Prediction with Expert Advice by Following the Perturbed Leader for General Weights

Marcus Hutter and **Jan Poland**

IDSIA, Galleria 2, CH-6928 Manno-Lugano, Switzerland*
{marcus,jan}@idsia.ch, http://www.idsia.ch/~{marcus,jan}

12 May 2004

Abstract

When applying aggregating strategies to Prediction with Expert Advice, the learning rate must be adaptively tuned. The natural choice of $\sqrt{\text{complexity/current loss}}$ renders the analysis of Weighted Majority 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 [KV03] (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, hierarchy of experts, expected and high probability bounds, general alphabet and loss, online sequential prediction,

*This work was supported by SNF grant 2100-67712.02.

1 Introduction

The theory of Prediction with Expert Advice (PEA) has rapidly developed in the recent past. Starting with the Weighted Majority (WM) algorithm of Littlestone and Warmuth [LW89, LW94] and the aggregating strategy of Vovk [Vov90], 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, [CB97] established the so-called doubling trick to make the learning rate coarsely adaptive. A little later, incrementally adaptive algorithms were developed [AG00, ACBG02, YEYS04, Gen03]. 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 [Gen03], see the discussion section). 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,... –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. in [FS97], 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 [Han57]. 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 [KV03]. 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 we study the FPL algorithm for PEA. The problems of WM algorithms mentioned above are addressed: We consider 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. Our result even seems to be the first for *equal weights* and *arbitrary losses*, however the proof technique from [ACBG02] is likely to carry over to this case.

This paper is structured as follows. In Section 2 we give the basic definitions. Sections 3 and 4 derive the main analysis tools, following the lines of [KV03], but with some important extensions. They are applied in order to prove various upper bounds in Section 5. Section 6 proposes a hierarchical procedure to improve the bounds for non-uniform weights. Section 7 treats some additional issues. Finally, in Section 8 we discuss our results, compare them to references, and state some open problems.

2 Setup & Notation

Setup. Prediction with Expert Advice proceeds as follows. We are asked to perform sequential predictions $y_t \in \mathcal{Y}$ at times $t=1,2,...$. At each time step t, we have access to the predictions $(y_t^i)_{1\leq i\leq n}$ of n experts $\{e_1,...,e_n\}$. 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 $n\in I\!N\cup\{\infty\}$ experts, each of which is assigned a known complexity $k^i\geq 0$. Usually we require $\sum_i \mathrm{e}^{-k^i}\leq 1$, 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 rather of complexities 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[1] $s_t^i=\mathrm{Loss}(x_t,y_t^i)\in[0,1]$, and $s_t=(s_t^i)_{1\leq i\leq n}$ is the vector of all losses at time t. Let $s_{<t}=s_1+...+s_{t-1}$ (respectively $s_{1:t}=s_1+...+s_t$) 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 $\mathcal{S}\subset I\!R^n$ be the *state space* and $\mathcal{D}\subset I\!R^n$ 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

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

[1] The setup, analysis and results easily scale to $s_t^i\in[0,S]$ for $S>0$ other than 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 a single expert, and $s_t\in[0,1]^n$, so states are identified with losses. Only Theorem 2 will be stated in terms of general decision space, where we require that all minima are attained.[2] Our main focus is $\mathcal{D}=\mathcal{E}$. However, 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 $s=\binom{0\,1\,0\,1\,0\,1...}{\frac{1}{2}\,0\,1\,0\,1\,0...}$, FL always chooses the wrong prediction [KV03]. 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 7). For the main analysis however, the latter is more convenient. Due to linearity of expectations, these two possibilities are equivalent when dealing with *expected losses*, so we can henceforth assume without loss of generality one initial perturbation q.

The FPL algorithm is defined as follows:

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

Other than $s_{<t}$, k and q, FPL depends on the *learning rate* ε_t. We will give choices for ε_t in Section 5, after having established the main tools for the analysis. The expected loss at time t of FPL is $\ell_t:=E[M(s_{<t}+\frac{k-q}{\varepsilon_t})\circ s_t]$. The key idea in the FPL analysis is the use of an intermediate predictor *IFPL* (for *Implicit or Infeasible FPL*). IFPL predicts according to $M(s_{1:t}+\frac{k-q}{\varepsilon_t})$, thus under the knowledge of s_t (which is of course not available in reality). By $r_t:=E[M(s_{1:t}+\frac{k-q}{\varepsilon_t})\circ s_t]$ 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.

