alpha}^{2} \sqrt{\frac{\log \ \mathcal{N}_2(\delta, \Psi, T)}{T}} d\delta \right\} + 24 \sqrt{\frac{2 \log \ T}{T}} \tag{17}$$

Proof. We start by applying the Triplex inequality in Theorem 1 along with Theorem 2 we get that :

$$\mathcal{V}_T \leq 2\mathfrak{R}_T(\ell, \mathcal{I}, B) + \sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \left\{ - \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{t:T} \\ x_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \right\} + 2\mathfrak{R}_T(\ell, \Phi_T, B)$$

$$= 0 + \sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \left\{ \mathop{\mathbb{E}}_{\substack{f_{1:T} \sim q_{t:T} \\ x_{1:T} \sim p_{1:T}}} \frac{1}{T} \sum_{t=1}^{T} (\text{loss}(f_t, x_t) - \text{loss}(\psi_t \circ f_t, x_t)) \mathbf{1}\{t \in I_t\} \right\} + 2\mathfrak{R}_T(\ell, \Phi_T, B)$$

where the last inequality above is because the first term of the triplex inequality is 0 as $\boldsymbol{B}$ is linear (see Corollary 9). If we use q_t to be point mass on $f_t = \operatorname*{argmin}_{f \in \mathcal{F}} \mathbb{E}_{x_t \sim p_t}[\text{loss}(f, x_t)]$ we see that the second term of the triplex inequality above is bounded above by 0. Hence we can conclude that

$$\mathcal{V}_T \leq 2\mathfrak{R}_T(\ell, \Phi_T, B) = 2 \sup_{\mathbf{f}, \mathbf{x}} \mathbb{E}_\epsilon \left[\sup_{\psi \in \Psi, [r,s] \subseteq [T]} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left(\text{loss}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) - \text{loss}(\psi \circ \mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right) \mathbf{1}\{t \in [r, s]\} \right]$$

To bound the above we use Theorem 8 (noting that $p = 2$, $\sigma = 0$ and $q = 1$) to get that

$$\mathcal{V}_T \leq 8 \inf_{\alpha > 0} \left\{ \alpha + 6 \int_{\alpha}^{2} \sqrt{\frac{\log \ \mathcal{N}_2(\delta, \Phi_T, T)}{T}} d\delta \right\}$$

$$\leq 8 \inf_{\alpha > 0} \left\{ \alpha + 6 \int_{\alpha}^{2} \sqrt{\frac{\log \ \mathcal{N}_2(\delta, \Psi, T) + \log(|\mathcal{I}_T|)}{T}} d\delta \right\} .$$

Now note that $|\mathcal{I}_T| \leq T^2$ and so we get that

$$\mathcal{V}_T \leq 8 \inf_{\alpha > 0} \left\{ \alpha + 6 \int_{\alpha}^{2} \sqrt{\frac{\log \ \mathcal{N}_2(\delta, \Psi, T)}{T}} d\delta \right\} + 48 \sqrt{\frac{2 \log \ T}{T}} .$$

We conclude that whenever covering number of Ψ can be bounded appropriately, adaptive regret can be bounded at the expense of an extra $O\left(\sqrt{\frac{\log T}{T}} \right)$ term. $\square$

Acknowledgements

References

[1] 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.

[2] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari. Optimal strategies and minimax lower bounds for online convex games. In *Proceedings of the 21st Annual Conference on Learning Theory*, pages 414–424. Omnipress, 2008.

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

[4] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pac. J. Math.*, 6:1–8, 1956.

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

[6] J. Borwein, A.J. Guirao, P. Hájek, and J. Vanderwerff. Uniformly convex functions on Banach spaces. *Proc. Amer. Math. Soc*, 137:1081–1091, 2009.

[7] 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.

[8] G.W. Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.

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

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

[11] 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.

[12] Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2):40–55, October 1997.

[13] D.P. Foster and R.V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379, 1998.

[14] E. Giné and J. Zinn. Some limit theorems for empirical processes. *Annals of Probability*, 12(4):929–989, 1984.

[15] 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.

[16] 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.

[17] E. Hazan and S. Kale. Computational equivalence of fixed points and no regret algorithms, and convergence to equilibria. In *Advances in Neural Information Processing Systems (NIPS)*, 2007.

[18] E. Hazan and C. Seshadhri. Efficient learning algorithms for changing environments. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 393–400, New York, NY, USA, 2009. ACM.

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

[20] W. S. Lee, P. L. Bartlett, and R. C. Williamson. The importance of convexity in learning with squared loss. *IEEE Transactions on Information Theory*, 44(5):1974–1980, 1998.

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

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

[23] S. Mannor and G. Stoltz. A geometric proof of calibration. *Arxiv preprint arXiv:0912.3604*, 2009.

[24] I. Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *The Annals of Probability*, 22(4):1679–1706, 1994.

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

[26] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Random averages, combinatorial parameters, and learnability. *Arxiv preprint arXiv:1006.1138*, 2010.

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

[28] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

Appendix

Proof of Theorem 1. The value of the game, defined in (2), is

$$\mathcal{V}_T(\ell, \Phi_T) = \inf_{q_1} \sup_{p_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \inf_{q_T} \sup_{p_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \sup_{\phi \in \Phi_T} \{\boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T))\}$$

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

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

$$\mathcal{V}_T(\ell, \Phi_T) = \sup_{p_1} \inf_{q_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \right.$$

$$\left. + \sup_{\phi \in \Phi_T} \left\{ \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \right\} \right]$$

$$\leq \sup_{p_1} \inf_{q_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \right.$$

$$+ \sup_{\phi \in \Phi_T} \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \left\{ \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) - \boldsymbol{B}(\ell_{\phi_1}(f'_1, x'_1), \ldots, \ell_{\phi_T}(f'_T, x'_T)) \right\}$$

$$\left. + \sup_{\phi \in \Phi_T} \left\{ \mathbb{E}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f'_1, x'_1), \ldots, \ell_{\phi_T}(f'_T, x'_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \right\} \right]$$

34

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

$\mathcal{V}_T(\ell, \Phi_T)$

$$\le \sup_{p_1} \sup_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \sup_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell(f'_1, x'_1), \ldots, \ell(f'_T, x'_T)) \right]$$

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

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

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

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

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

because the objective

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

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

Proof of Theorem 2. We turn to the third term in the Triplex Inequality. If $\boldsymbol{B}$ is subadditive,

$$\mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f'_1, x'_1), \ldots, \ell_{\phi_T}(f'_T, x'_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T))$$

$$\le \mathop{\mathbb{E}}_{\substack{f'_{1:T} \sim q_{1:T} \\ x'_{1:T} \sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f'_1, x'_1) - \ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f'_T, x'_T) - \ell_{\phi_T}(f_T, x_T)).$$

