Relax and Randomize: From Value to Algorithms

Alexander Rakhlin
University of Pennsylvania

Ohad Shamir
Microsoft Research

Karthik Sridharan
University of Pennsylvania

Abstract

We show a principled way of deriving online learning algorithms from a minimax analysis. Various upper bounds on the minimax value, previously thought to be non-constructive, are shown to yield algorithms. This allows us to seamlessly recover known methods and to derive new ones, also capturing such "unorthodox" methods as Follow the Perturbed Leader and the R^2 forecaster. Understanding the inherent complexity of the learning problem thus leads to the development of algorithms. To illustrate our approach, we present several new algorithms, including a family of randomized methods that use the idea of a "random playout". New versions of the Follow-the-Perturbed-Leader algorithms are presented, as well as methods based on the Littlestone's dimension, efficient methods for matrix completion with trace norm, and algorithms for the problems of transductive learning and prediction with static experts.

1 Introduction

This paper studies the online learning framework, where the goal of the player is to incur small regret while observing a sequence of data on which we place no distributional assumptions. Within this framework, many algorithms have been developed over the past two decades [6]. More recently, a non-algorithmic minimax approach has been developed to study the *inherent complexities* of sequential problems [2, 1, 15, 20]. It was shown that a theory in parallel to Statistical Learning can be developed, with random averages, combinatorial parameters, covering numbers, and other measures of complexity. Just as the classical learning theory is concerned with the study of the supremum of empirical or Rademacher process, online learning is concerned with the study of the supremum of martingale processes. While the tools introduced in [15, 17, 16] provide ways of studying the minimax value, no algorithms have been exhibited to achieve these non-constructive bounds in general.

In this paper, we show that algorithms can, in fact, be extracted from the minimax analysis. This observation leads to a unifying view of many of the methods known in the literature, and also gives a general recipe for developing new algorithms. We show that the potential method, which has been studied in various forms, naturally arises from the study of the minimax value as a certain *relaxation*. We further show that the sequential complexity tools introduced in [15] are, in fact, relaxations and can be used for constructing algorithms that enjoy the corresponding bounds. By choosing appropriate relaxations, we recover many known methods, improved variants of some known methods, and new algorithms. One can view our framework as one for converting a non-constructive proof of an upper bound on the value of the game into an algorithm. Surprisingly, this allows us to also study such "unorthodox" methods as Follow the Perturbed Leader [10], and the recent method of [7] under the same umbrella with others. We show that the idea of a random playout has a solid theoretical basis, and that Follow the Perturbed Leader algorithm is an example of such a method. Based on these developments, we exhibit an efficient method for the trace norm matrix completion problem, novel Follow the Perturbed Leader algorithms, and efficient methods for the problems of online transductive learning. The framework of this paper gives a recipe for developing algorithms. Throughout the paper, we stress that the notion of a relaxation, introduced below, is not appearing out of thin air but rather as an upper bound on the sequential Rademacher complexity. The understanding of *inherent complexity* thus leads to the *development of algorithms*.

Let us introduce some notation. The sequence $x_1, \ldots, x_t$ is often denoted by $x_{1:t}$, and the set of all distributions on some set $\mathcal{A}$ by $\Delta(\mathcal{A})$. Unless specified otherwise, ϵ denotes a vector $(\epsilon_1, \ldots, \epsilon_T)$ of i.i.d. Rademacher random variables. An $\mathcal{X}$-valued tree $\mathbf{x}$ of depth d is defined as a sequence $(\mathbf{x}_1, \ldots, \mathbf{x}_d)$ of mappings $\mathbf{x}_t : \{\pm 1\}^{t-1} \mapsto \mathcal{X}$ (see [15]). We often write $\mathbf{x}_t(\epsilon)$ instead of $\mathbf{x}_t(\epsilon_{1:t-1})$.

2 Value, The Minimax Algorithm, and Relaxations

Let $\mathcal{F}$ be the set of learner's moves and $\mathcal{X}$ the set of moves of Nature. The online protocol dictates that on every round $t = 1, \ldots, T$ the learner and Nature simultaneously choose $f_t \in \mathcal{F}$, $x_t \in \mathcal{X}$, and observe each other's actions. The learner aims to minimize regret $\mathbf{Reg}_T \triangleq \sum_{t=1}^T \ell(f_t, x_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^T \ell(f, x_t)$ where $\ell : \mathcal{F} \times \mathcal{X} \to \mathbb{R}$ is a known loss function. Our aim is to study this online learning problem at an abstract level without assuming convexity or other such properties of ℓ, $\mathcal{F}$ and $\mathcal{X}$. We do assume, however, that ℓ, $\mathcal{F}$, and $\mathcal{X}$ are such that the minimax theorem in the space of distributions over $\mathcal{F}$ and $\mathcal{X}$ holds. By studying the abstract setting, we are able to develop general algorithmic and non-algorithmic ideas that are common across various application areas. The starting point of our development is the minimax value of the associated online learning game:

$$\mathcal{V}_T(\mathcal{F}) = \inf_{q_1\in\Delta(\mathcal{F})} \sup_{x_1\in\mathcal{X}} \mathbb{E}_{f_1\sim q_1} \ldots \inf_{q_T\in\Delta(\mathcal{F})} \sup_{x_T\in\mathcal{X}} \mathbb{E}_{f_T\sim q_T} \left[\sum_{t=1}^T \ell(f_t, x_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^T \ell(f, x_t) \right] \quad (1)$$

where $\Delta(\mathcal{F})$ is the set of distributions on $\mathcal{F}$. The minimax formulation immediately gives rise to the optimal algorithm that solves the minimax expression at every round t and returns

$$\underset{q\in\Delta(\mathcal{F})}{\operatorname{argmin}} \left\{ \sup_{x_t} \mathbb{E}_{f_t\sim q} \left[\ell(f_t, x_t) + \inf_{q_{t+1}} \sup_{x_{t+1}} \mathbb{E}_{f_{t+1}} \ldots \inf_{q_T} \sup_{x_T} \mathbb{E}_{f_T} \left[\sum_{i=t+1}^T \ell(f_i, x_i) - \inf_{f\in\mathcal{F}} \sum_{i=1}^T \ell(f, x_i) \right] \right] \right\}$$

Henceforth, if the quantification in inf and sup is omitted, it will be understood that x_t, f_t, p_t, q_t range over $\mathcal{X}$, $\mathcal{F}$, $\Delta(\mathcal{X})$, $\Delta(\mathcal{F})$, respectively. Moreover, $\mathbb{E}_{x_t}$ is with respect to p_t while $\mathbb{E}_{f_t}$ is with respect to q_t. We now notice a recursive form for the value of the game. Define for any $t \in [T-1]$ and any given prefix $x_1, \ldots, x_t \in \mathcal{X}$ the *conditional value*

$$\mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_t) \triangleq \inf_{q\in\Delta(\mathcal{F})} \sup_{x\in\mathcal{X}} \left\{ \mathbb{E}_{f\sim q} [\ell(f, x)] + \mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_t, x) \right\}$$

with $\mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_T) \triangleq -\inf_{f\in\mathcal{F}} \sum_{t=1}^T \ell(f, x_t)$ and $\mathcal{V}_T(\mathcal{F}) = \mathcal{V}_T(\mathcal{F}|\{\})$. The minimax optimal algorithm specifying the mixed strategy of the player can be written succinctly as

$$q_t = \underset{q\in\Delta(\mathcal{F})}{\operatorname{argmin}} \sup_{x\in\mathcal{X}} \left\{ \mathbb{E}_{f\sim q} [\ell(f, x)] + \mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_{t-1}, x) \right\}. \quad (2)$$

Similar recursive formulations have appeared in the literature [8, 13, 19, 3], but now we have tools to study the conditional value of the game. We will show that various upper bounds on $\mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_{t-1}, x)$ yield an array of algorithms. In this way, the non-constructive approaches of [15, 16, 17] to upper bound the value of the game directly translate into algorithms. We note that the minimax algorithm in (2) can be interpreted as choosing the best decision that takes into account the present loss and the worst-case future. The first step in our analysis is to appeal to the minimax theorem and perform the same manipulation as in [1, 15], but only on the conditional values:

$$\mathcal{V}_T(\mathcal{F}|x_1, \ldots, x_t) = \sup_{p_{t+1}} \mathbb{E}_{x_{t+1}} \ldots \sup_{p_T} \mathbb{E}_{x_T} \left[\sum_{i=t+1}^T \inf_{f_i\in\mathcal{F}} \mathbb{E}_{x_i\sim p_i} \ell(f_i, x_i) - \inf_{f\in\mathcal{F}} \sum_{i=1}^T \ell(f, x_i) \right]. \quad (3)$$

The idea now is to come up with more manageable, yet tight, upper bounds on the conditional value. A *relaxation* $\mathbf{Rel}_T$ is a sequence of real-valued functions $\mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t)$ for each $t \in [T]$. We call a relaxation *admissible* if for any $x_1, \ldots, x_T \in \mathcal{X}$,

$$\mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \geq \inf_{q\in\Delta(\mathcal{F})} \sup_{x\in\mathcal{X}} \left\{ \mathbb{E}_{f\sim q} [\ell(f, x)] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t, x) \right\} \quad (4)$$

for all $t \in [T-1]$, and $\mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_T) \geq -\inf_{f\in\mathcal{F}} \sum_{t=1}^T \ell(f, x_t)$. We use the notation $\mathbf{Rel}_T(\mathcal{F})$ for $\mathbf{Rel}_T(\mathcal{F}|\{\})$. A strategy q that minimizes the expression in (4) defines an optimal *Meta-Algorithm* for an admissible relaxation $\mathbf{Rel}_T$:

on round t, compute

$$q_t = \arg\min_{q\in\Delta(\mathcal{F})} \sup_{x\in\mathcal{X}} \left\{ \mathbb{E}_{f\sim q} [\ell(f, x)] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_{t-1}, x) \right\}, \quad (5)$$

play $f_t \sim q_t$ and receive x_t from the opponent. Importantly, minimization need not be exact: any q_t that satisfies the admissibility condition (4) is a valid method, and we will say that such an algorithm is *admissible with respect to the relaxation* $\mathbf{Rel}_T$.

Proposition 1. *Let $\mathbf{Rel}_T$ be an admissible relaxation. For any admissible algorithm with respect to $\mathbf{Rel}_T$, (including the* Meta-Algorithm*), irrespective of the strategy of the adversary,*

$$\sum_{t=1}^{T} \mathbb{E}_{f_t \sim q_t} \ell(f_t, x_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f, x_t) \le \mathbf{Rel}_T(\mathcal{F}) \ , \tag{6}$$

and therefore, $\mathbb{E}[\mathbf{Reg}_T] \le \mathbf{Rel}_T(\mathcal{F})$. If $\ell(\cdot,\cdot)$ is bounded, the Hoeffding-Azuma inequality yields a high-probability bound on $\mathbf{Reg}_T$. We also have that $\mathcal{V}_T(\mathcal{F}) \le \mathbf{Rel}_T(\mathcal{F})$. Further, if for all $t \in [T]$, the admissible strategies q_t are deterministic, $\mathbf{Reg}_T \le \mathbf{Rel}_T(\mathcal{F})$.

The reader might recognize $\mathbf{Rel}_T$ as a potential function. It is known that one can derive regret bounds by coming up with a potential such that the current loss of the player is related to the difference in the potentials at successive steps, and that the regret can be extracted from the final potential. The origin/recipe for "good" potential functions has always been a mystery (at least to the authors). One of the key contributions of this paper is to show that they naturally arise as relaxations on the conditional value, and the conditional value is itself the tightest possible relaxation. In particular, for many problems a tight relaxation is achieved through symmetrization applied to the expression in (3). Define the *conditional Sequential Rademacher complexity*

$$\mathfrak{R}_T(\mathcal{F}|x_1,\dots,x_t) = \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{t+1:T}} \sup_{f \in \mathcal{F}} \left[2 \sum_{s=t+1}^{T} \epsilon_s \ell(f, \mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - \sum_{s=1}^{t} \ell(f, x_s) \right] . \tag{7}$$

Here the supremum is over all $\mathcal{X}$-valued binary trees of depth $T-t$. One may view this complexity as a partially symmetrized version of the sequential Rademacher complexity

$$\mathfrak{R}_T(\mathcal{F}) \triangleq \mathfrak{R}_T(\mathcal{F} \mid \{\}) = \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{1:T}} \sup_{f \in \mathcal{F}} \left[2 \sum_{s=1}^{T} \epsilon_s \ell(f, \mathbf{x}_s(\epsilon_{1:s-1})) \right] \tag{8}$$

defined in [15]. We shall refer to the term involving the tree $\mathbf{x}$ as the "future" and the term being subtracted off in (7) – as the "past". This indeed corresponds to the fact that the quantity is conditioned on the already observed $x_1,\dots,x_t$, while for the future we have the worst possible binary tree.[1]

Proposition 2. *The conditional Sequential Rademacher complexity is admissible.*

We now show that several well-known methods arise as further relaxations on $\mathfrak{R}_T$.

Exponential Weights [12, 21] Suppose $\mathcal{F}$ is a finite class and $|\ell(f,x)| \le 1$. In this case, a (tight) upper bound on sequential Rademacher complexity leads to the following relaxation:

$$\mathbf{Rel}_T(\mathcal{F}|x_1,\dots,x_t) = \inf_{\lambda > 0} \left\{ \frac{1}{\lambda} \log \left(\sum_{f \in \mathcal{F}} \exp\left(-\lambda \sum_{i=1}^{t} \ell(f, x_i) \right) \right) + 2\lambda(T-t) \right\} \tag{9}$$

Proposition 3. *The relaxation* (9) *is admissible and $\mathfrak{R}_T(\mathcal{F}|x_1,\dots,x_t) \le \mathbf{Rel}_T(\mathcal{F}|x_1,\dots,x_t)$. Furthermore, it leads to a parameter-free version of the Exponential Weights algorithm, defined on round $t+1$ by the mixed strategy $q_{t+1}(f) \propto \exp\left(-\lambda_t^* \sum_{s=1}^{t} \ell(f, x_s)\right)$ with λ_t^* the optimal value in* (9)*. The algorithm's regret is bounded by $\mathbf{Rel}_T(\mathcal{F}) \le 2\sqrt{2T \log |\mathcal{F}|}$.*

We point out that the exponential-weights algorithm arising from the relaxation (9) is a *parameter-free* algorithm. The learning rate λ^* can be optimized (via 1D line search) at each iteration.

Mirror Descent [4, 14] In the setting of online linear optimization [22], the loss is $\ell(f,x) = \langle f, x \rangle$. Suppose $\mathcal{F}$ is a unit ball in some Banach space and $\mathcal{X}$ is the dual. Let $\|\cdot\|$ be some $(2,C)$-smooth norm on $\mathcal{X}$ (in the Euclidean case, $C=2$). Using the notation $\tilde{x}_{t-1} = \sum_{s=1}^{t-1} x_s$, a straightforward upper bound on sequential Rademacher complexity is the following relaxation:

$$\mathbf{Rel}_T(\mathcal{F}|x_1,\dots,x_t) = \sqrt{\|\tilde{x}_{t-1}\|^2 + \left\langle \nabla \|\tilde{x}_{t-1}\|^2, x_t \right\rangle + C(T-t+1)} \tag{10}$$

[1] It is cumbersome to write out the indices on $\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})$ in (7), so we will instead use $\mathbf{x}_s(\epsilon)$ whenever this doesn't cause confusion.

