Journal of Machine Learning Research 6 (2005) 639–660 Submitted 10/04; Revised 3/05; Published 4/05

Adaptive Online Prediction by Following the Perturbed Leader

Marcus Hutter MARCHUS@IDSIA.CH
Jan Poland JAN@IDSIA.CH
IDSIA, Galleria 2
6928 Manno-Lugano, Switzerland

Editor: Manfred Warmuth

Abstract

When applying aggregating strategies to Prediction with Expert Advice (PEA), the learning rate must be adaptively tuned. The natural choice of $\sqrt{\text{complexity/current loss}}$ renders the analysis of Weighted Majority (WM) derivatives quite complicated. In particular, for arbitrary weights there have been no results proven so far. The analysis of the alternative Follow the Perturbed Leader (FPL) algorithm from Kalai and Vempala (2003) based on Hannan's algorithm is easier. We derive loss bounds for adaptive learning rate and both finite expert classes with uniform weights and countable expert classes with arbitrary weights. For the former setup, our loss bounds match the best known results so far, while for the latter our results are new.

Keywords: prediction with expert advice, follow the perturbed leader, general weights, adaptive learning rate, adaptive adversary, hierarchy of experts, expected and high probability bounds, general alphabet and loss, online sequential prediction

1. Introduction

In Prediction with Expert Advice (PEA) one considers an ensemble of sequential predictors (experts). A master algorithm is constructed based on the historical performance of the predictors. The goal of the master algorithm is to perform nearly as well as the best expert in the class, on any sequence of outcomes. This is achieved by making (randomized) predictions close to the better experts.

PEA theory has rapidly developed in the recent past. Starting with the Weighted Majority (WM) algorithm of Littlestone and Warmuth (1989, 1994) and the aggregating strategy of Vovk (1990), a vast variety of different algorithms and variants have been published. A key parameter in all these algorithms is the *learning rate*. While this parameter had to be fixed in the early algorithms such as WM, Cesa-Bianchi et al. (1997) established the so-called doubling trick to make the learning rate coarsely adaptive. A little later, incrementally adaptive algorithms were developed by Auer and Gentile (2000); Auer et al. (2002); Yaroshinsky et al. (2004); Gentile (2003), and others. In Section 10, we will compare our results with these works more in detail. Unfortunately, the loss bound proofs for the incrementally adaptive WM variants are quite complex and technical, despite the typically simple and elegant proofs for a static learning rate.

The complex growing proof techniques also had another consequence. While for the original WM algorithm, assertions are proven for countable classes of experts with arbitrary weights, the modern variants usually restrict to finite classes with uniform weights (an exception being Gentile

(2003); see the discussion section therein.) This might be sufficient for many practical purposes but it prevents the application to more general classes of predictors. Examples are extrapolating (=predicting) data points with the help of a polynomial (=expert) of degree $d = 1, 2, 3, \ldots$ –or– the (from a computational point of view largest) class of all computable predictors. Furthermore, most authors have concentrated on predicting *binary* sequences, often with the 0/1 loss for $\{0, 1\}$-valued and the absolute loss for $[0, 1]$-valued predictions. Arbitrary losses are less common. Nevertheless, it is easy to abstract completely from the predictions and consider the resulting losses only. Instead of predicting according to a "weighted majority" in each time step, one chooses one *single* expert with a probability depending on his past cumulated loss. This is done e.g. by Freund and Schapire (1997), where an elegant WM variant, the Hedge algorithm, is analyzed.

A different, general approach to achieve similar results is Follow the Perturbed Leader (FPL). The principle dates back to as early as 1957, now called Hannan's algorithm (Hannan, 1957). In 2003, Kalai and Vempala published a simpler proof of the main result of Hannan and also succeeded to improve the bound by modifying the distribution of the perturbation. The resulting algorithm (which they call FPL*) has the same performance guarantees as the WM-type algorithms for fixed learning rate, save for a factor of $\sqrt{2}$. A major advantage we will discover in this work is that its analysis remains easy for an adaptive learning rate, in contrast to the WM derivatives. Moreover, it generalizes to online decision problems other than PEA.

In this work,[1] we study the FPL algorithm for PEA. The problems of WM algorithms mentioned above are addressed. Bounds on the cumulative regret of the standard form $\sqrt{kL}$ (where k is the complexity and L is the cumulative loss of the best expert in hindsight) are shown for countable expert classes with arbitrary weights, adaptive learning rate, and arbitrary losses. Regarding the adaptive learning rate, we obtain proofs that are simpler and more elegant than for the corresponding WM algorithms. (In particular, the proof for a self-confident choice of the learning rate, Theorem 7, is less than half a page.) Further, we prove the first loss bounds for *arbitrary weights* and adaptive learning rate. In order to obtain the optimal $\sqrt{kL}$ bound in this case, we will need to introduce a hierarchical version of FPL, while without hierarchy we show a worse bound $k\sqrt{L}$. (For self-confident learning rate together with uniform weights and arbitrary losses, one can prove corresponding results for a variant of WM by adapting an argument by Auer et al. 2002.)

PEA usually refers to an *online worst case* setting: n experts that deliver sequential predictions over a time range $t = 1, \ldots, T$ are given. At each time t, we know the actual predictions and the *past* losses. The goal is to give a prediction such that the overall loss after T steps is "not much worse" than the best expert's loss *on any sequence of outcomes*. If the prediction is deterministic, then an adversary could choose a sequence which provokes maximal loss. So we have to *randomize* our predictions. Consequently, we ask for a prediction strategy such that the *expected* loss on any sequence is small.

This paper is structured as follows. In Section 2 we give the basic definitions. While Kalai and Vempala consider general online decision problems in finite-dimensional spaces, we focus on online prediction tasks based on a countable number of experts. Like Kalai and Vempala (2003) we exploit the infeasible FPL predictor (IFPL) in our analysis. Sections 3 and 4 derive the main analysis tools. In Section 3 we generalize (and marginally improve) the upper bound (Kalai and Vempala, 2003, Lem.3) on IFPL to arbitrary weights. The main difficulty we faced was to appropriately distribute the weights to the various terms. For the corresponding lower bound (Section 7) this

1. A shorter version appeared in the proceedings of the ALT 2004 conference (Hutter and Poland, 2004).

is an open problem. In Section 4 we exploit our restricted setup to significantly improve (Kalai and Vempala, 2003, Eq.(3)) allowing for bounds logarithmic rather than linear in the number of experts. The upper and lower bounds on IFPL are combined to derive various regret bounds on FPL in Section 5. Bounds for static and dynamic learning rate in terms of the sequence length follow straight-forwardly. The proof of our main bound in terms of the loss is much more elegant than the analysis of previous comparable results. Section 6 proposes a novel hierarchical procedure to improve the bounds for non-uniform weights. In Section 7, a lower bound is established. In Section 8, we consider the case of independent randomization more seriously. In particular, we show that the derived bounds also hold for an adaptive adversary. Section 9 treats some additional issues, including bounds with high probability, computational aspects, deterministic predictors, and the absolute loss. Finally, in Section 10 we discuss our results, compare them to references, and state some open problems.

2. Setup and Notation

Setup. Prediction with expert advice proceeds as follows. We are asked to perform sequential predictions $y_t \in \mathcal{Y}$ at times $t = 1,2,\ldots$. At each time step t, we have access to the predictions $(y_t^i)_{1\leq i\leq n}$ of n experts $\{e_1,...,e_n\}$, where the size of the expert pool is $n \in \boldsymbol{N}\cup\{\infty\}$. It is convenient to use the same notation for finite ($n \in \boldsymbol{N}$) and countably infinite ($n = \infty$) expert pool. After having made a prediction, we make some observation $x_t \in \mathcal{X}$, and a Loss is revealed for our and each expert's prediction. (E.g. the loss might be 1 if the expert made an erroneous prediction and 0 otherwise. This is the 0/1 loss.) Our goal is to achieve a total loss "not much worse" than the best expert, after t time steps.

We admit $n \in \boldsymbol{N}\cup\{\infty\}$ experts, each of which is assigned a known complexity $k^i \geq 0$. Usually we require $\sum_i \mathrm{e}^{-k^i} \leq 1$, which implies that the k^i are valid lengths of prefix code words, for instance $k^i = \ln n$ if $n < \infty$ or $k^i = \frac{1}{2} + 2\ln i$ if $n = \infty$. Each complexity defines a weight by means of e^{-k^i} and vice versa. In the following we will talk of complexities rather than of weights. If n is finite, then usually one sets $k^i = \ln n$ for all i; this is the case of *uniform complexities/weights*. If the set of experts is countably infinite ($n = \infty$), uniform complexities are not possible. The vector of all complexities is denoted by $k = (k^i)_{1\leq i\leq n}$. At each time t, each expert i suffers a loss[2] $s_t^i = \mathrm{Loss}(x_t, y_t^i) \in [0,1]$, and $s_t = (s_t^i)_{1\leq i\leq n}$ is the vector of all losses at time t. Let $s_{<t} = s_1 + \ldots + s_{t-1}$ (respectively $s_{1:t} = s_1 + \ldots + s_t$) be the total past loss vector (including current loss s_t) and $s_{1:t}^{min} = \min_i\{s_{1:t}^i\}$ be the loss of the *best expert in hindsight (BEH)*. Usually we do not know in advance the time $t \geq 0$ at which the performance of our predictions are evaluated.

General decision spaces. The setup can be generalized as follows. Let $\mathcal{S} \subset \boldsymbol{R}^n$ be the *state space* and $\mathcal{D} \subset \boldsymbol{R}^n$ the *decision space*. At time t the state is $s_t \in \mathcal{S}$, and a decision $d_t \in \mathcal{D}$ (which is made before the state is revealed) incurs a loss $d_t \circ s_t$, where "$\circ$" denotes the inner product. This implies that the loss function is *linear* in the states. Conversely, each linear loss function can be represented in this way. The decision which minimizes the loss in state $s \in \mathcal{S}$ is

$$M(s) := \arg\min_{d\in\mathcal{D}}\{d \circ s\} \tag{1}$$

if the minimum exists. The application of this general framework to PEA is straightforward: $\mathcal{D}$ is identified with the space of all unit vectors $\mathcal{E} = \{e_i : 1 \leq i \leq n\}$, since a decision consists of selecting

2. The setup, analysis and results easily scale to $s_t^i \in [0,S]$ for $S > 0$ other than 1.

a single expert, and $s_t\in[0,1]^n$, so states are identified with losses. Only Theorems 2 and 10 will be stated in terms of general decision space. Our main focus is $\mathcal{D}=\mathcal{E}$. (Even for this special case, the scalar product notation is not too heavy, but will turn out to be convenient.) All our results generalize to the simplex $\mathcal{D}=\Delta=\{v\in[0,1]^n:\sum_i v^i=1\}$, since the minimum of a linear function on Δ is always attained on $\mathcal{E}$.