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

$$\mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\ x'_{1:T}\sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f'_1,x'_1),\ldots,\ell_{\phi_T}(f'_T,x'_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_T}(f_T,x_T))$$

$$\leq - \mathop{\mathbb{E}}_{\substack{f'_{1:T}\sim q_{1:T}\\ x'_{1:T}\sim p_{1:T}}} \boldsymbol{B}(\ell_{\phi_1}(f_1,x_1) - \ell_{\phi_1}(f'_1,x'_1),\ldots,\ell_{\phi_T}(f_T,x_T) - \ell_{\phi_T}(f'_T,x'_T)). \tag{18}$$

Below assume that $\boldsymbol{B}$ 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 [26], applying the symmetrization technique from inside out. To this end, first note that,

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

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

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

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

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

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

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

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

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

where ϵ_T is a Rademacher random variable. Furthermore,

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

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

$$\leq \sup_{\substack{x_T,x'_T\in\mathcal{X}\\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T} \boldsymbol{B}\Big(\ell_{\phi_1}(f'_1,x'_1) - \ell_{\phi_1}(f_1,x_1),\ldots,\epsilon_T(\ell_{\phi_T}(f'_T,x'_T) - \ell_{\phi_T}(f_T,x_T))\Big)$$

36

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,

$$\sup_{\substack{p_{T-1},q_{T-1} \\ x_{T-1},x'_{T-1}\sim p_{T-1}}} \mathbb{E}_{\substack{f_{T-1},f'_{T-1}\sim q_{T-1}}} \sup_{\substack{x_T,x'_T\in\mathcal{X} \\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T}$$

$$\boldsymbol{B}\Big(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1),\ldots,\ell_{\phi_{T-1}}(f'_{T-1},x'_{T-1})-\ell_{\phi_{T-1}}(f_{T-1},x_{T-1}),\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

$$= \sup_{\substack{p_{T-1},q_{T-1} \\ x_{T-1},x'_{T-1}\sim p_{T-1}}} \mathbb{E}_{\substack{f_{T-1},f'_{T-1}\sim q_{T-1}}} \mathbb{E}_{\epsilon_{T-1}} \sup_{\substack{x_T,x'_T\in\mathcal{X} \\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T}$$

$$\boldsymbol{B}\Big(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1),\ldots,\epsilon_{T-1}(\ell_{\phi_T}(f'_{T-1},x'_{T-1})-\ell_{\phi_{T-1}}(f_{T-1},x_{T-1})),\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

$$\leq \sup_{\substack{x_{T-1},x'_{T-1}\in\mathcal{X} \\ f_{T-1},f'_{T-1}\in\mathcal{F}}} \mathbb{E}_{\epsilon_{T-1}} \sup_{\substack{x_T,x'_T\in\mathcal{X} \\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T}$$

$$\boldsymbol{B}\Big(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1),\ldots,\epsilon_{T-1}(\ell_{\phi_{T-1}}(f'_{T-1},x'_{T-1})-\ell_{\phi_{T-1}}(f_{T-1},x_{T-1})),\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

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

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

$$\leq \sup_{\substack{x_1,x'_1\in\mathcal{X} \\ f_1,f'_1\in\mathcal{F}}} \mathbb{E}_{\epsilon_1} \ldots \sup_{\substack{x_T,x'_T\in\mathcal{X} \\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T} \boldsymbol{B}\Big(\epsilon_1(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

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

$$\boldsymbol{B}\Big(\epsilon_1(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

$$\leq \boldsymbol{B}\Big(\epsilon_1\ell_{\phi_1}(f'_1,x'_1),\ldots,\epsilon_T\ell_{\phi_T}(f'_T,x'_T)\Big) + \boldsymbol{B}\Big(-\epsilon_1\ell_{\phi_1}(f_1,x_1),\ldots,-\epsilon_T\ell_{\phi_T}(f_T,x_T)\Big)$$

We, therefore, arrive at

$$\sup_{\substack{x_1,x'_1\in\mathcal{X} \\ f_1,f'_1\in\mathcal{F}}} \mathbb{E}_{\epsilon_1} \ldots \sup_{\substack{x_T,x'_T\in\mathcal{X} \\ f_T,f'_T\in\mathcal{F}}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T} \boldsymbol{B}\Big(\epsilon_1(\ell_{\phi_1}(f'_1,x'_1)-\ell_{\phi_1}(f_1,x_1)),\ldots,\epsilon_T(\ell_{\phi_T}(f'_T,x'_T)-\ell_{\phi_T}(f_T,x_T))\Big)$$

$$\leq 2 \sup_{f_1\in\mathcal{F},x_1\in\mathcal{X}} \mathbb{E}_{\epsilon_1} \ldots \sup_{f_T\in\mathcal{F},x_T\in\mathcal{X}} \mathbb{E}_{\epsilon_T} \sup_{\phi\in\Phi_T} \boldsymbol{B}\Big(\epsilon_1\ell_{\phi_1}(f_1,x_1),\ldots,\epsilon_T\ell_{\phi_T}(f_T,x_T)\Big)$$

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

where 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 $\boldsymbol{B}$ being subadditive. Starting from Eq. (18), the proof for the case of $-\boldsymbol{B}$ being subadditive and convex in each of its coordinates leads to the bound of

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

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

The following Proposition is immediate from the definition of a smooth function via successive expansions of each coordinate around zero.

Proposition 27. *Assume function $\boldsymbol{B} : \mathcal{H}^T \mapsto \mathbb{R}$ is (p, σ)-uniformly smooth in each of its arguments and that $\boldsymbol{B}(0, 0, \ldots, 0) = 0$. Then*

$$\boldsymbol{B}(z_1, \ldots, z_T) \leq \sum_{t=1}^{T} \langle \nabla_t \boldsymbol{B}(z_1, \ldots, z_{t-1}, 0, \ldots, 0), z_t \rangle + \sum_{t=1}^{T} \sigma \|z_t\|^p$$

Lemma 28. *Assume that for some $q \geq 1$, $\boldsymbol{B}^q$ is (σ, p)-uniformly smooth in each of its arguments and $\boldsymbol{B}(0, \ldots, 0) = 0$. Then we have that*

$$\sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \boldsymbol{B}(z_1 - z_1', \ldots, z_T - z_T') \leq ((2\eta)^p \sigma T)^{1/q}$$

where the maximization is over distributions p_t with support in the ball $\eta \cdot B_{\|\cdot\|}$ of radius η.

