HAL open science

Optimal learning with Bernstein Online Aggregation

Olivier Wintenberger

To cite this version:

Olivier Wintenberger. Optimal learning with Bernstein Online Aggregation. 2016. hal-00973918v4

HAL Id: hal-00973918

https://hal.science/hal-00973918v4

Preprint submitted on 29 Aug 2016 (v4), last revised 9 Sep 2016 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal learning with Bernstein Online Aggregation

Olivier Wintenberger
olivier.wintenberger@upmc.fr
Sorbonne Universités, UPMC Univ Paris 06
LSTA, Case 158, 4 place Jussieu
75005 Paris, FRANCE
&
University of Copenhagen, DENMARK

Abstract

We introduce a new recursive aggregation procedure called Bernstein Online Aggregation (BOA). Its exponential weights include a second order refinement. The procedure is optimal for the model selection aggregation problem in the bounded iid setting for the square loss: the excess of risk of its batch version achieves the fast rate of convergence $\log(M)/n$ in deviation. The BOA procedure is the first online algorithm that satisfies this optimal fast rate. The second order refinement is required to achieve the optimality in deviation as the classical exponential weights cannot be optimal, see [Aud09]. This refinement is settled thanks to a new stochastic conversion that estimates the cumulative predictive risk in any stochastic environment with observable second order terms. The observable second order term is shown to be sufficiently small to assert the fast rate in the iid setting when the loss is Lipschitz and strongly convex. We also introduce a multiple learning rates version of BOA. This fully adaptive BOA procedure is also optimal, up to a $\log\log(n)$ factor.

Keywords Exponential weighted averages, learning theory, individual sequences.

1 Introduction and main results

We consider the online setting where observations $\mathcal{D}_t = \{(X_1, Y_1), \ldots, (X_t, Y_t)\}$ are available recursively ($(X_0, Y_0) = (x_0, y_0)$ arbitrary). The goal of statistical learning is to predict $Y_{t+1} \in \mathbb{R}$ given $X_{t+1} \in \mathcal{X}$, for $\mathcal{X}$ a probability space, on the basis of $\mathcal{D}_t$. In this paper, we index with the subscript t any random element that is adapted to $\sigma(\mathcal{D}_t)$. A learner is a function $\mathcal{X} \mapsto \mathbb{R}$, denoted f_t, that depends only on the past observations $\mathcal{D}_t$ and such that $f_t(X_{t+1})$ is close to Y_{t+1}. This closeness at time $t+1$ is addressed by the predictive risk

$$\mathbb{E}[\ell(Y_{t+1}, f_t(X_{t+1})) \mid \mathcal{D}_t]$$

where $\ell : \mathbb{R}^2 \to \mathbb{R}$ is a loss function. We define an online learner f as a recursive algorithm that produces at each time $t \geq 0$ a learner: $f = (f_0, f_1, f_2, \ldots)$. The accuracy of an online

learner is quantified by the cumulative predictive risk

$$R_{n+1}(f) = \sum_{t=1}^{n+1} \mathbb{E}[\ell(Y_t, f_{t-1}(X_t)) \mid \mathcal{D}_{t-1}].$$

We will motivate the choice of this criteria later in the introduction.

Given a finite set $\mathcal{H} = \{f_1, \dots, f_M\}$ of online learners such that $f_j = (f_{j,t})_{t\geq 0}$, we aim at finding optimal online aggregation procedures

$$\hat{f} = \Big(\sum_{j=1}^{M} \pi_{j,0} f_{j,0}, \sum_{j=1}^{M} \pi_{j,1} f_{j,1}, \sum_{j=1}^{M} \pi_{j,2} f_{j,2} \dots \Big)$$

with $\sigma(\mathcal{D}_t)$-measurable weights $\pi_{j,t} \geq 0$, $\sum_{j=1}^{M} \pi_{j,t} = 1$, $t = 0, \dots, n$. We call deterministic aggregation procedures f_π any online learner of the form

$$f_\pi = \Big(\sum_{j=1}^{M} \pi_j f_{j,0}, \sum_{j=1}^{M} \pi_j f_{j,1}, \sum_{j=1}^{M} \pi_j f_{j,2}, \dots \Big)$$