Proposition 4. *The relaxation* (10) *is admissible and* $\mathfrak{R}_T(\mathcal{F}|x_1,\ldots,x_t) \le \mathbf{Rel}_T(\mathcal{F}|x_1,\ldots,x_t)$ *. It yields the update* $f_t = \frac{-\nabla\|\tilde{x}_{t-1}\|^2}{2\sqrt{\|\tilde{x}_{t-1}\|^2+C(T-t+1)}}$ *with regret bound* $\mathbf{Rel}_T(\mathcal{F}) \le \sqrt{2CT}$.

We would like to remark that the traditional mirror descent update can be shown to arise out of an appropriate relaxation. The algorithms proposed are parameter free as the step size is tuned automatically. We chose the popular methods of Exponential Weights and Mirror Descent for illustration. In the remainder of the paper, we develop new algorithms to show universality of our approach.

3 Classification

We start by considering the problem of supervised learning, where $\mathcal{X}$ is the space of instances and $\mathcal{Y}$ the space of responses (labels). There are two closely related protocols for the online interaction between the learner and Nature, so let us outline them. The "proper" version of supervised learning follows the protocol presented in Section 2: at time t, the learner selects $f_t \in \mathcal{F}$, Nature simultaneously selects $(x_t, y_t) \in \mathcal{X}\times\mathcal{Y}$, and the learner suffers the loss $\ell(f(x_t), y_t)$. The "improper" version is as follows: at time t, Nature chooses $x_t \in \mathcal{X}$ and presents it to the learner as "side information", the learner then picks $\hat{y}_t \in \mathcal{Y}$ and Nature simultaneously chooses $y_t \in \mathcal{Y}$. In the improper version, the loss of the learner is $\ell(\hat{y}_t, y_t)$, and it is easy to see that we may equivalently state this protocol as the learner choosing any function $f_t \in \mathcal{Y}^{\mathcal{X}}$ (not necessarily in $\mathcal{F}$), and Nature simultaneously choosing (x_t, y_t). We mostly focus on the "improper" version of supervised learning in this section. For the improper version, we may write the value in (1) as

$$\mathcal{V}_T(\mathcal{F}) = \sup_{x_1\in\mathcal{X}} \inf_{q_1\in\Delta(\mathcal{Y})} \sup_{y_1\in\mathcal{X}} \mathop{\mathbb{E}}_{\hat{y}_1\sim q_1} \ldots \sup_{x_T\in\mathcal{X}} \inf_{q_T\in\Delta(\mathcal{Y})} \sup_{y_T\in\mathcal{X}} \mathop{\mathbb{E}}_{\hat{y}_T\sim q_T} \left[\sum_{t=1}^T \ell(\hat{y}_t, y_t) - \inf_{f\in\mathcal{F}} \sum_{t=1}^T \ell(f(x_t), y_t)\right]$$

and a relaxation $\mathbf{Rel}_T$ is admissible if for any $(x_1,y_1)\ldots,(x_T,y_T) \in \mathcal{X}\times\mathcal{Y}$,

$$\sup_{x\in\mathcal{X}} \inf_{q\in\Delta(\mathcal{Y})} \sup_{y\in\mathcal{Y}} \left\{ \mathop{\mathbb{E}}_{\hat{y}\sim q} \ell(\hat{y}, y) + \mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t, (x,y)\right)\right\} \le \mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t\right) \tag{11}$$

Let us now focus on binary prediction, i.e. $\mathcal{Y} = \{\pm 1\}$. In this case, the supremum over y in (11) becomes a maximum over two values. Let us now take the absolute loss $\ell(\hat{y}, y) = |\hat{y} - y| = 1 - \hat{y}y$. We can see[2] that the optimal randomized strategy, given the side information x, is given by (11) as

$$q_t = \mathop{\mathrm{argmin}}_{q\in\Delta(\mathcal{Y})} \max\left\{1 - q + \mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t, (x,1)\right), 1 + q + \mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t, (x,-1)\right)\right\}$$

$$\text{or equivalently as : } \quad q_t = \frac{1}{2}\left\{\mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t, (x,1)\right) - \mathbf{Rel}_T\left(\mathcal{F}\middle|\{(x_i,y_i)\}_{i=1}^t, (x,-1)\right)\right\} \tag{12}$$

We now assume that $\mathcal{F}$ has a finite Littlestone's dimension $\mathrm{Ldim}(\mathcal{F})$ [11, 5]. Suppose the loss function is $\ell(\hat{y}, y) = |\hat{y} - y|$, and consider the "mixed" conditional Rademacher complexity

$$\sup_{\mathbf{x}} \mathbb{E}_\epsilon \sup_{f\in\mathcal{F}} \left\{2\sum_{i=1}^{T-t} \epsilon_i f(\mathbf{x}_i(\epsilon)) - \sum_{i=1}^t |f(x_i) - y_i|\right\} \tag{13}$$

as a possible relaxation. The admissibility condition (11) with the conditional sequential Rademacher (13) as a relaxation would require us to upper bound

$$\sup_{x_t} \inf_{q_t\in[-1,1]} \max_{y_t\in\{\pm1\}} \left\{ \mathop{\mathbb{E}}_{\hat{y}_t\sim q_t} |\hat{y}_t - y_t| + \sup_{\mathbf{x}} \mathbb{E}_\epsilon \sup_{f\in\mathcal{F}} \left\{2\sum_{i=1}^{T-t} \epsilon_i f(\mathbf{x}_i(\epsilon)) - \sum_{i=1}^t |f(x_i) - y_i|\right\}\right\} \tag{14}$$

However, the supremum over $\mathbf{x}$ is preventing us from obtaining a concise algorithm. We need to further "relax" this supremum, and the idea is to pass to a finite cover of $\mathcal{F}$ on the given tree $\mathbf{x}$ and then proceed as in the Exponential Weights example for a finite collection of experts. This leads to an upper bound on (13) and gives rise to algorithms similar in spirit to those developed in [5], but with more attractive computational properties and defined more concisely.

Define the function $g(d,t) = \sum_{i=0}^d \binom{t}{i}$, which is shown in [15] to be the maximum size of an exact (zero) cover for a function class with the Littlestone's dimension $\mathrm{Ldim} = d$. Given $\{(x_1,y_t),\ldots,(x_t,y_t)\}$ and $\sigma = (\sigma_1,\ldots,\sigma_t) \in \{\pm1\}^t$, let $\mathcal{F}_t(\sigma) = \{f \in \mathcal{F} : f(x_i) = \sigma_i \ \ \forall i \le t\}$, the subset of functions that agree with the signs given by σ on the "past" data and let $\mathcal{F}|_{x_1,\ldots,x_t} \triangleq \mathcal{F}|_{x^t} \triangleq \{(f(x_1),\ldots,f(x_t)) : f \in \mathcal{F}\}$ be the projection of $\mathcal{F}$ onto $x_1,\ldots,x_t$. Denote $L_t(f) = \sum_{i=1}^t |f(x_i) - y_i|$ and $L_t(\sigma) = \sum_{i=1}^t |\sigma_i - y_i|$ for $\sigma \in \{\pm1\}^t$. The following proposition gives a relaxation and an algorithm which achieves the $O(\sqrt{\mathrm{Ldim}(\mathcal{F})T\log T})$ regret bound. Unlike the algorithm of [5], we do not need to run an exponential number of experts in parallel and only require access to an oracle that computes the Littlestone's dimension.

[2] The extension to k-class prediction is immediate.

Proposition 5. *The relaxation*

$$\mathbf{Rel}_T\left(\mathcal{F}\big|(x^t,y^t)\right) = \frac{1}{\lambda}\log\left(\sum_{\sigma\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma)),T-t)\exp\left\{-\lambda L_t(\sigma)\right\}\right) + 2\lambda(T-t)\ .$$

is admissible and leads to an admissible algorithm which uses weights $q_t(-1) = 1 - q_t(+1)$ *and*

$$q_t(+1) = \frac{\sum_{(\sigma,+1)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,+1)),T-t)\exp\left\{-\lambda L_{t-1}(\sigma)\right\}}{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\left\{-\lambda L_{t-1}(\sigma)\right\}}, \tag{15}$$

There is a very close correspondence between the proof of Proposition 5 and the proof of the combinatorial lemma of [15], the analogue of the Vapnik-Chervonenkis-Sauer-Shelah result.

4 Randomized Algorithms and Follow the Perturbed Leader

We now develop a class of admissible randomized methods that arise through sampling. Consider the objective (5) given by a relaxation $\mathbf{Rel}_T$. If $\mathbf{Rel}_T$ is the sequential (or classical) Rademacher complexity, it involves an expectation over sequences of coin flips, and this computation (coupled with optimization for each sequence) can be prohibitively expensive. More generally, $\mathbf{Rel}_T$ might involve an expectation over possible ways in which the future might be realized. In such cases, we may consider a rather simple "random playout" strategy: draw the random sequence and solve only one optimization problem for that random sequence. The ideas of random playout have been discussed in previous literature for estimating the utility of a move in a game (see also [3]). We show that random playout strategy has a solid basis: for the examples we consider, it satisfies admissibility.

In many learning problems the sequential and the classical Rademacher complexities are within a constant factor of each other. This holds true, for instance, for linear functions in finite-dimensional spaces. In such cases, the relaxation $\mathbf{Rel}_T$ does not involve the supremum over a tree, and the randomized method only needs to draw a sequence of coin flips and compute a solution to an optimization problem slightly more complicated than ERM. We show that Follow the Perturbed Leader (FPL) algorithms [10] arise in this way. We note that FPL has been previously considered as a rather unorthodox algorithm providing some kind of regularization via randomization. Our analysis shows that it arises through a natural relaxation based on the sequential (and thus the classical) Rademacher complexity, coupled with the random playout idea. As a new algorithmic contribution, we provide a version of the FPL algorithm for the case of the decision sets being ℓ_2 balls, with a regret bound that is *independent of the dimension*. We also provide an FPL-style method for the combination of ℓ_1 and ℓ_∞ balls. To the best of our knowledge, these results are novel.

The assumption below implies that the sequential and classical Rademacher complexities are within constant factor C of each other. We later verify that it holds in the examples we consider.

Assumption 1. *There exists a distribution* $D\in\Delta(\mathcal{X})$ *and constant* $C\ge 2$ *such that for any* $t\in[T]$ *and given any* $x_1,\dots,x_{t-1},x_{t+1},\dots,x_T\in\mathcal{X}$ *and any* $\epsilon_{t+1},\dots,\epsilon_T\in\{\pm1\}$,

$$\sup_{p\in\Delta(\mathcal{X})}\mathop{\mathbb{E}}_{x_t\sim p}\sup_{f\in\mathcal{F}}\left[CA_{t+1}(f) - L_{t-1}(f) + \mathop{\mathbb{E}}_{x\sim p}\left[\ell(f,x)\right] - \ell(f,x_t)\right] \le \mathop{\mathbb{E}}_{\epsilon_t,x_t\sim D}\sup_{f\in\mathcal{F}}\left[CA_t(f) - L_{t-1}(f)\right]$$

where ϵ_t*'s are i.i.d. Rademacher,* $L_{t-1}(f) = \sum_{i=1}^{t-1}\ell(f,x_i)$*, and* $A_t(f) = \sum_{i=t}^{T}\epsilon_i\ell(f,x_i)$.

Under the above assumption one can use the following relaxation

$$\mathbf{Rel}_T\left(\mathcal{F}|x_1,\dots,x_t\right) = \mathop{\mathbb{E}}_{x_{t+1},\dots x_T\sim D}\mathbb{E}_\epsilon\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i) - \sum_{i=1}^{t}\ell(f,x_i)\right] \tag{16}$$

which is a partially symmetrized version of the classical Rademacher averages.

The proof of admissibility for the randomized methods is quite curious – the forecaster can be seen as mimicking the sequential Rademacher complexity by sampling from the "equivalently bad" classical Rademacher complexity under the specific distribution D specified by the above assumption.

Lemma 6. *Under Assumption 1, the relaxation in Eq.* (16) *is admissible and a randomized strategy that ensures admissibility is given by: at time t, draw* $x_{t+1},\dots,x_T\sim D$ *and* $\epsilon_{t+1},\dots,\epsilon_T$ *and then: (a) In the case the loss* ℓ *is convex in its first argument and set* $\mathcal{F}$ *is convex and compact, define*

$$f_t = \underset{g\in\mathcal{F}}{\operatorname{argmin}} \sup_{x\in\mathcal{X}} \left\{ \ell(g,x) + \sup_{f\in\mathcal{F}} \left\{ C \sum_{i=t+1}^{T} \epsilon_i \ell(f,x_i) - \sum_{i=1}^{t-1} \ell(f,x_i) - \ell(f,x) \right\} \right\} \tag{17}$$

(b) In the case of non-convex loss, sample f_t from the distribution

$$\hat{q}_t = \underset{\hat{q}\in\Delta(\mathcal{F})}{\operatorname{argmin}} \sup_{x\in\mathcal{X}} \left\{ \mathop{\mathbb{E}}_{f\sim\hat{q}} [\ell(f,x)] + \sup_{f\in\mathcal{F}} \left\{ C \sum_{i=t+1}^{T} \epsilon_i \ell(f,x_i) - \sum_{i=1}^{t-1} \ell(f,x_i) - \ell(f,x) \right\} \right\} \tag{18}$$

The expected regret for the method is bounded by the classical Rademacher complexity:

$$\mathbb{E}[\mathbf{Reg}_T] \le C\, \mathbb{E}_{x_{1:T}\sim D} \mathop{\mathbb{E}}_{\epsilon} \left[\sup_{f\in\mathcal{F}} \sum_{t=1}^{T} \epsilon_t \ell(f,x_t) \right],$$

Of particular interest are the settings of static experts and transductive learning, which we consider in Section 5. In the transductive case, the x_t's are pre-specified before the game, and in the static expert case – effectively absent. In these cases, as we show below, there is no explicit distribution D and we only need to sample the random signs ϵ's. We easily see that in these cases, the expected regret bound is simply two times the transductive Rademacher complexity.

The idea of sampling from a fixed distribution is particularly appealing in the case of linear loss, $\ell(f,x) = \langle f,x\rangle$. Suppose $\mathcal{X}$ is a unit ball in some norm $\|\cdot\|$ in a vector space B, and $\mathcal{F}$ is a unit ball in the dual norm $\|\cdot\|_*$. A sufficient condition implying Assumption 1 is then

Assumption 2. *There exists a distribution $D \in \Delta(\mathcal{X})$ and constant $C \ge 2$ such that for any $w \in B$,*

$$\sup_{x\in\mathcal{X}} \mathop{\mathbb{E}}_{x_t\sim p} \|w + 2\epsilon_t x_t\| \le \mathop{\mathbb{E}}_{x_t\sim D} \mathop{\mathbb{E}}_{\epsilon_t} \|w + C\epsilon_t x_t\| \tag{19}$$

At round t, the generic algorithm specified by Lemma 18 draws fresh Rademacher random variables ϵ and $x_{t+1},\ldots,x_T \sim D$ and picks

$$f_t = \underset{f\in\mathcal{F}}{\operatorname{argmin}} \sup_{x\in\mathcal{X}} \left\{ \langle f,x\rangle + \left\| C \sum_{i=t+1}^{T} \epsilon_i x_i - \sum_{i=1}^{t-1} x_i - x \right\| \right\} \tag{20}$$

We now look at ℓ_2/ℓ_2 and ℓ_1/ℓ_∞ cases and provide corresponding randomized algorithms.