Proof of Lemma 28. By Proposition 27 we have that

$$\sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \boldsymbol{B}^q(z_1 - z_1', \ldots, z_T - z_T')$$

$$\leq \sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \left\{ \sum_{t=1}^{T} \langle \nabla_t \boldsymbol{B}^q(z_1 - z_1', \ldots, z_{t-1} - z_{t-1}', 0, \ldots, 0), z_t - z_t' \rangle + \sigma \sum_{t=1}^{T} \|z_t - z_t'\|^p \right\}$$

$$\leq \sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \left\{ \sum_{t=1}^{T} \langle \nabla_t \boldsymbol{B}^q(z_1 - z_1', \ldots, z_{t-1} - z_{t-1}', 0, \ldots, 0), z_t - z_t' \rangle \right\}$$

$$+ \sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \left\{ \sigma \sum_{t=1}^{T} \|z_t - z_t'\|^p \right\}$$

$$= \sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \left\{ \sigma \sum_{t=1}^{T} \|z_t - z_t'\|^p \right\} \leq (2\eta)^p \sigma T$$

Since $q \geq 1$, by Jensen's inequality we conclude that

$$\sup_{p_1} \mathop{\mathbb{E}}_{z_1, z_1' \sim p_1} \ldots \sup_{p_T} \mathop{\mathbb{E}}_{z_T, z_T' \sim p_T} \boldsymbol{B}(z_1 - z_1', \ldots, z_T - z_T') \leq ((2\eta)^p \sigma T)^{1/q}$$

$\square$

Proof of Lemma 3. The proof follows immediately from Lemma 28. $\square$

Proof of Lemma 4. By Proposition 27 we have:

$$\boldsymbol{B}^q\Big(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_T \ell_{\phi_T}(\mathbf{f}_T(\epsilon), \mathbf{x}_T(\epsilon)) \Big)$$

$$\leq \sum_{t=1}^{T} \Big\langle \nabla_t \boldsymbol{B}^q\big(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_{t-1} \ell_{\phi_{t-1}}(\mathbf{f}_{t-1}(\epsilon), \mathbf{x}_{t-1}(\epsilon)), 0, \ldots, 0 \big), \epsilon_t \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \Big\rangle + \sigma \sum_{t=1}^{T} \|\ell_{\phi_t}(f_t, x_t)\|^p$$

$$\leq \sum_{t=1}^{T} \epsilon_t g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \big) + \sigma \eta^p T$$

where in the last line we used the definition of g_t as well as an upper bound on the norm. Now by Jensen's inequality we get

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

$$\leq \left(\sup_{\mathbf{f},\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi\in\Phi_T} \sum_{t=1}^{T}\epsilon_t g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big) + \sigma\eta^p T\right)^{1/q}$$

$$\leq \left(\sup_{\mathbf{f},\mathbf{x}} \ \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi\in\Phi_T} \sum_{t=1}^{T}\epsilon_t g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big)\right)^{1/q} + (\sigma\eta^p)^{1/q}T^{1/q}$$

$\square$

Proof of Proposition 6. Fix a $\mathcal{F}\times\mathcal{X}$-valued tree $(\mathbf{f},\mathbf{x})$. Note that

$$\Big|g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big)\Big|$$
$$\leq \big\|\nabla_t\boldsymbol{B}^q\big(\epsilon_1\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\epsilon_{t-1}\ell_{\phi_{t-1}}(\mathbf{f}_{t-1}(\epsilon),\mathbf{x}_{t-1}(\epsilon)),0,\ldots,0\big)\big\|_* \big\|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big\|$$
$$\leq R\cdot\eta$$

Using Lemma 5,

$$\mathbb{E}_{\epsilon_{1:T}} \max_{\phi\in\Phi_T} \sum_{t=1}^{T}\epsilon_t g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big)$$

$$\leq \sqrt{2\log(|\Phi_T|)\max_{\phi\in\Phi_T}\max_{\epsilon\in\{\pm1\}^T}\sum_{t=1}^{T}g_t\big(\ell_{\phi_1}(\mathbf{f}_1(\epsilon),\mathbf{x}_1(\epsilon)),\ldots,\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\big)^2}$$
$$\leq \sqrt{2\eta^2 R^2\log(|\Phi_T|)T}$$

Now using Lemma 4 we obtain the desired result. $\square$

Proof of Corollary 7. To appeal to Proposition 6, we need to specify smoothness parameters. It can be verified that if G is (γ,p)-smooth in its argument, then $\boldsymbol{B}^q$ is $(\gamma/T^p,p)$-smooth. Furthermore,

$$\|\nabla_t\boldsymbol{B}^q(z_1,\ldots,z_T)\|_* \leq \rho/T.$$

The bound of Proposition 6 then becomes

$$\mathfrak{R}_T(\ell,\Phi_T) \leq \left(\frac{2\eta^2\log(|\Phi_T|)}{T}\right)^{1/2q} + (\gamma\eta^p)^{1/q}T^{(1-p)/q}\ .$$

$\square$

Proof of Theorem 8. Since $\boldsymbol{B}^q(z_1,\ldots,z_T) = G\left(\frac{1}{T}\sum_{t=1}^{T}z_t\right)$, we see that

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

and, under our assumptions, $\|\nabla_t \boldsymbol{B}^q(z_1, \ldots, z_T)\|_* \leq \rho/T$. Further,

$$\left\langle \nabla_t \boldsymbol{B}^q \left(\epsilon_1 \ell_{\phi_1}(\mathbf{f}_1(\epsilon), \mathbf{x}_1(\epsilon)), \ldots, \epsilon_{t-1} \ell_{\phi_{t-1}}(\mathbf{f}_{t-1}(\epsilon), \mathbf{x}_{t-1}(\epsilon)), 0, \ldots, 0\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle$$

$$= \frac{1}{T} \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle .$$

Also observe that if G is (γ, p)-smooth, then $\boldsymbol{B}^q$ is $(\gamma/T^p, p)$-smooth in each coordinate. Applying Lemma 4,

$$\mathfrak{R}_T(\ell, \Phi_T) \leq \left(\sup_{\mathbf{f}, \mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle\right)^{1/q} + (\gamma \eta^p T^{1-p})^{1/q} .$$

$$(19)$$

Define $\beta_0 = \eta$ and $\beta_j = 2^{-j} \cdot \eta$. For a fixed tree $(\mathbf{f}, \mathbf{x})$ of depth T, let V_j be an ℓ_2-cover at scale β_j. For any path $\epsilon \in \{\pm 1\}^T$ and any $\phi \in \Phi_T$, let $\mathbf{v}[\phi, \epsilon]^j \in V_j$ the element of the cover such that

$$\sqrt{\frac{1}{T} \sum_{t=1}^{T} \left\| \mathbf{v}[\phi, \epsilon]_t^j(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\|^2} \leq \beta_j .$$