with $\pi = (\pi_j)_{1\leq j\leq M}$ with $\sum_{j=1}^{M} \pi_j = 1$. Notice that π can be viewed as a probability measure on the random index $J \in \{1, \dots, M\}$. We will also use the notation π_t for the probability measure $(\pi_{j,t})_{1\leq j\leq M}$ on $\{1, \dots, M\}$. Then, $f_\pi = \mathbb{E}_\pi[f_J]$ and $\hat{f} = (\mathbb{E}_{\pi_0}[f_{J,0}], \mathbb{E}_{\pi_1}[f_{J,1}], \mathbb{E}_{\pi_2}[f_{J,2}], \dots)$. The predictive performances of an online aggregation procedure $\hat{f}$ is compared with the best deterministic aggregation of $\mathcal{H}$ or the best element of $\mathcal{H}$. We refer to these two different objectives as, respectively, the convex aggregation Problem (C) or the model selection aggregation Problem (MS). The performance of online aggregation procedures is usually measured using the cumulative loss (in the context of individual sequences prediction, see the seminal book [CBL06]). The first aim of this paper is to use instead the cumulative predictive risk for any stochastic process (X_t, Y_t). However, to define properly the notion of optimality as in [Nem00, Tsy03], we will also consider the specific iid setting of independent identically distributed observations (X_t, Y_t) when the online learners are constants: $f_{j,t} = f_j$, $t \geq 0$. In the iid setting, we suppress the indexation with time t as much as possible and we define the usual risk $\bar{R}(f) = \mathbb{E}[\ell(Y, f(X))] = (n+1)^{-1} R_{n+1}(f)$ for any constant learner f. A batch learner is defined as $\bar{f} = f_\pi$ for $\sigma(\mathcal{D}_n)$-measurable weights $\pi_j \geq 0$, $\sum_{j=1}^{M} \pi_j = 1$. Lower bounds for the excesses of risk

$$\text{Problem (C):} \qquad \bar{R}(\bar{f}) - \inf_\pi \bar{R}(f_\pi), \qquad \text{Problem (MS):} \qquad \bar{R}(\bar{f}) - \min_j \bar{R}(f_j),$$

are provided for the expectation of the square loss in [Nem00, Tsy03]. Theorem 4.1 in [Rig12] provides sharper lower bounds in deviation for some strongly convex Lipschitz losses and deterministic X_ts. These lower bounds are called the optimal rates of convergence and a weaker version applies also in our more general setting with stochastic (but possibly degenerate) X_ts. For Problem (C), we retain the rate $\sqrt{\log(M)/n}$ that is optimal when $M > \sqrt{n}$ and $\log(M)/n$ for Problem (MS), see [Rig12] for details. We are now ready to define the notion of optimality:

Definition 1.1 (adapted from Theorem 4.1 in [Rig12]). *In the iid setting, a batch aggregation procedure is optimal for Problems (C) when $M > \sqrt{n}$ or (MS) if there exists $C > 0$ such that with probability $1 - e^{-x}$, $x > 0$, it holds*

$$\bar{R}(\bar{f}) - \inf_{\pi} \bar{R}(f_\pi) \leq C \frac{\sqrt{\log M} + x}{\sqrt{n}}, \qquad \text{or} \qquad \bar{R}(\bar{f}) - \min_j \bar{R}(f_j) \leq C \frac{\log M + x}{n}.$$

Very few known procedures achieve the fast rate $\log(M)/n$ in deviation and none of them are issued from an online procedure. In this article, we provide the Bernstein Online Aggregation (BOA) that is proved to be the first online aggregation procedure such that its batch version, defined as $\bar{f} = (n+1)^{-1} \sum_{t=0}^{n} \hat{f}_t$, is optimal. Before defining it properly, let us review the existing optimal procedures for Problem (MS).

The batch procedures in [Aud07, LM09, LR14] achieve the optimal rate in deviation. A priori, they face practical issues as they require a computational optimization technique to approximate the weights that are defined as an optimum. A step further has been done in the context of quadratic loss with gaussian noise in [DRXZ12] where an explicit iterative scheme is provided. We will now explain why the question of the existence of an online algorithm whose batch version achieves fast rate of convergence in deviations remained open (see the conclusion of [Aud09]) before our work. Optimal (for the regret) online aggregation procedures are exponential weights algorithms (EWAs), see [Vov90, HKW98]. The batch versions of EWAs coincides with the Progressive Mixture Rules (PRMs). In the iid setting, the properties of the excess of risk of such procedures have been extensively studied in [Cat04]. PRMs achieve the fast optimal rate $\log(M)/n$ in expectation (that follows from the expectation of the optimal regret bound by an application of Jensen's inequality, see [Cat04, JRT08]). However, PRMs are suboptimal in deviation, i.e. the optimal rate cannot hold with high probability, see [Aud07, DRXZ12]. It is because the optimality for the regret defined as in [HKW98] does not coincides with the notion of optimality for the risk in deviation used in Definition 1.1.