Example : ℓ_1/ℓ_∞ Follow the Perturbed Leader Here, we consider the setting similar to that in [10]. Let $\mathcal{F} \subset \mathbb{R}^N$ be the ℓ_1 unit ball and $\mathcal{X}$ the (dual) ℓ_∞ unit ball in $\mathbb{R}^N$. In [10], $\mathcal{F}$ is the probability simplex and $\mathcal{X} = [0,1]^N$ but these are subsumed by the ℓ_1/ℓ_∞ case. Next we show that any symmetric distribution satisfies Assumption 2.

Lemma 7. *If D is any symmetric distribution over $\mathbb{R}$, then Assumption 2 is satisfied by using the product distribution D^N and any $C \ge 6/\mathbb{E}_{x\sim D}|x|$. In particular, Assumption 2 is satisfied with a distribution D that is uniform on the vertices of the cube $\{\pm 1\}^N$ and $C = 6$.*

The above lemma is especially attractive with Gaussian perturbations as sum of normal random variables is again normal. Hence, instead of drawing $x_{t+1},\ldots,x_T \sim N(0,1)$ on round t, one can simply draw one vector $X_t \sim N(0, T-t)$ as the perturbation. In this case, $C \le 8$.

The form of update in Equation (20), however, is not in a convenient form, and the following lemma shows a simple Follow the Perturbed Leader type algorithm with the associated regret bound.

Lemma 8. *Suppose $\mathcal{F}$ is the ℓ_1^N unit ball and $\mathcal{X}$ is the dual ℓ_∞^N unit ball, and let D be any symmetric distribution. Consider the randomized algorithm that at each round t, freshly draws Rademacher random variables $\epsilon_{t+1},\ldots,\epsilon_T$ and $x_{t+1},\ldots,x_T \sim D^N$ and picks $f_t = \underset{f\in\mathcal{F}}{\operatorname{argmin}} \left\langle f, \sum_{i=1}^{t-1} x_i - C\sum_{i=t+1}^{T} \epsilon_i x_i \right\rangle$ where $C = 6/\mathbb{E}_{x\sim D}|x|$. The expected regret is bounded as :*

$$\mathbb{E}[\mathbf{Reg}_T] \le C \mathop{\mathbb{E}}_{x_{1:T}\sim D^N} \mathbb{E}_\epsilon \left\| \sum_{t=1}^{T} \epsilon_t x_t \right\|_\infty + 4\sum_{t=1}^{T} \mathbf{P}_{y_{t+1:T}\sim D} \left(C \left| \sum_{i=t+1}^{T} y_i \right| \le 4 \right)$$

For instance, for the case of coin flips (with $C = 6$) or the Gaussian distribution (with $C = 3\sqrt{2\pi}$) the bound above is $4C\sqrt{T\log N}$, as the second term is bounded by a constant.

Example : ℓ_2/ℓ_2 Follow the Perturbed Leader We now consider the case when $\mathcal{F}$ and $\mathcal{X}$ are both the unit ℓ_2 ball. We can use as perturbation the uniform distribution on the surface of unit sphere, as the following lemma shows. This result was hinted at in [2], as in high dimensional case, the random draw from the unit sphere is likely to produce orthogonal directions. However, we do not require dimensionality to be high for our result.

Lemma 9. *Let $\mathcal{X}$ and $\mathcal{F}$ be unit balls in Euclidean norm. Then Assumption 2 is satisfied with a uniform distribution D on the surface of the unit sphere with constant $C = 4\sqrt{2}$.*

As in the previous example the update in (20) is not in a convenient form and this is addressed below.

Lemma 10. *Let $\mathcal{X}$ and $\mathcal{F}$ be unit balls in Euclidean norm, and D be the uniform distribution on the surface of the unit sphere. Consider the randomized algorithm that at each round (say round t) freshly draws $x_{t+1}, \ldots, x_T \sim D$ and picks $f_t = \left(-\sum_{i=1}^{t-1} x_i + C \sum_{i=t+1}^{T} x_i\right)/L$ where $C = 4\sqrt{2}$ and scaling factor $L = \left(\left\|-\sum_{i=1}^{t-1} x_i + C \sum_{i=t+1}^{T} \epsilon_i x_i\right\|_2^2 + 1\right)^{1/2}$. The randomized algorithm enjoys a bound on the expected regret given by $\mathbb{E}\left[\mathbf{Reg}_T\right] \le C\, \mathbb{E}_{x_1,\ldots,x_T \sim D} \left\|\sum_{t=1}^{T} x_t\right\|_2 \le 4\sqrt{2T}$.*

Importantly, the bound does not depend on the dimensionality of the space. To the best of our knowledge, this is the first such result for Follow the Perturbed Leader style algorithms. Further, unlike [10, 6], we directly deal with the adaptive adversary.

5 Static Experts with Convex Losses and Transductive Online Learning

We show how to recover a variant of the R^2 forecaster of [7], for static experts and transductive online learning. At each round, the learner makes a prediction $q_t \in [-1, 1]$, observes the outcome $y_t \in [-1, 1]$, and suffers convex L-Lipschitz loss $\ell(q_t, y_t)$. Regret is defined as the difference between learner's cumulative loss and $\inf_{f \in F} \sum_{t=1}^{T} \ell(f[t], y_t)$, where $F \subset [-1, 1]^T$ can be seen as a set of static experts. The transductive setting is equivalent to this: the sequence of x_t's is known before the game starts, and hence the *effective* function class is once again a subset of $[-1, 1]^T$. It turns out that in these cases, sequential Rademacher complexity becomes the classical Rademacher complexity (see [17]), which can thus be taken as a relaxation. This is also the reason that an efficient implementation by sampling is possible. For general convex loss, one possible admissible relaxation is just a conditional version of the classical Rademacher averages:

$$\mathbf{Rel}_T\left(F | y_1, \ldots, y_t\right) = \mathbb{E}_{\epsilon_{t+1:T}} \sup_{f \in F} \left[2L \sum_{s=t+1}^{T} \epsilon_s f[s] - L_t(f)\right] \tag{21}$$

where $L_t(f) = \sum_{s=1}^{t} \ell(f[s], y_s)$. If (21) is used as a relaxation, the calculation of prediction $\hat{y}_t$ involves a supremum over $f \in F$ with (potentially nonlinear) loss functions of instances seen so far. In some cases this optimization might be hard and it might be preferable if the supremum only involves terms *linear* in f. To this end we start by noting that by convexity

$$\sum_{t=1}^{T} \ell(\hat{y}_t, y_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), y_t) \le \sum_{t=1}^{T} \partial \ell(\hat{y}_t, y_t) \cdot \hat{y}_t - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \partial \ell(\hat{y}_t, y_t) \cdot f[t] \tag{22}$$

One can now consider an alternative online learning problem which, if we solve, also solves the original problem. More precisely, the new loss is $\ell'(\hat{y}, r) = r \cdot \hat{y}$; we first pick prediction $\hat{y}_t$ (deterministically) and the adversary picks r_t (corresponding to $r_t = \partial \ell(\hat{y}_t, y_t)$ for choice of y_t picked by adversary). Now note that ℓ' is indeed convex in its first argument and is L Lipschitz because $|\partial \ell(\hat{y}_t, y_t)| \le L$. This is a one dimensional convex learning game where we pick $\hat{y}_t$ and regret is given by the right hand side of (22). Hence, we can consider the relaxation

$$\mathbf{Rel}_T\left(F | \partial \ell(\hat{y}_1, y_1), \ldots, \partial \ell(\hat{y}_t, y_t)\right) = \mathbb{E}_{\epsilon_{t+1:T}} \sup_{f \in F} \left[2L \sum_{i=t+1}^{T} \epsilon_i f[t] - \sum_{i=1}^{t} \partial \ell(\hat{y}_i, y_i) \cdot f[i]\right] \tag{23}$$

as a linearized form of (21). At round t, the prediction of the algorithm is then

$$\hat{y}_t = \mathbb{E}_{\epsilon}\left[\sup_{f \in F}\left\{\sum_{i=t+1}^{T} \epsilon_i f[i] - \frac{1}{2L} \sum_{i=1}^{t-1} \partial \ell(\hat{y}_i, y_i) f[i] + \frac{1}{2} f[t]\right\} - \sup_{f \in F}\left\{\sum_{i=t+1}^{T} \epsilon_i f[i] - \frac{1}{2L} \sum_{i=1}^{t-1} \partial \ell(\hat{y}_i, y_i) f[i] - \frac{1}{2} f[t]\right\}\right] \tag{24}$$

Lemma 11. *The relaxation in Eq.* (23) *is admissible w.r.t. the prediction strategy specified in Equation* (24)*. Further the regret of the strategy is bounded as $\mathbf{Reg}_T \le 2L\ \mathbb{E}_{\epsilon}\left[\sup_{f \in F} \sum_{t=1}^{T} \epsilon_t f[t]\right]$.*

This algorithm is similar to R^2, with the main difference that R^2 computes the infima over a sum of absolute losses, while here we have a more manageable linearized objective. While we need to evaluate the expectation over ϵ's on each round, we can estimate $\hat{y}_t$ by sampling ϵ's and using McDiarmid's inequality argue that the estimate is close to $\hat{y}_t$ with high probability. The randomized prediction is now given simply as: on round t, draw $\epsilon_{t+1}, \ldots, \epsilon_T$ and predict

$$\hat{y}_t(\epsilon) = \inf_{f \in F}\left\{-\sum_{i=t+1}^{T} \epsilon_i f[i] + \frac{1}{2L} \sum_{i=1}^{t-1} \ell(f[i], y_i) + \frac{1}{2} f[t]\right\} - \inf_{f \in F}\left\{-\sum_{i=t+1}^{T} \epsilon_i f[i] + \frac{1}{2L} \sum_{i=1}^{t-1} \ell(f[i], y_i) - \frac{1}{2} f[t]\right\} \tag{25}$$

We now show that this predictor enjoys regret bound of the transductive Rademacher complexity :

Lemma 12. *The relaxation specified in Equation* (21) *is admissible w.r.t. the randomized prediction strategy specified in Equation* (25)*, and enjoys bound* $\mathbb{E}\left[\mathbf{Reg}_T\right] \le 2L\ \mathbb{E}_\epsilon\left[\sup_{f\in F}\sum_{t=1}^T \epsilon_t f[t]\right]$.

6 Matrix Completion

Consider the problem of predicting unknown entries in a matrix, in an online fashion. At each round t the adversary picks an entry in an $m \times n$ matrix and a value y_t for that entry. The learner then chooses a predicted value $\hat{y}_t$, and suffers loss $\ell(y_t, \hat{y}_t)$, assumed to be ρ-Lipschitz. We define our regret with respect to the class $\mathcal{F}$ of all matrices whose trace-norm is at most B (namely, we can use any such matrix to predict just by returning its relevant entry at each round). Usually, one has $B = \Theta(\sqrt{mn})$. Consider a transductive version, where we know in advance the location of all entries we need to predict. We show how to develop an algorithm whose regret is bounded by the (transductive) Rademacher complexity of $\mathcal{F}$, which by Theorem 6 of [18], is $O(B\sqrt{n})$ independent of T. Moreover, in [7], it was shown how one can convert algorithms with such guarantees to obtain the same regret even in a "fully" online case, where the set of entry locations is unknown in advance. In this section we use the two alternatives provided for transductive learning problem in the previous subsection, and provide two alternatives for the matrix completion problem. Both variants proposed here improve on the one provided by the R^2 forecaster in [7], since that algorithm competes against the smaller class $\mathcal{F}'$ of matrices with bounded trace-norm *and* bounded individual entries, and our variants are also computationally more efficient. Our first variant also improves on the recently proposed method in [9] in terms of memory requirements, and each iteration is simpler: Whereas that method requires storing and optimizing full $m \times n$ matrices every iteration, our algorithm only requires computing spectral norms of sparse matrices (assuming $T \ll mn$, which is usually the case). This can be done very efficiently, e.g. with power iterations or the Lanczos method.

Our first algorithm follows from Eq. (24), which for our setting gives the following prediction rule:

$$\hat{y}_t = B\ \mathbb{E}_{\epsilon}\left[\left(\left\|\sum_{i=t+1}^{T}\epsilon_i x_i - \frac{1}{2\rho}\sum_{i=1}^{t-1}\partial\ell(\hat{y}_i, y_i)x_i + \frac{1}{2}x_t\right\|_\sigma - \left\|\sum_{i=t+1}^{T}\epsilon_i x_i - \frac{1}{2\rho}\sum_{i=1}^{t-1}\partial\ell(\hat{y}_i, y_i)x_i - \frac{1}{2}x_t\right\|_\sigma\right)\right] \quad (26)$$

In the above $\|\cdot\|_\sigma$ stands for the spectral norm and each x_i is a matrix with a 1 at some specific position and 0 elsewhere. Notice that the algorithm only involves calculation of spectral norms on each round, which can be done efficiently. As mentioned in previous subsection, one can approximately evaluate the expectation by sampling several ϵ's on each round and averaging. The second algorithm follows (25), and is given by first drawing ϵ at random and then predicting

$$\hat{y}_t(\epsilon) = \sup_{\|f\|_\Sigma \le B}\left\{\sum_{i=t+1}^{T}\epsilon_i f[x_i] - \frac{1}{2\rho}\sum_{i=1}^{t-1}\ell(f[x_i], y_i) + \frac{1}{2}f[x_t]\right\} - \sup_{\|f\|_\Sigma \le B}\left\{\sum_{i=t+1}^{T}\epsilon_i f[x_i] - \frac{1}{2\rho}\sum_{i=1}^{t-1}\ell(f[x_i], y_i) - \frac{1}{2}f[x_t]\right\}$$

where $\|f\|_\Sigma$ is the trace norm of the $m \times n$ f, and $f[x_i]$ is the entry of the matrix f at the position x_i. Notice that the above involves solving two trace norm constrained convex optimization problems per round. As a simple corollary of Lemma 12, together with the bound on the Rademacher complexity mentioned earlier, we get that the expected regret of either variant is $O\left(B\ \rho\ (\sqrt{m} + \sqrt{n})\right)$.

7 Conclusion

In [2, 1, 15, 20] the minimax value of the online learning game has been analyzed and *non-constructive* bounds on the value have been provided. In this paper, we provide a general *constructive* recipe for deriving new (and old) online learning algorithms, using techniques from the apparently non-constructive minimax analysis. The recipe is rather simple: we start with the notion of conditional sequential Rademacher complexity, and find an "admissible" relaxation which upper bounds it. This relaxation immediately leads to an online learning algorithm, as well as to an associated regret guarantee. In addition to the development of a unified algorithmic framework, our contributions include (1) a new algorithm for online binary classification whenever the Littlestone dimension of the class is finite; (2) a family of randomized online learning algorithms based on the idea of a random playout, with new Follow the Perturbed Leader style algorithms arising as special cases; and (3) efficient algorithms for trace norm based online matrix completion problem which improve over currently known methods.

Acknowledgements: We gratefully acknowledge the support of NSF under grants CAREER DMS-0954737 and CCF-1116928.

References

[1] J. Abernethy, A. Agarwal, P. L. Bartlett, and A. Rakhlin. A stochastic view of optimal regret through minimax duality. In *COLT*, 2009.