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.

[2]Apparently, there is no natural condition on $\mathcal{D}$ and/or $\mathcal{S}$ which guarantees the existence of all minima for $n=\infty$.

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 $s_t = \text{Loss}(x_t, y_t^i)_{1\leq i\leq n}$. 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 to FPL's past decisions, then FPL must use independent randomization at each time step in order to achieve good regret bounds.

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 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* $k^i \in I\!R$ *be real numbers with* $u := \sum_{i=1}^n \mathrm{e}^{-k^i}$*. Then*

$$E[\max_i \{q^i - k^i\}] \;\;\leq\;\; 1 + \ln u.$$

Proof. Using $P[q^i \geq b] \leq \mathrm{e}^{-b}$ for $b \in I\!R$ we get

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

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 assertion. □

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 [KV03, Lem.3] to arbitrary weights.

Theorem 2 (IFPL bounded by BEH) *Let* $\mathcal{D} \subseteq I\!R^n$*,* $s_t \in I\!R^n$ *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* $q,k \in I\!R^n$. *If* $\varepsilon_t > 0$ *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:*

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

Proof. For notational convenience, let $\varepsilon_0 = \infty$ and $\tilde{s}_{1:t} = s_{1:t} + \frac{k-q}{\varepsilon_t}$. Consider the losses $\tilde{s}_t = s_t + (k-q)(\frac{1}{\varepsilon_t} - \frac{1}{\varepsilon_{t-1}})$ for the moment. We first show by induction on T that the infeasible predictor $M(\tilde{s}_{1:t})$ has zero regret, i.e.

$$\sum_{t=1}^{T} M(\tilde{s}_{1:t}) \circ \tilde{s}_t \leq M(\tilde{s}_{1:T}) \circ \tilde{s}_{1:T}. \tag{2}$$

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

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

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 (2), we obtain

$$\sum_{t=1}^{T} M(\tilde{s}_{1:t}) \circ s_t \;\leq\; M(\tilde{s}_{1:T}) \circ \tilde{s}_{1:T} - \sum_{t=1}^{T} M(\tilde{s}_{1:t}) \circ (k-q)\Big(\frac{1}{\varepsilon_t} - \frac{1}{\varepsilon_{t-1}}\Big) \tag{3}$$

Moreover, by minimality of M,

$$\begin{aligned} M(\tilde{s}_{1:T}) \circ \tilde{s}_{1:T} &\leq M\Big(s_{1:T} + \frac{k}{\varepsilon_T}\Big) \circ \Big(s_{1:T} + \frac{k-q}{\varepsilon_T}\Big) \\ &= \min_{d\in\mathcal{D}}\Big\{d \circ (s_{1:T} + \frac{k}{\varepsilon_T})\Big\} - M\Big(s_{1:T} + \frac{k}{\varepsilon_T}\Big) \circ \frac{q}{\varepsilon_T} \end{aligned} \tag{4}$$

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

$$\begin{aligned} \sum_{t=1}^{T}(\frac{1}{\varepsilon_t} - \frac{1}{\varepsilon_{t-1}}) M(\tilde{s}_{1:t}) \circ (q-k) &\leq \sum_{t=1}^{T}(\frac{1}{\varepsilon_t} - \frac{1}{\varepsilon_{t-1}}) M(k-q) \circ (q-k) \\ &= \frac{1}{\varepsilon_T} M(k-q)\circ(q-k) = \frac{1}{\varepsilon_T}\max_{d\in\mathcal{D}}\{d\circ(q-k)\} \end{aligned} \tag{5}$$

Inserting (4) and (5) back into (3) we obtain the assertion. □

Assuming q random with $E[q^i] = 1$ and taking the expectation in Theorem 2, the last term reduces to $-\frac{1}{\varepsilon_T}\sum_{i=1}^{n} M(s_{1:T} + \frac{k}{\varepsilon_T})^i$. 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 $-\frac{1}{\varepsilon_T}$ (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 $\varepsilon_t>0$, the expected loss of the infeasible FPL exceeds the loss of expert i by at most k^i/ε_T:*

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

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

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

where $\varepsilon^{min}_T := \min_i\{\varepsilon^i_T\}$. For example, for $\varepsilon^i_t = \sqrt{k^i/t}$, additionally assuming $k^i \geq 1$ $\forall i$, we get the desired bound $s^i_{1:T} + \sqrt{T\cdot(k^i+1)}$. Unfortunately we were not able to generalize Theorem 4 to expert-dependent ε, 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 $\ell_t = E[M(s_{<t} + \frac{k-q}{\varepsilon_t})\circ s_t]$ is the expected loss of FPL at time t and $r_t = E[M(s_{1:t} + \frac{k-q}{\varepsilon_t})\circ s_t]$ is the expected loss of IFPL at time t.