The optimal BOA procedure is obtained using a necessary second order refinement of EWA. Figure 1 describes the computation of the weights in the BOA procedure where $\ell_{j,t}$ denotes the opposite of the instantaneous regret $\ell(Y_t, f_{j,t-1}(X_t)) - \mathbb{E}_{\pi_{t-1}}[\ell(Y_t, f_{J,t-1}(X_t))]$ (linearized when the loss ℓ is convex). Other procedures already exist with different second

Parameters: Learning rate $\eta > 0$.
Initialization: Set $\pi_{j,0} > 0$ such that $\sum_{j=1}^{M} \pi_{j,0} = 1$.
For: Each time round $1 \leq t \leq n$, compute the weight vector $\pi_t = (\pi_{j,t})_{1 \leq j \leq M}$:

$$\pi_{j,t} = \frac{\exp(-\eta \ell(Y_t, f_{j,t-1}(X_t)) - \eta^2 \ell_{j,t}^2)\pi_{j,t-1}}{\mathbb{E}_{\pi_{t-1}}[\exp(-\eta \ell(Y_t, f_{J,t-1}(X_t)) - \eta^2 \ell_{J,t}^2)]} = \frac{\exp(-\eta \ell_{j,t}(1 + \eta \ell_{j,t}))\pi_{j,t-1}}{\mathbb{E}_{\pi_{t-1}}[\exp(-\eta \ell_{J,t}(1 + \eta \ell_{J,t}))]}.$$

Figure 1: The BOA algorithm

order refinements, see [Aud09, HK10]. None of them have been proved to be optimal for (MS) in deviation. The choice of the second order refinement is crucial. In this paper, the

second order refinement is chosen as

$$\ell_{j,t}^2 = (\ell(Y_t, f_{j,t-1}(X_t)) - \mathbb{E}_{\pi_{t-1}}[\ell(Y_t, f_{J,t-1}(X_t))])^2$$

thanks to the new stochastic conversion that we describe below. Notice that the second order refinement $\ell_{j,t}^2$ tends to stabilize the procedure as the distances between the losses of the learners and the aggregation procedure are costly.

We achieve an upper bound for the excess of the cumulative predictive risk by first deriving a second order bound on the regret:

$$\mathrm{Err}_{n+1}(\hat{f}) - \mathrm{Err}_{n+1}(f_\pi) \qquad \text{where} \qquad \mathrm{Err}_{n+1}(f) = \sum_{t=1}^{n+1} \ell(Y_t, f_{t-1}(X_t)).$$

Second we extend it to an upper bound on the excess of the cumulative predictive risk $R_{n+1}(\hat{f}) - R_{n+1}(f_\pi)$ in any stochastic environment. In previous works, the online to batch conversion follows from an application of a Bernstein inequality for martingales. It provides a control of the deviations in the stochastic environment via the predictable quadratic variation, see for instance [Fre75, Zha05, KT08, GSVE14]. Here we prefer to use an empirical counterpart of the classical Bernstein inequality, based on the quadratic variation instead of the predictive quadratic variation. For any martingale (M_t), we denote $\Delta M_t = M_t - M_{t-1}$ its difference ($\Delta M_0 = 0$ by convention) and $[M]_t = \sum_{j=1}^t \Delta M_j^2$ its quadratic variation. We will use the following new empirical Bernstein inequality:

Theorem 1.1. *Let (M_t) be a martingale such that $\Delta M_t \geq -1/2$ a.s. for all $t \geq 0$. Then for any $n \geq 0$ we have $\mathbb{E}[\exp(M_n - [M]_n)] \leq 1$. Without any boundedness assumption, we still have*

$$\mathbb{E}\left[\exp\left(2^{-1}\left(M_n - [M]_n - \sum_{t=1}^n \Delta M_t \mathbb{1}_{\Delta M_t < -1/2}\right)\right)\right] \leq 1. \tag{1}$$

Empirical Bernstein's inequalities have already been developed in [AMS06, MP09] and use in the multi-armed bandit and penalized ERM problems. Applying Theorem 1.1, we estimate successively the deviations of two different martingales

1. $\Delta M_{j,t} = -\eta \ell_{j,t}$ as a function of j distributed conditionally as π_{t-1} on $\{1, \ldots, M\}$ in the deterministic setting (given all the observations $\mathcal{D}_{n+1}$),
2. $M_{j,t} = \eta(R_t(\hat{f}) - R_t(f_j) - \mathrm{Err}_t(\hat{f}) + \mathrm{Err}_t(f_j))$ such that $\Delta M_{j,t} = \eta(\mathbb{E}_{t-1}[\ell_{j,t}] - \ell_{j,t})$ where $\mathbb{E}_{t-1}$ denotes the expectation of (X_t, Y_t) conditionally on $\mathcal{D}_{t-1}$.

