On Semi-Probabilistic Universal Prediction

Alexander Rakhlin
Department of Statistics
University of Pennsylvania

Karthik Sridharan
Department of Statistics
University of Pennsylvania

***Abstract*—We discuss two scenarios of universal prediction, as well as some recent advances in the study of minimax regret and algorithmic development. We then propose an intermediate scenario, the Semi-Probabilistic Setting, and make progress towards understanding the associated minimax regret.**

I. Introduction

Suppose a learner is observing a sequence $z_1, \dots, z_n, \dots$ from some alphabet $\mathcal{Z}$. At each time instant $t \in \{1, 2, \dots, \}$, the learner is required to make a prediction, denoted by f_t, based on the accumulated knowledge $z^{t-1} \triangleq (z_1, \dots, z_{t-1})$. The quality of this decision is measured by the loss function $\ell(f_t, z_t)$, known to the learner. We are interested in prediction methods that ensure good performance, as measured by the average loss

$$\frac{1}{n}\sum_{t=1}^{n} \ell(f_t, z_t) \ . \tag{1}$$

As written, the problem is uninteresting: it decomposes into n identical rounds, with z^{t-1} providing no information about the next element z_t of the sequence. For the prediction method to be applicable, we need a way of injecting prior knowledge into the problem, thereby making it possible to "learn" from the sequence. This is typically done in two distinct ways, as described, for instance, in [1].

The first approach (termed *The Probabilistic Setting*) is to place assumptions on the data-generating mechanism. Such a prior knowledge comes in the form of a set of possible sources

$$\mathcal{P} = \{P^\theta : \theta \in \Theta\},$$

made available to the learner. The sequence evolves according to conditional probabilities $P^{\theta^*}(z_t|z^{t-1})$, yet the identity of θ^* is not available to the decision-maker. Her task then is to incur small average cost, relative to the best possible performance obtained with the full knowledge of the process:

$$R_{\text{pr}} \triangleq \frac{1}{n}\sum_{t=1}^{n} \ell(f_t, Z_t) - \frac{1}{n}\sum_{t=1}^{n} \inf_{f_t^*} \mathbb{E}\left[\ell(f_t^*, Z_t)|Z^{t-1}\right] \ . \tag{2}$$

Here, the expectation is with respect to the unknown distribution P^{θ^*}, and we replaced z_t with Z_t to indicate the random nature of the sequence. Since (2) is a random variable, its size is measured e.g. in expectation. The choice of f_t^* for each P^θ defines[1] a set of *strategies* $\Pi = \{\pi^\theta : \theta \in \Theta\}$ via

$$\pi_t^\theta(z^{t-1}) = \arg\min_{f_t^*} \mathbb{E}\left[\ell(f_t^*, Z_t) \mid Z^{t-1} = z^{t-1}\right] \tag{3}$$

[1]For the sake of brevity, we are omitting the details pertaining to achievability of the infimum, as well as all the measurability issues.

and we may write (2) as

$$\frac{1}{n}\sum_{t=1}^{n} \ell(f_t, Z_t) - \inf_{\pi\in\Pi} \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[\ell(\pi_t(z^{t-1}), Z_t)|Z^{t-1} = z^{t-1}\right] \ .$$

An approach that is "dual" to the Probabilistic Setting is to directly model the set of strategies Π and to require good performance for *all sequences*. Such an approach, termed *The Deterministic Setting*, places an assumption on "what types of strategies are expected to perform well" rather than on the probabilistic nature of the sequence. Let Π be a set of strategies, each $\pi \in \Pi$ specifying decisions $\pi_t(z^{t-1})$ for all possible prefixes z^{t-1}, and incurring an instantaneous loss of $\ell(\pi_t(z^{t-1}), z_t)$. The learner is tasked with making predictions in such a fashion that keeps *regret*

$$R_{\text{det}} \triangleq \frac{1}{n}\sum_{t=1}^{n} \ell(f_t, z_t) - \inf_{\pi\in\Pi} \frac{1}{n}\sum_{t=1}^{n} \ell(\pi_t(z^{t-1}), z_t) \tag{4}$$

small *for any* sequence $z_1, \dots, z_n$. We may say that the learner "competes" in its performance with the set of strategies Π.