Follow the Perturbed Leader. Given $s_{<t}$ at time t, an immediate idea to solve the expert problem is to "Follow the Leader" (FL), i.e. selecting the expert e_i which performed best in the past (minimizes $s^i_{<t}$), that is predict according to expert $M(s_{<t})$. This approach fails for two reasons. First, for $n=\infty$ the minimum in (1) may not exist. Second, for $n=2$ and $s=\binom{0\,1\,0\,1\,0\,1\ldots}{\frac{1}{2}\,0\,1\,0\,1\,0\ldots}$, FL always chooses the wrong prediction (Kalai and Vempala, 2003). We solve the first problem by penalizing each expert by its complexity, i.e. predicting according to expert $M(s_{<t}+k)$. The *FPL (Follow the Perturbed Leader)* approach solves the second problem by adding to each expert's loss $s^i_{<t}$ a random perturbation. We choose this perturbation to be negative *exponentially distributed*, either independent in each time step or once and for all at the very beginning at time $t=0$. The former choice is preferable in order to protect against an adaptive adversary who generates the s_t, and in order to get bounds with high probability (Section 9). For the main analysis however, the latter choice is more convenient. Due to linearity of expectations, these two possibilities are equivalent when dealing with *expected losses* (this is straightforward for oblivious adversary, for adaptive adversary see Section 8), so we can henceforth assume without loss of generality one initial perturbation q.

The FPL algorithm is defined as follows:

Choose random vector $q\overset{d.}{\sim}\exp$, i.e. $P[q^1...q^n]=\mathrm{e}^{-q^1}\cdot\ldots\cdot\mathrm{e}^{-q^n}$ for $q\geq 0$.
For $t=1,...,T$
- Choose learning rate η_t.
- Output prediction of expert i which minimizes $s^i_{<t}+(k^i-q^i)/\eta_t$.
- Receive loss s^i_t for all experts i.

Other than $s_{<t}$, k and q, FPL depends on the *learning rate* η_t. We will give choices for η_t in Section 5, after having established the main tools for the analysis. The expected loss at time t of FPL is $\ell_t:=E[M(s_{<t}+\frac{k-q}{\eta_t})\circ s_t]$. The key idea in the FPL analysis is the use of an intermediate predictor *IFPL* (for *Implicit or Infeasible FPL*). IFPL predicts according to $M(s_{1:t}+\frac{k-q}{\eta_t})$, thus under the knowledge of s_t (which is of course not available in reality). By $r_t:=E[M(s_{1:t}+\frac{k-q}{\eta_t})\circ s_t]$ we denote the expected loss of IFPL at time t. The losses of IFPL will be upper-bounded by BEH in Section 3 and lower-bounded by FPL in Section 4. Note that our definition of the FPL algorithm deviates from that of Kalai and Vempala. It uses an exponentially distributed perturbation similar to their FPL* but one-sided and a non-stationary learning rate like Hannan's algorithm.

Notes. Observe that we have stated the FPL algorithm regardless of the actual *predictions* of the experts and possible *observations*, only the *losses* are relevant. Note also that an expert can implement a highly complicated strategy depending on past outcomes, despite its trivializing identification with a constant unit vector. The complex expert's (and environment's) behavior is summarized and hidden in the state vector $s_t=\text{Loss}(x_t,y^i_t)_{1\leq i\leq n}$. Our results therefore apply to *arbitrary prediction and observation spaces $\mathcal{Y}$ and $\mathcal{X}$ and arbitrary bounded loss functions*. This is in contrast to the major part of PEA work developed for binary alphabet and 0/1 or absolute loss only. Finally note that the setup allows for losses generated by an adversary who tries to maximize the regret of FPL and knows the FPL algorithm and all experts' past predictions/losses. If the adversary also has access

Symbol	**Definition / Explanation**
n	$\in \boldsymbol{N}\cup\{\infty\}$ ($n=\infty$ means countably infinite $\mathcal{E}$). Number of experts.
x^i	= ith component of vector $x\in\boldsymbol{R}^n$.
$\mathcal{E}$	$:=\{e_i:1\leq i\leq n\}=$ set of unit vectors ($e_i^j=\delta_{ij}$).
Δ	$:=\{v\in[0,1]^n:\sum_i v^i=1\}=$ simplex.
$s_t\in[0,1]^n$	= environmental state/loss vector at time t.
$s_{1:t}$	$:=s_1+...+s_t=$ state/loss (similar for ℓ_t and r_t).
$s_{1:T}^{min}$	$=\min_i\{s_{1:T}^i\}=$ loss of Best Expert in Hindsight (BEH).
$s_{<t}$	$:=s_1+...+s_{t-1}=$ state/loss summary ($s_{<0}=0$).
$M(s)$	$:=\text{argmin}_{d\in\mathcal{D}}\{d\circ s\}=$ best decision on s.
$T\in\boldsymbol{N}_0$	= total time=step, $t\in\boldsymbol{N}=$ current time=step.
$k^i\geq 0$	= penalization = complexity of expert i.
$q\in\boldsymbol{R}^n$	= random vector with independent exponentially distributed components.
I_t	$:=\text{argmin}_{i\in\mathcal{E}}\{s_{<t}^i+\frac{k^i-q^i}{\eta_t}\}=$ randomized prediction of FPL.
ℓ_t	$:=E[M(s_{<t}+\frac{k-q}{\eta_t})\circ s_t]=$ expected loss at time t of FPL ($=E[s_t^{I_t}]$ for $\mathcal{D}=\mathcal{E}$).
r_t	$:=E[M(s_{1:t}+\frac{k-q}{\eta_t})\circ s_t]=$ expected loss at time t of IFPL.
u_t	$:=M(s_{<t}+\frac{k-q}{\eta_t})\circ s_t=$ actual loss at time t of FPL ($=s_t^{I_t}$ for $\mathcal{D}=\mathcal{E}$).

Table 1: List of notation.

to FPL's past decisions, then FPL must use independent randomization at each time step in order to achieve good regret bounds. Table 1 summarizes notation.

Motivation of FPL. Let $d(s_{<t})$ be any predictor with decision based on $s_{<t}$. The following identity is easy to show:

$$\underbrace{\sum_{t=1}^T d(s_{<t})\circ s_t}_{\text{“FPL”}} \;\equiv\; \underbrace{d(s_{1:T})\circ s_{1:T}}_{\text{“BEH”}} + \overbrace{\underbrace{\sum_{t=1}^T [d(s_{<t})-d(s_{1:t})]\circ s_{<t}}_{\text{“IFPL}-\text{BEH”}}}^{\leq 0 \text{ if } d\approx M} + \overbrace{\underbrace{\sum_{t=1}^T [d(s_{<t})-d(s_{1:t})]\circ s_t}_{\text{“FPL}-\text{IFPL”}}}^{\text{small if } d(\cdot) \text{ is continuous}}. \tag{2}$$

For a good bound of FPL in terms of BEH we need the first term on the r.h.s. to be close to BEH and the last two terms to be small. The first term is close to BEH if $d\approx M$. The second to last term is even negative if $d=M$, hence small if $d\approx M$. The last term is small if $d(s_{<t})\approx d(s_{1:t})$, which is the case if $d(\cdot)$ is a sufficiently smooth function. Randomization smoothes the discontinuous function M: The function $d(s):=E[M(s-q)]$, where $q\in\boldsymbol{R}^n$ is some random perturbation, is a continuous function in s. If the mean and variance of q are small, then $d\approx M$, if the variance of q is large, then $d(s_{<t})\approx d(s_{1:t})$. An intermediate variance makes the last two terms of (2) simultaneously small enough, leading to excellent bounds for FPL.

3. IFPL bounded by Best Expert in Hindsight

In this section we provide tools for comparing the loss of IFPL to the loss of the best expert in hindsight. The first result bounds the expected error induced by the exponentially distributed perturbation.

Lemma 1 (Maximum of Shifted Exponential Distributions) *Let $q^1,...,q^n$ be (not necessarily independent) exponentially distributed random variables, i.e. $P[q^i]=\mathrm{e}^{-q^i}$ for $q^i\geq 0$ and $1\leq i\leq n\leq\infty$, and $k^i\in\boldsymbol{R}$ be real numbers with $u:=\sum_{i=1}^n \mathrm{e}^{-k^i}$. Then*

$$\begin{aligned} P[\max_i\{q^i-k^i\}\geq a] &= 1-\prod_{i=1}^n \max\{0,1-\mathrm{e}^{-a-k^i}\} \quad \textit{if} \quad q^1,...,q^n \textit{ are independent}, \\ P[\max_i\{q^i-k^i\}\geq a] &\leq \min\{1,u\,\mathrm{e}^{-a}\}, \\ E[\max_i\{q^i-k^i\}] &\leq 1+\ln u. \end{aligned}$$

Proof. Using

$$P[q^i<a]=\max\{0,1-\mathrm{e}^{-a}\}\geq 1-\mathrm{e}^{-a} \quad \text{and} \quad P[q^i\geq a]=\min\{1,\mathrm{e}^{-a}\}\leq \mathrm{e}^{-a},$$

valid for any $a\in\boldsymbol{R}$, the exact expression for $P[\max]$ in Lemma 1 follows from

$$P[\max_i\{q^i-k^i\}<a]=P[q^i-k^i<a\;\forall i]=\prod_{i=1}^n P[q^i<a+k^i]=\prod_{i=1}^n \max\{0,\mathrm{e}^{-a-k^i}\},$$

where the second equality follows from the independence of the q^i. The bound on $P[\max]$ for any $a\in\boldsymbol{R}$ (including negative a) follows from

$$P[\max_i\{q^i-k^i\}\geq a]=P[\exists i:q^i-k^i\geq a]\leq\sum_{i=1}^n P[q^i-k^i\geq a]\leq\sum_{i=1}^n \mathrm{e}^{-a-k^i}=u\cdot\mathrm{e}^{-a}$$

where the first inequality is the union bound. Using $E[z]\leq E[\max\{0,z\}]=\int_0^\infty P[\max\{0,z\}\geq y]dy=\int_0^\infty P[z\geq y]dy$ (valid for any real-valued random variable z) for $z=\max_i\{q^i-k^i\}-\ln u$, this implies

$$E[\max_i\{q^i-k^i\}-\ln u]\leq\int_0^\infty P[\max_i\{q^i-k^i\}\geq y+\ln u]dy\leq\int_0^\infty \mathrm{e}^{-y}dy \;=\; 1,$$

which proves the bound on $E[\max]$. □

If n is finite, a lower bound $E[\max_i q^i]\geq 0.57721+\ln n$ can be derived, showing that the upper bound on $E[\max]$ is quite tight (at least) for $k^i=0\;\forall i$. The following bound generalizes (Kalai and Vempala, 2003, Lem.3) to arbitrary weights, establishing a relation between IFPL and the best expert in hindsight.

Theorem 2 (IFPL bounded by BEH) *Let $\mathcal{D}\subseteq\boldsymbol{R}^n$, $s_t\in\boldsymbol{R}^n$ for $1\leq t\leq T$ (both $\mathcal{D}$ and s may even have negative components, but we assume that all required extrema are attained), and $q,k\in\boldsymbol{R}^n$. If $\eta_t>0$ is decreasing in t, then the loss of the infeasible FPL knowing s_t at time t in advance (l.h.s.) can be bounded in terms of the best predictor in hindsight (first term on r.h.s.) plus additive corrections:*

$$\sum_{t=1}^T M(s_{1:t}+\frac{k-q}{\eta_t})\circ s_t\leq\min_{d\in\mathcal{D}}\{d\circ(s_{1:T}+\frac{k}{\eta_T})\}+\frac{1}{\eta_T}\max_{d\in\mathcal{D}}\{d\circ(q-k)\}-\frac{1}{\eta_T}M(s_{1:T}+\frac{k}{\eta_T})\circ q.$$

Note that if $\mathcal{D}=\mathcal{E}$ (or $\mathcal{D}=\Delta$) and $s_t\geq 0$, then all extrema in the theorem are attained almost surely. The same holds for all subsequent extrema in the proof and throughout the paper.