The first application 1. of Theorem 1.1 will provide a second order bound on the regret in the deterministic setting whereas the second application 2. will provide the new stochastic conversion. In both cases, the second order term will be equal to $\eta^{-1}[M_j]_{n+1} = \eta \sum_{t=1}^{n+1} \ell_{j,t}^2$ after renormalization. It is the main motivation of BOA; as our notion of optimality requires a stochastic conversion, a second order term necessarily appears in the bound of the excess of the cumulative predictive risk. An online procedure will achieve good performances in the batch setting if the second order bound of the regret is similar than

the necessary cost due to the stochastic conversion. The BOA procedure achieves this aim by incorporating this second order term in the computation of the weights.

In the first application 1. of Theorem 1.1, we have $\mathbb{E}_{\pi_{t-1}}[\Delta M_{J,t}] = 0$ and an application of Theorem 1.1 yields the regret bound of Theorem 3.1:

$$\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] \leq \inf_{\pi} \Big\{ \mathbb{E}_{\pi} \Big[\mathrm{Err}_{n+1}(f_J) + \eta \sum_{t=1}^{n+1} \ell_{J,t}^2 + \frac{\log(\pi_J/\pi_{J,0})}{\eta} \Big] \Big\},$$

where $\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] = \sum_{t=1}^{n+1} \mathbb{E}_{\hat{\pi}_{t-1}}[\ell(Y_t f_{J,t}(X_t))]$. Such second order bounds also hold for the regret of other algorithms, see [CBMS07, GSVE14, LS15, KVE15]. Using the new stochastic conversion based on the application 2. of Theorem 1.1, this second order regret bound is converted to a one on the cumulative predictive risk (see Theorem 4.2). With probability $1 - e^{-x}$, $x > 0$, we have

$$\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)] \leq \inf_{\pi} \Big\{ \mathbb{E}_{\pi} \Big[R_{n+1}(f_J) + 2\eta \sum_{t=1}^{n+1} \ell_{J,t}^2 + \frac{\log(\pi_J/\pi_{J,0}) + x}{\eta} \Big] \Big\}.$$

Thanks to the use of the cumulative predictive risk, this bound is valid in any stochastic environment. We will extend it in various directions. We will introduce

- the "gradient trick" to bound the excess of the cumulative predictive risk in Problem (C),
- the multiple learning rates for adapting the procedure and
- the batch version of BOA to achieve the fast rate of convergence in Problem (MS).

The "gradient trick" is a standard argument to solve Problem (C), see [CBL06]. When the loss ℓ is convex with respect to its second argument, its sub-gradient is denoted ℓ'. In this case, we consider a convex version of the BOA procedure described in Figure 1. The original loss ℓ is replaced with its linearized version

$$\ell'(Y_t, \hat{f}_{t-1}(X_t)) f_{j,t-1}(X_t), \tag{2}$$

and we denote, with some abuse of notation,

$$\ell_{j,t} = \ell'(Y_t, \hat{f}_{t-1}(X_t))(f_{j,t-1}(X_t) - \hat{f}_{t-1}(X_t)).$$

Linearizing the loss, we can compare the regret of the (sub-gradient version of the) BOA procedure $\hat{f} = \mathbb{E}_{\hat{\pi}}[f_J]$ with the best deterministic aggregation of the elements in the dictionary. We obtain in Theorem 3.1 a second order regret bound for Problem (C)

$$\mathrm{Err}_{n+1}(\hat{f}) \leq \inf_{\pi} \Big\{ \mathrm{Err}_{n+1}(f_\pi) + \eta \sum_{t=1}^{n+1} \mathbb{E}_{\pi}[\ell_{J,t}^2] + \frac{\mathbb{E}_{\pi}[\log(\pi_J/\pi_{J,0})]}{\eta} \Big\}.$$

When trying to optimize the regret bound in the learning rate $\eta > 0$, we obtain

$$\sqrt{\frac{\mathbb{E}_\pi[\log(\pi_J/\pi_{J,0})]}{\sum_{t=1}^{n+1}\mathbb{E}_\pi[\ell_{J,t}^2]}} \leq \mathbb{E}_\pi\left[\sqrt{\frac{\log(\pi_{J,0}^{-1})}{\sum_{t=1}^{n+1}\ell_{J,t}^2}}\right],$$

As π is unknown, this tuning parameter is not tractable in practice. Its worst case version $\max_j \sqrt{\log(\pi_{j,0}^{-1})}/\sqrt{\sum_{t=1}^{n+1}\ell_{j,t}^2}$ is not satisfactory. Multiple learning rates have been introduced by [BM05] to solve this issue (see also [GSVE14]). We introduce the multiple learning rates version of BOA in Figure 2 which is also fully adaptative as it adapts to any possible range of observations.