The set Π chosen in the Deterministic Setting may or may not correspond to solutions for some well-defined set of data-generating sources, giving us great flexibility in modeling *solutions* to the problem at hand and relieving us from having to solve (3). On the downside, the goal of ensuring small regret (4) for all sequences is more demanding than that in the Probabilistic Setting. A grand result would establish properties of the set of sources $\mathcal{P}$ such that the performance in the Probabilistic Setting is of the same order of magnitude as that in the Deterministic Setting with the corresponding ("dual") set of solutions Π. For instance, one may consider the set of Markov predictors as solutions to a set of Markov processes; or, one may consider the set of linear predictors as solutions to a Gaussian source (see [1]). Such a comparison of optimal performance can be done by studying the minimax formulations. The *minimax redundancy* in the Probabilistic Setting is

$$\mathcal{V}_n^{\text{pr}}(\mathcal{P}) = \inf_{\mathcal{A}} \sup_{P_\theta\in\mathcal{P}} \mathbb{E}\left\{R_{\text{pr}}\right\} \tag{5}$$

where the infimum is over algorithms $\mathcal{A}$ that produce predictions f_t in the causal manner described above and the expectation is under P_θ. The *minimax regret*, on the other hand, can be defined as

$$\mathcal{V}_n^{\text{det}}(\Pi) = \inf_{\mathcal{A}} \sup_{z^n} \mathbb{E}\left\{R_{\text{det}}\right\}, \tag{6}$$

with expectation over the internal randomization of the prediction method. While the two minimax values are of the same order for some simple "parametric" cases [2], there exist also examples on which the two differ [3]. Therefore, the "duality"

is, at the moment, philosophical rather than quantitative; it is an interesting open question whether the relationship can be made more precise [1].

The gaps between minimax redundancy and minimax regret in the two scenarios can be attributed to the pessimistic nature of the regret definition: the learner might alter the prediction method significantly to go after a particular sequence that is unlikely to appear in practice. The aim of this paper is to formalize a setting in-between Probabilistic and Deterministic settings. By doing so, the hope is to alleviate the pessimistic requirement while still preserving the "distribution-free" quality of the results.

Back to the discussion about the two scenarios, we remark that a similar "duality" has led to the Vapnik-Chervonenkis theory in the realm of Statistical Learning (that is, learning from a batch of i.i.d. data). It is being argued that if prediction is the end goal, one may attempt to construct algorithms that compete in prediction performance with a reference set of decision rules $\mathcal{G} = \{g : \mathcal{X} \mapsto \{0,1\}\}$ (solutions to a classification problem), as opposed to modeling the data-generating distributions. This point of view has led to major developments in the theory and practice of machine learning. The classical results in this area are of a distribution-free nature, and the possibility of the corresponding minimax value being small is completely determined by complexity of the set $\mathcal{G}$ of decision rules with which the learner "competes". It is also recognized within Statistical Learning that the distribution-free approach is overly pessimistic and additional assumptions (such as large margin, Tsybakov noise condition, etc) lead to better results and suggest better algorithms. Yet, the setting augmented with these assumptions is still far away from modeling the data-generating process. In some sense, our attempt in this paper is to propose a similar middle ground between Deterministic and Probabilistic scenarios.

In the past few years, a theory similar to the distribution-free Statistical Learning theory has been developed for the case of universal prediction in the Deterministic Setting with Π being a set of *constant* strategies

$$\left\{\pi^\theta : \theta \in \Theta,\ \pi_t^\theta(z^{t-1}) = \theta \quad \forall z^{t-1}, \forall t\right\} \ .$$

In this case, the complexities of the set of strategies (equivalently, Θ) that dictate the behavior of $\mathcal{V}_n^{\text{det}}(\Pi)$ can be seen as temporal extensions of covering numbers, random averages, combinatorial parameters, and other complexity measures from Statistical Learning [4]. For the case of real-valued prediction with absolute loss, these "sequential" complexities also yield matching lower bounds. In addition to providing non-constructive analysis of $\mathcal{V}_n^{\text{det}}(\Pi)$, sequential complexities give a generic recipe for deriving computationally-feasible algorithms. This is achieved through the notion of a relaxation [5]. Both constructive and non-constructive results have been extended in [6] to the case of non-constant strategies, where the authors developed computationally attractive methods that are not achieved with standard techniques. We are interested in extending these results to the following "Semi-Probabilistic" setting.