Theorem 4 (FPL bounded by IFPL) *For $\mathcal{D}=\mathcal{E}$ and $0\leq s^i_t \leq 1$ $\forall i$ 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 $\mathrm{e}^{\varepsilon_t}>1$ larger than for the infeasible FPL:*

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

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

Proof. Let $s = s_{<t} + \frac{1}{\varepsilon}k$ 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 $\varepsilon=\varepsilon_t$. We now define the random variables $I := \mathrm{argmin}_i\{s^i - \frac{1}{\varepsilon}q^i\}$ and $J := \mathrm{argmin}_i\{s^i + s^i_t - \frac{1}{\varepsilon}q^i\}$, where $0\leq s^i_t \leq 1$ $\forall i$. Furthermore, for fixed vector $x\in I\!R^n$ and fixed j we define $m := \min_{i\neq j}\{s^i - \frac{1}{\varepsilon}x^i\} \leq \min_{i\neq j}\{s^i + s^i_t - \frac{1}{\varepsilon}x^i\} =: m'$. With this notation and using the independence of q^j from q^i for all $i\neq j$, we get

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

$$\leq\ e^\varepsilon P[q^j \geq \varepsilon(s^j - m + 1)]\ \leq\ e^\varepsilon P[q^j \geq \varepsilon(s^j + s_t^j - m')]$$

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

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

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

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

Remark. As in [KV03], 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 $\ell_t \leq e^{\varepsilon_t A} r_t$. 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 ε_t

Throughout this section, we assume

$$\mathcal{D} = \mathcal{E}, \quad s_t \in [0,1]^n\ \forall t, \quad P[q] = e^{-\sum_i q^i} \text{ for } q \geq 0, \quad \text{and} \quad \sum_i e^{-k^i} \leq 1. \qquad (6)$$

We distinguish *static* and *dynamic* bounds. Static bounds refer to a constant $\varepsilon_t \equiv \varepsilon$. Since this value has to be chosen in advance, a static choice of ε_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 ε_t depending only on t and/or past observations, is necessary.

Theorem 5 (FPL bound for static $\varepsilon_t = \varepsilon \propto 1/\sqrt{L}$) *Assume (6) holds, then the expected loss ℓ_t of feasible FPL, which employs the prediction of the expert i minimizing $s_{<t}^i + \frac{k^i - q^i}{\varepsilon_t}$, is bounded by the loss of the best expert in hindsight in the following way:*

i) For $\varepsilon_t = \varepsilon = 1/\sqrt{L}$ *with* $L \geq \ell_{1:T}$ *we have*
$\ell_{1:T}\ \leq\ s_{1:T}^i + \sqrt{L}(k^i+1) \quad \forall i$

ii) For $\varepsilon_t = \sqrt{K/L}$ *with* $L \geq \ell_{1:T}$ *and* $k^i \leq K\ \forall i$ *we have*
$\ell_{1:T}\ \leq\ s_{1:T}^i + 2\sqrt{LK} \quad \forall i$

iii) For $\varepsilon_t = \sqrt{k^i/L}$ *with* $L \geq \max\{s_{1:T}^i, k^i\}$ *we have*
$\ell_{1:T}\ \leq\ s_{1:T}^i + 2\sqrt{Lk^i} + 3k^i$

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 $\varepsilon_t = \sqrt{K/L}$ and $L \geq \ell_{1:T}$, from Theorem 4 and Corollary 3, we get

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

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 $\varepsilon = \sqrt{k^i/L} \leq 1$ we get

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

□

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 ε_t is necessary. We first present bounds with regret $\propto \sqrt{T}$, thereafter with regret $\propto \sqrt{s_{1:T}^i}$.