[2] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari. Optimal strategies and minimax lower bounds for online convex games. In *COLT*, 2008.

[3] J. Abernethy, M.K. Warmuth, and J. Yellin. Optimal strategies from random walks. In *COLT*, pages 437–445, 2008.

[4] A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[5] S. Ben-David, D. Pál, and S. Shalev-Shwartz. Agnostic online learning. In *COLT*, 2009.

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

[7] N. Cesa-Bianchi and O. Shamir. Efficient online learning via randomized rounding. In *NIPS*, 2011.

[8] J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

[9] E. Hazan, S. Kale, and S. Shalev-Shwartz. Near-optimal algorithms for online matrix prediction. In *COLT*, 2012.

[10] A. Kalai and S. Vempala. Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.*, 71(3):291–307, 2005.

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

[12] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.

[13] J.F. Mertens, S. Sorin, and S. Zamir. *Repeated games*. Univ. Catholique de Louvain, Center for Operations Research & Econometrics, 1994.

[14] A.S. Nemirovsky and D.B. Yudin. Problem complexity and method efficiency in optimization. 1983.

[15] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Random averages, combinatorial parameters, and learnability. In *NIPS*, 2010. Available at http://arxiv.org/abs/1006.1138.

[16] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Beyond regret. In *COLT*, 2011. Available at http://arxiv.org/abs/1011.3168.

[17] A. Rakhlin, K. Sridharan, and A. Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. In *NIPS*, 2011. Available at http://arxiv.org/abs/1104.5070.

[18] O. Shamir and S. Shalev-Shwartz. Collaborative filtering with the trace norm: Learning, bounding, and transducing. In *COLT*, 2011.

[19] S. Sorin. The operator approach to zero-sum stochastic games. *Stochastic Games and Applications, NATO Science Series C, Mathematical and Physical Sciences*, 570:417–426, 2003.

[20] K. Sridharan and A. Tewari. Convex games in banach spaces. In *COLT*, 2010.

[21] V.G. Vovk. Aggregating strategies. In *Proc. Third Workshop on Computational Learning Theory*, pages 371–383. Morgan Kaufmann, 1990.

[22] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. 2003.

A PROOFS

Proof of Proposition 1. By definition,

$$\sum_{t=1}^T \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) - \inf_{f\in\mathcal{F}}\sum_{t=1}^T \ell(f,x_t) \le \sum_{t=1}^T \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) + \mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_T\right) \ .$$

Peeling off the T-th expected loss, we have

$$\begin{aligned}\sum_{t=1}^T \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) + \mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_T\right) &\le \sum_{t=1}^{T-1} \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) + \left\{\mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) + \mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_T\right)\right\} \\ &\le \sum_{t=1}^{T-1} \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) + \mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_{T-1}\right)\end{aligned}$$

where we used the fact that q_T is an admissible algorithm for this relaxation, and thus the last inequality holds for any choice x_T of the opponent. Repeating the process, we obtain

$$\sum_{t=1}^T \mathbb{E}_{f_t\sim q_t}\ell(f_t,x_t) - \inf_{f\in\mathcal{F}}\sum_{t=1}^T \ell(f,x_t) \le \mathbf{Rel}_T\left(\mathcal{F}\right) \ .$$

We remark that the left-hand side of this inequality is random, while the right-hand side is not. Since the inequality holds for any realization of the process, it also holds in expectation. The inequality

$$\mathcal{V}_T(\mathcal{F}) \le \mathbf{Rel}_T\left(\mathcal{F}\right)$$

holds by unwinding the value recursively and using admissibility of the relaxation. The high-probability bound is an immediate consequences of (6) and the Hoeffding-Azuma inequality for bounded martingales. The last statement is immediate. □

Proof of Proposition 2. Denote $L_t(f) = \sum_{s=1}^t \ell(f,x_s)$. The first step of the proof is an application of the minimax theorem (we assume the necessary conditions hold):

$$\begin{aligned}&\inf_{q_t\in\Delta(\mathcal{F})}\sup_{x_t\in\mathcal{X}}\left\{\mathop{\mathbb{E}}_{f_t\sim q_t}\left[\ell(f_t,x_t)\right] + \sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_t(f)\right]\right\} \\ &= \sup_{p_t\in\Delta(\mathcal{X})}\inf_{f_t\in\mathcal{F}}\left\{\mathop{\mathbb{E}}_{x_t\sim p_t}\left[\ell(f_t,x_t)\right] + \mathop{\mathbb{E}}_{x_t\sim p_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_t(f)\right]\right\}\end{aligned}$$

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

$$\begin{aligned}&\mathop{\mathbb{E}}_{x_t\sim p_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_{t-1}(f) + \inf_{f_t\in\mathcal{F}}\mathop{\mathbb{E}}_{x_t\sim p_t}\left[\ell(f_t,x_t)\right] - \ell(f,x_t)\right] \\ &\le \mathop{\mathbb{E}}_{x_t\sim p_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_{t-1}(f) + \mathop{\mathbb{E}}_{x_t\sim p_t}\left[\ell(f,x_t)\right] - \ell(f,x_t)\right] \\ &\le \mathop{\mathbb{E}}_{x_t,x'_t\sim p_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_{t-1}(f) + \ell(f,x'_t) - \ell(f,x_t)\right]\end{aligned}$$

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

$$\begin{aligned}&\mathop{\mathbb{E}}_{x_t,x'_t\sim p_t}\mathop{\mathbb{E}}_{\epsilon_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_{t-1}(f) + \epsilon_t(\ell(f,x'_t) - \ell(f,x_t))\right] \\ &\le \sup_{x_t,x'_t\in\mathcal{X}}\mathop{\mathbb{E}}_{\epsilon_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[2\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - L_{t-1}(f) + \epsilon_t(\ell(f,x'_t) - \ell(f,x_t))\right]\end{aligned}$$

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

$$2\sup_{x_t\in\mathcal{X}}\mathop{\mathbb{E}}_{\epsilon_t}\sup_{\mathbf{x}}\mathbb{E}_{\epsilon_{t+1:T}}\sup_{f\in\mathcal{F}}\left[\sum_{s=t+1}^T \epsilon_s\ell(f,\mathbf{x}_{s-t}(\epsilon_{t+1:s-1})) - \frac{1}{2}L_{t-1}(f) + \epsilon_t\ell(f,x_t)\right] = \mathfrak{R}_T(\mathcal{F}|x_1,\ldots,x_{t-1}) \ .$$

The last equality is easy to verify, as we are effectively adding a root x_t to the two subtrees, for $\epsilon_t = +1$ and $\epsilon_t = -1$, respectively.

One can see that the proof of admissibility corresponds to one step minimax swap and symmetrization in the proof of [15]. In contrast, in the latter paper, all T minimax swaps are performed at once, followed by T symmetrization steps. □

Proof of Proposition 3. Let us first prove that the relaxation is admissible with the Exponential Weights algorithm as an admissible algorithm. Let $L_t(f) = \sum_{i=1}^{t} \ell(f, x_i)$. Let λ^* be the optimal value in the definition of $\mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_{t-1})$. Then

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{x_t \in \mathcal{X}} \left\{ \mathbb{E}_{f \sim q_t} [\ell(f, x_t)] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \right\}$$
$$\le \inf_{q_t \in \Delta(\mathcal{F})} \sup_{x_t \in \mathcal{X}} \left\{ \mathbb{E}_{f \sim q_t} [\ell(f, x_t)] + \frac{1}{\lambda^*} \log \left(\sum_{f \in \mathcal{F}} \exp\left(-\lambda^* L_t(f)\right) \right) + 2\lambda^*(T - t) \right\}$$

Let us upper bound the infimum by a particular choice of q which is the exponential weights distribution

$$q_t(f) = \exp(-\lambda^* L_{t-1}(f))/Z_{t-1}$$

where $Z_{t-1} = \sum_{f \in \mathcal{F}} \exp\left(-\lambda^* L_{t-1}(f)\right)$. By [6, Lemma A.1],

$$\frac{1}{\lambda^*} \log \left(\sum_{f \in \mathcal{F}} \exp\left(-\lambda^* L_t(f)\right) \right) = \frac{1}{\lambda^*} \log \left(\mathbb{E}_{f \sim q_t} \exp\left(-\lambda^* \ell(f, x_t)\right) \right) + \frac{1}{\lambda^*} \log Z_{t-1}$$
$$\le -\mathbb{E}_{f \sim q_t} \ell(f, x_t) + \frac{\lambda^*}{2} + \frac{1}{\lambda^*} \log Z_{t-1}$$

Hence,

$$\inf_{q_t \in \Delta(\mathcal{F})} \sup_{x_t \in \mathcal{X}} \left\{ \mathbb{E}_{f \sim q_t} [\ell(f, x_t)] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \right\}$$
$$\le \frac{1}{\lambda^*} \log \left(\sum_{f \in \mathcal{F}} \exp\left(-\lambda^* L_{t-1}(f)\right) \right) + 2\lambda^*(T - t + 1)$$
$$= \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_{t-1})$$

by the optimality of λ^*. The bound can be improved by a factor of 2 for some loss functions, since it will disappear from the definition of sequential Rademacher complexity.

We conclude that the Exponential Weights algorithm is an admissible strategy for the relaxation (9). The final regret bound follows immediately from the bound on sequential Rademacher complexity (which, in this case, is simply the supremum of a martingale difference process indexed by N elements – see e.g. [15]).

Arriving at the relaxation We now show that the Exponential Weights relaxation arises naturally as an upper bound on sequential Rademacher complexity of a finite class. For any $\lambda > 0$,

$$\mathbb{E}_{\epsilon} \left[\sup_{f \in \mathcal{F}} \left\{ 2 \sum_{i=1}^{T-t} \epsilon_i \ell(f, \mathbf{x}_i(\epsilon)) - L_t(f) \right\} \right] \le \frac{1}{\lambda} \log \left(\mathbb{E}_{\epsilon} \left[\sup_{f \in \mathcal{F}} \exp \left(2\lambda \sum_{i=1}^{T-t} \epsilon_i \ell(f, \mathbf{x}_i(\epsilon)) - \lambda L_t(f) \right) \right] \right)$$
$$\le \frac{1}{\lambda} \log \left(\mathbb{E}_{\epsilon} \left[\sum_{f \in \mathcal{F}} \exp \left(2\lambda \sum_{i=1}^{T-t} \epsilon_i \ell(f, \mathbf{x}_i(\epsilon)) - \lambda L_t(f) \right) \right] \right)$$
$$= \frac{1}{\lambda} \log \left(\sum_{f \in \mathcal{F}} \exp\left(-\lambda L_t(f)\right) \mathbb{E}_{\epsilon} \left[\prod_{i=1}^{T-t} \exp\left(2\lambda \epsilon_i \ell(f, \mathbf{x}_i(\epsilon))\right) \right] \right)$$

Since, conditioned on $\epsilon_1, \ldots, \epsilon_{i-1}$, the random variable $\epsilon_i \ell(f, \mathbf{x}_i(\epsilon))$ is subgaussian, we can upper bound the expected value of the product, peeling one random variable at a time from the end (see

[15] for the proof). We arrive at the upper bound

$$\frac{1}{\lambda}\log\left(\sum_{f\in\mathcal{F}}\exp\left(-\lambda L_t(f)\right)\times\exp\left(2\lambda^2\max_{\epsilon_1,\ldots\epsilon_{T-t}\in\{\pm1\}}\sum_{i=1}^{T-t}\ell(f,\mathbf{x}_i(\epsilon))^2\right)\right)$$
$$\le\frac{1}{\lambda}\log\left(\sum_{f\in\mathcal{F}}\exp\left(-\lambda L_t(f)+2\lambda^2\max_{\epsilon_1,\ldots\epsilon_{T-t}\in\{\pm1\}}\sum_{i=1}^{T-t}\ell(f,\mathbf{x}_i(\epsilon))^2\right)\right)$$
$$\le\frac{1}{\lambda}\log\left(\sum_{f\in\mathcal{F}}\exp\left(-\lambda L_t(f)\right)\right)+2\lambda\sup_{\mathbf{x}}\sup_{f\in\mathcal{F}}\max_{\epsilon_1,\ldots\epsilon_{T-t}\in\{\pm1\}}\sum_{i=1}^{T-t}\ell(f,\mathbf{x}_i(\epsilon))^2$$

The last term, representing the "worst future", is upper bounded by $2\lambda(T-t)$, assuming that the losses are bounded by 1. This removes the $\mathbf{x}$ tree and leads to the relaxation (9) and a computationally tractable algorithm. □

Proof of Proposition 4. The argument can be seen as a generalization of the Euclidean proof in [2] to general smooth norms. Let $\tilde{x}_{t-1}=\sum_{i=1}^{t-1}x_i$. The optimal algorithm for the relaxation (10) is

$$f_t^*=\operatorname*{argmin}_{f\in\mathcal{F}}\left\{\sup_{x_t\in\mathcal{X}}\left\{\langle f,x_t\rangle+\sqrt{\|\tilde{x}_{t-1}\|^2+\left\langle\nabla\|\tilde{x}_{t-1}\|^2,x_t\right\rangle+C(T-t+1)}\right\}\right\} \tag{27}$$

We shall show admissibility instead using

$$f_t=-\frac{\nabla\|\tilde{x}_{t-1}\|^2}{2\sqrt{\|\tilde{x}_{t-1}\|^2+C(T-t+1)}}\ . \tag{28}$$

Plugging this choice into the admissibility condition (4), we get

$$\sup_{x_t\in\mathcal{X}}\left\{-\frac{\left\langle\nabla\|\tilde{x}_{t-1}\|^2,x_t\right\rangle}{2\sqrt{A}}+\sqrt{A+\left\langle\nabla\|\tilde{x}_{t-1}\|^2,x_t\right\rangle}\right\}$$

where $A=\|\tilde{x}_{t-1}\|^2+C(T-t+1)$. It can be easily verified that an expression of the form $-\frac{x}{2\sqrt{y}}+\sqrt{y+x}$ is maximized at $x=0$ for a positive y. Hence,

$$\inf_{f_t\in\mathcal{F}}\left\{\sup_{x_t\in\mathcal{X}}\left\{\langle f_t,x_t\rangle+\left(\|\tilde{x}_{t-1}\|^2+\left\langle\nabla\|\tilde{x}_{t-1}\|^2,x_t\right\rangle+C(T-t+1)\right)^{1/2}\right\}\right\}\le\left(\|\tilde{x}_{t-1}\|^2+C(T-t+1)\right)^{1/2}$$
$$\le\left(\|\bar{x}_{t-2}\|^2+\left\langle\nabla\|\bar{x}_{t-2}\|^2,x_{t-1}\right\rangle+C(T-t+2)\right)^{1/2}=\mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_{t-1}\right)$$

Hence, the choice (28) is an admissible algorithm for the relaxation (10). Evidently, the above proof of admissibility is very simple, but it might seem that we pulled the algorithm (28) out of a hat. We now show that in fact the choice of f_t above is the optimal choice f_t^*. The proof below is not required for admissibility, and we only include it for completeness. The proof uses the fact that for *any* norm $\|\cdot\|$,

$$\left\langle\nabla\tfrac{1}{2}\|x\|^2,x\right\rangle=\|x\|^2\ . \tag{29}$$

To prove this, observe that by convexity $\|0\|\ge\|x\|+\langle\nabla\|x\|,0-x\rangle$ and $\|2x\|\ge\|x\|+\langle\nabla\|x\|,2x-x\rangle$ implying $\langle\nabla\|x\|,x\rangle=\|x\|$. On the other hand, by the chain rule, $\nabla\frac{1}{2}\|x\|^2=\|x\|\cdot\nabla\|x\|$, thus implying (29).