By the definition such a $\mathbf{v}[\phi, \epsilon]^j \in V_j$ exists, and we assume for simplicity this element is unique (ties can be broken in an arbitrary manner). Thus, $\phi \mapsto \mathbf{v}[\phi, \epsilon]^j$ is a well-defined mapping for any fixed ϵ and j. As before, $\mathbf{v}[\phi, \epsilon]_t^j$ denotes the t-th mapping of $\mathbf{v}[\phi, \epsilon]^j$. For any $t \in [T]$, we have

$$\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle$$

$$= \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^N(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^N(\epsilon) \right\rangle$$

$$+ \sum_{j=1}^{N} \left(\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^j(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^j(\epsilon) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^{j-1}(\epsilon) \right\rangle\right)$$

where $\mathbf{v}[\phi, \epsilon]_t^0(\epsilon) = 0$. Hence,

$$\mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle\right]$$

$$\leq \mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left(\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^N(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^N(\epsilon) \right\rangle\right)\right]$$

$$+ \mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \sum_{j=1}^{N} \left(\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^j(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^j(\epsilon) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^{j-1}(\epsilon) \right\rangle\right)\right]$$

$$\leq \mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left(\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon))\right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^N(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^N(\epsilon) \right\rangle\right)\right]$$

$$+ \sum_{j=1}^{N} \mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left(\left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^j(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^j(\epsilon) \right\rangle - \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \mathbf{v}[\phi, \epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi, \epsilon]_t^{j-1}(\epsilon) \right\rangle\right)\right]$$

$$(20)$$

We now consider the first term in the sum above for a fixed ϵ and ϕ:

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right),\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right\rangle - \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle\right)$$

$$=\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right),\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))-\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle$$

$$+\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)-\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle$$

$$\leq\frac{1}{T}\sum_{t=1}^{T}\left|\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right),\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))-\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle\right|$$

$$+\frac{1}{T}\sum_{t=1}^{T}\left|\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)-\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle\right|$$

$$\leq\frac{1}{T}\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|_*\left\|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))-\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\|$$

$$+\frac{1}{T}\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)-\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right)\right\|_*\left\|\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\|$$

Using a result from [6], for a (γ,p)-smooth function G, we have that

$$\|\nabla G(a)-\nabla G(b)\|_* \leq 2\gamma\|a-b\|^{p-1}. \tag{21}$$

Since $\|\ell_{\phi_t}(f,x)\|\leq\eta$ and $\|\nabla G(a)\|_*\leq\rho$ uniformly, we can appeal to (21) and conclude that

$$\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right),\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right\rangle - \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\right\rangle\right)$$

$$\leq\rho\sqrt{\frac{1}{T}\sum_{t=1}^{T}\|\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))-\mathbf{v}[\phi,\epsilon]_t^N(\epsilon)\|^2}+2\gamma\eta\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\left(\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))-\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right)\right\|^{p-1}$$

$$\leq\rho\,\beta_N+2\gamma\eta\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{i=1}^{t-1}\left\|\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))-\mathbf{v}[\phi,\epsilon]_i^N(\epsilon)\right\|\right)^{p-1}$$

$$\leq\rho\,\beta_N+2\gamma\eta\beta_N^{p-1}. \tag{22}$$

The second term in (20) is bounded by considering successive refinements of the cover. The argument is delicate as the trees $\mathbf{v}[\phi,\epsilon]^j$, $\mathbf{v}[\phi,\epsilon]^{j-1}$ depend on the particular path. Another layer of difficulty over the proof in [26] arises from the fact that gradients of the objective function G also depend on ϕ. Here, the smoothness of the function plays a key role: for near-by elements of Φ_T, the gradients of G are close, and so a covering of Φ_T induces a covering of the gradients.

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 an $\mathcal{H}$-valued tree $\mathbf{w}^{(s,r)}$ by

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

for all $t\in[T]$ and $\epsilon\in\{\pm1\}^T$. 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* $\phi\in\Phi_T$. It

is easy to see that $\mathbf{w}^{(s,r)}$ is well-defined. Let the set of trees W_j be defined as

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

On similar lines also define for all possible pairs $(\mathbf{v}^s, \mathbf{v}^r) \in V_j \times V_{j-1}$, define a $\mathcal{B}^*$-valued tree $\mathbf{u}^{(s,r)}$ by

$$\mathbf{u}_t^{(s,r)}(\epsilon) = \begin{cases} \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i^s(\epsilon)\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i^r(\epsilon)\right) & \text{if exists } \phi \in \Phi_T \text{ s.t. } \mathbf{v}^s = \mathbf{v}[\phi,\epsilon]^j, \mathbf{v}^r = \mathbf{v}[\phi,\epsilon]^{j-1} \\ 0 & \text{otherwise.} \end{cases}$$

for all $t \in [T]$ and $\epsilon \in \{\pm 1\}^T$. It is easy to see that $\mathbf{u}^{(s,r)}$ is well-defined. Let the set of trees U_j be defined as

$$U_j = \left\{ \mathbf{u}^{(s,r)} : 1 \le s \le |V_j|, 1 \le r \le |V_{j-1}| \right\}$$

Now, consider the second term in (20). Fix $j \in [N]$ and consider the jth summand of the second term in (20).

$$\mathbb{E}_\epsilon\left[\sup_{\phi \in \Phi_T} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^j(\epsilon)\right), \mathbf{v}[\phi,\epsilon]_t^j(\epsilon)\right\rangle - \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi,\epsilon]_t^{j-1}(\epsilon)\right\rangle\right)\right]$$

$$\le \mathbb{E}_\epsilon\left[\sup_{\phi \in \Phi_T} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^j(\epsilon)\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi,\epsilon]_t^j(\epsilon)\right\rangle\right]$$

$$+ \mathbb{E}_\epsilon\left[\sup_{\phi \in \Phi_T} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^{j-1}(\epsilon)\right), \mathbf{v}[\phi,\epsilon]_t^j(\epsilon) - \mathbf{v}[\phi,\epsilon]_t^{j-1}(\epsilon)\right\rangle\right]$$

$$\le \mathbb{E}_\epsilon\left[\max_{\substack{\mathbf{u}\in U_j \\ \mathbf{v}\in V_j}} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t \left\langle \mathbf{u}_t(\epsilon), \mathbf{v}_t(\epsilon)\right\rangle\right] + \mathbb{E}_\epsilon\left[\max_{\substack{\mathbf{w}\in W_j \\ \mathbf{v}\in V_{j-1}}} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t(\epsilon)\right\rangle\right] \qquad (23)$$

The last inequality holds because for any $j \in [N]$, $\epsilon \in \{\pm 1\}^T$ and $\phi \in \Phi_T$ there is some $\mathbf{w}^{(s,r)} \in W_j$ with $\mathbf{v}[\phi,\epsilon]^j = \mathbf{v}^s$, $\mathbf{v}[\phi,\epsilon]^{j-1} = \mathbf{v}^r$ and

$$\mathbf{v}_t^s(\epsilon) - \mathbf{v}_t^r(\epsilon) = \mathbf{w}_t^{(s,r)}(\epsilon) \quad \forall t \le T.$$

Similarly, by construction, for any difference of the gradients, there exists a representative $\mathbf{u}^{(s,r)} \in U_j$.