II. Summary of Results

A. Minimax Regret

Let $\mathcal{F}$ denote the set of decisions of the learner and $\Delta(\mathcal{F})$ the set of distributions on $\mathcal{F}$. Formally, a (deterministic) strategy is a mapping $\pi_t : \mathcal{Z}^{t-1} \mapsto \mathcal{F}$, and $\pi = (\pi_1, \ldots, \pi_n)$. Fix a set Π of strategies and let us write a more explicit version of the minimax regret in (6) by sequentially writing out the optimal choices of predictions and observations in a sequence of infima, suprema, and expectations:

$$\inf_{q_1} \sup_{z_1} \mathop{\mathbb{E}}_{f_1 \sim q_1} \ \ldots \ \inf_{q_n} \sup_{z_n} \mathop{\mathbb{E}}_{f_n \sim q_n} \left\{ \sum_{t=1}^n \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^n \ell(\pi_t(z^{t-1}), z_t) \right\},$$

where we omitted the normalization factor $1/n$. Here, the infima range over distributions $q_t \in \Delta(\mathcal{F})$ and the notation $\mathbb{E}_{f_t \sim q_t}$ denotes integration with respect to f_t with distribution q_t. This formulation, in fact, allows the sequence $z_1, \ldots, z_n$ to evolve based on the predictions of the learner, a setting known as "non-oblivious adversary" in the online learning literature. Crucially, z_t does not depend on f_t, but only on the mixed strategy q_t, as well as on f^{t-1}.

In the above minimax regret definition, each z_t is chosen in an arbitrary manner, and thus an upper bound on this minimax value ensures existence of a method (that is, a choice of q_t at each step) that enjoy small regret for all sequences $z_1, \ldots, z_n$. As argued in the previous section, we would like to restrict the choices of z_t, incorporating some additional prior knowledge about the process. Unlike the Probabilistic Setting, however, we are not matching the strategy set Π to the (possibly mild) restrictions on the choice of the sequence.

Formally, following [7], we define subsets of allowed distributions for z_t on each round.

Definition 1. A *restriction* $\mathcal{P}_{1:n}$ is a sequence $\mathcal{P}_1, \ldots, \mathcal{P}_n$ of mappings $\mathcal{P}_t : \mathcal{Z}^{t-1} \mapsto 2^{\Delta(\mathcal{Z})}$ such that $\mathcal{P}_t(z^{t-1})$ is a *convex* subset of $\Delta(\mathcal{Z})$ for any $z^{t-1} \in \mathcal{Z}^{t-1}$.

We now define the (unnormalized) minimax regret with respect to the set of strategies Π and the set of restrictions $\mathcal{P}_{1:n}$, denoted henceforth by $\mathcal{V}(\Pi, \mathcal{P}_{1:n})$ or, simply, $\mathcal{V}$:

$$\left\langle\!\!\left\langle \inf_{q_t} \sup_{p_t \in \mathcal{P}_t(z^{t-1})} \mathop{\mathbb{E}}_{f_t \sim q_t} \mathop{\mathbb{E}}_{z_t \sim p_t} \right\rangle\!\!\right\rangle_{t=1}^n \left\{ \sum_{t=1}^n \ell(f_t, z_t) - \inf_{\pi \in \Pi} \sum_{t=1}^n \ell(\pi_t(z^{t-1}), z_t) \right\} .$$

We have used the notation $\langle\!\langle \ldots \rangle\!\rangle_{t=1}^n$ as a shorthand to denote n repeated applications of the operators within the angular bracket pairs. If $\mathcal{P}_t(z^{t-1}) = \Delta(\mathcal{Z})$, the supremum over p_t is achieved at a point mass on some z_t. Thus, $\mathcal{P}_{1:n} = \Delta(\mathcal{Z})_{1:n}$ models the Deterministic Setting with arbitrary sequences.

The goal of universal prediction in the Semi-Probabilistic setting is now formulated as finding methods that choose the mixed strategy q_t in a near-optimal manner with respect to the above minimax regret definition. Clearly, the long sequence of infima and suprema is, in general, impossible to compute in closed form, and it is not clear how "near-optimal" solutions

can be found. The first step to resolving this problem is the following reformulation in the spirit of dynamic programming.