Let

$$K\triangleq\mathrm{Kernel}(\nabla\|\tilde{x}_{t-1}\|^2)=\left\{h:\left\langle\nabla\|\tilde{x}_{t-1}\|^2,h\right\rangle=0\right\}\ ,\ K'\triangleq\mathrm{Kernel}(\tilde{x}_{t-1})=\{h:\langle h,\tilde{x}_{t-1}\rangle=0\}\,.$$

We first claim that x_t can always be written as $x_t=\beta\tilde{x}_{t-1}+\gamma y$ for some $y\in K$ and for some scalars β,γ. Indeed, suppose that $x_t=\beta\tilde{x}_{t-1}+\gamma y+z$ for some $y\in K$ and $z\notin K$. Then we can rewrite x_t as

$$x_t=(\beta+\delta)\,\tilde{x}_{t-1}+(\gamma y-\delta\tilde{x}_{t-1}+z)$$

where $\delta = \frac{\langle \nabla \|\tilde{x}_{t-1}\|^2, z\rangle}{2\|\tilde{x}_{t-1}\|^2}$. It is enough to check that $(\gamma y - \delta \tilde{x}_{t-1} + z) \in K$. Indeed, using (29),

$$\left\langle \nabla \|\tilde{x}_{t-1}\|^2, -\delta \tilde{x}_{t-1} + z \right\rangle = -2\delta \|\tilde{x}_{t-1}\|^2 + \left\langle \nabla \|\tilde{x}_{t-1}\|^2, z \right\rangle = 0 \, .$$

An analogous proof shows that we may always decompose any f_t as $f_t = -\alpha \nabla \frac{1}{2}\|\tilde{x}_{t-1}\|^2 + g$ for some $g \in K'$ and a scalar α. Hence,

$$\begin{aligned} &\langle f_t, x_t\rangle + \left(\|\tilde{x}_{t-1}\|^2 + \left\langle \nabla\|\tilde{x}_{t-1}\|^2, x_t\right\rangle + C(T-t+1)\right)^{1/2} \\ &= -\alpha\beta\|\tilde{x}_{t-1}\|^2 + \gamma\langle g, y\rangle + \left(\|\tilde{x}_{t-1}\|^2 + 2\beta\|\tilde{x}_{t-1}\|^2 + C(T-t+1)\right)^{1/2} \end{aligned} \tag{30}$$

Given any $f_t = -\alpha \nabla \frac{1}{2} \|\tilde{x}_{t-1}\|^2 + g$, x_t can be picked with $y \in K$ that satisfies $\langle g, y\rangle \geq 0$. One can always do this because if for some y', $\langle g, y'\rangle < 0$ by picking $y = -y'$ we can ensure that $\langle g, y\rangle \geq 0$. Hence the minimizer f_t must be once such that $f_t = -\alpha \nabla \frac{1}{2} \|\tilde{x}_{t-1}\|^2$ and thus $\langle g, y\rangle = 0$. Now, it must be that $\alpha \geq 0$ so that x_t either increases the first term or second term but not both. Hence we conclude that $f_t = -\alpha \nabla \frac{1}{2} \|\tilde{x}_{t-1}\|^2$ for some $\alpha \geq 0$. It remains to determine the optimal α. Given such an f_t, the sup over x_t can be written as supremum over β of a concave function, which gives rise to the derivative condition

$$-\alpha \|\tilde{x}_{t-1}\|^2 + \frac{\|\tilde{x}_{t-1}\|^2}{\sqrt{\|\tilde{x}_{t-1}\|^2 + 2\beta \|\tilde{x}_{t-1}\|^2 + C(T-t+1)}} = 0$$

Hence we can conclude that for any $f_t = -\alpha \nabla \frac{1}{2} \|\tilde{x}_{t-1}\|^2$,

$$\begin{aligned} &\sup_{x_t} \left\{ \langle f_t, x_t\rangle + \left(\|\tilde{x}_{t-1}\|^2 + \left\langle \nabla\|\tilde{x}_{t-1}\|^2, x_t\right\rangle + C(T-t+1)\right)^{1/2} \right\} \\ &= \sup_{\beta} \left\{ -\alpha\beta\|\tilde{x}_{t-1}\|^2 + \left(\|\tilde{x}_{t-1}\|^2 + 2\beta\|\tilde{x}_{t-1}\|^2 + C(T-t+1)\right)^{1/2} \right\} \\ &= \frac{\alpha}{2}\left(\|\tilde{x}_{t-1}\|^2 + C(T-t+1)\right) + \frac{1}{2\alpha} \end{aligned}$$

Hence optimizing the above over α we get

$$\alpha = \frac{1}{\sqrt{\|\tilde{x}_{t-1}\|^2 + C(T-t+1)}} \, .$$

Hence we can conclude that

$$f_t^* = -\frac{\nabla \frac{1}{2} \|\tilde{x}_{t-1}\|^2}{\sqrt{\|\tilde{x}_{t-1}\|^2 + C(T-t+1)}}$$

Plugging back the value of α, we find that $\beta = 0$. Hence we conclude that f_t defined in (28) in fact coincides with the optimal solution (27).

Arriving at the Relaxation The derivation of the relaxation is immediate:

$$\mathfrak{R}_T(\mathcal{F}|x_1, \ldots, x_t) = \sup_{\mathbf{x}} \mathbb{E}_{\epsilon_{t+1:T}} \left\| \sum_{s=t+1}^{T} \epsilon_s \mathbf{x}_{s-t}(\epsilon_{t+1:s-1}) - \sum_{s=1}^{t} x_s \right\| \tag{31}$$

$$\leq \sup_{\mathbf{x}} \sqrt{ \mathbb{E}_{\epsilon_{t+1:T}} \left\| \sum_{s=t+1}^{T} \epsilon_s \mathbf{x}_{s-t}(\epsilon_{t+1:s-1}) - \sum_{s=1}^{t} x_s \right\|^2 } \tag{32}$$

$$\leq \sup_{\mathbf{x}} \sqrt{ \left\| \sum_{s=1}^{t} x_s \right\|^2 + C\mathbb{E}_{\epsilon_{t+1:T}} \sum_{s=t+1}^{T} \left\| \epsilon_s \mathbf{x}_{s-t}(\epsilon_{t+1:s-1}) \right\|^2 } \tag{33}$$

where the last step is due to the smoothness of the norm and the fact that the first-order terms disappear under the expectation. The sum of norms is now upper bounded by $T-t$, thus removing the dependence on the "future", and we arrive at

$$\sqrt{ \left\| \sum_{s=1}^{t} x_s \right\|^2 + C(T-t) } \leq \sqrt{ \left\| \sum_{s=1}^{t-1} x_s \right\|^2 + \left\langle \nabla \frac{1}{2} \left\| \sum_{s=1}^{t-1} x_s \right\|^2, x_t \right\rangle + C(T-t+1) }$$

as a relaxation on the sequential Rademacher complexity. □

Proof of Proposition 5. We would like to show that, with the distribution q_t^* defined in (15),

$$\max_{y_t\in\{\pm1\}}\left\{\mathop{\mathbb{E}}_{\hat{y}_t\sim q_t^*}|\hat{y}_t-y_t|+\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^t,y^t)\right)\right\}\le\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^{t-1},y^{t-1})\right)$$

for any $x_t\in\mathcal{X}$. Let $\sigma\in\{\pm1\}^{t-1}$ and $\sigma_t\in\{\pm1\}$. We have

$$\begin{aligned}
&\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^t,y^t)\right)-2\lambda(T-t)\\
&=\frac{1}{\lambda}\log\left(\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda|\sigma_t-y_t|\}\right)\\
&\le\frac{1}{\lambda}\log\left(\sum_{\sigma_t\in\{\pm1\}}\exp\{-\lambda|\sigma_t-y_t|\}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right)
\end{aligned}$$

Just as in the proof of Proposition 3, we may think of the two choices σ_t as the two experts whose weighting q_t^* is given by the sum involving the Littlestone's dimension of subsets of $\mathcal{F}$. Introducing the normalization term, we arrive at the upper bound

$$\frac{1}{\lambda}\log\left(\mathbb{E}_{\sigma_t\sim q_t^*}\exp\{-\lambda|\sigma_t-y_t|\}\right)+\frac{1}{\lambda}\log\left(\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right)$$

$$\le-\mathbb{E}_{\sigma_t\sim q_t^*}|\sigma_t-y_t|+2\lambda+\frac{1}{\lambda}\log\left(\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right)$$

The last step is due to Lemma A.1 in [6]. It remains to show that the log normalization term is upper bounded by the relaxation at the previous step:

$$\begin{aligned}
&\frac{1}{\lambda}\log\left(\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right)\\
&\le\frac{1}{\lambda}\log\left(\sum_{\sigma\in\mathcal{F}|_{x^{t-1}}}\exp\{-\lambda L_{t-1}(\sigma)\}\sum_{\sigma_t\in\{\pm1\}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\right)\\
&\le\frac{1}{\lambda}\log\left(\sum_{\sigma\in\mathcal{F}|_{x^{t-1}}}\exp\{-\lambda L_{t-1}(\sigma)\}\,g(\mathrm{Ldim}(\mathcal{F}_{t-1}(\sigma)),T-t+1)\right)\\
&=\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^{t-1},y^{t-1})\right)
\end{aligned}$$

To justify the last inequality, note that $\mathcal{F}_{t-1}(\sigma)=\mathcal{F}_t(\sigma,+1)\cup\mathcal{F}_t(\sigma,-1)$ and at most one of $\mathcal{F}_t(\sigma,+1)$ or $\mathcal{F}_t(\sigma,-1)$ can have Littlestone's dimension $\mathrm{Ldim}(\mathcal{F}_{t-1}(\sigma))$. We now appeal to the recursion

$$g(d,T-t)+g(d-1,T-t)\le g(d,T-t+1)$$

where $g(d,T-t)$ is the size of the zero cover for a class with Littlestone's dimension d on the worst-case tree of depth $T-t$ (see [15]). This completes the proof of admissibility.

Alternative Method Let us now derive the algorithm. Once again, consider the optimization problem

$$\max_{y_t\in\{\pm1\}}\left\{\mathop{\mathbb{E}}_{\hat{y}_t\sim q_t^*}|\hat{y}_t-y_t|+\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^t,y^t)\right)\right\}$$

with the relaxation

$$\mathbf{Rel}_T\left(\mathcal{F}\middle|(x^t,y^t)\right)=\frac{1}{\lambda}\log\left(\sum_{\sigma\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma)),T-t)\exp\{-\lambda L_t(\sigma)\}\right)+\frac{\lambda}{2}(T-t)$$

The maximum can be written explicitly, as in Section 3:

$$\begin{aligned}
\max\Bigg\{&1-q_t^*+\frac{1}{\lambda}\log\left(\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1-\sigma_t)\}\right),\\
&1+q_t^*+\frac{1}{\lambda}\log\left(\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1+\sigma_t)\}\right)\Bigg\}
\end{aligned}$$

where we have dropped the $\frac{\lambda}{2}(T-t)$ term from both sides. Equating the two values, we obtain

$$2q_t^* = \frac{1}{\lambda}\log\frac{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1-\sigma_t)\}}{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1+\sigma_t)\}}$$

The resulting value becomes

$$1+\frac{\lambda}{2}(T-t)+\frac{1}{2\lambda}\log\left\{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1-\sigma_t)\}\right\}$$

$$+\frac{1}{2\lambda}\log\left\{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1+\sigma_t)\}\right\}$$

$$=1+\frac{\lambda}{2}(T-t)+\frac{1}{\lambda}\mathbb{E}_\epsilon\log\left\{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\exp\{-\lambda(1-\epsilon\sigma_t)\}\right\}$$

$$\le 1+\frac{\lambda}{2}(T-t)+\frac{1}{\lambda}\log\left\{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\mathbb{E}_\epsilon\exp\{-\lambda(1-\epsilon\sigma_t)\}\right\}$$

for a Rademacher random variable $\epsilon\in\{\pm1\}$. Now,

$$\mathbb{E}_\epsilon\exp\{-\lambda(1-\epsilon\sigma_t)\} = e^{-\lambda}\mathbb{E}_\epsilon e^{\lambda\epsilon\sigma_t} \le e^{-\lambda}e^{\lambda^2/2}$$

Substituting this into the above expression, we obtain an upper bound of

$$\frac{\lambda}{2}(T-t+1)+\frac{1}{\lambda}\log\left\{\sum_{(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}_t(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right\}$$

which completes the proof of admissibility using the same combinatorial argument as in the earlier part of the proof.

Arriving at the Relaxation Finally, we show that the relaxation we use arises naturally as an upper bound on the sequential Rademacher complexity. Fix a tree $\mathbf{x}$. Let $\sigma\in\{\pm1\}^{t-1}$ be a sequence of signs. Observe that given history $x^t=(x_1,\ldots,x_t)$, the signs $\epsilon\in\{\pm1\}^{T-t}$, and a tree $\mathbf{x}$, the function class $\mathcal{F}$ takes on only a finite number of possible values (σ,σ_t,ω) on $(x^t,\mathbf{x}(\epsilon))$. Here, $\mathbf{x}(\epsilon)$ denotes the sequences of values along the path ϵ. We have,

$$\sup_{\mathbf{x}}\mathbb{E}_\epsilon\sup_{f\in\mathcal{F}}\left\{2\sum_{i=1}^{T-t}\epsilon_i f(\mathbf{x}_i(\epsilon))-\sum_{i=1}^{t}|f(x_i)-y_i|\right\}$$

$$=\sup_{\mathbf{x}}\mathbb{E}_\epsilon\max_{\sigma_t\in\{\pm1\}}\max_{(\sigma,\omega):(\sigma,\sigma_t,\omega)\in\mathcal{F}|_{(x^t,\mathbf{x}(\epsilon))}}\left\{2\sum_{i=1}^{T-t}\epsilon_i\omega_i-\sum_{i=1}^{t}|\sigma_i-y_i|\right\}$$

$$\le\sup_{\mathbf{x}}\mathbb{E}_\epsilon\max_{\sigma_t\in\{\pm1\}}\max_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}\max_{\mathbf{v}\in V(\mathcal{F}(\sigma,\sigma_t),\mathbf{x})}\left\{2\sum_{i=1}^{T-t}\epsilon_i\mathbf{v}_i(\epsilon)-\sum_{i=1}^{t}|\sigma_i-y_i|\right\}$$

where $\mathcal{F}|_{(x^t,\mathbf{x}(\epsilon))}$ is the projection of $\mathcal{F}$ onto $(x^t,\mathbf{x}(\epsilon))$, $\mathcal{F}(\sigma,\sigma_t)=\{f\in\mathcal{F}: f(x^t)=(\sigma,\sigma_t)\}$, and $V(\mathcal{F}(\sigma,\sigma_t),\mathbf{x})$ is the zero-cover of the set $\mathcal{F}(\sigma,\sigma_t)$ on the tree $\mathbf{x}$. We then have the following relaxation:

$$\frac{1}{\lambda}\log\left(\sup_{\mathbf{x}}\mathbb{E}_\epsilon\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}\sum_{\mathbf{v}\in V(\mathcal{F}(\sigma,\sigma_t),\mathbf{x})}\exp\left\{2\lambda\sum_{i=1}^{T-t}\epsilon_i\mathbf{v}_i(\epsilon)-\lambda L_t(\sigma,\sigma_t)\right\}\right)$$

where $L_t(\sigma,\sigma_t)=\sum_{i=1}^t|\sigma_i-y_i|$. The latter quantity can be factorized:

$$\frac{1}{\lambda}\log\left(\sup_{\mathbf{x}}\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}\exp\{-\lambda L_t(\sigma,\sigma_t)\}\mathbb{E}_\epsilon\sum_{\mathbf{v}\in V(\mathcal{F}(\sigma,\sigma_t),\mathbf{x})}\exp\left\{2\lambda\sum_{i=1}^{T-t}\epsilon_i\mathbf{v}_i(\epsilon)\right\}\right)$$

$$\le\frac{1}{\lambda}\log\left(\sup_{\mathbf{x}}\sum_{\sigma_t\in\{\pm1\}}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}}\exp\{-\lambda L_t(\sigma,\sigma_t)\}\,\mathrm{card}(V(\mathcal{F}(\sigma,\sigma_t),\mathbf{x}))\exp\left\{2\lambda^2(T-t)\right\}\right)$$

$$\le\frac{1}{\lambda}\log\left(\sum_{\sigma_t\in\{\pm1\}}\exp\{-\lambda|\sigma_t-y_t|\}\sum_{\sigma:(\sigma,\sigma_t)\in\mathcal{F}|_{x^t}} g(\mathrm{Ldim}(\mathcal{F}(\sigma,\sigma_t)),T-t)\exp\{-\lambda L_{t-1}(\sigma)\}\right)+2\lambda(T-t)\ .$$

This concludes the derivation of the relaxation.