Proof. For notational convenience, let $\eta_0=\infty$ and $\tilde{s}_{1:t}=s_{1:t}+\frac{k-q}{\eta_t}$. Consider the losses $\tilde{s}_t=s_t+(k-q)(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}})$ for the moment. We first show by induction on T that the infeasible predictor $M(\tilde{s}_{1:t})$ has zero regret for any loss $\tilde{s}$, i.e.

$$\sum_{t=1}^{T} M(\tilde{s}_{1:t})\circ\tilde{s}_t \leq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T}. \tag{3}$$

For $T=1$ this is obvious. For the induction step from $T-1$ to T we need to show

$$M(\tilde{s}_{1:T})\circ\tilde{s}_T \leq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} - M(\tilde{s}_{<T})\circ\tilde{s}_{<T}. \tag{4}$$

This follows from $\tilde{s}_{1:T}=\tilde{s}_{<T}+\tilde{s}_T$ and $M(\tilde{s}_{1:T})\circ\tilde{s}_{<T}\geq M(\tilde{s}_{<T})\circ\tilde{s}_{<T}$ by minimality of M. Rearranging terms in (3), we obtain

$$\sum_{t=1}^{T} M(\tilde{s}_{1:t})\circ s_t \;\leq\; M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} - \sum_{t=1}^{T} M(\tilde{s}_{1:t})\circ(k-q)\Big(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\Big) \tag{5}$$

Moreover, by minimality of M,

$$\begin{aligned} M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} &\leq M\Big(s_{1:T}+\frac{k}{\eta_T}\Big)\circ\Big(s_{1:T}+\frac{k-q}{\eta_T}\Big) \\ &= \min_{d\in\mathcal{D}}\Big\{d\circ(s_{1:T}+\frac{k}{\eta_T})\Big\} - M\Big(s_{1:T}+\frac{k}{\eta_T}\Big)\circ\frac{q}{\eta_T} \end{aligned} \tag{6}$$

holds. Using $\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\geq 0$ and again minimality of M, we have

$$\begin{aligned} \sum_{t=1}^{T}(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}})M(\tilde{s}_{1:t})\circ(q-k) &\leq \sum_{t=1}^{T}(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}})M(k-q)\circ(q-k) \\ &= \frac{1}{\eta_T}M(k-q)\circ(q-k) = \frac{1}{\eta_T}\max_{d\in\mathcal{D}}\{d\circ(q-k)\}. \end{aligned} \tag{7}$$

Inserting (6) and (7) back into (5) we obtain the assertion. □

Assuming q random with $E[q^i]=1$ and taking the expectation in Theorem 2, the last term reduces to $-\frac{1}{\eta_T}\sum_{i=1}^{n}M(s_{1:T}+\frac{k}{\eta_T})^i$. If $\mathcal{D}\geq 0$, the term is negative and may be dropped. In case of $\mathcal{D}=\mathcal{E}$ or Δ, the last term is identical to $-\frac{1}{\eta_T}$ (since $\sum_i d^i=1$) and keeping it improves the bound. Furthermore, we need to evaluate the expectation of the second to last term in Theorem 2, namely $E[\max_{d\in\mathcal{D}}\{d\circ(q-k)\}]$. For $\mathcal{D}=\mathcal{E}$ and q being exponentially distributed, using Lemma 1, the expectation is bounded by $1+\ln u$. We hence get the following bound:

Corollary 3 (IFPL bounded by BEH) *For $\mathcal{D}=\mathcal{E}$ and $\sum_i \mathrm{e}^{-k^i}\leq 1$ and $P[q^i]=\mathrm{e}^{-q^i}$ for $q\geq 0$ and decreasing $\eta_t>0$, the expected loss of the infeasible FPL exceeds the loss of expert i by at most k^i/η_T:*

$$r_{1:T} \;\leq\; s^i_{1:T}+\frac{1}{\eta_T}k^i \quad \forall i.$$

Theorem 2 can be generalized to expert dependent factorizable $\eta_t \rightsquigarrow \eta_t^i = \eta_t \cdot \eta^i$ by scaling $k^i \rightsquigarrow k^i/\eta^i$ and $q^i \rightsquigarrow q^i/\eta^i$. Using $E[\max_i\{\frac{q^i-k^i}{\eta^i}\}] \leq E[\max_i\{q^i-k^i\}]/\min_i\{\eta^i\}$, Corollary 3, generalizes to

$$E[\sum_{t=1}^T M(s_{1:t} + \frac{k-q}{\eta_t^i})\circ s_t] \;\leq\; s_{1:T}^i + \frac{1}{\eta_T^i}k^i + \frac{1}{\eta_T^{min}} \quad \forall i,$$

where $\eta_T^{min} := \min_i\{\eta_T^i\}$. For example, for $\eta_t^i = \sqrt{k^i/t}$ we get the desired bound $s_{1:T}^i + \sqrt{T\cdot(k^i+4)}$. Unfortunately we were not able to generalize Theorem 4 to expert-dependent η, necessary for the final bound on FPL. In Section 6 we solve this problem by a hierarchy of experts.

4. Feasible FPL bounded by Infeasible FPL

This section establishes the relation between the FPL and IFPL losses. Recall that $\ell_t = E[M(s_{<t}+\frac{k-q}{\eta_t})\circ s_t]$ is the expected loss of FPL at time t and $r_t = E[M(s_{1:t}+\frac{k-q}{\eta_t})\circ s_t]$ is the expected loss of IFPL at time t.

Theorem 4 (FPL bounded by IFPL) *For $\mathcal{D}=\mathcal{E}$ and $0\leq s_t^i \leq 1\ \forall i$ and arbitrary $s_{<t}$ and $P[q] = \mathrm{e}^{-\sum_i q^i}$ for $q\geq 0$, the expected loss of the feasible FPL is at most a factor $\mathrm{e}^{\eta_t} > 1$ larger than for the infeasible FPL:*

$$\ell_t \leq \mathrm{e}^{\eta_t} r_t, \quad \textit{which implies} \quad \ell_{1:T} - r_{1:T} \leq \sum_{t=1}^T \eta_t \ell_t.$$

Furthermore, if $\eta_t \leq 1$, then also $\ell_t \leq (1+\eta_t+\eta_t^2) r_t \leq (1+2\eta_t) r_t$.

Proof. Let $s = s_{<t} + \frac{1}{\eta}k$ be the past cumulative penalized state vector, q be a vector of independent exponential distributions, i.e. $P[q^i] = \mathrm{e}^{-q^i}$, and $\eta = \eta_t$. Then

$$\frac{P[q^j \geq \eta(s^j - m + 1)]}{P[q^j \geq \eta(s^j - m)]} = \left\{ \begin{array}{ccc} \mathrm{e}^{-\eta} & \text{if} & s^j \geq m \\ \mathrm{e}^{-\eta(s^j-m+1)} & \text{if} & m-1 \leq s^j \leq m \\ 1 & \text{if} & s^j \leq m-1 \end{array} \right\} \geq \mathrm{e}^{-\eta}$$

We now define the random variables $I := \arg\min_i\{s^i - \frac{1}{\eta}q^i\}$ and $J := \arg\min_i\{s^i + s_t^i - \frac{1}{\eta}q^i\}$, where $0 \leq s_t^i \leq 1\ \forall i$. Furthermore, for fixed vector $x \in \boldsymbol{R}^n$ and fixed j we define $m := \min_{i\neq j}\{s^i - \frac{1}{\eta}x^i\} \leq \min_{i\neq j}\{s^i + s_t^i - \frac{1}{\eta}x^i\} =: m'$. With this notation and using the independence of q^j from q^i for all $i\neq j$, we get

$$P[I=j|q^i = x^i\,\forall i\neq j] \;=\; P[s^j - \tfrac{1}{\eta}q^j \leq m | q^i = x^i\,\forall i\neq j] \;=\; P[q^j \geq \eta(s^j - m)]$$

$$\leq\; \mathrm{e}^{\eta} P[q^j \geq \eta(s^j - m + 1)] \;\leq\; \mathrm{e}^{\eta} P[q^j \geq \eta(s^j + s_t^j - m')]$$

$$=\; \mathrm{e}^{\eta} P[s^j + s_t^j - \tfrac{1}{\eta}q^j \leq m' | q^i = x^i\,\forall i\neq j] \;=\; \mathrm{e}^{\eta} P[J=j|q^i = x^i\,\forall i\neq j].$$

Since this bound holds under any condition x, it also holds unconditionally, i.e. $P[I=j] \leq \mathrm{e}^{\eta} P[J=j]$. For $\mathcal{D}=\mathcal{E}$ we have $s_t^I = M(s_{<t}+\frac{k-q}{\eta})\circ s_t$ and $s_t^J = M(s_{1:t}+\frac{k-q}{\eta})\circ s_t$, which implies

$$\ell_t \;=\; E[s_t^I] \;=\; \sum_{j=1}^n s_t^j \cdot P[I=j] \;\leq\; \mathrm{e}^{\eta} \sum_{j=1}^n s_t^j \cdot P[J=j] \;=\; \mathrm{e}^{\eta} E[s_t^J] \;=\; \mathrm{e}^{\eta} r_t.$$

Finally, $\ell_t - r_t \leq \eta_t \ell_t$ follows from $r_t \geq \mathrm{e}^{-\eta_t}\ell_t \geq (1-\eta_t)\ell_t$, and $\ell_t \leq \mathrm{e}^{\eta_t} r_t \leq (1+\eta_t+\eta_t^2) r_t \leq (1+2\eta_t) r_t$ for $\eta_t \leq 1$ is elementary. □

Remark. As done by Kalai and Vempala (2003), one can prove a similar statement for general decision space $\mathcal{D}$ as long as $\sum_i |s_t^i| \leq A$ is guaranteed for some $A>0$: In this case, we have $\ell_t \leq \mathrm{e}^{\eta_t A} r_t$. If n is finite, then the bound holds for $A=n$. For $n=\infty$, the assertion holds under the somewhat unnatural assumption that $\mathcal{S}$ is l^1-bounded.

5. Combination of Bounds and Choices for η_t

Throughout this section, we assume

$$\mathcal{D} = \mathcal{E}, \quad s_t \in [0,1]^n \ \forall t, \quad P[q] = \mathrm{e}^{-\sum_i q^i} \text{ for } q \geq 0, \quad \text{and} \quad \sum_i \mathrm{e}^{-k^i} \leq 1. \tag{8}$$

We distinguish *static* and *dynamic* bounds. Static bounds refer to a constant $\eta_t \equiv \eta$. Since this value has to be chosen in advance, a static choice of η_t requires certain prior information and therefore is not practical in many cases. However, the static bounds are very easy to derive, and they provide a good means to compare different PEA algorithms. If on the other hand the algorithm shall be applied without appropriate prior knowledge, a dynamic choice of η_t depending only on t and/or past observations, is necessary.