Parameter: an integer $c > 0$ big enough to consider that 2^{-c} is negligible.
Initialization: Set $L_{j,0} = 0$, $\eta_{j,0} = 0$, $\pi_{j,0} > 0$ such that $\sum_{j=1}^M \pi_{j,0} = 1$.
For: each time round $t \geq 1$,

1. Compute recursively

$$L_{j,t} = L_{j,t-1} + 2^{-1}(\ell_{j,t}(1+\eta_{j,t-1}\ell_{j,t}) + E_{j,t}1_{\eta_{j,t-1}\ell_{j,t}>1/2}),$$

2. Estimate the ranges $E_{j,t} = 2^{k+1}$ where $k \geq -c$ is the smallest integer such that $\max_{1\leq s\leq t}|\ell_{j,t}| \leq 2^k$, $1 \leq j \leq M$,

3. Compute the adaptative learning rate

$$\eta_{j,t} = \min\left\{\frac{1}{E_{j,t}}, \sqrt{\frac{\log(\pi_{j,0}^{-1})}{\sum_{s=1}^t \ell_{j,s}^2}}\right\}, \qquad 1 \leq t \leq n,\ 1 \leq j \leq M, \tag{3}$$

4. Compute the weights vector $\pi_t = (\pi_{j,t})_{1\leq j\leq M}$:

$$\pi_{j,t} = \frac{\eta_{j,t}\exp(-\eta_{j,t}L_{j,t})\pi_{j,0}}{\mathbb{E}_{\pi_0}[\eta_{J,t}\exp(-\eta_{J,t}L_{J,t})]}.$$

Figure 2: The fully adaptive BOA procedure

The novelty, compared with the "doubling trick" developed in [CBMS07], is the dependence of the learning rates and the estimated ranges with respect to j and the expression of the weights with respect to the learning rates and the ranges estimators. For this adaptive BOA procedure we obtain regret bounds such as

$$\mathrm{Err}_{n+1}(\hat{f}) \leq \inf_\pi\left\{\mathrm{Err}_{n+1}(f_\pi) + C\mathbb{E}_\pi\left[\sqrt{\sum_{t=1}^{n+1}\ell_{J,t}^2\log(\pi_{J,0}^{-1})} + CE_J\log(\pi_{J,0}^{-1})\right]\right\},$$

for some "constant" $C > 0$ that grows as $\log\log(n)$, see Theorem 3.3 for details. Such second order bounds involving excess losses terms as the $\ell_{j,t}$s have been proved for other

algorithms in [GSVE14, LS15, KVE15]. We refer to these articles for nice consequences of such bounds in the individual sequences framework. Here again, the stochastic conversion holds in any stochastic environment and we obtain with probability $1-e^{-x}$, $x>0$

$$R_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ R_{n+1}(f_\pi) + C\mathbb{E}_\pi \left[\sqrt{\sum_{t=1}^{n+1} \ell_{J,t}^2 (\log(\pi_{J,0}^{-1}) + x)} + C(E_J \log(\pi_{J,0}^{-1}) + x) \right] \right\},$$

The optimal bound for Problem (C) is proved in the very general setting: under a boundedness assumption, the rate of convergence of $R_{n+1}(\hat{f})/(n+1)$ (an upper bound of $\bar{R}(\bar{f})$ in the iid setting), is smaller than $\sqrt{\log(M)/n}$. Notice that the classical online-to-batch conversion leads to similar results, see [CBMS07, Ger13, GSVE14], and that our second order refinement is not necessary to obtain such upper bounds. Notice also that such upper bounds were already derived in non iid settings under restrictive dependent assumptions for the excess of risk (and not the cumulative predictive risk), see [ALW13, MR10, AD13]. It is remarkable to extend the optimal bound of Problem (C) to any stochastic environment thanks to the use of the cumulative predictive risk. It is because the cumulative predictive risk of f_π in the upper bound takes into account the dependence as it is a random variable in non iid settings. We believe that the cumulative predictive risk is the correct criteria to assert the prediction accuracy of online algorithms in stochastic environment as it coincides with the regret in the deterministic setting and with the classical risk for batch procedures in the iid setting. Moreover, it appears naturally when using the minimax theory approach, see [AABR09]. However, up to our knowledge, it is the first time that the cumulative predictive risk is used to compare an online procedure with deterministic aggregation procedures.

The fast rate of convergence $\log(M)/n$ in Problem (MS) is achieved thanks to a careful study of the second order terms $\sum_{t=1}^{n+1} \ell_{j,t}^2$. It also requires more conditions on the loss in order to compare it locally with the square loss, see [Aud09]. We restrict us to losses ℓ that are strongly convex and Lipschitz functions in the iid setting, see [KT08] for an extensive study of this context. We fix the initial weights uniformly $\pi_{j,0} = M^{-1}$. We obtain in Theorem 4.4 the fast rate of convergence for the batch version of the BOA procedure; with probability $1-e^{-x}$, $x>0$,

$$\bar{R}(\bar{f}) \leq \frac{R_{n+1}(\hat{f})}{n+1} \leq \min_{1\leq j\leq M} R(f_j) + \frac{2\log(M)+3x}{\eta(n+1)},$$

for η smaller than a constant depending on the range. The second order term is estimated with the excess of risk using the strong convexity assumption on the loss. We conclude by providing the fast rate bound on the excess of risk of the batch version of the adaptive BOA at the price of larger "constants" that grows at the rate $\log\log(n)$.