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

i) For $\varepsilon_t = 1/\sqrt{t}$ *we have* $\ell_{1:T} \leq s_{1:T}^i + \sqrt{T}(k^i+2) \quad \forall i$

ii) For $\varepsilon_t = \sqrt{K/2t}$ *and* $k^i \leq K \; \forall i$ *we have* $\ell_{1:T} \leq s_{1:T}^i + 2\sqrt{2TK} \quad \forall i$

Proof. For $\varepsilon_t = \sqrt{K/2t}$, 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

$$\ell_{1:T} - r_{1:T} \leq \sum_{t=1}^{T} \varepsilon_t \leq \sqrt{2TK} \quad \text{and} \quad r_{1:T} - s_{1:T}^i \leq k^i/\varepsilon_T = k^i\sqrt{\frac{2T}{K}}$$

Combining both, we get $\ell_{1:T} - s_{1:T}^i \leq \sqrt{2T}(\sqrt{K} + k^i/\sqrt{K})$. (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 $\varepsilon_t \propto 1/\sqrt{L_t}$. Such choice of ε_t is also called *self-confident* [ACBG02].

Theorem 7 (FPL bound for self-confident $\varepsilon_t \propto 1/\sqrt{\ell_{<t}}$) *Assume (6) holds.*

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

Proof. Using $\varepsilon_t = \sqrt{K/2(\ell_{<t}{+}1)} \leq \sqrt{K/2\ell_{1:t}}$ 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

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

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

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

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

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

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

The proofs of results similar to (ii) for WM for 0/1 loss all fill several pages [ACBG02, YEYS04]. 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 7).

Theorem 8 (FPL bound for adaptive $\varepsilon_t \propto 1/\sqrt{s^{min}_{<t}}$) *Assume (6) holds.*

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

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

Proof. Similar to the proof of the previous theorem, but more technical. □

The bound (i) is a complete square, and also the bounds of Theorem 7 when adding 1 to them. Hence the bounds can be written as $\sqrt{\ell_{1:T}}\leq\sqrt{s_{1:T}^i}+\sqrt{2}(k^i+2)$ and $\sqrt{\ell_{1:T}}\leq\sqrt{s_{1:T}^i+1}+\sqrt{8K}$ and $\sqrt{\ell_{1:T}}\leq\sqrt{s_{1:T}^i+1}+\sqrt{2}(k^i+1)$, respectively, hence the $\sqrt{\text{Loss}}$-regrets are bounded for $T\to\infty$.

Remark. The same analysis as for Theorems [5-8](ii) applies to general $\mathcal{D}$, using $\ell_t\leq e^{\varepsilon_t n}r_t$ instead of $\ell_t\leq e^{\varepsilon_t}r_t$, 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 $\propto\sqrt{s_{1:T}^iK}$ 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. It is likely that improved regret bounds $\propto\sqrt{Tk^i}$ and $\propto\sqrt{s_{1:T}^ik^i}$ as in the finite case 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 $I_t^K:=\text{argmin}_{i\in\mathcal{E}^K}\{s_{<t}^i+\frac{k^i-q^i}{\varepsilon_t^K}\}$ with $\varepsilon_t^K:=\sqrt{K/2t}$ 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:

$$E[u_{1:T}^K]\;=\;\ell_{1:T}^K\;\leq\;s_{1:T}^i+2\sqrt{2TK}\quad\forall i\in\mathcal{E}^K\quad\forall K\in\mathbb{N}. \tag{7}$$

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 $\tilde{I}_t := \arg\min_{K\in I\!N}\{\tilde{s}^K_{<t} + \frac{\tilde{k}^K - \tilde{q}^K}{\tilde{\varepsilon}_t}\}$ with $\tilde{\varepsilon}_t = 1/\sqrt{t}$ and $\tilde{k}^K = \frac{1}{2} + 2\ln K$ (implying $\sum_{K=1}^\infty \mathrm{e}^{-\tilde{k}^K} \leq 1$). 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 (7) 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 ε_t) *The hierarchical $\widetilde{FPL}$ employs at time t the prediction of expert $i_t := I_t^{\tilde{I}_t}$, where*

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

Under assumptions (6) and $P[\tilde{q}] = \mathrm{e}^{-\sum_K \tilde{q}^K}$, 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* $\varepsilon^K_t = \sqrt{K/2t}$ *and* $\tilde{\varepsilon}_t = 1/\sqrt{t}$ *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 $\tilde{\varepsilon}_t$ as in (i) and ε^K_t 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 ε.