For any z^n, let us recursively define $n+1$ functions

$$\mathcal{V}(z^n) \;\triangleq\; -\inf_{\pi\in\Pi}\sum_{t=1}^{n}\boldsymbol{\ell}(\pi_t(z^{t-1}),z_t)$$

and

$$\mathcal{V}(z^{t-1}) \triangleq \inf_{q_t}\sup_{p_t\in\mathcal{P}_t(z^{t-1})}\mathbb{E}_{f_t\sim q_t}\mathbb{E}_{z_t\sim p_t}\Big\{\boldsymbol{\ell}(f_t,z_t)+\mathcal{V}(z^t)\Big\} \quad (7)$$

for $t \geq 1$. We call the function $\mathcal{V}(z^t)$ *a conditional value* since it can be thought of as minimax regret given the observed sequence z^t. It is not difficult to verify that the final unconditional value $\mathcal{V}(z^0)$ (that is, $\mathcal{V}(\varnothing)$) is precisely the minimax regret $\mathcal{V}(\Pi,\mathcal{P}_{1:n})$, justifying the overloading of the notation. A minimax optimal q_t is obtained by solving (7).

B. Relaxations

While the recursive reformulation does not shed light on a minimax optimal solution, it does give hope for deriving near-optimal computationally feasible methods. The idea is to replace each $\mathcal{V}(z^t)$ with a good upper bound. Formally, a sequence of functions $\mathbf{Rel}$ is an *admissible* relaxation if for any sequence z^n,

$$\mathbf{Rel}(z^n) \geq -\inf_{\pi\in\Pi}\sum_{t=1}^{n}\boldsymbol{\ell}(\pi_t(z^{t-1}),z_t)$$

and

$$\mathbf{Rel}(z^{t-1}) \geq \inf_{q_t}\sup_{p_t\in\mathcal{P}_t(z^{t-1})}\mathbb{E}_{f_t\sim q_t}\mathbb{E}_{z_t\sim p_t}\Big\{\boldsymbol{\ell}(f_t,z_t)+\mathbf{Rel}(z^t)\Big\}$$

Any choice of q_t's, for each t, that ensures the above inequalities is an algorithm with a regret bound of $\mathbf{Rel}(\varnothing)$ (see [5] for a short proof).

The problem of finding low-regret algorithms now boils down to coming up with a relaxation and a choice q_t that ensures admissibility. For instance, for any admissible relaxation $\mathbf{Rel}$ one can obtain an associated low regret algorithm which at time t picks the distribution q_t over set of decisions $\mathcal{F}$ as follows:

$$q_t = \operatorname*{argmin}_{q\in\Delta(\mathcal{F})}\sup_{p_t\in\mathcal{P}_t(z^{t-1})}\mathbb{E}_{f_t\sim q}\mathbb{E}_{z_t\sim p_t}\Big\{\boldsymbol{\ell}(f_t,z_t)+\mathbf{Rel}(z^t)\Big\}$$

For appropriately chosen relaxations, this algorithm can often be written in a closed form, or at least q_t can be obtained by solving a tractable numerical optimization procedure.

What are the candidates for an admissible relaxation $\mathbf{Rel}$? This is where minimax duality comes into play. We may start with the conditional value $\mathcal{V}(z^{t-1})$ written in the extended form with multiple infima and suprema, and, given the appropriate conditions, exchange them. The resulting quantity is often easier to upper bound. This is the idea employed in [5], and it led the authors to recover many known methods, but also to develop novel prediction algorithms with superior computational performance and guarantees on regret. Before proceeding to develop relaxations for the Semi-Probabilistic setting, let us recap this approach for the case of constant strategies in the Deterministic setting.

C. Constant Strategies, Deterministic Setting

Consider the case of constant strategies with decisions in Θ and no restrictions on the sequence z^n. This is the most studied scenario (see [8]), often termed "Online Learning". Most results in this literature are obtained by exhibiting an algorithm and proving a bound on its regret for all sequences. Here, we would like to mention an approach that directly studies complexity of the set Θ by upper and lower-bounding $\mathcal{V}(\Theta)$.