Now we would like to apply Lemma 5 to the two terms in Equation (23). To this end first consider the second term and note that if there exists $\phi \in \Phi_T$ for which $\mathbf{v}^s = \mathbf{v}[\phi,\epsilon]^j$ and $\mathbf{v}^r = \mathbf{v}[\phi,\epsilon]^{j-1}$, then $\mathbf{w}_t^{(s,r)}(\epsilon) = \mathbf{v}[\phi,\epsilon]_t^j - \mathbf{v}[\phi,\epsilon]_t^{j-1}$ for any $t \in [T]$. In this case,

$$\sqrt{\sum_{t=1}^{T}\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t^{(s,r)}(\epsilon)\right\rangle^2} \le \sqrt{\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right)\right\|_*^2 \left\|\mathbf{w}_t^{(s,r)}(\epsilon)\right\|^2} \le \rho\sqrt{\sum_{t=1}^{T}\left\|\mathbf{w}_t^{(s,r)}(\epsilon)\right\|^2},$$

which, in turn, is bounded via triangle inequality by

$$\rho\sqrt{\sum_{t=1}^{T}\left(\left\|\mathbf{v}[\phi,\epsilon]_t^j(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\right\| + \left\|\mathbf{v}[\phi,\epsilon]_t^{j-1}(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\right\|\right)^2}$$

$$\le \sqrt{2}\rho\left(\sqrt{\sum_{t=1}^{T}\left\|\mathbf{v}[\phi,\epsilon]_t^j(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\right\|^2} + \sqrt{\sum_{t=1}^{T}\left\|\mathbf{v}[\phi,\epsilon]_t^{j-1}(\epsilon) - \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon))\right\|^2}\right)$$

$$\le \rho\sqrt{2T}\left(\beta_j + \beta_{j-1}\right) = 3\rho\sqrt{2T}\beta_j$$

If there exists no such $\phi \in \Phi_T$ for the given ϵ and (s,r), then $\mathbf{w}_t^{(s,r)}(\epsilon)$ is zero for all $t \geq t_o$, for some $1 \leq t_o < T$, and thus

$$\sqrt{\sum_{t=1}^{T} \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t^{(s,r)}(\epsilon)\right\rangle^2} \leq \sqrt{\sum_{t=1}^{T} \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t^{(s,r)}(\epsilon')\right\rangle^2}$$

$$\leq \sqrt{\sum_{t=1}^{T} \left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right)\right\|_*^2 \left\|\mathbf{w}_t^{(s,r)}(\epsilon')\right\|^2}$$

$$\leq \rho\sqrt{\sum_{t=1}^{T}\left\|\mathbf{w}_t^{(s,r)}(\epsilon')\right\|^2} \leq 3\rho\sqrt{2T}\beta_j$$

where ϵ' is any path that agrees with ϵ up to t_o. Hence we can conclude that for all $\epsilon \in \{\pm 1\}^T$ and all $\mathbf{w} \in W_j$

$$\sqrt{\sum_{t=1}^{T} \left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t(\epsilon)\right\rangle^2} \leq 3\rho\sqrt{2T}\beta_j$$

We can now invoke Lemma 5. For the second term of (23) we then get

$$\mathbb{E}_\epsilon\left[\sup_{\substack{\mathbf{w}\in W_j \\ \mathbf{v}\in V_{j-1}}} \frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}_i(\epsilon)\right), \mathbf{w}_t(\epsilon)\right\rangle\right] \leq 3\sqrt{2}\rho\,\beta_j\sqrt{\frac{2(\log(|W_j|)+\log(|V_{j-1}|))}{T}}$$

$$= 6\rho\,\beta_j\sqrt{\frac{\log(|V_j|\,|V_{j-1}|)+\log(|V_{j-1}|)}{T}}$$

$$\leq 6\rho\,\beta_j\sqrt{\frac{3\log(|V_j|)}{T}} \tag{24}$$

The first term of (23) is bounded in a similar fashion. To this end, if there exists $\phi \in \Phi_T$ for which $\mathbf{v}^s = \mathbf{v}[\phi,\epsilon]^j$ and $\mathbf{v}^r = \mathbf{v}[\phi,\epsilon]^{j-1}$, then $\mathbf{u}_t^{(s,r)}(\epsilon) = \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^j\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^{j-1}\right)$ for any $t \in [T]$. In this case,

$$\sqrt{\sum_{t=1}^{T}\left\langle \mathbf{u}_t^{(s,r)}(\epsilon), \mathbf{v}_t(\epsilon)\right\rangle^2} \leq \sqrt{\sum_{t=1}^{T}\left\|\mathbf{u}_t^{(s,r)}\right\|_*^2 \|\mathbf{v}_t^s(\epsilon)\|^2} \leq \eta\sqrt{\sum_{t=1}^{T}\left\|\mathbf{u}_t^{(s,r)}(\epsilon)\right\|_*^2}$$

$$\leq \eta\sqrt{2}\left(\sqrt{\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^j\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|_*^2}\right. \tag{25}$$

$$\left. + \sqrt{\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^{j-1}\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|_*^2}\right)$$

Using the assumption that G is (γ,p)-smooth,

$$\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \mathbf{v}[\phi,\epsilon]_i^j\right) - \nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|_*^2 \leq 4\gamma^2\sum_{t=1}^{T}\left\|\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\left(\mathbf{v}[\phi,\epsilon]_i^j - \ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|^{2(p-1)}$$

$$\leq 4\gamma^2\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{i=1}^{t-1}\left\|\mathbf{v}[\phi,\epsilon]_i^j - \ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right\|\right)^{2(p-1)}$$

$$\leq 4\gamma^2 T\beta_j^{2(p-1)}$$

43

Similarly,

$$\sum_{t=1}^{T}\left\|\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^{j-1}\right)-\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right)\right\|_*^2\leq 4\gamma^2 T\beta_{j-1}^{2(p-1)}$$

The bound on (25) then becomes

$$\sqrt{\sum_{t=1}^{T}\left\langle\mathbf{u}_t^{(s,r)}(\epsilon),\mathbf{v}_t(\epsilon)\right\rangle^2}\leq 2\gamma\eta\sqrt{2T}(\beta_j^{(p-1)}+\beta_{j-1}^{(p-1)})\leq 6\gamma\eta\sqrt{2T}\beta_j^{(p-1)}\ . \tag{26}$$

Invoking Lemma 5 for the first term in (23) we get

$$\mathbb{E}_\epsilon\left[\sup_{\substack{\mathbf{v}\in V_j\\\mathbf{u}\in U_j}}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle\mathbf{u}_t(\epsilon),\mathbf{v}_t(\epsilon)\right\rangle\right]\leq 6\sqrt{2}\gamma\eta\beta_j^{p-1}\sqrt{\frac{2(\log(|U_j|)+\log(|V_j|))}{T}}$$

$$\leq 12\gamma\eta\beta_j^{p-1}\sqrt{\frac{\log(|V_j|\ |V_{j-1}|)+\log(|V_j|)}{T}}\leq 12\gamma\eta\beta_j^{p-1}\sqrt{\frac{3\log(|V_j|)}{T}}\ .$$