Theorem 5 (FPL bound for static $\eta_t = \eta \propto 1/\sqrt{L}$) *Assume (8) holds, then the expected loss ℓ_t of feasible FPL, which employs the prediction of the expert i minimizing $s_{<t}^i + \frac{k^i - q^i}{\eta_t}$, is bounded by the loss of the best expert in hindsight in the following way:*

$$\begin{array}{rl} i) & \textit{For} \quad \eta_t = \eta = 1/\sqrt{L} \quad \textit{with} \quad L \geq \ell_{1:T} \quad \textit{we have} \\ & \ell_{1:T} \ \leq \ s_{1:T}^i + \sqrt{L}(k^i+1) \quad \forall i. \\ ii) & \textit{For} \quad \eta_t = \sqrt{K/L} \quad \textit{with} \quad L \geq \ell_{1:T} \quad \textit{and} \quad k^i \leq K \ \forall i \quad \textit{we have} \\ & \ell_{1:T} \ \leq \ s_{1:T}^i + 2\sqrt{LK} \quad \forall i. \\ iii) & \textit{For} \quad \eta_t = \sqrt{k^i/L} \quad \textit{with} \quad L \geq \max\{s_{1:T}^i, k^i\} \quad \textit{we have} \\ & \ell_{1:T} \ \leq \ s_{1:T}^i + 2\sqrt{Lk^i} + 3k^i. \end{array}$$

Note that according to assertion (iii), knowledge of only the *ratio* of the complexity and the loss of the best expert is sufficient in order to obtain good static bounds, even for non-uniform complexities.

Proof. (i,ii) For $\eta_t = \sqrt{K/L}$ and $L \geq \ell_{1:T}$, from Theorem 4 and Corollary 3, we get

$$\ell_{1:T} - r_{1:T} \leq \sum_{t=1}^T \eta_t \ell_t = \ell_{1:T}\sqrt{K/L} \leq \sqrt{LK} \quad \text{and} \quad r_{1:T} - s_{1:T}^i \leq k^i/\eta_T = k^i\sqrt{L/K}.$$

Combining both, we get $\ell_{1:T} - s_{1:T}^i \leq \sqrt{L}(\sqrt{K} + k^i/\sqrt{K})$. (i) follows from $K=1$ and (ii) from $k^i \leq K$. (iii) For $\eta = \sqrt{k^i/L} \leq 1$ we get

$$\begin{aligned} \ell_{1:T} \ &\leq \ \mathrm{e}^{\eta} r_{1:T} \leq (1+\eta+\eta^2) r_{1:T} \leq (1 + \sqrt{\frac{k^i}{L}} + \frac{k^i}{L})(s_{1:T}^i + \sqrt{\frac{L}{k^i}}k^i) \\ &\leq \ s_{1:T}^i + \sqrt{Lk^i} + (\sqrt{\frac{k^i}{L}} + \frac{k^i}{L})(L + \sqrt{Lk^i}) = s_{1:T}^i + 2\sqrt{Lk^i} + (2 + \sqrt{\frac{k^i}{L}})k^i. \end{aligned}$$

□

The static bounds require knowledge of an upper bound L on the loss (or the ratio of the complexity of the best expert and its loss). Since the instantaneous loss is bounded by 1, one may set $L=T$ if T is known in advance. For finite n and $k^i=K=\ln n$, bound (ii) gives the classic regret $\propto\sqrt{T\ln n}$. If neither T nor L is known, a dynamic choice of η_t is necessary. We first present bounds with regret $\propto\sqrt{T}$, thereafter with regret $\propto\sqrt{s^i_{1:T}}$.

Theorem 6 (FPL bound for dynamic $\eta_t\propto 1/\sqrt{t}$) *Assume (8) holds.*

$$\begin{array}{ll} i) & \textit{For}\quad \eta_t=1/\sqrt{t}\quad \textit{we have}\quad \ell_{1:T}\;\leq\; s^i_{1:T}+\sqrt{T}(k^i+2)\quad \forall i. \\ ii) & \textit{For}\quad \eta_t=\sqrt{K/2t}\;\; \textit{and}\;\; k^i\leq K\;\forall i\;\; \textit{we have}\;\; \ell_{1:T}\;\leq\; s^i_{1:T}+2\sqrt{2TK}\quad \forall i. \end{array}$$

Proof. For $\eta_t=\sqrt{K/2t}$, using $\sum_{t=1}^T\frac{1}{\sqrt{t}}\leq\int_0^T\frac{dt}{\sqrt{t}}=2\sqrt{T}$ and $\ell_t\leq 1$ we get

$$\ell_{1:T}-r_{1:T}\leq\sum_{t=1}^T\eta_t\leq\sqrt{2TK}\quad\text{and}\quad r_{1:T}-s^i_{1:T}\leq k^i/\eta_T=k^i\sqrt{\frac{2T}{K}}.$$

Combining both, we get $\ell_{1:T}-s^i_{1:T}\leq\sqrt{2T}(\sqrt{K}+k^i/\sqrt{K})$. (i) follows from $K=2$ and (ii) from $k^i\leq K$. □

In Theorem 5 we assumed knowledge of an upper bound L on $\ell_{1:T}$. In an adaptive form, $L_t:=\ell_{<t}+1$, known at the beginning of time t, could be used as an upper bound on $\ell_{1:t}$ with corresponding adaptive $\eta_t\propto 1/\sqrt{L_t}$. Such choice of η_t is also called *self-confident* (Auer et al., 2002).

Theorem 7 (FPL bound for self-confident $\eta_t\propto 1/\sqrt{\ell_{<t}}$) *Assume (8) holds.*

$$\begin{array}{ll} i) & \textit{For}\quad \eta_t=1/\sqrt{2(\ell_{<t}+1)}\quad \textit{we have} \\ & \ell_{1:T}\;\leq\; s^i_{1:T}+(k^i+1)\sqrt{2(s^i_{1:T}+1)}+2(k^i+1)^2\quad\forall i. \\ ii) & \textit{For}\quad \eta_t=\sqrt{K/2(\ell_{<t}+1)}\quad \textit{and}\quad k^i\leq K\;\forall i\quad \textit{we have} \\ & \ell_{1:T}\;\leq\; s^i_{1:T}+2\sqrt{2(s^i_{1:T}+1)K}+8K\quad\forall i. \end{array}$$

Proof. Using $\eta_t=\sqrt{K/2(\ell_{<t}+1)}\leq\sqrt{K/2\ell_{1:t}}$ and $\frac{b-a}{\sqrt{b}}=(\sqrt{b}-\sqrt{a})(\sqrt{b}+\sqrt{a})\frac{1}{\sqrt{b}}\leq 2(\sqrt{b}-\sqrt{a})$ for $a\leq b$ and $t_0:=\min\{t:\ell_{1:t}>0\}$ we get

$$\ell_{1:T}-r_{1:T}\leq\sum_{t=t_0}^T\eta_t\ell_t\leq\sqrt{\frac{K}{2}}\sum_{t=t_0}^T\frac{\ell_{1:t}-\ell_{<t}}{\sqrt{\ell_{1:t}}}\leq\sqrt{2K}\sum_{t=t_0}^T[\sqrt{\ell_{1:t}}-\sqrt{\ell_{<t}}]=\sqrt{2K}\sqrt{\ell_{1:T}}.$$

Adding $r_{1:T}-s^i_{1:T}\leq\frac{k^i}{\eta_T}\leq k^i\sqrt{2(\ell_{1:T}+1)/K}$ we get

$$\ell_{1:T}-s^i_{1:T}\leq\sqrt{2\overline{\kappa}^i(\ell_{1:T}+1)},\quad\text{where}\quad\sqrt{\overline{\kappa}^i}:=\sqrt{K}+k^i/\sqrt{K}.$$

Taking the square and solving the resulting quadratic inequality w.r.t. $\ell_{1:T}$ we get

$$\ell_{1:T}\leq s^i_{1:T}+\overline{\kappa}^i+\sqrt{2(s^i_{1:T}+1)\overline{\kappa}^i+(\overline{\kappa}^i)^2}\leq s^i_{1:T}+\sqrt{2(s^i_{1:T}+1)\overline{\kappa}^i}+2\overline{\kappa}^i.$$

For $K=1$ we get $\sqrt{\overline{\kappa}^i}=k^i+1$ which yields (i). For $k^i\leq K$ we get $\overline{\kappa}^i\leq 4K$ which yields (ii). $\square$

The proofs of results similar to (ii) for WM for 0/1 loss all fill several pages (Auer et al., 2002; Yaroshinsky et al., 2004). The next result establishes a similar bound, but instead of using the *expected* value $\ell_{<t}$, the *best loss so far* $s^{min}_{<t}$ is used. This may have computational advantages, since $s^{min}_{<t}$ is immediately available, while $\ell_{<t}$ needs to be evaluated (see discussion in Section 9).

Theorem 8 (FPL bound for adaptive $\eta_t\propto 1/\sqrt{s^{min}_{<t}}$) *Assume (8) holds.*

$$\begin{aligned} i)\quad & \textit{For}\quad \eta_t = 1/\min_i\{k^i+\sqrt{(k^i)^2+2s^i_{<t}+2}\}\quad \textit{we have}\\ & \ell_{1:T} \;\leq\; s^i_{1:T}+(k^i+2)\sqrt{2s^i_{1:T}}+2(k^i+2)^2 \quad \forall i.\\ ii)\quad & \textit{For}\quad \eta_t = \sqrt{\tfrac{1}{2}}\cdot\min\{1,\sqrt{K/s^{min}_{<t}}\}\quad \textit{and}\quad k^i\leq K\ \forall i \quad \textit{we have}\\ & \ell_{1:T} \;\leq\; s^i_{1:T}+2\sqrt{2Ks^i_{1:T}}+5K\ln(s^i_{1:T})+3K+6 \quad \forall i. \end{aligned}$$

We briefly motivate the strange looking choice for η_t in (i). The first naive candidate, $\eta_t\propto 1/\sqrt{s^{min}_{<t}}$, turns out too large. The next natural trial is requesting $\eta_t = 1/\sqrt{2\min\{s^i_{<t}+\frac{k^i}{\eta_t}\}}$. Solving this equation results in $\eta_t=1/(k^i+\sqrt{(k^i)^2+2s^i_{<t}})$, where i be the index for which $s^i_{<t}+\frac{k^i}{\eta_t}$ is minimal.