7 Miscellaneous

Lower Bound on FPL. For finite n, a lower bound on FPL similar to the upper bound in Theorem 2 can also be proven. For any $\mathcal{D} \subseteq I\!R^n$ and $s_t \in I\!R$ such that the required extrema exist, $q\in I\!R^n$, and $\varepsilon_t>0$ decreasing, the loss of FPL for uniform complexities can be lower bounded in terms of the best predictor in hindsight plus/minus additive corrections:

$$\sum_{t=1}^T M(s_{<t}-\frac{q}{\varepsilon_t})\circ s_t \geq \min_{d\in\mathcal{D}}\{d\circ s_{1:T}\} - \frac{1}{\varepsilon_T}\max_{d\in\mathcal{D}}\{d\circ q\} + \sum_{t=1}^T(\frac{1}{\varepsilon_t}-\frac{1}{\varepsilon_{t-1}})M(s_{<t})\circ q \quad (8)$$

For $\mathcal{D}=\mathcal{E}$ and any $\mathcal{S}$ and all k^i equal and $P[q^i]=\mathrm{e}^{-q^i}$ for $q\geq 0$ and decreasing $\varepsilon_t>0$, this reduces to

$$\ell_{1:T} \;\geq\; s_{1:T}^{min} - \frac{\ln n}{\varepsilon_T} \quad (9)$$

The upper and lower bounds on $\ell_{1:T}$ (Theorem 4 and Corollary 3 and (9)) together show that

$$\frac{\ell_{1:t}}{s_{1:t}^{min}}\to 1 \qquad \text{if} \qquad \varepsilon_t\to 0 \quad\text{and}\quad \varepsilon_t\!\cdot\! s_{1:t}^{min}\to\infty \quad\text{and}\quad k^i=K\;\forall i. \quad (10)$$

For instance, $\varepsilon_t=\sqrt{K/2s_{<t}^{min}}$. For $\varepsilon_t=\sqrt{K/2(\ell_{<t}+1)}$ 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 (10), we can derive a bound similar to Theorem 7(ii) for $\varepsilon_t=\sqrt{K/2s_{<t}^{min}}$. This choice for ε_t has the advantage that we do not have to compute $\ell_{<t}$ (see below), as also achieved by Theorem 8(ii). We do not know whether (8) can be generalized to expert dependent complexities k^i.

Initial versus independent randomization. So far we assumed that the perturbations 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, if the losses are generated by an adaptive adversary, then he may after some time figure out the initial random perturbation and use it to force FPL to have a large loss. On the other hand, for independent randomization, one can show that our bounds remain valid, even if the environment has access to FPL's past predictions. Second, repeated sampling of the perturbations guarantees better bounds with high probability.

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 $u_t=M(s_{<t}+\frac{k-q}{\varepsilon_t})\circ s_t$. 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 paragraph, the actual loss is $u_t=M(s_{<t}+\frac{k-q_t}{\varepsilon_t})\circ s_t$ 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 [McD89, Cor.5.2b], valid for arbitrary independent random variables $0\leq u_t\leq 1$ for $t=1,...,T$:

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

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

$$P[|u_{1:T}-\ell_{1:T}|\geq\sqrt{3c\ell_{1:T}}] \;\leq\; 2\mathrm{e}^{-c} \quad \text{as soon as} \quad \ell_{1:T}\geq 3c. \tag{11}$$

Using (11), 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, (11) shows that with high probability, $u_{1:T}/\ell_{1:T}$ converges rapidly to 1 for $\ell_{1:T}\to\infty$. Hence we may use the easier to compute $\varepsilon_t=\sqrt{K/2u_{<t}}$ instead of $\varepsilon_t=\sqrt{K/2(\ell_{<t}+1)}$, 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 $q\in I\!R^n$ is needed. Only for adaptive $\varepsilon_t\propto 1/\sqrt{\ell_{<t}}$ (see Theorem 7) we need to compute expectations explicitly. Given ε_t, from $t\rightsquigarrow t+1$ we need to compute ℓ_t in order to update ε_t. Note that $\ell_t=w_t\circ s_t$, where $w_t^i=P[I_t=i]$ and $I_t:=\mathrm{argmin}_{i\in\mathcal{E}}\{s^i_{<t}+\frac{k^i-q^i}{\varepsilon_t}\}$ is the actual (randomized) prediction of FPL. With $s:=s_{<t}+k/\varepsilon_t$, $P[I_t=i]$ has the following representation:

$$P[I_t\!=i] \;=\int_{-\infty}^{s^{min}}\!\!\varepsilon_t e^{-\varepsilon_t(s^i-m)}\prod_{j\neq i}(1-e^{-\varepsilon_t(s^j-m)})dm \;=\sum_{\mathcal{M}:\{i\}\subseteq\mathcal{M}\subseteq\mathcal{N}}\frac{(-)^{|\mathcal{M}|-1}}{|\mathcal{M}|}e^{-\varepsilon_t\sum_{j\in\mathcal{M}}(s^j-s^{min})}$$