After appealing to minimax duality for the conditional value $\mathcal{V}(z^t)$, one can employ the process of *sequential symmetrization* [4]. Roughly speaking, it allows us to replace the randomness arising from the mixed strategies with a σ-algebra generated by coin flips. In particular, if the symmetrization process is performed for $\mathcal{V}(\Theta)$, we arrive at *sequential Rademacher complexity*, defined for the set of constant strategies as

$$\mathfrak{R}(\Theta) \;\triangleq\; \sup_{\mathbf{z}}\mathbb{E}_\epsilon\sup_{\theta\in\Theta}\left[\sum_{t=1}^{n}\epsilon_t\boldsymbol{\ell}(\theta,\mathbf{z}_t(\epsilon))\right] \quad (8)$$

where the supremum is over $\mathcal{Z}$-valued trees $\mathbf{z}$ of depth n and expectation is over a sequence of n independent Rademacher random variables $\epsilon = (\epsilon_1,\dots,\epsilon_n)$. The definition of a $\mathcal{Z}$-valued tree is as follows: it is a complete binary tree of depth n with nodes labeled by elements of $\mathcal{Z}$. For level t, the labeling function $\mathbf{z}_t : \{\pm1\}^{t-1} \mapsto \mathcal{Z}$ specifies the label given the path from the root, with -1 denoting "left" and $+1$ denoting "right". We write $\mathbf{z}$ for the collection $(\mathbf{z}_1,\dots,\mathbf{z}_n)$ and also $\mathbf{z}_t(\epsilon)$ for $\mathbf{z}_t(\epsilon^{t-1})$. It is proved in [4] that

$$\mathcal{V}(\Theta) \;\leq\; 2\,\mathfrak{R}(\Theta),$$

and the analogous bound for $\mathcal{V}(z^t)$ is twice

$$\mathfrak{R}(z^t) \;\triangleq\; \sup_{\mathbf{z}}\mathbb{E}_\epsilon\sup_{\theta\in\Theta}\left[\sum_{s=t+1}^{n}\epsilon_s\boldsymbol{\ell}(\theta,\mathbf{z}_s(\epsilon))-\sum_{s=1}^{t}\boldsymbol{\ell}(\theta,z_s)\right] \quad (9)$$

This *conditional* version of sequential Rademacher complexity can be shown to be admissible, and provides a starting point for the development of computationally feasible algorithms. In particular, this complexity facilitates the development of randomized prediction methods that roll out randomly into the future, yielding a solid footing for the Smoothed Fictitious Play method from Game Theory and the Follow the Perturbed Leader method studied by the Online Learning community. The Exponential Weights (or, Weighted Majority) algorithm [9], [10] follows from a relaxation that is an immediate upper bound on $\mathfrak{R}(z^t)$, and so do many other methods studied within Online Learning. We refer to [5] for more details, and to [11] for an extension to adaptive data-driven methods.

Example 1 (Binary Prediction with Absolute Loss)**.** Let $\mathcal{Z} = \{-1,1\}$ and $\mathcal{F} = [-1,1]$. The learner observes a sequence of bits and predicts an $f_t \in \mathcal{F}$, where $(f_t+1)/2$ can be interpreted as the probability the next bit is 1. The incurred loss is

$$\boldsymbol{\ell}(f_t,z_t) = |f_t - z_t| = 1 - z_t f_t \;.$$

Let $\Theta = \{-1,1\}$; that is, the learner competes with the set of

two constant strategies (see e.g. Cover [12]). From Eq. (9),

$$\mathfrak{R}(z^t) = \sup_{\mathbf{z}} \mathbb{E}_\epsilon \max_{\theta\in\{-1,1\}} \left[\sum_{s=t+1}^{n} \epsilon_s(1 - \mathbf{z}_s(\epsilon)\theta) - \sum_{s=1}^{t}(1 - z_s\theta) \right]$$
$$= \sup_{\mathbf{z}} \mathbb{E}_\epsilon \left| \sum_{s=1}^{t} z_s - \sum_{s=t+1}^{n} \epsilon_s \mathbf{z}_s(\epsilon) \right| - t = \mathbb{E}_\epsilon \left| \sum_{s=1}^{t} z_s + \sum_{s=t+1}^{n} \epsilon_s \right| - t$$