Proof. Define the minimum of a vector as its minimum component, e.g. $\min(k)=k^{min}$. For notational convenience, let $\eta_0=\infty$ and $\tilde{s}_{1:t}=s_{1:t}+\frac{k-q}{\eta_t}$. Like in the proof of Theorem 2, we consider one exponentially distributed perturbation q. Since $M(\tilde{s}_{1:t})\circ\tilde{s}_t\leq M(\tilde{s}_{1:t})\circ\tilde{s}_{1:t}-M(\tilde{s}_{<t})\circ\tilde{s}_{<t}$ by (4), we have

$$M(\tilde{s}_{1:t})\circ s_t \leq M(\tilde{s}_{1:t})\circ\tilde{s}_{1:t}-M(\tilde{s}_{<t})\circ\tilde{s}_{<t}-M(\tilde{s}_{1:t})\circ\left(\frac{k-q}{\eta_t}-\frac{k-q}{\eta_{t-1}}\right)$$

Since $\eta_t\leq\sqrt{1/2}$, Theorem 4 asserts $\ell_t\leq E[(1+\eta_t+\eta_t^2)M(\tilde{s}_{1:t})\circ s_t]$, thus $\ell_{1:T}\leq A+B$, where

$$\begin{aligned} A &= \sum_{t=1}^{T} E\left[(1+\eta_t+\eta_t^2)(M(\tilde{s}_{1:t})\circ\tilde{s}_{1:t}-M(\tilde{s}_{<t})\circ\tilde{s}_{<t})\right]\\ &= E[(1+\eta_T+\eta_T^2)M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T}]-E[(1+\eta_1+\eta_1^2)\min(\frac{k-q}{\eta_1})]\\ &\quad +\sum_{t=1}^{T-1} E\left[(\eta_t-\eta_{t+1}+\eta_t^2-\eta_{t+1}^2)M(\tilde{s}_{1:t})\circ\tilde{s}_{1:t}\right] \quad \text{and}\\ B &= \sum_{t=1}^{T} E\left[(1+\eta_t+\eta_t^2)M(\tilde{s}_{1:t})\circ\left(\frac{q-k}{\eta_t}-\frac{q-k}{\eta_{t-1}}\right)\right]\\ &\leq \sum_{t=1}^{T}(1+\eta_t+\eta_t^2)\left(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\right) = \frac{1+\eta_T+\eta_T^2}{\eta_T}+\sum_{t=1}^{T-1}\frac{\eta_t-\eta_{t+1}+\eta_t^2-\eta_{t+1}^2}{\eta_t}. \end{aligned}$$

Here, the estimate for B follows from $\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\geq 0$ and $E[M(\eta_t s_{1:t}+k-q)\circ(q-k)]\leq E[\max_i\{q^i-k^i\}]\leq 1$, which in turn holds by minimality of M, $\sum_i e^{-k^i}\leq 1$ and Lemma 1. In order to estimate A, we set $\bar{s}_{1:t}=s_{1:t}+\frac{k}{\eta_t}$ and observe $M(\tilde{s}_{1:t})\circ\tilde{s}_{1:t}\leq M(\bar{s}_{1:t})\circ(\bar{s}_{1:t}-\frac{q}{\eta_t})$ by minimality of M. The expectations of q can then be evaluated to $E[M(\bar{s}_{1:t})\circ q]=1$, and as before we have $E[-\min(k-q)]\leq 1$. Hence

$$\begin{aligned}\ell_{1:T} &\leq A+B \leq (1+\eta_T+\eta_T^2)\left(M(\bar{s}_{1:T})\circ\bar{s}_{1:T}-\frac{1}{\eta_T}\right)+\frac{1+\eta_1+\eta_1^2}{\eta_1}\\ &\quad+\sum_{t=1}^{T-1}(\eta_t-\eta_{t+1}+\eta_t^2-\eta_{t+1}^2)\left(M(\bar{s}_{1:t})\circ\bar{s}_{1:t}-\frac{1}{\eta_t}\right)+B \qquad (9)\\ &\leq (1+\eta_T+\eta_T^2)\min(\bar{s}_{1:T})+\sum_{t=1}^{T-1}(\eta_t-\eta_{t+1}+\eta_t^2-\eta_{t+1}^2)\min(\bar{s}_{1:t})+\frac{1}{\eta_1}+2.\end{aligned}$$

We now proceed by considering the two parts of the theorem separately.

(i) Here, $\eta_t=1/\min(k+\sqrt{k^2+2s_{<t}+2})$. Fix $t\leq T$ and choose m such that $k^m+\sqrt{(k^m)^2+2s_{<t}^m+2}$ is minimal. Then

$$\min(s_{1:t}+\frac{k}{\eta_t})\leq s_{<t}^m+1+\frac{k^m}{\eta_t}=\tfrac{1}{2}(k^m+\sqrt{(k^m)^2+2s_{<t}^m+2})^2=\frac{1}{2\eta_t^2}\leq\frac{1}{2\eta_t\eta_{t+1}}.$$

We may overestimate the quadratic terms η_t^2 in (9) by η_t – the easiest justification is that we could have started with the cruder estimate $\ell_t\leq(1+2\eta_t)r_t$ from Theorem 4. Then

$$\begin{aligned}\ell_{1:T} &\leq (1+2\eta_T)\min(s_{1:T}+\frac{k}{\eta_T})+2\sum_{t=1}^{T-1}(\eta_t-\eta_{t+1})\min(s_{1:t}+\frac{k}{\eta_t})+\frac{1}{\eta_1}+2\\ &\leq (1+2\eta_T)\frac{1}{2\eta_T^2}+2\sum_{t=1}^{T-1}(\eta_t-\eta_{t+1})\frac{1}{2\eta_t^2}+\frac{1}{\eta_1}+2\\ &\leq \frac{1}{2\eta_T^2}+\frac{1}{\eta_T}+\sum_{t=1}^{T-1}\left(\frac{1}{\eta_{t+1}}-\frac{1}{\eta_t}\right)+\frac{1}{\eta_1}+2\\ &\leq \tfrac{1}{2}\min(k+\sqrt{k^2+2s_{<T}+2})^2+2\min(k+\sqrt{k^2+2s_{<T}+2})+2\\ &\leq s_{1:T}^i+(k^i+2)\sqrt{2s_{1:T}^i}+2(k^i)^2+6k^i+6 \quad \text{for all } i.\end{aligned}$$

This proves the first part of the theorem.

(ii) Here we have $K\geq k^i$ for all i. Abbreviate $a_t=\max\{K,s_{1:t}^{min}\}$ for $1\leq t\leq T$, then $\eta_t=\sqrt{\frac{K}{2a_{t-1}}}$, $a_t\geq K$, and $a_t-a_{t-1}\leq 1$ for all t. Observe $M(\bar{s}_{1:t})=M(s_{1:t})$, $\eta_t-\eta_{t+1}=\frac{\sqrt{K}(a_t-a_{t-1})}{\sqrt{2}\sqrt{a_t}\sqrt{a_{t-1}}(\sqrt{a_t}+\sqrt{a_{t-1}})}$, $\eta_t^2-\eta_{t+1}^2=\frac{K(a_t-a_{t-1})}{2a_ta_{t-1}}$, and $\frac{a_t-a_{t-1}}{2a_{t-1}}\leq\ln(1+\frac{a_t-a_{t-1}}{a_{t-1}})=\ln(a_t)-\ln(a_{t-1})$ which is true for $\frac{a_t-a_{t-1}}{a_{t-1}}\leq\frac{1}{K}\leq\frac{1}{\ln 2}$. This implies

$$\begin{aligned}\frac{(\eta_t-\eta_{t+1})K}{\eta_t} &\leq \frac{K(a_t-a_{t-1})}{2a_{t-1}}\leq K\ln\left(1+\frac{a_t-a_{t-1}}{a_{t-1}}\right)=K(\ln(a_t)-\ln(a_{t-1})),\\ (\eta_t-\eta_{t+1})s_{1:t}^{min} &\leq \frac{\sqrt{K}(a_t-a_{t-1})(\sqrt{a_{t-1}}+\sqrt{a_t}-\sqrt{a_{t-1}})}{\sqrt{2}\sqrt{a_{t-1}}(\sqrt{a_t}+\sqrt{a_{t-1}})}\\ &= \sqrt{\frac{K}{2}}(\sqrt{a_t}-\sqrt{a_{t-1}})+\frac{\sqrt{K}(a_t-a_{t-1})^2}{\sqrt{2a_{t-1}}(\sqrt{a_t}+\sqrt{a_{t-1}})^2}\\ &\overset{\text{use } a_t-a_{t-1}\leq 1 \text{ and } a_{t-1}\geq K}{\leq} \sqrt{\frac{K}{2}}(\sqrt{a_t}-\sqrt{a_{t-1}})+\frac{1}{2\sqrt{2}}(\ln(a_t)-\ln(a_{t-1})),\end{aligned}$$

$$\frac{(\eta_t^2-\eta_{t+1}^2)K}{\eta_t} \;=\; \frac{K\sqrt{K}(a_t-a_{t-1})}{\sqrt{2}a_t\sqrt{a_{t-1}}} \overset{\boxed{a_{t-1}\geq K}}{\leq} \sqrt{2}K(\ln(a_t)-\ln(a_{t-1})), \text{ and}$$
$$(\eta_t^2-\eta_{t+1}^2)s_{1:t}^{min} \;\leq\; \frac{K(a_t-a_{t-1})}{2a_{t-1}} \leq K(\ln(a_t)-\ln(a_{t-1})).$$

The logarithmic estimate in the second and third bound is unnecessarily rough and for convenience only. Therefore, the coefficient of the log-term in the final bound of the theorem can be reduced to $2K$ without much effort. Plugging the above estimates back into (9) yields

$$\begin{aligned}\ell_{1:T} \;\leq\; & s_{1:T}^{min}+\sqrt{\frac{K}{2}s_{1:T}^{min}}+\sqrt{2Ks_{1:T}^{min}}+3K+2+\sqrt{\frac{K}{2}s_{1:T}^{min}}+(\tfrac{7}{2}K+\tfrac{1}{2\sqrt{2}})\ln(s_{1:T}^{min})\\ &+\frac{1}{\eta_1}+2 \leq s_{1:T}^{min}+2\sqrt{2Ks_{1:T}^{min}}+5K\ln(s_{1:T}^{min})+3K+6.\end{aligned}$$

This completes the proof. □

Theorem 7 and Theorem 8 (i) immediately imply the following bounds on the $\sqrt{\text{Loss}}$-regrets: $\sqrt{\ell_{1:T}} \leq \sqrt{s_{1:T}^i+1}+\sqrt{8K}$, $\sqrt{\ell_{1:T}} \leq \sqrt{s_{1:T}^i+1}+\sqrt{2}(k^i+1)$, and $\sqrt{\ell_{1:T}} \leq \sqrt{s_{1:T}^i}+\sqrt{2}(k^i+2)$, respectively.

Remark. The same analysis as for Theorems [5–8](ii) applies to general $\mathcal{D}$, using $\ell_t \leq e^{\eta_t n} r_t$ instead of $\ell_t \leq e^{\eta_t} r_t$, and leading to an additional factor $\sqrt{n}$ in the regret. Compare the remark at the end of Section 4.

6. Hierarchy of Experts

We derived bounds which do not need prior knowledge of L with regret $\propto \sqrt{TK}$ and $\propto \sqrt{s_{1:T}^i K}$ for a finite number of experts with equal penalty $K=k^i=\ln n$. For an infinite number of experts, unbounded expert-dependent complexity penalties k^i are necessary (due to constraint $\sum_i e^{-k^i} \leq 1$). Bounds for this case (without prior knowledge of T) with regret $\propto k^i\sqrt{T}$ and $\propto k^i\sqrt{s_{1:T}^i}$ have been derived. In this case, the complexity k^i is no longer under the square root. Although this already implies Hannan consistency, i.e. the average per round regret tends to zero as $t\to\infty$, improved regret bounds $\propto \sqrt{Tk^i}$ and $\propto \sqrt{s_{1:T}^i k^i}$ are desirable and likely to hold. We were not able to derive such improved bounds for FPL, but for a (slight) modification. We consider a two-level hierarchy of experts. First consider an FPL for the subclass of experts of complexity K, for each $K\in\mathbb{N}$. Regard these FPL^K as (meta) experts and use them to form a (meta) FPL. The class of meta experts now contains for each complexity only one (meta) expert, which allows us to derive good bounds. In the following, quantities referring to complexity class K are superscripted by K, and meta quantities are superscripted by $\tilde{}$.