□

Proof of Lemma 6. We first exhibit the proof for the convex loss case. To show admissibility using the particular randomized strategy q_t given in the lemma, we need to show that

$$\sup_{x_t}\left\{\mathbb{E}_{f\sim q_t}\left[\ell(f,x_t)\right]+\mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_t\right)\right\}\le \mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_{t-1}\right)$$

The strategy q_t proposed by the lemma is such that we first draw $x_{t+1},\ldots,x_T\sim D$ and $\epsilon_{t+1},\ldots\epsilon_T$ Rademacher random variables, and then based on this sample pick $f_t=f_t(x_{t+1:T},\epsilon_{t+1:T})$ as in (17). Hence,

$$\begin{aligned}&\sup_{x_t}\left\{\mathbb{E}_{f\sim q_t}\left[\ell(f,x_t)\right]+\mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_t\right)\right\}\\&=\sup_{x_t}\left\{\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\ell(f_t,x)+\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}\\&\le\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\sup_{x_t}\left\{\ell(f_t,x)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}\end{aligned}$$

where $L_t(f)=\sum_{i=1}^{t}\ell(f,x_i)$. Observe that our strategy "matched the randomness" arising from the relaxation! Now, with f_t defined as

$$f_t=\underset{g\in\mathcal{F}}{\operatorname{argmin}}\sup_{x_t\in\mathcal{X}}\left\{\ell(g,x_t)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}$$

for any given $x_{t+1:T},\epsilon_{t+1:T}$, we have

$$\sup_{x_t}\left\{\ell(f_t,x_t)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}=\inf_{g\in\mathcal{F}}\sup_{x_t}\left\{\ell(g,x_t)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}$$

We can conclude that for this choice of q_t,

$$\begin{aligned}&\sup_{x_t}\left\{\mathbb{E}_{f\sim q_t}\left[\ell(f,x_t)\right]+\mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_t\right)\right\}\le\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\inf_{g\in\mathcal{F}}\sup_{x_t}\left\{\ell(g,x_t)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}\\&=\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\inf_{g\in\mathcal{F}}\sup_{p_t\in\Delta(\mathcal{X})}\mathbb{E}_{x_t\sim p_t}\left[\ell(g,x_t)+\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right]\\&=\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\sup_{p\in\Delta(\mathcal{X})}\inf_{g\in\mathcal{F}}\left\{\mathbb{E}_{x_t\sim p}\left[\ell(g,x_t)\right]+\mathbb{E}_{x_t\sim p}\left[\sup_{f\in\mathcal{F}}C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_t(f)\right]\right\}\end{aligned}$$

In the last step we appealed to the minimax theorem which holds as loss is convex in g and $\mathcal{F}$ is a compact convex set and the term in the expectation is linear in p_t, as it is an expectation. The last expression can be written as

$$\begin{aligned}&\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\sup_{p\in\Delta(\mathcal{X})}\mathbb{E}_{x_t\sim p}\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_{t-1}(f)+\inf_{g\in\mathcal{F}}\mathbb{E}_{x_t\sim p}\left[\ell(g,x_t)\right]-\ell(f,x_t)\right]\\&\le\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\sup_{p\in\Delta(\mathcal{X})}\mathbb{E}_{x_t\sim p}\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_{t-1}(f)+\mathbb{E}_{x_t\sim p}\left[\ell(f,x_t)\right]-\ell(f,x_t)\right]\\&\le\mathbb{E}_{\substack{\epsilon_{t+1:T}\\x_{t+1:T}}}\mathbb{E}_{x_t\sim D}\mathbb{E}_{\epsilon_t}\sup_{f\in\mathcal{F}}\left[C\sum_{i=t+1}^{T}\epsilon_i\ell(f,x_i)-L_{t-1}(f)+C\epsilon_t\ell(f,x_t)\right]\\&=\mathbf{Rel}_T\left(\mathcal{F}|x_1,\ldots,x_{t-1}\right)\end{aligned}$$

Last inequality is by Assumption 1, using which we can replace a draw from supremum over distributions by a draw from the "equivalently bad" fixed distribution D by suffering an extra factor of C multiplied to that random instance.

The key step where we needed convexity was to use minimax theorem to swap infimum and supremum inside the expectation. In general the minimax theorem need not hold. In the non-convex scenario this is the reason we add the extra randomization through $\hat{q}_t$. The non-convex case has a similar proof except that we have expectation w.r.t. $\hat{q}_t$ extra on each round which essentially convexifies our loss and thus allows us to appeal to the minimax theorem. □

Proof of Lemma 7. Let $w \in \mathbb{R}^N$ be arbitrary. We need to show

$$\max_{x\in\{\pm1\}^N} \mathbb{E}_\epsilon \max_{i\in[N]} |w_i + 2\epsilon x_i| \le \mathop{\mathbb{E}}_{x\sim D} \mathop{\mathbb{E}}_{\epsilon} \max_{i\in[N]} |w_i + C\epsilon x_i| \tag{34}$$

Let $i^* = \operatorname*{argmax}_i |w_i|$ and $j^* = \operatorname*{argmax}_{i\neq i^*} |w_i|$ be the coordinates with largest and second-largest magnitude. If $|w_{i^*}| - |w_{j^*}| \ge 4$, the statement is immediate as the top coordinate stays at the top. It remains to consider the case when $|w_{i^*}| - |w_{j^*}| < 4$. In this case first note that,

$$\max_{x\in\{\pm1\}^N} \mathbb{E}_\epsilon \max_{i\in[N]} |w_i + 2\epsilon x_i| \le |w_{i^*}| + 2$$

On the other hand, since the distribution we consider is symmetric, with probability $1/2$ its sign is negative and with remaining probability positive. Define $\sigma_{i^*} = \text{sign}(x_{i^*})$, $\sigma_{j^*} = \text{sign}(x_{j^*})$, $\tau_{i^*} = \text{sign}(w_{i^*})$, and $\tau_{j^*} = \text{sign}(w_{j^*})$. Since each coordinate is drawn i.i.d., using conditional expectations we have,

$$\begin{aligned}
&\mathbb{E}_{x,\epsilon} \max_i |w_i + C\epsilon x_i| = \mathbb{E}_x \max_i |w_i + Cx_i| \\
&\ge \frac{\mathbb{E}_x\left[|w_{i^*} + Cx_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{2} + \frac{\mathbb{E}_x\left[|w_{j^*} + Cx_{j^*}| \mid \sigma_{i^*} \neq \tau_{i^*}, \sigma_{j^*} = \tau_{j^*}\right]}{4} + \frac{\mathbb{E}\left[|w_{i^*} + Cx_{i^*}| \mid \sigma_{i^*} \neq \tau_{i^*}, \sigma_{j^*} \neq \tau_{j^*}\right]}{4} \\
&\ge \frac{\mathbb{E}_x\left[|w_{i^*}| + C|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{2} + \frac{\mathbb{E}_x\left[|w_{j^*}| + C|x_{j^*}| \mid \sigma_{i^*} \neq \tau_{i^*}, \sigma_{j^*} = \tau_{j^*}\right]}{4} + \frac{\mathbb{E}\left[|w_{i^*}| - C|x_{i^*}| \mid \sigma_{i^*} \neq \tau_{i^*}, \sigma_{j^*} \neq \tau_{j^*}\right]}{4} \\
&= \frac{\mathbb{E}\left[|w_{i^*}| + C|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{2} + \frac{\mathbb{E}\left[|w_{j^*}| + C|x_{j^*}| \mid \sigma_{j^*} = \tau_{j^*}\right]}{4} + \frac{\mathbb{E}\left[|w_{i^*}| - C|x_{i^*}| \mid \sigma_{i^*} \neq \tau_{i^*}\right]}{4} \\
&= \frac{|w_{i^*}| + C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{2} + \frac{|w_{j^*}| + C\mathbb{E}\left[|x_{j^*}| \mid \sigma_{j^*} = \tau_{j^*}\right]}{4} + \frac{|w_{i^*}| - C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} \neq \tau_{i^*}\right]}{4} \\
&= \frac{2|w_{i^*}| + |w_{j^*}| + 3C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{4} + \frac{|w_{i^*}| - C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} \neq \tau_{i^*}\right]}{4} \\
&= \frac{3|w_{i^*}| + |w_{j^*}| + 2C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{4}
\end{aligned}$$

Now since we are in the case when $|w_{i^*}| - |w_{j^*}| < 4$ we see that

$$\mathbb{E}_{x,\epsilon} \max_i |w_i + C\epsilon x_i| \ge \frac{3|w_{i^*}| + |w_{j^*}| + 2C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right]}{4} \ge \frac{4|w_{i^*}| + 2C\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right] - 4}{4}$$

On the other hand, as we already argued,

$$\max_{x\in\{\pm1\}^N} \mathbb{E}_\epsilon \max_{i\in[N]} |w_i + 2\epsilon x_i| \le |w_{i^*}| + 2$$

Hence, as long as

$$\frac{C\,\mathbb{E}\left[|x_{i^*}| \mid \sigma_{i^*} = \tau_{i^*}\right] - 2}{2} \ge 2$$

or, in other words, as long as

$$C \ge 6/\mathbb{E}\left[|x_i| \mid \text{sign}(x_i) = \text{sign}(w_i)\right] = 6/\mathop{\mathbb{E}}_{x}\left[|x|\right] \ ,$$

we have that

$$\max_{x\in\{\pm1\}^N} \mathbb{E}_\epsilon \max_{i\in[N]} |w_i + 2\epsilon x_i| \le \mathbb{E}_{x,\epsilon} \max_i |w_i + C\epsilon x_i| \ .$$

This concludes the proof. □

Lemma 13. *Consider the case when $\mathcal{X}$ is the ℓ_∞^N ball and $\mathcal{F}$ is the ℓ_1^N unit ball. Let $f^* = \underset{f\in\mathcal{F}}{\operatorname{argmin}} \langle f, R\rangle$, then for any random vector R,*

$$\mathbb{E}_R\left[\sup_{x\in\mathcal{X}}\{\langle f^*, x\rangle + \|R+x\|_\infty\}\right] \le \mathbb{E}_R\left[\inf_{f\in\mathcal{F}}\sup_x\{\langle f, x\rangle + \|R+x\|_\infty\}\right] + 4\,\mathbf{P}\left(\|R\|_\infty \le 4\right)$$

Proof. Let $f^* = \underset{f\in\mathcal{F}}{\operatorname{argmin}} \langle f, R\rangle$. We start by noting that for any $f' \in \mathcal{F}$,

$$\begin{aligned}
\sup_{x\in\mathcal{X}}\{\langle f', x\rangle + \|R+x\|_\infty\} &= \sup_{x\in\mathcal{X}}\left\{\langle f', x\rangle + \sup_{f\in\mathcal{F}}\langle f, R+x\rangle\right\} \\
&= \sup_{f\in\mathcal{F}}\sup_{x\in\mathcal{X}}\{\langle f', x\rangle + \langle f, R+x\rangle\} \\
&= \sup_{f\in\mathcal{F}}\left\{\sup_{x\in\mathcal{X}}\langle f'+f, x\rangle + \langle f, R\rangle\right\} \\
&= \sup_{f\in\mathcal{F}}\{\|f'+f\|_1 + \langle f, R\rangle\}
\end{aligned}$$

Hence note that

$$\inf_{f'\in\mathcal{F}}\sup_{x\in\mathcal{X}}\{\langle f', x\rangle + \|R+x\|_\infty\} = \inf_{f'\in\mathcal{F}}\sup_{f\in\mathcal{F}}\{\|f'+f\|_1 + \langle f, R\rangle\} \tag{35}$$

$$\ge \inf_{f'\in\mathcal{F}}\{\|f'-f^*\|_1 - \langle f^*, R\rangle\} \ge \inf_{f'\in\mathcal{F}}\{\|f'-f^*\|_1 + \|R\|_\infty\} = \|R\|_\infty \tag{36}$$

On the other hand note that, f^* is the vertex of the ℓ_1 ball (any one which given by $\underset{i\in[d]}{\operatorname{argmin}} |R[i]|$ with sign opposite as sign of $R[i]$ on that vertex). Since the ℓ_1 ball is the convex hull of the $2d$ vertices, any vector $f \in \mathcal{F}$ can be written as $f = \alpha h - \beta f^*$ some $h \in \mathcal{F}$ such that $\|h\|_1 = 1$ and $\langle h, R\rangle = 0$ (which means that h is 0 on the maximal co-ordinate of R specified by f^*) and for some $\beta \in [-1,1]$, $\alpha \in [0,1]$ s.t. $\|\alpha h - \beta f^*\|_1 \le 1$. Further note that the constraint on α, β imposed by requiring that $\|\alpha h - \beta f^*\|_1 \le 1$ can be written as $\alpha + |\beta| \le 1$. Hence,

$$\begin{aligned}
\sup_{x\in\mathcal{X}}\{\langle f^*, x\rangle + \|R+x\|_\infty\} &= \sup_{f\in\mathcal{F}}\{\|f^*+f\|_1 + \langle f, R\rangle\} \\
&= \sup_{\alpha\in[0,1]}\ \sup_{h\perp f^*, \|h\|_1=1}\ \sup_{\beta\in[-1,1], \|\alpha h-\beta f^*\|_1\le 1}\{\|(1-\beta)f^* + \alpha h\|_1 + \beta\langle f^*, R\rangle + \alpha\langle h, R\rangle\} \\
&= \sup_{\alpha\in[0,1]}\ \sup_{h\perp f^*, \|h\|_1=1}\ \sup_{\beta\in[-1,1], \|\alpha h-\beta f^*\|_1\le 1}\{|1-\beta|\,\|f^*\|_1 + \alpha\|h\|_1 + \beta\|R\|_\infty\} \\
&= \sup_{\alpha\in[0,1]}\ \sup_{\beta\in[-1,1]:|\beta|+\alpha\le 1}\{|1-\beta| + \alpha + \beta\|R\|_\infty\} \\
&\le \sup_{\beta\in[-1,1]}\{|1-\beta| + 1 - |\beta| + \beta\|R\|_\infty\} \\
&\le \sup_{\beta\in[-1,1]}\{2|1-\beta| + \beta\|R\|_\infty\} \\
&= \sup_{\beta\in\{-1,1\}}\{2|1-\beta| + \beta\|R\|_\infty\} \\
&= \max\{\|R\|_\infty, 4 - \|R\|_\infty\} \\
&\le \|R\|_\infty + 4\,\mathbf{1}\{\|R\|_\infty \le 4\}
\end{aligned}$$