In other words, the complexity is given by the expected length of a random walk starting from the observed bits z^t. Solving the recursive problem with this relaxation, we get

$$\operatorname*{argmin}_{q_t} \max_{z_t\in\{\pm1\}} \left\{ \mathbb{E}(1 - f_t z_t) + \mathfrak{R}(z^t) \right\}$$
$$= \operatorname*{argmin}_{q_t} \max \left\{ 1 - q_t + \mathfrak{R}((z^{t-1}, +1)), 1 + q_t + \mathfrak{R}((z^{t-1}, -1)) \right\}$$
$$= \frac{1}{2} \left[\mathfrak{R}((z^{t-1}, +1)) - \mathfrak{R}((z^{t-1}, -1)) \right]$$

This recovers a well-known solution [8], [13]. Furthermore, one may simulate the random walk, resulting in an efficient randomized algorithm. A similar argument yields a prediction method for any class of *static experts* with a regret bound of classical Rademacher complexity [14]. However, when the experts are not static, the *sequential* Rademacher generalization is the correct notion.

Further upper bounds on sequential Rademacher complexity can be obtained in terms of sequential covering numbers, combinatorial parameters (a generalization of the Vapnik-Chervonenkis dimension), etc. These complexity measures of the set of decisions (later – strategies) can be used to derive further relaxations and, hence, prediction methods [5].

D. Arbitrary Strategies, Deterministic Setting

The notion of sequential Rademacher complexity has been extended in [6] to the case of non-constant strategies Π:

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

and the value of prediction problem with a set Π of strategies and no restrictions on the sequences (see definition of $\mathcal{V}(\Pi, \mathcal{P}_{1:n})$ on page 2) is shown to be upper bounded as

$$\mathcal{V}_T(\Pi, \Delta(\mathcal{Z})_{1:n}) \;\le\; 2\, \mathfrak{R}(\Pi, \Delta(\mathcal{Z})_{1:n}) \;.$$

Similarly to (9), we obtain a version of conditional sequential Rademacher complexity, and this can be used as a starting point for the development of computationally feasible methods. We refer to [6] for examples of this approach.

E. Arbitrary Strategies, Semi-Stochastic Setting

We first consider the case when restrictions $\mathcal{P}_{1:n}$ on the set of allowed distributions come in the form of an allowed supports of the distributions p_t. For instance, such restrictions can model sequences with a bound $\|z_t - z_{t-1}\| \le \sigma_t$. More precisely, let $\mathcal{Z}_t(z^{t-1}) \subseteq \mathcal{Z}$ be a data-dependent subset, and let $\mathcal{P}_t(z^{t-1})$ be the set of all distributions with support on $\mathcal{Z}_t(z^{t-1})$, that is $\mathcal{P}_t(z^{t-1}) = \Delta(\mathcal{Z}_t(z^{t-1}))$. Then the infimum over $p_t \in \mathcal{P}_t(z^{t-1})$ in the definition of $\mathcal{V}(\Pi, \mathcal{P}_{1:n})$ is attained at some $z_t \in \mathcal{Z}_t(z^{t-1})$. In this setting, the sequences are *deterministically constrained* since each element z_t is required to be chosen from a (history dependent) subset $\mathcal{Z}_t(z^{t-1})$. We may then write minimax regret as

$$\left\langle\!\!\left\langle \inf_{q_t} \sup_{z_t\in\mathcal{Z}_t(z^{t-1})} \mathbb{E}_{f_t\sim q_t} \right\rangle\!\!\right\rangle_{t=1}^{n} \left\{ \sum_{t=1}^{n} \ell(f_t, z_t) - \inf_{\pi\in\Pi} \sum_{t=1}^{n} \ell(\pi_t(z^{t-1}), z_t) \right\}.$$

We may then define the sequential Rademacher complexity $\mathfrak{R}(\Pi, \mathcal{P}_{1:n})$ as

$$\sup_{(\mathbf{w},\mathbf{z})\in\mathcal{C}} \mathbb{E}_\epsilon \sup_{\pi\in\Pi} \left[\sum_{t=1}^{T} \epsilon_t \ell(\pi_t(\mathbf{w}_0(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon)), \mathbf{z}_t(\epsilon)) \right]$$

where $\mathbf{w}$ is a $\mathcal{Z}$-valued tree that starts at level 0 (this modification is the easiest to keep the indices consistent) and has depth $n+1$, while $\mathbf{z}$ is the usual $\mathcal{Z}$-valued tree of depth n. The above supremum ranges over pairs $(\mathbf{w}, \mathbf{z})$ satisfying the constraints $\mathcal{Z}_{1:n}$ as follows:

$$\forall t \in \{1, \ldots, n\}, \quad \mathbf{w}_t(\epsilon), \mathbf{z}_t(\epsilon) \in \mathcal{Z}_t(\mathbf{w}_0(\epsilon), \ldots, \mathbf{w}_{t-1}(\epsilon)).$$