In the last equality we expanded the product and performed the resulting exponential integrals. For finite n, the one-dimensional integral should be numerically feasible. Once the product $\prod_{j=1}^n(1-\mathrm{e}^{-\varepsilon_t(s^j-m)})$ 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[3] n, the last sum may be approximated by the dominant contributions. 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.

Deterministic prediction and absolute loss. Another use of w_t from the 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 $\mathrm{Loss}(x,y)$ are convex.

[3]For practical realizations in case of infinite n, one must use finite subclasses of increasing size, compare [LW94].

8 Discussion and Open Problems

How does FPL compare with other expert advice algorithms? We briefly discuss four issues.

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 [FS97], which is sharp in some sense [Vov95]. For special loss functions, the bounds can sometimes be improved, e.g. to a leading constant of 1 in the static WM case with 0/1 loss [CB97].

Dynamic bounds. Not knowing the right learning rate in advance usually costs a factor of $\sqrt{2}$. This is true for Hannan's algorithm [KV03] as well as in all our cases. Also for binary prediction with uniform complexities and 0/1 loss, this result has been established recently – [YEYS04] show a dynamic regret bound with leading constant $\sqrt{2}(1+\varepsilon)$. Remarkably, the best dynamic bound for a WM variant proven in [ACBG02] 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 result for non-uniform weights we know of is [Gen03, Cor.16], which gives a dynamic bound for a p-norm algorithm for the absolute loss if the weights are rapidly decaying. Our hierarchical FPL bound in Theorem 9 (b) generalizes it to arbitrary weights and losses and strengthens it, since both, asymptotic order and leading constant, are smaller. Also the FPL analysis gets more complicated for general weights. 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 ε^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 ($\sum_i 2^{-k^{\nu_i}}\leq 1$). Then it is known that the Bayes-optimal predictor based on the $2^{-k^{\nu_i}}$-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 [Hut03a, Thm.2]. 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 [Hut03b, Sec.6.3] for a more detailed comparison of Bayes bounds with PEA bounds.

References

[ACBG02] P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002.

[AG00] P. Auer and C. Gentile. Adaptive and self-confident on-line learning algorithms. In *Proceedings of the 13th Conference on Computational Learning Theory*, pages 107–117. Morgan Kaufmann, San Francisco, 2000.

[CB97] N. Cesa-Bianchi et al. How to use expert advice. *Journal of the ACM*, 44(3):427–485, 1997.

[FS97] 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.

[Gen03] C. Gentile. The robustness of the p-norm algorithm. *Machine Learning*, 53(3):265–299, 2003.

[Han57] 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.

[Hut03a] M. Hutter. Convergence and loss bounds for Bayesian sequence prediction. *IEEE Transactions on Information Theory*, 49(8):2061–2067, 2003.

[Hut03b] M. Hutter. Optimality of universal Bayesian prediction for general loss and alphabet. *Journal of Machine Learning Research*, 4:971–1000, 2003.

[KV03] A. Kalai and S. Vempala. Efficient algorithms for online decision. In *Proceedings of the 16th Annual Conference on Learning Theory (COLT-2003)*, Lecture Notes in Artificial Intelligence, pages 506–521, Berlin, 2003. Springer.

[LW89] 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.

[LW94] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.

[McD89] C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics*, 141, London Mathematical Society Lecture Notes Series:148–188, 1989.

[Vov90] V. G. Vovk. Aggregating strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory*, pages 371–383, Rochester, New York, 1990. ACM Press.

[Vov95] V. G. Vovk. A game of prediction with expert advice. In *Proceedings of the 8th Annual Conference on Computational Learning Theory*, pages 51–60. ACM Press, New York, NY, 1995.

[YEYS04] R. Yaroshinsky, R. El-Yaniv, and S. Seiden. How to better use expert advice. *Machine Learning*, 2004.