Combining Eq. (24) and Eq. (26) with Eq. (23) we see that

$$\mathbb{E}_\epsilon\left[\sup_{\phi\in\Phi_T}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left(\left\langle\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^{j}(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^{j}(\epsilon)\right\rangle-\left\langle\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\mathbf{v}[\phi,\epsilon]_i^{j-1}(\epsilon)\right),\mathbf{v}[\phi,\epsilon]_t^{j-1}(\epsilon)\right\rangle\right)\right]$$

$$\leq\frac{6\rho}{\sqrt{T}}\beta_j\sqrt{3\log(|V_j|)}+\frac{12\gamma\eta}{\sqrt{T}}\beta_j^{p-1}\sqrt{3\log(|V_j|)}\ .$$

Since above result holds for any $j\in[N]$, we can sum over all $j\in[N]$, obtaining the following upper bound on the second term in (23):

$$\sum_{j=1}^{N}\left(\frac{6\rho}{\sqrt{T}}\beta_j\sqrt{3\log(|V_j|)}+\frac{12\gamma\eta}{\sqrt{T}}\beta_j^{p-1}\sqrt{3\log(|V_j|)}\right)$$

$$=\frac{12\sqrt{3}\rho}{\sqrt{T}}\sum_{j=1}^{N}(\beta_j-\beta_{j+1})\sqrt{\log\mathcal{N}_2(\beta_j,\Phi_T,\mathbf{f},\mathbf{x})}+\frac{2^{p-1}}{2^{p-1}-1}\frac{12\sqrt{3}\gamma\eta}{\sqrt{T}}\sum_{j=1}^{N}(\beta_j^{p-1}-\beta_{j+1}^{p-1})\sqrt{\log\mathcal{N}_2(\beta_j,\Phi_T,\mathbf{f},\mathbf{x})}$$

$$\leq\frac{12\sqrt{3}\rho}{\sqrt{T}}\int_{\beta_{N+1}}^{\eta}\sqrt{\log\mathcal{N}_2(\delta,\Phi_T,\mathbf{f},\mathbf{x})}\,d\delta+\frac{2^{p-1}}{2^{p-1}-1}\frac{12\sqrt{3}\gamma\eta}{\sqrt{T}}\int_{\beta_{N+1}^{p-1}}^{\eta}\sqrt{\log\mathcal{N}_2\left(\delta^{1/(p-1)},\Phi_T,\mathbf{f},\mathbf{x}\right)}\,d\delta$$

where the first equality is because $\beta_j=2(\beta_j-\beta_{j+1})$ and $\beta_j^{p-1}=\frac{2^{p-1}}{2^{p-1}-1}(\beta_j^{p-1}-\beta_{j+1}^{p-1})$. Using the above upper bound together with Eq. (22) in Eq. (20), we conclude that

$$\mathbb{E}_\epsilon\left[\sup_{\phi\in\Phi_T}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t\left\langle\nabla G\left(\frac{1}{T}\sum_{i=1}^{t-1}\epsilon_i\ell_{\phi_i}(\mathbf{f}_i(\epsilon),\mathbf{x}_i(\epsilon))\right),\ell_{\phi_t}(\mathbf{f}_t(\epsilon),\mathbf{x}_t(\epsilon))\right\rangle\right]$$

$$\leq\rho\beta_N+2\gamma\eta\beta_N^{p-1}+\frac{12\sqrt{3}\rho}{\sqrt{T}}\int_{\beta_{N+1}}^{\eta}\sqrt{\log\mathcal{N}_2(\delta,\Phi_T,T)}\,d\delta+\frac{24\sqrt{3}\gamma\eta}{(2^{p-1}-1)\sqrt{T}}\int_{\beta_{N+1}^{p-1}}^{\eta}\sqrt{\log\mathcal{N}_2\left(\delta^{\frac{1}{p-1}},\Phi_T,T\right)}\,d\delta\ .$$

We used the assumption that $1<p\leq 2$ and the fact that our choice of $\mathbf{f},\mathbf{x}$ was arbitrary. Now for any $\alpha>0$, pick $N=\sup\{j:\beta_j>2\alpha\}$. In this case we see that by our choice of N, $\beta_{N+1}\leq 2\alpha$ and so

44

$\beta_N = 2\beta_{N+1} \leq 4\alpha$. Also note that since $\beta_N > 2\alpha$, $\beta_{N+1} = \frac{\beta_N}{2} > \alpha$. Therefore,

$$\mathbb{E}_\epsilon \left[\sup_{\phi \in \Phi_T} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \left\langle \nabla G \left(\frac{1}{T} \sum_{i=1}^{t-1} \epsilon_i \ell_{\phi_i}(\mathbf{f}_i(\epsilon), \mathbf{x}_i(\epsilon)) \right), \ell_{\phi_t}(\mathbf{f}_t(\epsilon), \mathbf{x}_t(\epsilon)) \right\rangle \right]$$

$$\leq 4\rho\alpha + 32\gamma\eta\alpha^{p-1} + \frac{12\sqrt{3}\rho}{\sqrt{T}} \int_\alpha^\eta \sqrt{\log \mathcal{N}_2(\delta, \Phi_T, T)} \, d\delta + \frac{24\sqrt{3}\gamma\eta}{(2^{p-1}-1)\sqrt{T}} \int_{\alpha^{p-1}}^\eta \sqrt{\log \mathcal{N}_2 \left(\delta^{\frac{1}{p-1}}, \Phi_T, T \right)} \, d\delta .$$

Since choice of α is arbitrary we can take the infimum over it and hence using this in Equation (19) we conclude that

$$\mathfrak{R}_T(\ell, \Phi_T) \leq$$

$$\inf_{\alpha > 0} \left(4\rho\alpha + 32\gamma\eta\alpha^{p-1} + 12\sqrt{3}\rho \int_\alpha^\eta \sqrt{\frac{\log \mathcal{N}_2(\delta, \Phi_T, T)}{T}} \, d\delta + \frac{24\sqrt{3}\gamma\eta}{(2^{p-1}-1)} \int_{\alpha^{p-1}}^\eta \sqrt{\frac{\log \mathcal{N}_2 \left(\delta^{\frac{1}{p-1}}, \Phi_T, T \right)}{T}} \, d\delta \right)^{1/q}$$

$$+ (\gamma\eta^p T^{1-p})^{1/q}$$

$\square$

Proof of Corollary 9. The first statement is trivially verified. In fact, for this to hold we only require that $\boldsymbol{B}$ is subadditive, affine in its arguments, and $\boldsymbol{B}(0, \ldots, 0) = 0$. Indeed, the expectations can be sequentially moved inside of $\boldsymbol{B}$, making the coordinates of $\boldsymbol{B}$ zero, and making the suprema over the distributions irrelevant.