The paper is organized as follows: We present the second order regret bounds for different versions of BOA in Section 3. The new stochastic conversion and the excess of cumulative predictive risk bounds in a stochastic environment are provided in Section 4. In the next Section, we introduce some useful probabilistic preliminaries.

2 Preliminaries

Similarly than in [Aud09], the recursive argument for supermartingales will be at the core of the proofs developed in this paper. It will be used jointly with the variational form of the entropy to provide second order regret bounds.

2.1 The proof of the martingale inequality in Theorem 1.1

The proof of the first empirical Bernstein inequality for martingales of Theorem 1.1 follows from an exponential inequality and by a classical recursive supermartingales argument, see [Fre75]. As $X = \Delta M_t \geq -1/2$ a.s., from the inequality $\log(1+x) \geq x - x^2$ for $x > -1/2$ (stated as Lemma 1 in [CBMS07]), we have

$$X - X^2 \leq \log(1+X) \Leftrightarrow \exp(X - X^2) \leq 1 + X \Rightarrow \mathbb{E}_{t-1}[\exp(X - X^2)] \leq 1. \tag{4}$$

Here we used that $\mathbb{E}_{t-1}[X] = 0$ as $X = \Delta M_t$ is a difference of martingale. The proof ends by using the classical recursive argument for supermartingales; from the definition of the difference of martingale $X = \Delta M_t$, we obtain as a consequence of (4) that

$$\mathbb{E}[\exp(M_t - [M]_t^2)] \leq \mathbb{E}[\exp(M_{t-1} - [M]_{t-1}^2)].$$

As $\mathbb{E}[\exp(M_0 - [M]_0^2)] = 1$, applying a recursion for $t = 1, \ldots, n$ provides the desired result.

Without any boundedness assumption, we have

$$\begin{aligned}\mathbb{E}[\exp(X - X^2)] &= \mathbb{E}[\exp(X - X^2)\mathbb{1}_{X \geq -1/2}] + \mathbb{E}[\exp(X - X^2)\mathbb{1}_{X < -1/2}] \\ &\leq \mathbb{E}[(1+X)\mathbb{1}_{X \geq -1/2}] + \mathbb{E}[\mathbb{1}_{X < -1/2}] \leq 1 + \mathbb{E}[X\mathbb{1}_{X \geq -1/2}].\end{aligned}$$

As X is centered, we can bound

$$1 + \mathbb{E}[X\mathbb{1}_{X \geq -1/2}] = 1 - \mathbb{E}[X\mathbb{1}_{X < -1/2}] \leq \mathbb{E}[1 + X\mathbb{1}_{X < -1/2}] \leq \mathbb{E}[\exp(X\mathbb{1}_{X < -1/2})].$$

Using Cauchy-Schwarz inequality and the preceding arguments, we obtain

$$\begin{aligned}\mathbb{E}[\exp(2^{-1}(X - X^2 - X\mathbb{1}_{X < -1/2}))] &\leq \sqrt{\mathbb{E}[\exp(X - X^2)]\mathbb{E}[\exp(-X\mathbb{1}_{X < -1/2})]} \\ &\leq \sqrt{\mathbb{E}[\exp(X\mathbb{1}_{X < -1/2})]\mathbb{E}[\exp(-X\mathbb{1}_{X < -1/2})]}.\end{aligned}$$

The desired result follows from the Jensen's inequality followed by the same recursive argument for supermartingales as above.

2.2 The variational form of the entropy

The relative entropy (or Kullback-Leibler divergence) $\mathcal{K}(Q, P) = \mathbb{E}_Q[\log(dQ/dP)]$ is a pseudo-distance between any probability measures P and Q. Let us remind the basic property of the entropy: the variational formula of the entropy originally proved in full generality in [DV75]. We consider here a version well adapted for obtaining second order regret bounds:

Lemma 2.1. *For any probability measure P on $\mathcal{X}$ and any measurable functions h, $g : \mathcal{X} \to \mathbb{R}$ we have:*

$$\mathbb{E}_P[\exp(h - \mathbb{E}_P[h] - g)] \leq 1$$
$$\iff \mathbb{E}_Q[h] - \mathbb{E}_P[h] \leq \mathbb{E}_Q[g] + \mathcal{K}(Q, P), \qquad \text{for any probability measure } Q. \quad (5)$$

The left hand side corresponds to the right hand side with Q equals the Gibbs measure $\mathbb{E}_P[e^{h-g}]dQ = e^{h-g}dP$.