Consider the class of experts $\mathcal{E}^K := \{i : K-1 < k^i \leq K\}$ of complexity K, for each $K\in\mathbb{N}$. FPL^K makes randomized prediction $I_t^K := \text{argmin}_{i\in\mathcal{E}^K}\{s_{<t}^i+\frac{k^i-q^i}{\eta_t^K}\}$ with $\eta_t^K := \sqrt{K/2t}$ and suffers loss $u_t^K := s_t^{I_t^K}$ at time t. Since $k^i \leq K\ \forall i\in\mathcal{E}^k$ we can apply Theorem 6(ii) to FPL^K:

$$E[u_{1:T}^K] \;=\; \ell_{1:T}^K \;\leq\; s_{1:T}^i+2\sqrt{2TK} \quad \forall i\in\mathcal{E}^K \quad \forall K\in\mathbb{N}. \tag{10}$$

We now define a meta state $\tilde{s}_t^K = u_t^K$ and regard FPL^K for $K\in\mathbb{N}$ as meta experts, so meta expert K suffers loss $\tilde{s}_t^K$. (Assigning expected loss $\tilde{s}_t^K = E[u_t^K] = \ell_t^K$ to FPL^K would also work.) Hence the setting is again an expert setting and we define the meta $\widetilde{\text{FPL}}$ to predict $\tilde{I}_t := \arg\min_{K\in N}\{\tilde{s}^K_{<t} + \frac{\tilde{k}^K - \tilde{q}^K}{\tilde{\eta}_t}\}$ with $\tilde{\eta}_t = 1/\sqrt{t}$ and $\tilde{k}^K = \frac{1}{2} + 2\ln K$ (implying $\sum_{K=1}^{\infty} e^{-\tilde{k}^K} \leq 1$). Note that $\tilde{s}^K_{1:t} = \tilde{s}^K_1 + ... + \tilde{s}^K_t = s_1^{I_1^K} + ... + s_t^{I_t^K}$ sums over the same meta state components K, but over different components I_t^K in normal state representation.

By Theorem 6(i) the $\tilde{q}$-expected loss of $\widetilde{\text{FPL}}$ is bounded by $\tilde{s}^K_{1:T} + \sqrt{T}(\tilde{k}^K + 2)$. As this bound holds for all q it also holds in q-expectation. So if we define $\tilde{\ell}_{1:T}$ to be the q *and* $\tilde{q}$ expected loss of $\widetilde{\text{FPL}}$, and chain this bound with (10) for $i \in \mathcal{E}^K$ we get:

$$\begin{aligned} \tilde{\ell}_{1:T} &\leq E[\tilde{s}^K_{1:T} + \sqrt{T}(\tilde{k}^K + 2)] \;=\; \ell^K_{1:T} + \sqrt{T}(\tilde{k}^K + 2) \\ &\leq s^i_{1:T} + \sqrt{T}[2\sqrt{2(k^i + 1)} + \tfrac{1}{2} + 2\ln(k^i + 1) + 2], \end{aligned}$$

where we have used $K \leq k^i + 1$. This bound is valid for all i and has the desired regret $\propto \sqrt{Tk^i}$. Similarly we can derive regret bounds $\propto \sqrt{s^i_{1:T} k^i}$ by exploiting that the bounds in Theorems 7 and 8 are concave in $s^i_{1:T}$ and using Jensen's inequality.

Theorem 9 (Hierarchical FPL bound for dynamic η_t) *The hierarchical $\widetilde{FPL}$ employs at time t the prediction of expert $i_t := I_t^{\tilde{I}_t}$, where*

$$I_t^K := \arg\min_{i:\lceil k^i\rceil = K}\left\{s^i_{<t} + \frac{k^i - q^i}{\eta_t^K}\right\} \quad \textit{and} \quad \tilde{I}_t := \arg\min_{K\in N}\left\{s_1^{I_1^K} + ... + s_{t-1}^{I_{t-1}^K} + \frac{\frac{1}{2} + 2\ln K - \tilde{q}^K}{\tilde{\eta}_t}\right\}$$

Under assumptions (8) and independent $P[\tilde{q}^K] = e^{-\tilde{q}^K}$ $\forall K \in N$, the expected loss $\tilde{\ell}_{1:T} = E[s_1^{i_1} + ... + s_T^{i_T}]$ of $\widetilde{FPL}$ is bounded as follows:

$a)$ *For* $\eta_t^K = \sqrt{K/2t}$ *and* $\tilde{\eta}_t = 1/\sqrt{t}$ *we have*

$$\tilde{\ell}_{1:T} \;\leq\; s^i_{1:T} + 2\sqrt{2Tk^i}\cdot(1 + O(\tfrac{\ln k^i}{\sqrt{k^i}})) \quad \forall i.$$

$b)$ *For $\tilde{\eta}_t$ as in (i) and η_t^K as in (ii) of Theorem* $\{{7 \atop 8}\}$ *we have*

$$\tilde{\ell}_{1:T} \;\leq\; s^i_{1:T} + 2\sqrt{2s^i_{1:T}k^i}\cdot(1 + O(\tfrac{\ln k^i}{\sqrt{k^i}})) + \left\{{O(k^i) \atop O(k^i \ln s^i_{1:T})}\right\} \quad \forall i.$$

The hierarchical $\widetilde{\text{FPL}}$ differs from a direct FPL over all experts $\mathcal{E}$. One potential way to prove a bound on direct FPL may be to show (if it holds) that FPL performs better than $\widetilde{\text{FPL}}$, i.e. $\ell_{1:T} \leq \tilde{\ell}_{1:T}$. Another way may be to suitably generalize Theorem 4 to expert dependent η.

7. Lower Bound on FPL

A lower bound on FPL similar to the upper bound in Theorem 2 can also be proven.

Theorem 10 (FPL lower-bounded by BEH) *Let n be finite. Assume $\mathcal{D} \subseteq \boldsymbol{R}^n$ and $s_t \in \boldsymbol{R}^n$ are chosen such that the required extrema exist (possibly negative), $q \in \boldsymbol{R}^n$, and $\eta_t > 0$ is a decreasing sequence. Then the loss of FPL for uniform complexities (l.h.s.) can be lower-bounded in terms of the best predictor in hindsight (first term on r.h.s.) plus/minus additive corrections:*

$$\sum_{t=1}^{T} M(s_{<t} - \frac{q}{\eta_t})\circ s_t \geq \min_{d\in\mathcal{D}}\{d\circ s_{1:T}\} - \frac{1}{\eta_T}\max_{d\in\mathcal{D}}\{d\circ q\} + \sum_{t=1}^{T}(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}})M(s_{<t})\circ q$$

Proof. For notational convenience, let $\eta_0=\infty$ and $\tilde{s}_{1:t}=s_{1:t}-\frac{q}{\eta_t}$. Consider the losses $\tilde{s}_t=s_t-q(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}})$ for the moment. We first show by induction on T that the predictor $M(\tilde{s}_{<t})$ has nonnegative regret, i.e.

$$\sum_{t=1}^{T} M(\tilde{s}_{<t})\circ\tilde{s}_t \geq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T}. \tag{11}$$

For $T=1$ this follows immediately from minimality of M ($\tilde{s}_{<1}:=0$). For the induction step from $T-1$ to T we need to show

$$M(\tilde{s}_{<T})\circ\tilde{s}_T \geq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} - M(\tilde{s}_{<T})\circ\tilde{s}_{<T}.$$

Due to $\tilde{s}_{1:T}=\tilde{s}_{<T}+\tilde{s}_T$, this is equivalent to $M(\tilde{s}_{<T})\circ\tilde{s}_{1:T}\geq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T}$, which holds by minimality of M. Rearranging terms in (11) we obtain

$$\sum_{t=1}^{T} M(\tilde{s}_{<t})\circ s_t \geq M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} + \sum_{t=1}^{T} M(\tilde{s}_{<t})\circ q\Big(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\Big), \quad \text{with} \tag{12}$$

$$M(\tilde{s}_{1:T})\circ\tilde{s}_{1:T} = M(s_{1:T}-\frac{q}{\eta_T})\circ s_{1:T} - M(s_{1:T}-\frac{q}{\eta_T})\circ\frac{q}{\eta_T} \geq \min_{d\in\mathcal{D}}\{d\circ s_{1:T}\} - \frac{1}{\eta_T}\max_{d\in\mathcal{D}}\{d\circ q\}$$

$$\text{and} \quad \sum_{t=1}^{T} M(\tilde{s}_{<t})\circ q\Big(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\Big) \;\geq\; \sum_{t=1}^{T}\Big(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}}\Big)M(s_{<t})\circ q.$$

Again, the last bound follows from the minimality of M, which asserts that $[M(s-q)-M(s)]\circ s\geq 0\geq[M(s-q)-M(s)]\circ(s-q)$ and thus implies that $M(s-q)\circ q\geq M(s)\circ q$. So Theorem 10 follows from (12). $\square$

Assuming q random with $E[q^i]=1$ and taking the expectation in Theorem 10, the last term reduces to $\sum_t(\frac{1}{\eta_t}-\frac{1}{\eta_{t-1}})\sum_i M(s_{<t})^i$. If $\mathcal{D}\geq 0$, the term is positive and may be dropped. In case of $\mathcal{D}=\mathcal{E}$ or Δ, the last term is identical to $\frac{1}{\eta_T}$ (since $\sum_i d^i=1$) and keeping it improves the bound. Furthermore, we need to evaluate the expectation of the second to last term in Theorem 10, namely $E[\max_{d\in\mathcal{D}}\{d\circ q\}]$. For $\mathcal{D}=\mathcal{E}$ and q being exponentially distributed, using Lemma 1 with $k^i=0\ \forall i$, the expectation is bounded by $1+\ln n$. We hence get the following lower bound:

Corollary 11 (FPL lower-bounded by BEH) *For $\mathcal{D}=\mathcal{E}$ and any $\mathcal{S}$ and all k^i equal and $P[q^i]=e^{-q^i}$ for $q\geq 0$ and decreasing $\eta_t>0$, the expected loss of FPL is at most $\ln n/\eta_T$ lower than the loss of the best expert in hindsight:*

$$\ell_{1:T} \;\geq\; s_{1:T}^{min} - \frac{\ln n}{\eta_T}$$

The upper and lower bounds on $\ell_{1:T}$ (Theorem 4 and Corollaries 3 and 11) together show that

$$\frac{\ell_{1:t}}{s_{1:t}^{min}} \to 1 \qquad \text{if} \qquad \eta_t\to 0 \quad \text{and} \quad \eta_t\cdot s_{1:t}^{min}\to\infty \quad \text{and} \quad k^i=K\ \forall i. \tag{13}$$

For instance, $\eta_t=\sqrt{K/2s_{<t}^{min}}$. For $\eta_t=\sqrt{K/2(\ell_{<t}+1)}$ we proved the bound in Theorem 7(ii). Knowing that $\sqrt{K/2(\ell_{<t}+1)}$ converges to $\sqrt{K/2s_{<t}^{min}}$ due to (13), we can derive a bound similar to Theorem 7(ii) for $\eta_t=\sqrt{K/2s_{<t}^{min}}$. This choice for η_t has the advantage that we do not have to compute $\ell_{<t}$ (cf. Section 9), as also achieved by Theorem 8(ii).

We do not know whether Theorem 10 can be generalized to expert dependent complexities k^i.

8. Adaptive Adversary

In this section we show that bounds that hold against an oblivious adversary automatically also hold against an adaptive one.