Hence combining with equation 35 we can conclude that

$$\begin{aligned}
\mathbb{E}_R\left[\sup_x\{\langle f^*, x\rangle + \|R+x\|_\infty\}\right] &\le \mathbb{E}_R\left[\inf_{f\in\mathcal{F}}\sup_x\{\langle f, x\rangle + \|R+x\|_\infty\}\right] + 4\ \mathbb{E}_R\left[\mathbf{1}\{\|R\|_\infty \le 4\}\right] \\
&= \mathbb{E}_R\left[\inf_{f\in\mathcal{F}}\sup_x\{\langle f, x\rangle + \|R+x\|_\infty\}\right] + 4\,\mathbf{P}\left(\|R\|_\infty \le 4\right)
\end{aligned}$$

□

Proof of Lemma 8. On any round t, the algorithm draws $\epsilon_{t+1}, \ldots, \epsilon_T$ and $x_{t+1}, \ldots, x_T \sim D^N$ and plays

$$f_t = \operatorname*{argmin}_{f \in \mathcal{F}} \left\langle f, \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i \right\rangle$$

We shall show that this randomized algorithm is (almost) admissible w.r.t. the relaxation (with some small additional term at each step). We define the relaxation as

$$\mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) = \mathop{\mathbb{E}}_{x_{t+1},\ldots x_T \sim D} \left[\left\| \sum_{i=1}^{t} x_i - C \sum_{i=t+1}^{T} x_i \right\|_\infty \right]$$

Proceeding just as in the proof of Lemma 6 note that, for our randomized strategy,

$$\begin{aligned}
&\sup_x \left\{ \mathop{\mathbb{E}}_{f \sim q_t} [\langle f, x \rangle] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \right\} \\
&= \sup_x \left\{ \mathop{\mathbb{E}}_{x_{t+1:T} \sim D^N} [\langle f_t, x \rangle] + \mathop{\mathbb{E}}_{x_{t+1:T} \sim D^N} \left[\left\| \sum_{i=1}^{t-1} x_i + x - C \sum_{i=t+1}^{T} x_i \right\|_\infty \right] \right\} \\
&\le \mathop{\mathbb{E}}_{x_{t+1:T} \sim D^N} \left[\sup_x \left\{ \langle f_t, x \rangle + \left\| \sum_{i=1}^{t-1} x_i + x - C \sum_{i=t+1}^{T} x_i \right\|_\infty \right\} \right] \qquad (37)
\end{aligned}$$

In view of Lemma 13 (with $R = \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} \epsilon_i x_i$) we conclude that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\sup_{x \in \mathcal{X}} \left\{ \langle f_t, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] \\
&\le \mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\inf_{f \in \mathcal{F}} \sup_x \left\{ \langle f, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] + 4\,\mathbf{P}\left(\left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i \right\|_\infty \le 4 \right) \\
&= \mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\sup_x \left\{ \langle f_t^*, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] + 4\,\mathbf{P}\left(\left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i \right\|_\infty \le 4 \right)
\end{aligned}$$

where

$$f_t^* = \operatorname*{argmin}_{f \in \mathcal{F}} \sup_x \left\{ \langle f, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\}$$

Combining with Equation (37) we conclude that

$$\begin{aligned}
&\sup_x \left\{ \mathop{\mathbb{E}}_{f \sim q_t} [\langle f, x \rangle] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \right\} \\
&\le \mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\sup_x \left\{ \langle f_t^*, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] + 4\,\mathbf{P}\left(\left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i \right\|_\infty \le 4 \right)
\end{aligned}$$

Now, since

$$4\,\mathbf{P}\left(\left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i \right\|_\infty \le 4 \right) \le 4\,\mathbf{P}\left(C \left\| \sum_{i=t+1}^{T} x_i \right\|_\infty \le 4 \right) \le 4\,\mathbf{P}_{y_{t+1},\ldots,y_T \sim D}\left(C \left| \sum_{i=t+1}^{T} y_i \right| \le 4 \right)$$

we have

$$\begin{aligned}
&\sup_x \left\{ \mathop{\mathbb{E}}_{f \sim q_t} [\langle f, x \rangle] + \mathbf{Rel}_T(\mathcal{F}|x_1, \ldots, x_t) \right\} \qquad (38) \\
&\le \mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\sup_x \left\{ \langle f_t^*, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] + 4\,\mathbf{P}_{y_{t+1},\ldots,y_T \sim D}\left(C \left| \sum_{i=t+1}^{T} y_i \right| \le 4 \right) \qquad (39)
\end{aligned}$$

In view of Lemma 7, Assumption 2 is satisfied by D^N with constant C. Further in the proof of Lemma 6 we already showed that whenever Assumption 2 is satisfied, the randomized strategy specified by f_t^* is admissible. More specifically we showed that

$$\mathop{\mathbb{E}}_{x_{t+1},\ldots,x_T} \left[\sup_x \left\{ \langle f_t^*, x \rangle + \left\| \sum_{i=1}^{t-1} x_i - C \sum_{i=t+1}^{T} x_i + x \right\|_\infty \right\} \right] \le \mathbf{Rel}_T(F|x_1, \ldots, x_{t-1})$$

and so using this in Equation (38) we conclude that for the randomized strategy in the statement of the lemma,

$$\sup_x \left\{ \mathop{\mathbb{E}}_{f\sim q_t} [\langle f, x\rangle] + \mathbf{Rel}_T (\mathcal{F}|x_1, \ldots, x_t) \right\}$$
$$\leq \mathbf{Rel}_T (F|x_1, \ldots, x_{t-1}) + 4\,\mathbf{P}_{y_{t+1},\ldots,y_T\sim D} \left(C \left| \sum_{i=t+1}^{T} y_i \right| \leq 4 \right)$$

Or in other words the randomized strategy proposed is admissible with an additional additive factor of $4\,\mathbf{P}_{y_{t+1},\ldots,y_T\sim D} \left(C \left|\sum_{i=t+1}^{T} y_i\right| \leq 4 \right)$ at each time step t. Hence by Proposition 1 we have that for the randomized algorithm specified in the lemma,

$$\mathbb{E}[\mathbf{Reg}_T] \leq \mathbf{Rel}_T (F) + 4 \sum_{t=1}^{T} \mathbf{P}_{y_{t+1},\ldots,y_T\sim D} \left(C \left| \sum_{i=t+1}^{T} y_i \right| \leq 4 \right)$$
$$= C \mathop{\mathbb{E}}_{x_1,\ldots,x_T\sim D^N} \left[\left\| \sum_{t=1}^{T} x_t \right\|_\infty \right] + 4 \sum_{t=1}^{T} \mathbf{P}_{y_{t+1},\ldots,y_T\sim D} \left(C \left| \sum_{i=t+1}^{T} y_i \right| \leq 4 \right)$$

This concludes the proof. □

Proof of Lemma 9. Instead of using $C = 4\sqrt{2}$ and drawing uniformly from surface of unit sphere we can equivalently think of the constant as being 1 and drawing uniformly from surface of sphere of radius $4\sqrt{2}$. Let $\|\cdot\|$ stand for the Euclidean norm. To prove (19), first observe that

$$\sup_{p\in\Delta(\mathcal{X})} \mathop{\mathbb{E}}_{x_t\sim p} \left\| w + \mathop{\mathbb{E}}_{x\sim p} [x] - x_t \right\| \leq \sup_{x\in\mathcal{X}} \mathop{\mathbb{E}}_{\epsilon} \|w + 2\epsilon x\| \tag{40}$$

for any $w \in B$. Further, using Jensen's inequality

$$\sup_{x\in\mathcal{X}} \mathop{\mathbb{E}}_{\epsilon} \|w + 2\epsilon x\| \leq \sup_{x\in\mathcal{X}} \sqrt{\mathop{\mathbb{E}}_{\epsilon} \|w + 2\epsilon x\|^2} \leq \sup_{x\in\mathcal{X}} \sqrt{\|w\|^2 + \mathop{\mathbb{E}}_{\epsilon} \|2\epsilon x\|^2} = \sqrt{\|w\|^2 + 4}$$

To prove the lemma, it is then enough to show that for $r = 4\sqrt{2}$

$$\mathbb{E}_{x\sim D} \|w + rx\| \geq \sqrt{\|w\|^2 + 4} \tag{41}$$

for any w, where we omitted ϵ since D is symmetric. This fact can be proved with the following geometric argument.

We define quadruplets $(w + z_1, w + z_2, w - z_1, w - z_2)$ of points on the sphere of radius r. Each quadruplets will have the property that

$$\frac{\|w + z_1\| + \|w + z_2\| + \|w - z_1\| + \|w - z_2\|}{4} \geq \sqrt{\|w\|^2 + 4} \tag{42}$$

for any w. We then argue that the uniform distribution can be decomposed into these quadruplets such that each point on the sphere occurs in only one quadruplet (except for a measure zero set when z_1 is aligned with $-w$), thus concluding that (41) holds true.

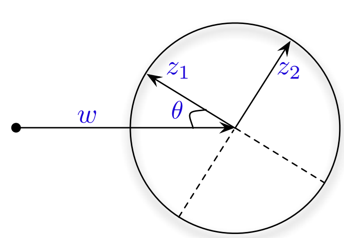

Figure 1: The two-dimensional construction for the proof of Lemma 9.

Pick any direction $w^\perp$ perpendicular to w. A quadruplet is defined by perpendicular vectors z_1 and z_2 which have length r and which lie in the plane spanned by $w, w^\perp$. Let θ be the angle between $-w$ and z_1. Since we are now dealing with a two dimensional plane spanned by w and $w^\perp$, we may as well assume that w is aligned with the positive x-axis, as in Figure 1. We write w for $\|w\|$. The coordinates of the quadruplet are

$$(w - r\cos(\theta), r\sin(\theta)),\ (w + r\cos(\theta), -r\sin(\theta)),\ (w + r\sin(\theta), r\cos(\theta)),\ (w - r\sin(\theta), -r\cos(\theta))$$

For brevity, let $s = \sin(\theta), c = \cos(\theta)$. The desired inequality (42) then reads

$$\sqrt{w^2 - 8wc + r^2} + \sqrt{w^2 + 8wc + r^2} + \sqrt{w^2 + 8ws + r^2} + \sqrt{w^2 - 8ws + r^2} \ge 4\sqrt{w^2 + 4}$$

To prove that this inequality holds, we square both sides, keeping in mind that the terms are non-negative. The sum of four squares on the left hand side gives $4w^2 + 4r^2$. For the six cross terms, we can pass to a lower bound by replacing r^2 in each square root by r^2c^2 or r^2s^2, whichever completes the square. Then observe that

$$|w + rs| \cdot |w - rs| + |w + rc| \cdot |w - rc| = 2w^2 - r^2$$

while the other four cross terms

$$(|w + rs| \cdot |w - rc| + |w + rs| \cdot |w + rc|) + (|w - rs| \cdot |w + rc| + |w - rs| \cdot |w - rc|) \ge |w + rs| \cdot 2w + |w - rs| \cdot 2w \ge 4w^2$$

Doubling the cross terms gives a contribution of $2(6w^2 - r^2)$, while the sum of squares yielded $4w^2 + 4r^2$. The desired inequality is satisfied as long as $16w^2 + 2r^2 \ge 16(w^2 + 4)$, or $r \ge 4\sqrt{2}$.

□

Proof of Lemma 10. By Lemma 9, Assumption 2 is satisfied by distribution D with constant $C = 4\sqrt{2}$. Hence by Lemma 7 we can conclude that for the randomized algorithm which at round t freshly draws $x_{t+1}, \ldots, x_T \sim D$ and picks

$$f_t^* = \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x \in \mathcal{X}} \left\{ \langle f, x \rangle + \left\| -\sum_{i=1}^{t-1} x_i + 4\sqrt{2} \sum_{i=t+1}^{T} x_i - x \right\|_2 \right\}$$

(we dropped the ϵ's as the distribution is symmetric to start with) the expected regret is bounded as

$$\mathbb{E}[\mathbf{Reg}_T] \le 4\sqrt{2} \mathop{\mathbb{E}}_{x_1, \ldots, x_T \sim D} \left[\left\| \sum_{t=1}^{T} x_t \right\|_2 \right] \le 4\sqrt{2T}$$

We claim that the strategy specified in the lemma that chooses

$$f_t = \frac{-\sum_{i=1}^{t-1} x_i + 4\sqrt{2} \sum_{i=t+1}^{T} x_i}{\sqrt{\left\| -\sum_{i=1}^{t-1} x_i + 4\sqrt{2} \sum_{i=t+1}^{T} \epsilon_i x_i \right\|_2^2 + 1}}$$

is the same as choosing f_t^*. To see this let us start by defining

$$\bar{x}_t = -\sum_{i=1}^{t-1} x_i + 4\sqrt{2} \sum_{i=t+1}^{T} x_i$$

Now note that

$$\begin{aligned}
f_t^* &= \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x \in \mathcal{X}} \left\{ \langle f, x \rangle + \left\| -\sum_{i=1}^{t-1} x_i + 4\sqrt{2} \sum_{i=t+1}^{T} x_i - x \right\|_2 \right\} \\
&= \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x \in \mathcal{X}} \left\{ \langle f, x \rangle + \|\bar{x}_t - x\|_2 \right\} \\
&= \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x \in \mathcal{X}} \left\{ \langle f, x \rangle + \sqrt{\|\bar{x}_t - x\|_2^2} \right\} \\
&= \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x : \|x\|_2 \le 1} \left\{ \langle f, x \rangle + \sqrt{\|\bar{x}_t\|^2 - 2\langle \bar{x}_t, x \rangle + \|x\|_2^2} \right\} \\
&= \operatorname*{argmin}_{f \in \mathcal{F}} \sup_{x : \|x\|_2 = 1} \left\{ \langle f, x \rangle + \sqrt{\|\bar{x}_t\|^2 - 2\langle \bar{x}_t, x \rangle + 1} \right\}
\end{aligned}$$

However this argmin calculation is identical to the one in the proof of Proposition 4 (with $C = 1$ and $T - t = 0$) and the solution is given by

$$f_t^* = f_t = \frac{-\sum_{i=1}^{t-1} x_i + 4\sqrt{2}\sum_{i=t+1}^{T} x_i}{\sqrt{\left\|-\sum_{i=1}^{t-1} x_i + 4\sqrt{2}\sum_{i=t+1}^{T} \epsilon_i x_i\right\|_2^2 + 1}}$$