That the Gibbs measure realizes the dual identity is at the core of the PAC-bayesian approach. Exponential weights aggregation procedures arise naturally as they can be considered as Gibbs measures, see [Cat07].

3 Second order regret bounds for the BOA procedure

3.1 First regret bounds and link with the individual sequences framework

We work conditionally on $\mathcal{D}_{n+1}$; it is the deterministic setting, similar than in [Ger13], where $(X_t, Y_t) = (x_t, y_t)$ are provided recursively for $1 \leq t \leq n$. In that case, the cumulative loss $\mathrm{Err}_{n+1}(f)$ quantify the prediction of $f = (f_0, f_1, f_2, \ldots)$. We state first a regret bound for non convex losses, and then move to the case of convex losses combined with the "gradient trick" as in the Appendix of [GSVE14]. Recall that $\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] = \sum_{t=1}^{n+1} \mathbb{E}_{\hat{\pi}_{t-1}}[\ell(Y_t f_{J,t}(X_t))]$.

Theorem 3.1. *Assume that $\eta > 0$ satisfies*

$$\eta \max_{1 \leq t \leq n+1} \max_{1 \leq j \leq M} \ell_{j,t} \leq 1/2, \quad (6)$$

then the cumulative loss of the BOA procedure with $\ell_{j,t} = \ell(Y_t, f_{j,t-1}(X_t)) - \ell(Y_t, \hat{f}_{t-1}(X_t))$ satisfies

$$\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] \leq \inf_{\pi} \Big\{ \mathbb{E}_\pi \Big[\mathrm{Err}_{n+1}(f_J) + \eta \sum_{t=1}^{n+1} \ell_{J,t}^2 \Big] + \frac{\mathcal{K}(\pi, \pi_0)}{\eta} \Big\}.$$

If ℓ is convex with respect to its second argument, the cumulative loss of the BOA procedure with $\ell_{j,t} = \ell'(Y_t, \hat{f}_{t-1}(X_t))(f_{j,t-1}(X_t) - \hat{f}_{t-1}(X_t))$ also satisfies

$$\mathrm{Err}_{n+1}(\hat{f}) \leq \min_{\pi} \Big\{ \mathrm{Err}_{n+1}(f_\pi) + \eta \sum_{t=1}^{n+1} \mathbb{E}_\pi[\ell_{J,t}^2] + \frac{\mathcal{K}(\pi, \pi_0)}{\eta} \Big\}.$$

Proof. We consider $\Delta M_{j,t+1} = -\eta \ell_{j,t+1}$ that is a centered random variable on $\{1, \ldots, M\}$ when distributed as π_t. Under the assumption (6), $\Delta M_{j,t+1} \geq -1/2$ for any $0 \leq t \leq n$ a.s.. An application of the inequality (4) provides the inequality

$$\mathbb{E}_{\pi_t}[\exp(-\eta \ell_{J,t+1}(1 + \eta \ell_{J,t+1}))] \leq 1. \quad (7)$$

We have from the recursive definition of the BOA procedure provided in Figure 1 the expression

$$\pi_{j,t} = \frac{\exp(-\eta \sum_{s=1}^{t} \ell_{j,s}(1+\eta\ell_{j,s}))\pi_{j,0}}{\mathbb{E}_{\pi_0}[\exp(-\eta \sum_{s=1}^{t} \ell_{J,s}(1+\eta\ell_{J,s}))]}.$$

Plugging the expression of the weights $\pi_{j,t}$ in the inequality (7) provides

$$\mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta\sum_{s=1}^{t+1}\ell_{J,s}(1+\eta\ell_{J,s})\Big)\Big] \le \mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta\sum_{s=1}^{t}\ell_{J,s}(1+\eta\ell_{J,s})\Big)\Big].$$

By the recursive argument we obtain

$$\mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta\sum_{t=1}^{n+1}\ell_{J,t}(1+\eta\ell_{J,t})\Big)\Big] \le 1.$$

Equivalently, using the variational form of the entropy (5),

$$0 \le \inf_{\pi}\Big\{\mathbb{E}_{\pi}\Big[\eta\sum_{t=1}^{n+1}\ell_{J,t} + \eta^2\sum_{t=1}^{n+1}\ell_{J,t}^2\Big] + \mathcal{K}(\pi,\pi_0)\Big\}, \tag{8}$$