Initial versus independent randomization. So far we assumed that the perturbations q are sampled only once at time $t=0$. As already indicated, under the expectation this is equivalent to generating a new perturbation q_t at each time step t, i.e. Theorems 4–9 remain valid for this case. While the former choice was favorable for the analysis, the latter has two advantages. First, repeated sampling of the perturbations guarantees better bounds with high probability (see next section). Second, if the losses are generated by an adaptive adversary (not to be confused with an adaptive learning rate) which has access to FPL's past decisions, then he may after some time figure out the initial random perturbation and use it to force FPL to have a large loss. We now show that the bounds for FPL remain valid, even in case of an adaptive adversary, if independent randomization $q \rightsquigarrow q_t$ is used.

Oblivious versus adaptive adversary. Recall the protocol for FPL: After each expert i made its prediction y_t^i, and FPL combined them to form its own prediction y_t^{FPL}, we observe x_t, and Loss(x_t,$y_t^{\cdots}$) is revealed for FPL's and each expert's prediction. For independent randomization, we have $y_t^{\text{FPL}}=y_t^{\text{FPL}}(x_{<t},y_{1:t},q_t)$. For an oblivious (non-adaptive) adversary, $x_t=x_t(x_{<t},y_{<t})$. Recursively inserting and eliminating the experts $y_t^i=y_t^i(x_{<t},y_{<t})$ and y_t^{FPL}, we get the dependencies

$$u_t := \text{Loss}(x_t, y_t^{\text{FPL}}) = u_t(x_{1:t}, q_t) \quad \text{and} \quad s_t^i := \text{Loss}(x_t, y_t^i) = s_t^i(x_{1:t}), \tag{14}$$

where $x_{1:t}$ is a "fixed" sequence. With this notation, Theorems 5–8 read $\ell_{1:T} \equiv E[\sum_{t=1}^T u_t(x_{1:t},q_t)] \leq f(x_{1:T})$ for all $x_{1:T} \in \mathcal{X}^T$, where $f(x_{1:T})$ is one of the r.h.s. in Theorems 5–8. Noting that f is independent of $q_{1:T}$, we can write this as

$$A_1 \leq 0, \quad \text{where} \quad A_t(x_{<t}, q_{<t}) := \max_{x_{t:T}} E_{q_{t:T}}\Big[\sum_{\tau=1}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T})\Big], \tag{15}$$

where $E_{q_{t:T}}$ is the expectation w.r.t. $q_t...q_T$ (keeping $q_{<t}$ fixed).

For an adaptive adversary, $x_t=x_t(x_{<t},y_{<t},y_{<t}^{\text{FPL}})$ can additionally depend on $y_{<t}^{\text{FPL}}$. Eliminating y_t^i and y_t^{FPL} we get, again, (14), but $x_t=x_t(x_{<t},q_{<t})$ is no longer fixed, but an (arbitrary) random function. So we have to replace x_t by $x_t(x_{<t},q_{<t})$ in (15) for $t=1..T$. The maximization is now a functional maximization over all functions $x_t(\cdot,\cdot)...x_T(\cdot,\cdot)$. Using "$\max_{x(\cdot)}E_q[g(x(q),q)] = E_q\max_x[g(x,q)]$," we can write this as

$$B_1 \overset{?}{\leq} 0, \quad \text{where} \quad B_t(x_{<t}, q_{<t}) := \max_{x_t} E_{q_t} ... \max_{x_T} E_{q_T}\Big[\sum_{\tau=1}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T})\Big]. \tag{16}$$

So, establishing $B_1 \leq 0$ would show that all bounds also hold in the adaptive case.

Lemma 12 (Adaptive=Oblivious) *Let $q_1...q_T \in \mathbb{R}^T$ be independent random variables, E_{q_t} be the expectation w.r.t. q_t, f any function of $x_{1:T} \in \mathcal{X}^T$, and u_t arbitrary functions of $x_{1:t}$ and q_t. Then, $A_t(x_{<t},q_{<t})=B_t(x_{<t},q_{<t})$ for all $1 \leq t \leq T$, where A_t and B_t are defined in (15) and (16). In particular, $A_1 \leq 0$ implies $B_1 \leq 0$.*

Proof. We prove $B_t = A_t$ by induction on t, which establishes the theorem. $B_T = A_T$ is obvious. Assume $B_t = A_t$. Then

$$\begin{aligned} B_{t-1} &= \max_{x_{t-1}} E_{q_{t-1}} B_t \;=\; \max_{x_{t-1}} E_{q_{t-1}} A_t \\ &= \max_{x_{t-1}} E_{q_{t-1}} \Big[\max_{x_{t:T}} E_{q_{t:T}} \Big[\sum_{\tau=1}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T}) \Big]\Big] \\ &= \max_{x_{t-1}} E_{q_{t-1}} \Big[\underbrace{\sum_{\tau=1}^{t-1} u_\tau(x_{1:\tau}, q_\tau)}_{\text{independent } x_{t:T} \text{ and } q_{t:T}} + \underbrace{\max_{x_{t:T}} E_{q_{t:T}} \Big[\sum_{\tau=t}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T}) \Big]}_{\text{independent } q_{t-1}\text{, since the } q_t \text{ are i.d.}} \Big] \\ &= \max_{x_{t-1}} \Big[\overbrace{E_{q_{t-1}} \Big[\sum_{\tau=1}^{t-1} u_\tau(x_{1:\tau}, q_\tau) \Big]} + \overbrace{\max_{x_{t:T}} E_{q_{t:T}} \Big[\sum_{\tau=t}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T}) \Big]} \Big] \\ &= \max_{x_{t-1}} \max_{x_{t:T}} E_{q_{t:T}} \Big[E_{q_{t-1}} \sum_{\tau=1}^{t-1} u_\tau(x_{1:\tau}, q_\tau) + \sum_{\tau=t}^{T} u_\tau(x_{1:\tau}, q_\tau) - f(x_{1:T}) \Big] \;=\; A_{t-1}. \end{aligned}$$

□

Corollary 13 (FPL Bounds for adaptive adversary) *Theorems 5–8 also hold for an adaptive adversary in case of independent randomization* $q \rightsquigarrow q_t$.

Lemma 12 shows that every bound of the form $A_1 \leq 0$ proven for an oblivious adversary, implies an analogous bound $B_1 \leq 0$ for an adaptive adversary. Note that this strong statement holds only for the *full observation game*, i.e. if after each time step we learn all losses. In partial observation games such as the Bandit case (Auer et al., 1995), our actual action may depend on our past action by means of our past observation, and the assertion no longer holds. In this case, FPL with an adaptive adversary can be analyzed as shown by McMahan and Blum (2004); Poland and Hutter (2005). Finally, y_t^{IFPL} can additionally depend on x_t, but the "reduced" dependencies (14) are the same as for FPL, hence, IFPL bounds also hold for adaptive adversary.

9. Miscellaneous

Bounds with high probability. We have derived several bounds for the expected loss $\ell_{1:T}$ of FPL. The *actual* loss at time t is $u_t = M(s_{<t} + \frac{k-q}{\eta_t}) \circ s_t$. A simple Markov inequality shows that the total actual loss $u_{1:T}$ exceeds the total expected loss $\ell_{1:T} = E[u_{1:T}]$ by a factor of $c > 1$ with probability at most $1/c$:

$$P[u_{1:T} \geq c \cdot \ell_{1:T}] \;\leq\; 1/c.$$

Randomizing independently for each t as described in the previous Section, the actual loss is $u_t = M(s_{<t} + \frac{k-q_t}{\eta_t}) \circ s_t$ with the same expected loss $\ell_{1:T} = E[u_{1:T}]$ as before. The advantage of independent randomization is that we can get a much better high-probability bound. We can exploit a Chernoff-Hoeffding bound (McDiarmid, 1989, Cor.5.2b), valid for arbitrary independent random variables $0 \leq u_t \leq 1$ for $t = 1, \ldots, T$:

$$P\Big[|u_{1:T} - E[u_{1:T}]| \geq \delta E[u_{1:T}] \Big] \;\leq\; 2\exp(-\tfrac{1}{3}\delta^2 E[u_{1:T}]), \qquad 0 \leq \delta \leq 1.$$

For $\delta=\sqrt{3c/\ell_{1:T}}$ we get

$$P[|u_{1:T}-\ell_{1:T}| \geq \sqrt{3c\ell_{1:T}}] \;\leq\; 2\mathrm{e}^{-c} \quad \text{as soon as} \quad \ell_{1:T} \geq 3c. \tag{17}$$

Using (17), the bounds for $\ell_{1:T}$ of Theorems 5–8 can be rewritten to yield similar bounds with high probability $(1-2\mathrm{e}^{-c})$ for $u_{1:T}$ with small extra regret $\propto\sqrt{c\cdot L}$ or $\propto\sqrt{c\cdot s^i_{1:T}}$. Furthermore, (17) shows that with probability 1, $u_{1:T}/\ell_{1:T}$ converges rapidly to 1 for $\ell_{1:T}\to\infty$. Hence we may use the easier to compute $\eta_t=\sqrt{K/2u_{<t}}$ instead of $\eta_t=\sqrt{K/2(\ell_{<t}+1)}$, likely with similar bounds on the regret.

Computational Aspects. It is easy to generate the randomized decision of FPL. Indeed, only a single initial exponentially distributed vector $q\in\boldsymbol{R}^n$ is needed. Only for self-confident $\eta_t\propto 1/\sqrt{\ell_{<t}}$ (see Theorem 7) we need to compute expectations explicitly. Given η_t, from $t\rightsquigarrow t+1$ we need to compute ℓ_t in order to update η_t. Note that $\ell_t=w_t\circ s_t$, where $w^i_t=P[I_t=i]$ and $I_t:=\mathrm{argmin}_{i\in\mathcal{E}}\{s^i_{<t}+\frac{k^i-q^i}{\eta_t}\}$ is the actual (randomized) prediction of FPL. With $s:=s_{<t}+k/\eta_t$, $P[I_t=i]$ has the following representation:

$$\begin{aligned}
P[I_t=i] &= P[s-\frac{q^i}{\eta_t}\leq s-\frac{q^j}{\eta_t}\;\forall j\neq i]\\
&= \int P[s-\frac{q^i}{\eta_t}=m \;\wedge\; s-\frac{q^j}{\eta_t}\geq m\;\forall j\neq i]dm\\
&= \int P[q^i=\eta_t(s^i-m)]\cdot\prod_{j\neq i}P[q^j\leq\eta_t(s^j-m)]dm\\
&= \int_{-\infty}^{s^{min}}\eta_t\mathrm{e}^{-\eta_t(s^i-m)}\prod_{j\neq i}(1-\mathrm{e}^{-\eta_t(s^j-m)})dm\\
&= \sum_{\mathcal{M}:\{i\}\subseteq\mathcal{M}\subseteq\mathcal{N}}\frac{(-)^{|\mathcal{M}|-1}}{|\mathcal{M}|}\mathrm{e}^{-\eta_t\sum_{j\in\mathcal{M}}(s^j-s^{min})}.
\end{aligned}$$

In the last equality we expanded the product and performed the resulting exponential integrals. For finite n, the second to last one-dimensional integral should be numerically feasible. Once the product $\prod_{j=1}^n(1-\mathrm{e}^{-\eta_t(s^j-m)})$ has been computed in time $O(n)$, the argument of the integral can be computed for each i in time $O(1)$, hence the overall time to compute ℓ_t is $O(c\cdot n)$, where c is the time to numerically compute one integral. For infinite n, the last sum may be approximated by the dominant contributions. Alternatively, one can modify the algorithm by considering only a finite pool of experts in each time step; see next paragraph. The expectation may also be approximated by (Monte Carlo) sampling I_t several times.