Thus we conclude the proof. □

Proof of Lemma 11. We shall start by showing that the relaxation is admissible for the game where we pick prediction $\hat{y}_t$ and the adversary then directly picks the gradient $\partial\ell(\hat{y}_t, y_t)$. To this end note that

$$\begin{aligned}
&\inf_{\hat{y}_t} \sup_{\partial\ell(\hat{y}_t,y_t)} \left\{\partial\ell(\hat{y}_t, y_t)\cdot \hat{y}_t + \mathbf{Rel}_T\left(\mathcal{F}|\partial\ell(\hat{y}_1,y_1),\ldots,\partial\ell(\hat{y}_t,y_t)\right)\right\} \\
&\qquad = \inf_{\hat{y}_t} \sup_{\partial\ell(\hat{y}_t,y_t)} \left\{\partial\ell(\hat{y}_t, y_t)\cdot \hat{y}_t + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}} 2L\sum_{i=t+1}^{T}\epsilon_i f[t] - \sum_{i=1}^{t}\partial\ell(\hat{y}_i,y_i)\cdot f[i]\right]\right\} \\
&\qquad \le \inf_{\hat{y}_t} \sup_{r_t\in[-L,L]} \left\{r_t\cdot\hat{y}_t + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}} 2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) - r_t\cdot f[t]\right]\right\}
\end{aligned}$$

Let us use the notation $L_{t-1}(f) = \sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)\cdot f[i]$ for the present proof. The supremum over $r_t \in [-L, L]$ is achieved at the endpoints since the expression is convex in r_t. Therefore, the last expression is equal to

$$\begin{aligned}
&\inf_{\hat{y}_t} \sup_{r_t\in\{-L,L\}} \left\{r_t\cdot\hat{y}_t + \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) - r_t\cdot f[t]\right]\right\} \\
&= \inf_{\hat{y}_t} \sup_{p_t\in\Delta(\{-L,L\})} \mathbb{E}_{r_t\sim p_t}\left[r_t\cdot\hat{y}_t + \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) - r_t\cdot f[t]\right]\right] \\
&= \sup_{p_t\in\Delta(\{-L,L\})} \inf_{\hat{y}_t} \mathbb{E}_{r_t\sim p_t}\left[r_t\cdot\hat{y}_t + \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) - r_t\cdot f[t]\right]\right]
\end{aligned}$$

where the last step is due to the minimax theorem. The last quantity is equal to

$$\begin{aligned}
&\sup_{p_t\in\Delta(\{-L,L\})} \mathbb{E}_{\epsilon}\left[\mathbb{E}_{r_t\sim p_t}\left[\inf_{\hat{y}_t}\mathbb{E}_{r_t\sim p_t}[r_t]\cdot\hat{y}_t + \sup_{f\in\mathcal{F}}\left(2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) - r_t\cdot f[t]\right)\right]\right] \\
&\le \sup_{p_t\in\Delta(\{-L,L\})} \mathbb{E}_{\epsilon}\left[\mathbb{E}_{r_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left(2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) + (\mathbb{E}_{r_t\sim p_t}[r_t] - r_t)\cdot f[t]\right)\right]\right] \\
&\le \sup_{p_t\in\Delta(\{-L,L\})} \mathbb{E}_{r_t,r_t'\sim p_t}\left[\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) + (r_t' - r_t)\cdot f[t]\right]\right] \\
&= \sup_{p_t\in\Delta(\{-L,L\})} \mathbb{E}_{r_t,r_t'\sim p_t}\left[\mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) + \epsilon_t(r_t' - r_t)\cdot f[t]\right]\right]
\end{aligned}$$

By passing to the worst-case choice of r_t, r_t' (which is achieved at the endpoints because of convexity), we obtain a further upper bound

$$\begin{aligned}
&\sup_{r_t,r_t'\in\{L,-L\}} \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) + \epsilon_t(r_t' - r_t)\cdot f[t]\right] \\
&\le \sup_{r_t\in\{L,-L\}} \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t+1}^{T}\epsilon_i f[t] - L_{t-1}(f) + 2\epsilon_t r_t\cdot f[t]\right] \\
&= \sup_{r_t\in\{L,-L\}} \mathbb{E}_{\epsilon}\sup_{f\in\mathcal{F}}\left[2L\sum_{i=t}^{T}\epsilon_i f[t] - L_{t-1}(f)\right] \\
&= \mathbf{Rel}_T\left(\mathcal{F}|\partial\ell(\hat{y}_1,y_1),\ldots,\partial\ell(\hat{y}_{t-1},y_{t-1})\right)
\end{aligned}$$

Thus we see that the relaxation is admissible. Now the corresponding prediction is given by

$$\begin{aligned}\hat{y}_t &= \underset{\hat{y}}{\text{argmin}} \sup_{r_t \in [-L,L]} \left\{ r_t \hat{y} + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - \sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)f[i] - r_t f[t]\right\}\right]\right\}\\ &= \underset{\hat{y}}{\text{argmin}} \sup_{r_t \in [-L,L]} \left\{ r_t \hat{y} + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - \sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)f[i] - r_t f[t]\right\}\right]\right\}\\ &= \underset{\hat{y}}{\text{argmin}} \sup_{r_t \in \{-L,L\}} \left\{ r_t \hat{y} + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - \sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)f[i] - r_t f[t]\right\}\right]\right\}\end{aligned}$$

The last step holds because of convexity of the term inside the supremum over r_t is convex in r_t and so the suprema is attained at the endpoints of the interval. The $\hat{y}_t$ above is attained when both terms of the supremum are equalized, that is for $\hat{y}_t$ is the prediction that satisfies :

$$\hat{y}_t = \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}\sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)f[i] + \frac{1}{2}f[t]\right\} - \sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}\sum_{i=1}^{t-1}\partial\ell(\hat{y}_i,y_i)f[i] - \frac{1}{2}f[t]\right\}\right]$$

Finally since the relaxation is admissible we can conclude that the regret of the algorithm is bounded as

$$\mathbf{Reg}_T \le \mathbf{Rel}_T(\mathcal{F}) = 2\,L\,\mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\sum_{t=1}^{T}\epsilon_t f[t]\right].$$

This concludes the proof. □

Proof of Lemma 12. The proof is similar to that of Lemma 11, with a few more twists. We want to establish admissibility of the relaxation given in (21) w.r.t. the randomized strategy q_t we provided. To this end note that

$$\begin{aligned}&\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}[\ell(\hat{y}_t,y_t)] + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_t(f)\right\}\right]\right\}\\ &= \sup_{y_t}\left\{\mathbb{E}_{\epsilon}[\ell(\hat{y}_t(\epsilon),y_t)] + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_t(f)\right\}\right]\right\}\\ &\le \mathbb{E}_{\epsilon}\left[\sup_{y_t}\left\{\ell(\hat{y}_t(\epsilon),y_t) + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_t(f)\right\}\right\}\right]\end{aligned}$$

by Jensen's inequality, with the usual notation $L_t(f) = \sum_{i=1}^{t}\ell(f[i],y_i)$. Further, by convexity of the loss, we may pass to the upper bound

$$\begin{aligned}&\mathbb{E}_{\epsilon}\left[\sup_{y_t}\left\{\partial\ell(\hat{y}_t(\epsilon),y_t)\hat{y}_t(\epsilon) + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - \partial\ell(\hat{y}_t(\epsilon),y_t)f[t]\right\}\right\}\right]\\ &\le \mathbb{E}_{\epsilon}\left[\sup_{y_t}\left\{\mathbb{E}_{r_t}[r_t\cdot\hat{y}_t(\epsilon)] + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - \mathbb{E}_{r_t}[r_t\cdot f[t]]\right\}\right\}\right]\end{aligned}$$

where r_t is a $\{\pm L\}$-valued random variable with the mean $\partial\ell(\hat{y}_t(\epsilon),y_t)$. With the help of Jensen's inequality, and passing to the worst-case r_t (observe that this is legal for any given ϵ), we have an upper bound

$$\begin{aligned}&\mathbb{E}_{\epsilon}\left[\sup_{y_t}\left\{\mathbb{E}_{r_t\sim\partial\ell(\hat{y}_t(\epsilon),y_t)}[r_t\cdot\hat{y}_t(\epsilon)] + \mathbb{E}_{r_t\sim\partial\ell(\hat{y}_t(\epsilon),y_t)}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right]\right\}\right]\\ &\le \mathbb{E}_{\epsilon}\left[\sup_{r_t\in\{\pm L\}}\left\{r_t\cdot\hat{y}_t(\epsilon) + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right\}\right] \qquad (43)\end{aligned}$$

Now the strategy we defined is

$$\hat{y}_t(\epsilon) = \underset{\hat{y}_t}{\text{argmin}} \sup_{r_t\in\{\pm L\}}\left\{r_t\cdot\hat{y}_t(\epsilon) + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - \sum_{i=1}^{t-1}\ell(f[i],y_i) - r_t\cdot f[t]\right\}\right\}$$

which can be re-written as

$$\hat{y}_t(\epsilon) = \left(\sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}L_{t-1}(f) + \frac{1}{2}f[t]\right\} - \sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}L_{t-1}(f) - \frac{1}{2}f[t]\right\}\right)$$

By this choice of $\hat{y}_t(\epsilon)$, plugging back in Equation (43) we see that

$$\begin{aligned}
&\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}\left[\ell(\hat{y}_t,y_t)\right] + \mathbb{E}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_t(f)\right\}\right]\right\}\\
&\quad\le \mathbb{E}_{\epsilon}\left[\sup_{r_t\in\{\pm L\}}\left\{r_t\cdot\hat{y}_t(\epsilon) + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right\}\right]\\
&\quad= \mathbb{E}_{\epsilon}\left[\inf_{\hat{y}_t}\sup_{r_t\in\{\pm L\}}\left\{r_t\cdot\hat{y}_t + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right\}\right]\\
&\quad= \mathbb{E}_{\epsilon}\left[\inf_{\hat{y}_t}\sup_{p_t\in\Delta(\{\pm L\})}\mathbb{E}_{r_t\sim p_t}\left\{r_t\cdot\hat{y}_t + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right\}\right]
\end{aligned}$$

The expression inside the supremum is linear in p_t, as it is an expectation. Also note that the term is convex in $\hat{y}_t$, and the domain $\hat{y}_t \in [-\sup_{f\in\mathcal{F}}|f[t]|, \sup_{f\in\mathcal{F}}|f[t]|]$ is a bounded interval (hence, compact). We conclude that we can use the minimax theorem, yielding

$$\begin{aligned}
&\mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\inf_{\hat{y}_t}\mathbb{E}_{r_t\sim p_t}\left[r_t\cdot\hat{y}_t + \sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right]\right]\\
&= \mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\inf_{\hat{y}_t}\mathbb{E}_{r_t\sim p_t}\left[r_t\cdot\hat{y}_t\right] + \mathbb{E}_{r_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right]\right\}\right]\\
&= \mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\mathbb{E}_{r_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{\inf_{\hat{y}_t}\mathbb{E}_{r_t\sim p_t}\left[r_t\cdot\hat{y}_t\right] + 2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right]\right\}\right]\\
&\le \mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\mathbb{E}_{r_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{\mathbb{E}_{r_t\sim p_t}\left[r_t\cdot f[t]\right] + 2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) - r_t\cdot f[t]\right\}\right]\right\}\right]
\end{aligned}$$

In the last step, we replaced the infimum over $\hat{y}_t$ with $f[t]$, only increasing the quantity. Introducing an i.i.d. copy r'_t of r_t,

$$\begin{aligned}
&= \mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\mathbb{E}_{r_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) + \left(\mathbb{E}_{r_t\sim p_t}[r_t] - r_t\right)\cdot f[t]\right\}\right]\right\}\right]\\
&\le \mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\mathbb{E}_{r_t,r'_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) + (r'_t - r_t)\cdot f[t]\right\}\right]\right\}\right]
\end{aligned}$$

Introducing the random sign ϵ_t and passing to the supremum over r_t, r'_t, yields the upper bound

$$\begin{aligned}
&\mathbb{E}_{\epsilon}\left[\sup_{p_t\in\Delta(\{\pm L\})}\left\{\mathbb{E}_{r_t,r'_t\sim p_t}\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) + (r'_t - r_t)\cdot f[t]\right\}\right]\right\}\right]\\
&\le \mathbb{E}_{\epsilon}\left[\sup_{r_t,r'_t\in\{\pm L\}}\left\{\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) + \epsilon_t\,(r'_t - r_t)\cdot f[t]\right\}\right]\right\}\right]\\
&\le \mathbb{E}_{\epsilon}\left[\sup_{r_t,r'_t\in\{\pm L\}}\left\{\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{L\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2}L_{t-1}(f) + \epsilon_t r'_t\cdot f[t]\right\}\right]\right\}\right]\\
&\qquad + \mathbb{E}_{\epsilon}\left[\sup_{r_t,r'_t\in\{\pm L\}}\left\{\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{L\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2}L_{t-1}(f) - \epsilon_t r_t\cdot f[t]\right\}\right]\right\}\right]
\end{aligned}$$

In the above we split the term in the supremum as the sum of two terms one involving r_t and other r'_t (other terms are equally split by dividing by 2), yielding

$$\mathbb{E}_{\epsilon}\left[\sup_{r_t\in\{\pm L\}}\left\{\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_{t-1}(f) + 2\,\epsilon_t\,r_t\cdot f[t]\right\}\right]\right\}\right]$$

The above step used the fact that the first term only involved r'_t and second only r_t and further ϵ_t and $-\epsilon_t$ have the same distribution. Now finally noting that irrespective of whether r_t in the above supremum is L or $-L$, since it is multiplied by ϵ_t we obtain an upper bound

$$\mathop{\mathbb{E}}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t}^{T}\epsilon_i f[i] - L_{t-1}(f)\right\}\right]$$

We conclude that the relaxation

$$\mathbf{Rel}_T\left(\mathcal{F}|y_1,\ldots,y_t\right) = \mathop{\mathbb{E}}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\left\{2L\sum_{i=t+1}^{T}\epsilon_i f[i] - L_t(f)\right\}\right]$$

is admissible and further the randomized strategy where on each round we first draw ϵ's and then set

$$\begin{aligned}\hat{y}_t(\epsilon) &= \left(\sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}L_{t-1}(f) + \frac{1}{2}f[t]\right\} - \sup_{f\in\mathcal{F}}\left\{\sum_{i=t+1}^{T}\epsilon_i f[i] - \frac{1}{2L}L_{t-1}(f) - \frac{1}{2}f[t]\right\}\right)\\ &= \left(\inf_{f\in\mathcal{F}}\left\{-\sum_{i=t+1}^{T}\epsilon_i f[i] + \frac{1}{2L}L_{t-1}(f) + \frac{1}{2}f[t]\right\} - \inf_{f\in\mathcal{F}}\left\{-\sum_{i=t+1}^{T}\epsilon_i f[i] + \frac{1}{2L}L_{t-1}(f) - \frac{1}{2}f[t]\right\}\right)\end{aligned}$$

is an admissible strategy. Hence, the expected regret under the strategy is bounded as

$$\mathbb{E}\left[\mathbf{Reg}_T\right] \le \mathbf{Rel}_T\left(\mathcal{F}\right) = 2L\ \mathop{\mathbb{E}}_{\epsilon}\left[\sup_{f\in\mathcal{F}}\sum_{i=1}^{T}\epsilon_i f[i]\right]$$

which concludes the proof. □