In plain words, sequential Rademacher complexity has a very similar form to that in Deterministic Setting, but the constraints we put on the sequences are reflected in the trees. We can prove the following.

Theorem 1. Let $\mathcal{P}_{1:n}$ be restrictions that take the form of deterministic constraints on the supports of probability distributions p_t. Then the minimax regret is upper bounded as

$$\mathcal{V}(\Pi, \mathcal{P}_{1:n}) \;\le\; 2\, \mathfrak{R}(\Pi, \mathcal{P}_{1:n})$$

for any set of strategies Π.

While the above theorem handles the deterministically constrained scenario, one can also obtain, on similar lines, bounds for the general stochastic restriction case in Definition 1. However the sequential Rademacher complexity associated with the general case is more involved, with a complex stochastic tree generating procedure. For both the deterministically constrained sequences and the case with stochastic restrictions, one can use the corresponding sequential Rademacher complexities as a starting point for deriving relaxations and algorithms.

The goal of this paper was to pose the problem of studying regret in the Semi-Probabilistic model. We have started to explore one direction – through minimax duality and sequential symmetrization – but a full understanding of how the various restrictions and sets of strategies interact is yet to be determined. More specifically, we may ask

- Under which conditions on the "size" of sources $\{P^\theta\}$ is the minimax redundancy of the same order of magnitude as minimax regret with the corresponding set of solutions Π?
- How does the gap between the two change as we introduce additional restrictions in the Semi-Probabilistic setting?
- What are the natural restrictions $\mathcal{P}_{1:n}$ in various applications of sequential prediction methods?
- When can computationally feasible methods be derived in the Semi-Probabilistic model? Can the relaxation approach prove to be as fruitful here as in the Deterministic Setting?

ACKNOWLEDGMENT

We gratefully acknowledge the support of NSF under grants CAREER DMS-0954737 and CCF-1116928, as well as Dean's Research Fund.

REFERENCES

[1] N. Merhav and M. Feder, "Universal prediction," *IEEE Transactions on Information Theory*, vol. 44, pp. 2124–2147, 1998.

[2] J. Rissanen, "Fisher information and stochastic complexity," *Information Theory, IEEE Transactions on*, vol. 42, no. 1, pp. 40–47, 1996.

[3] N. Merhav and M. Feder, "A strong version of the redundancy-capacity theorem of universal coding," *Information Theory, IEEE Transactions on*, vol. 41, no. 3, pp. 714–722, 1995.

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

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

[6] W. Han, A. Rakhlin, and K. Sridharan, "Competing with strategies," in *Journal of Machine Learning, Workshop and Conference Proceedings*, 2013.

[7] A. Rakhlin, K. Sridharan, and A. Tewari, "Online learning: Stochastic, constrained, and smoothed adversaries," in *NIPS*, 2011, pp. 1764–1772.

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

[9] N. Littlestone and M. K. Warmuth, "The weighted majority algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212–261, 1994.

[10] V. Vovk, "Aggregating strategies," in *Proceedings of the Third Annual Workshop on Computational Learning Theory*. Morgan Kaufmann, 1990, pp. 372–383.

[11] A. Rakhlin, O. Shamir, and K. Sridharan, "Localization and adaptation in online learning," in *AISTATS*, 2013.

[12] T. Cover, "Behaviour of sequential predictors of binary sequences," in *Proc. 4th Prague Conf. Inform. Theory, Statistical Decision Functions, Random Processes*, 1965.

[13] N. Cesa-Bianchi and O. Shamir, "Efficient online learning via randomized rounding," in *NIPS*, 2011, pp. 343–351.

[14] N. Cesa-Bianchi and G. Lugosi, "On prediction of individual sequences," *The Annals of Statistics*, vol. 27, no. 6, pp. pp. 1865–1895, 1999.