Recall that approximating $\ell_{<t}$ can be avoided by using $s^{min}_{<t}$ (Theorem 8) or $u_{<t}$ (bounds with high probability) instead.

Finitized expert pool. In the case of an infinite expert class, FPL has to compute a minimum over an infinite set in each time step, which is not directly feasible. One possibility to address this is to choose the experts from a *finite pool* in each time step. This is the case in the algorithm of Gentile (2003), and also discussed by Littlestone and Warmuth (1994). For FPL, we can obtain this behavior by introducing an *entering time* $\tau^i\geq 1$ for each expert. Then expert i is not considered for $i<\tau^i$. In the bounds, this leads to an additional $\frac{1}{\eta_T}$ in Theorem 2 and Corollary 3 and a further additional τ^i in the final bounds (Theorems 5–8), since we must add the regret of the best expert in hindsight which has already entered the game and the best expert in hindsight at all. Selecting $\tau^i = k^i$ implies bounds for FPL with entering times similar to the ones we derived here. The details and proofs for this construction can be found in (Poland and Hutter, 2005).

Deterministic prediction and absolute loss. Another use of w_t from the second last paragraph is the following: If the decision space is $\mathcal{D}=\Delta$, then FPL may make a deterministic decision $d=w_t\in\Delta$ at time t with bounds now holding for sure, instead of selecting e_i with probability w_t^i. For example for the absolute loss $s_t^i = |x_t - y_t^i|$ with observation $x_t \in [0,1]$ and predictions $y_t^i \in [0,1]$, a master algorithm predicting deterministically $w_t{\circ}y_t\in[0,1]$ suffers absolute loss $|x_t-w_t{\circ}y_t|\leq\sum_i w_t^i|x_t-y_t^i|=\ell_t$, and hence has the same (or better) performance guarantees as FPL. In general, masters can be chosen deterministic if prediction space $\mathcal{Y}$ and loss-function $\text{Loss}(x,y)$ are convex. For $x_t, y_t^i \in \{0,1\}$, the absolute loss $|x_t - p_t|$ of a master deterministically predicting $p_t \in [0,1]$ actually coincides with the p_t-expected 0/1 loss of a master predicting 1 with probability p_t. Hence a regret bound for the absolute loss also implies the same regret for the 0/1 loss.

10. Discussion and Open Problems

How does FPL compare with other expert advice algorithms? We briefly discuss four issues, summarized in Table 2.

Static bounds. Here the coefficient of the regret term $\sqrt{KL}$, referred to as the *leading constant* in the sequel, is 2 for FPL (Theorem 5). It is thus a factor of $\sqrt{2}$ worse than the Hedge bound for arbitrary loss by Freund and Schapire (1997), which is sharp in some sense (Vovk, 1995). This is the price one pays for the elegance of FPL. There is evidence that this (worst-case) difference really exists and is not only a proof artifact. For special loss functions, the bounds can sometimes be improved, e.g. to a leading constant of 1 in the static (randomized) WM case with 0/1 loss (Cesa-Bianchi et al., 1997)[3]. Because of the structure of the FPL algorithm however, it is questionable if corresponding bounds hold there.

Dynamic bounds. Not knowing the right learning rate in advance usually costs a factor of $\sqrt{2}$. This is true for Hannan's algorithm (Kalai and Vempala, 2003) as well as in all our cases. Also for binary prediction with uniform complexities and 0/1 loss, this result has been established recently – Yaroshinsky et al. (2004) show a dynamic regret bound with leading constant $\sqrt{2}(1+\varepsilon)$. Remarkably, the best dynamic bound for a WM variant proven by Auer et al. (2002) has a leading constant $2\sqrt{2}$, which matches ours. Considering the difference in the static case, we therefore conjecture that a bound with leading constant of 2 holds for a dynamic Hedge algorithm.

General weights. While there are several dynamic bounds for uniform weights, the only previous result for non-uniform weights we know of is (Gentile, 2003, Cor.16), which gives the dynamic bound $\ell_{1:T}^{\text{Gentile}} \leq s_{1:T}^i + i + O\left[\sqrt{(s_{1:T}^i + i)\ln(s_{1:T}^i + i)}\right]$ for a p-norm algorithm for the absolute loss. This is comparable to our bound for rapidly decaying weights $w^i = \exp(-i)$, i.e. $k^i = i$. Our hierarchical FPL bound in Theorem 9 (b) generalizes this to arbitrary weights and losses and strengthens it, since both, asymptotic order and leading constant, are smaller.

It seems that the analysis of all experts algorithms, including Weighted Majority variants and FPL, gets more complicated for general weights together with adaptive learning rate, because the

3. While FPL and Hedge and WMR (Littlestone and Warmuth, 1994) can sample an expert without knowing its prediction, Cesa-Bianchi et al. (1997) need to know the experts' predictions. Note also that for many (smooth) loss-functions like the quadratic loss, finite regret can be achieved (Vovk, 1990).

η	Loss	conjecture	Lower Bound	Upper Bound
static	0/1	1	1?	1 (Cesa-Bianchi et al., 1997)
static	any	$\sqrt{2}$!	$\sqrt{2}$ (Vovk, 1995)	$\sqrt{2}$ (Hedge), 2 (FPL)
dynamic	0/1	$\sqrt{2}$	1? (Hutter, 2003b)	$\sqrt{2}$ (Yaroshinsky) , $2\sqrt{2}$ (Auer 2002)
dynamic	any	2	$\sqrt{2}$ (Vovk, 1995)	$2\sqrt{2}$ (FPL), 2 (Hutter, 2003b, Bayes)

Table 2: Comparison of the constants c in regrets $c\sqrt{\text{Loss}\times\ln n}$ for various settings and algorithms.

choice of the learning rate must account for both the weight of the best expert (in hindsight) and its loss. Both quantities are not known in advance, but may have a different impact on the learning rate: While increasing the current loss estimate always decreases η_t, the optimal learning rate for an expert with higher complexity would be larger. On the other hand, all analyses known so far require decreasing η_t. Nevertheless we conjecture that the bounds $\propto\sqrt{Tk^i}$ and $\propto\sqrt{s^i_{1:T}k^i}$ also hold without the hierarchy trick, probably by using expert dependent learning rate η^i_t.

Comparison to Bayesian sequence prediction. We can also compare the *worst-case* bounds for FPL obtained in this work to similar bounds for *Bayesian sequence prediction*. Let $\{\nu_i\}$ be a class of probability distributions over sequences and assume that the true sequence is sampled from $\mu\in\{\nu_i\}$ with complexity k^μ ($\sum_i e^{-k^{\nu_i}}\leq 1$). Then it is known that the Bayes optimal predictor based on the $e^{-k^{\nu_i}}$-weighted mixture of ν_i's has an expected total loss of at most $L^\mu+2\sqrt{L^\mu k^\mu}+2k^\mu$, where L^μ is the expected total loss of the Bayes optimal predictor based on μ (Hutter, 2003a, Thm.2), (Hutter, 2004b, Thm.3.48). Using FPL, we obtained the same bound except for the leading order constant, but for any sequence independently of the assumption that it is generated by μ. This is another indication that a PEA bound with leading constant 2 could hold. See Hutter (2004a), Hutter (2003b, Sec.6.3) and Hutter (2004b, Sec.3.7.4) for a more detailed comparison of Bayes bounds with PEA bounds.

Acknowledgments

This work was supported by SNF grant 2100-67712.02.

References

P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proc. 36th Annual Symposium on Foundations of Computer Science (FOCS 1995)*, pages 322–331, Los Alamitos, CA, 1995. IEEE Computer Society Press.

P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64:48–75, 2002.

P. Auer and C. Gentile. Adaptive and self-confident on-line learning algorithms. In *Proc. 13th Conference on Computational Learning Theory*, pages 107–117. Morgan Kaufmann, San Francisco, 2000.

N. Cesa-Bianchi, Y. Freund, D. Haussler, D. Helmbold, R. Schapire, and M. K. Warmuth. How to use expert advice. *Journal of the ACM*, 44(3):427–485, 1997.

Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

C. Gentile. The robustness of the p-norm algorithm. *Machine Learning*, 53(3):265–299, 2003.

J. Hannan. Approximation to Bayes risk in repeated plays. In M. Dresher, A. W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games 3*, pages 97–139. Princeton University Press, 1957.

M. Hutter. Convergence and loss bounds for Bayesian sequence prediction. *IEEE Trans. on Information Theory*, 49(8):2061–2067, 2003a. URL `http://arxiv.org/abs/cs.LG/0301014`.

M. Hutter. Optimality of universal Bayesian prediction for general loss and alphabet. *Journal of Machine Learning Research*, 4:971–1000, 2003b. URL `http://arxiv.org/abs/cs.LG/0311014`.

M. Hutter. Online prediction – Bayes versus experts. Technical report, July 2004a. URL `http://www.idsia.ch/~marcus/ai/bayespea.htm`. Presented at the EU PASCAL Workshop on Learning Theoretic and Bayesian Inductive Principles (LTBIP-2004).

M. Hutter. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Springer, Berlin. 300 pages, 2004b. URL `http://www.idsia.ch/~marcus/ai/uaibook.htm`.

M. Hutter and J. Poland. Prediction with expert advice by following the perturbed leader for general weights. In *Proc. 15th International Conf. on Algorithmic Learning Theory (ALT-2004)*, volume 3244 of *LNAI*, pages 279–293, Padova, 2004. Springer, Berlin. URL `http://arxiv.org/abs/cs.LG/0405043`.

A. Kalai and S. Vempala. Efficient algorithms for online decision. In *Proc. 16th Annual Conference on Learning Theory (COLT-2003)*, Lecture Notes in Artificial Intelligence, pages 506–521, Berlin, 2003. Springer.

N. Littlestone and M. K. Warmuth. The weighted majority algorithm. In *30th Annual Symposium on Foundations of Computer Science*, pages 256–261, Research Triangle Park, North Carolina, 1989. IEEE.

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

C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics*, 141, London Mathematical Society Lecture Notes Series:148–188, 1989.

H. B. McMahan and A. Blum. Online geometric optimization in the bandit setting against an adaptive adversary. In *17th Annual Conference on Learning Theory (COLT)*, volume 3120 of *LNCS*, pages 109–123. Springer, 2004.

J. Poland and M. Hutter. Master algorithms for active experts problems based on increasing loss values. In *Annual Machine Learning Conference of Belgium and the Netherlands (Benelearn-2005)*, Enschede, 2005. URL `http://arxiv.org/abs/cs.LG/0502067`.

V. G. Vovk. Aggregating strategies. In *Proc. Third Annual Workshop on Computational Learning Theory*, pages 371–383, Rochester, New York, 1990. ACM Press.

V. G. Vovk. A game of prediction with expert advice. In *Proc. 8th Annual Conference on Computational Learning Theory*, pages 51–60. ACM Press, New York, NY, 1995.

R. Yaroshinsky, R. El-Yaniv, and S. Seiden. How to better use expert advice. *Machine Learning*, 55 (3):271–309, 2004.