For the second claim, consider the second term in (4), specialized to the case of departure mappings:

$$\sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \mathbb{E}_{\substack{f_{1:T} \sim q_{1:T} \\ x_{1:T} \sim p_{1:T}}} \left\{ \frac{1}{T} \sum_{t=1}^{T} \ell(f_t, x_t) - \ell(\phi_t(f_t), x_t) \right\} \tag{27}$$

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 (27) is non-positive.

The third claim is a straightforward consequence of Theorem 8. Indeed, $\mathcal{H} \subset [-1, 1]$ and G is the identity mapping, hence $\rho = 1$, $\gamma = 0$, $q = 1$ and $p = 2$. $\square$

Proof of Lemma 13. 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 $\epsilon \in \{\pm 1\}^T$. 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 $b = i_{j+1}$,

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

Hence,

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

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

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

45

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

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

and $\mathbf{v}' = \mathbf{v}'\left(\mathbf{v}^0, \ldots, \mathbf{v}^k, i_0, \ldots, i_k\right)$ 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 $I_a = \{i_j, \ldots, i_{j+1} - 1\}$ and $a(t)$ is the index of the interval to which t belongs. In plain words, we consider all ways of partitioning $\{1, \ldots, T\}$ 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 $\Phi_T^{k,\alpha}$. $\qquad\square$

Lemma 29. *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, $1/Z_g + 1/Z_h \le 1$. Thus,

$$\|f \odot (x+y)\| \ge \frac{1}{Z_g}\|f \odot (x+y)\| + \frac{1}{Z_h}\|f \odot (x+y)\|$$
$$= \|g \odot x\| + \|h \odot y\|$$
$$\ge \inf_{f \in \mathcal{F}} \|f \odot x\| + \inf_{f \in \mathcal{F}} \|f \odot y\| \,.$$

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

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

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. □

Proof of Proposition 14. Consider any equalizer strategy $\{p_t^*\}$ for the adversary. Note that

$$\mathcal{V}_T(\ell, \Phi_T) = \inf_{q_1} \sup_{p_1} \mathbb{E}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \inf_{q_T} \sup_{p_T} \mathbb{E}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \}$$

$$\geq \inf_{q_1} \mathbb{E}_{\substack{x_1 \sim p_1^* \\ f_1 \sim q_1}} \inf_{q_2} \mathbb{E}_{\substack{x_2 \sim p_2^* \\ f_2 \sim q_2}} \ldots \inf_{q_T} \mathbb{E}_{\substack{x_T \sim p_T^* \\ f_T \sim q_T}} \left\{ \boldsymbol{B}\left(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)\right) - \inf_{\phi \in \Phi_T} \boldsymbol{B}\left(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)\right) \right\}$$

$$= \mathbb{E}_{x_1 \sim p_1} \ldots \mathbb{E}_{x_T \sim p_T} \left\{ \boldsymbol{B}\left(\ell(f, x_1), \ldots, \ell(f, x_T)\right) - \inf_{\phi \in \Phi_T} \boldsymbol{B}\left(\ell_{\phi_1}(f, x_1), \ldots, \ell_{\phi_T}(f, x_T)\right) \right\}$$

where $f \in \mathcal{F}$ is any arbitrary choice fixed before starting the game and $p_t = p_t^*\left(\{f_s = f, x_s\}_{s=1}^{t-1}\right)$ is defined by the equalizer strategy. □

Lemma 30. *For any departure mapping Φ_T and any $L > 0$ we have that*

$$\mathcal{V}_T(\mathcal{C}_{\mathcal{F}}, \mathcal{F}, \Phi_T) = \mathcal{V}_T(\mathcal{L}_{\mathcal{F}}, \mathcal{F}, \Phi_T)$$

Proof. Note that for any convex $x_1, \ldots, x_T$ we have that

$$\sum_{t=1}^T x_t(f_t) - \inf_{\phi \in \Phi} \sum_{t=1}^T x_t(\phi \circ f_t) = \sup_{\phi \in \Phi} \sum_{t=1}^T (x_t(f_t) - x_t(\phi \circ f_t))$$

$$\leq \sup_{\phi \in \Phi} \sum_{t=1}^T \langle \nabla x_t(f_t), f_t - \phi \circ f_t \rangle$$

$$= \sum_{t=1}^T \langle \nabla x_t(f_t), f_t \rangle - \inf_{\phi \in \Phi} \sum_{t=1}^T \langle \nabla x_t(f_t), \phi \circ f_t \rangle \qquad (28)$$

For any adversary strategy $x^* = (x_1^*, \ldots, x_T^*)$ where each $x_t^* : \mathcal{F}^t \mapsto \mathcal{X}$ and any player strategy $f^* = (f_1^*, \ldots, f_T^*)$ where each $f_t^* : \mathcal{X}^{t-1} \mapsto \mathcal{F}$, by Equation (28) we have that

$$\sum_{t=1}^T \langle \nabla x_t(f_t), f_t \rangle - \inf_{\phi \in \Phi} \sum_{t=1}^T \langle \nabla x_t(f_t), \phi \circ f_t \rangle \geq \sum_{t=1}^T x_t(f_t) - \inf_{\phi \in \Phi} \sum_{t=1}^T x_t(\phi \circ f_t)$$

where in the above, $f_t = f_t^*(\langle \nabla x_1(f_1), \cdot \rangle, \ldots, \langle \nabla x_{t-1}(f_{t-1}), \cdot \rangle)$ and $x_t = x_t^*(f_1, \ldots, f_t)$. Now if we take f^* and x^* to be the minimax optimal strategies then we see that

$$
\begin{aligned}
\mathcal{V}_T(\mathcal{L}_\mathcal{F}, \mathcal{F}, \Phi_T) &\geq \sum_{t=1}^T \langle \nabla x_t(f_t), f_t \rangle - \inf_{\phi \in \Phi} \sum_{t=1}^T \langle \nabla x_t(f_t), \phi \circ f_t \rangle \\
&\geq \sum_{t=1}^T x_t(f_t) - \inf_{\phi \in \Phi} \sum_{t=1}^T x_t(\phi \circ f_t) \\
&\geq \mathcal{V}_T(\mathcal{C}_\mathcal{F}, \mathcal{F}, \Phi_T)
\end{aligned}
$$

Thus we see that the value of the linear game upper bounds the value of the Lipschitz convex game. In fact the above argument shows that any strategy that provides vanishing regret guarantee against linear adversary provides vanishing regret gaurantee (with same rate) against convex Lipschitz adversary. This means that all that one needs to do to solve convex Lipschitz optimization optimally is to be able to solve online linear optimization optimally and also be able to calculate sub-gradient of a given function at any desired point.

Further since the set of linear functions is a subset of the set of convex Lipschitz functions we can conclude that

$$
\mathcal{V}_T(\mathcal{L}_\mathcal{F}, \mathcal{F}, \Phi_T) \leq \mathcal{V}_T(\mathcal{C}_\mathcal{F}, \mathcal{F}, \Phi_T)
$$

Hence we conclude the required statement that the value of the linear game is equal to the value of the convex Lipschitz game. □

Lemma 31. *Consider a game where player plays from set $\mathcal{F}$ adversary from set $\mathcal{X}$ and we are give a linear $\boldsymbol{B}$, loss ℓ and transformation set Φ_T. Assume that there exists a set $\mathcal{X}'$, loss function ℓ' and transformation set Φ'_T such that for any $\phi \in \Phi_T$ there exists $\phi' \in \Phi'_T$ such that for $x \in \mathcal{X}$ and $f \in \mathcal{F}$ there exists an $x' \in \mathcal{X}'$ such that for any $t \in [T]$,*

$$
\ell(f, x) - \ell_{\phi_t}(f, x) \leq \ell'(f, x') - \ell_{\phi'_t}(f, x')
$$

In that case we can conclude that value of the first game is bounded by value of the second game played with $\mathcal{F}$, $\mathcal{X}'$, $\boldsymbol{B}$, ℓ', Φ'_T, that is

$$
\mathcal{V}_T(\ell, \Phi_T, \mathcal{F}, \mathcal{X}) \leq \mathcal{V}_T(\ell', \Phi'_T, \mathcal{F}, \mathcal{X}')
$$

Proof. By assumption that for any $\phi \in \Phi_T$ there exists $\phi' \in \Phi'_T$ such that for $x \in \mathcal{X}$ and $f \in \mathcal{F}$ there exists an $x' \in \mathcal{X}'$ such that for any $t \in [T]$,

$$
\ell(f, x) - \ell_{\phi_t}(f, x) \leq \ell'(f, x') - \ell_{\phi'_t}(f, x')
$$

We can conclude that since $\boldsymbol{B}$ is linear, for any $\phi \in \Phi_T$ there exists $\phi' \in \Phi'_T$ such that for any $f_1, \ldots, f_T$ and $x_1, \ldots, x_T$ we have that for the corresponding $x'_1, \ldots, x'_T$ given by our assumption, we have that

$$
\begin{aligned}
\boldsymbol{B}(\ell(f_1, x_1), \ldots, &\ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \\
&\leq \boldsymbol{B}(\ell'(f_1, x'_1), \ldots, \ell'(f_T, x'_T)) - \boldsymbol{B}(\ell_{\phi'_1}(f_1, x'_1), \ldots, \ell_{\phi'_T}(f_T, x'_T))
\end{aligned}
$$

Hence we can conclude that

$$
\begin{aligned}
\sup_{\phi \in \Phi_T} &\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} \\
&\leq \sup_{\phi' \in \Phi'_T} \{ \boldsymbol{B}(\ell'(f_1, x'_1), \ldots, \ell'(f_T, x'_T)) - \boldsymbol{B}(\ell_{\phi'_1}(f_1, x'_1), \ldots, \ell_{\phi'_T}(f_T, x'_T)) \}
\end{aligned}
$$

Now say $q^* = (q_1^*, \ldots, q_T^*)$ where each $q_t^* : (\mathcal{F} \times \mathcal{X}')^{t-1} \mapsto \Delta(\mathcal{F})$ is the minimax optimal strategy for the player while playing the second game. Also let $p^* = (p_1^*, \ldots, p_T^*)$ where each $p_t^* : (\mathcal{F} \times \mathcal{X})^t \mapsto \Delta(\mathcal{X}')$ be the minimax optimal strategy for the player while playing the first game. In this case we see that

$$\mathcal{V}_T(\ell, \Phi_T, \mathcal{F}, \mathcal{X}) = \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1^* \\ x_1 \sim p_1^*}} \ldots \mathop{\mathbb{E}}_{\substack{f_T \sim q_T^* \\ x_T \sim p_T^*}} \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \}$$

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

$$\leq \mathcal{V}_T(\ell', \Phi_T', \mathcal{F}, \mathcal{X}')$$

$\square$

Proof of Theorem 1. The value of the game, defined in (2), is

$$\mathcal{V}_T(\ell, \Phi_T) = \inf_{q_1} \sup_{p_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \inf_{q_T} \sup_{p_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\mathbf{1} \left\{ \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} > A \right\} \right]$$

$$= \sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\mathbf{1} \left\{ \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} > A \right\} \right]$$

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

$$\mathcal{V}_T(\ell, \Phi_T) = \sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \quad [\mathbf{1}\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T))$$

$$+ \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} > A \}]$$

$$\leq \sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \quad [\mathbf{1}\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T))$$

$$+ \sup_{\phi \in \Phi_T} \left\{ \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) \right\}$$

$$+ \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} > A \}]$$

$$\leq \sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \quad [\mathbf{1}\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) > A/3 \}$$

$$+ \mathbf{1}\left\{ \sup_{\phi \in \Phi_T} \left\{ \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) \right\} > A/3 \right\}$$

$$+ \mathbf{1}\left\{ \sup_{\phi \in \Phi_T} \{ \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \} > A/3 \right\}]$$

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,

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

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

$$\mathcal{V}_T(\ell, \Phi_T)$$

$$\leq \sup_{p_1} \sup_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \sup_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\mathbf{1} \left\{ \boldsymbol{B}(\ell(f_1, x_1), \ldots, \ell(f_T, x_T)) - \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) > A/3 \right\} \right]$$

$$+ \sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\mathbf{1} \left\{ \sup_{\phi \in \Phi_T} \left\{ \boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) \right\} > A/3 \right\} \right]$$

$$+ \sup_{p_1} \sup_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \sup_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\mathbf{1} \left\{ \sup_{\phi \in \Phi_T} \left\{ \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(f_1, x_1), \ldots, \ell_{\phi_T}(f_T, x_T)) \right\} > A/3 \right\} \right]$$

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

$$\sup_{p_1} \inf_{q_1} \mathop{\mathbb{E}}_{\substack{f_1 \sim q_1 \\ x_1 \sim p_1}} \ldots \sup_{p_T} \inf_{q_T} \mathop{\mathbb{E}}_{\substack{f_T \sim q_T \\ x_T \sim p_T}} \left[\sup_{\phi \in \Phi_T} \left[\boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) \right] \right]$$

$$= \sup_{p_1} \inf_{q_1} \ldots \sup_{p_T} \inf_{q_T} \sup_{\phi \in \Phi_T} \left[\boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T)) \right]$$

because the objective

$$\boldsymbol{B}(\ell(q_1, p_1), \ldots, \ell(q_T, p_T)) - \boldsymbol{B}(\ell_{\phi_1}(q_1, p_1), \ldots, \ell_{\phi_T}(q_T, p_T))$$

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