π denoting this time any probability measure on $\{1,\dots,M\}$. The first regret bound in Theorem 3.1 follows from the identity $\sum_{t=1}^{n+1}\ell_{j,t} = \mathrm{Err}_{n+1}(f_j) - \mathbb{E}_{\hat\pi}[\mathrm{Err}_{n+1}(f_j)]$. The second result follows by an application of the "gradient trick", i.e. noticing that

$$\begin{aligned}
\mathrm{Err}_{n+1}(\hat f) - \mathrm{Err}_{n+1}(f_\pi) &= \sum_{t=1}^{n+1}\ell(Y_t,\hat f_{t-1}(X_t)) - \ell(Y_t,\mathbb{E}_\pi[f_{J,t-1}](X_t)) \\
&\le \sum_{t=1}^{n+1}\ell'(Y_t,\hat f_{t-1}(X_t))(\hat f_{t-1}(X_t) - \mathbb{E}_\pi[f_{J,t-1}](X_t)) \\
&= \mathbb{E}_\pi\Big[\sum_{t=1}^{n+1}\ell'(Y_t,\hat f_{t-1}(X_t))(\hat f_{t-1}(X_t) - f_{J,t-1}(X_t))\Big] \\
&= -\mathbb{E}_\pi\Big[\sum_{t=1}^{n+1}\ell_{J,t}\Big].
\end{aligned}$$

□

The second order term in the last regret bound is equal to

$$\sum_{t=1}^{n+1}\mathbb{E}_\pi[\eta\ell_{J,t}^2] = \sum_{t=1}^{n+1}\mathbb{E}_\pi[\eta\ell'(Y_t,\hat f_{t-1}(X_t))^2(\hat f_{t-1}(X_t) - f_{J,t-1}(X_t))^2].$$

This term can be small because the sub-gradient is small and because the deterministic aggregation π is close to the BOA procedure. Such second order upper bounds can heavily

depend on the behaviors of the different learners f_j. Thus, a unique learning rate cannot be efficient in cases where the learners have different second order properties. To solve this issue, we consider the multiple learning rates version of BOA

$$\pi_{j,t} = \frac{\exp(-\eta_j \ell_{j,t}(1+\eta_j \ell_{j,t}))\pi_{j,t-1}}{\mathbb{E}_{\pi_{t-1}}[\exp(-\eta_J \ell_{J,t}(1+\eta_J \ell_{J,t}))]}. \tag{9}$$

We can extend the preceding regret bound to this more sophisticated procedure:

Theorem 3.2. *Consider that ℓ is convex with respect to its second argument and multiple learning rates η_j, $1 \le j \le M$, that are positive. If*

$$\max_{1\le t\le n+1} \max_{1\le j\le M} \eta_j \ell_{j,t} \le 1/2, \qquad a.s.,$$

then the cumulative loss of the BOA procedure with multiple learning rates satisfies

$$\mathrm{Err}_{n+1}(\hat{f}) \le \inf_{\pi} \Big\{ \mathrm{Err}_{n+1}(f_\pi) + \mathbb{E}_\pi \Big[\eta_J \sum_{t=1}^{n+1} \ell_{J,t}^2 + \frac{\log(\pi_J/\pi_{J,0}) + \log(\mathbb{E}_{\pi_0}[\eta_J^{-1}]/\mathbb{E}_\pi[\eta_J^{-1}])}{\eta_J} \Big] \Big\}.$$

Proof. Let us consider the weights $\pi'_{i,t} = \eta_i^{-1}\pi_{i,t}/\mathbb{E}_{\pi_t}[\eta_j^{-1}]$, for all $1 \le i \le M$ and $0 \le t \le n+1$. Then, for any function $j \to h_j$ measurable on $\{1,\dots,M\}$ we have the relation

$$\mathbb{E}_{\pi'_t}[\eta_j h_j] = \mathbb{E}_{\pi_t}[h_J]/\mathbb{E}_{\pi_t}[\eta_J^{-1}], \qquad 1 \le t \le n+1. \tag{10}$$

We consider $\Delta M_t = -\eta_j \ell_{j,t+1}$. Thanks to relation (10), it is a centered random variable on $\{1,\dots,M\}$ when distributed as π'_t. Moreover, the weights (π'_t) also satisfy the recursive relation (9). Thus, one can apply the same reasoning than in the proof of Theorem 3.1. We obtain, equivalently than the inequality (8), that

$$0 \le \inf_{\pi'} \Big\{ \mathbb{E}_{\pi'} \Big[\eta_J \sum_{t=1}^{n+1} \ell_{J,t} + \eta_J^2 \sum_{t=1}^{n+1} \ell_{J,t}^2 \Big] + \mathcal{K}(\pi', \pi'_0) \Big\},$$

for π' denoting any probability measure on $\{1,\dots,M\}$. Using the identity (10) to define π from π', and multiplying the above inequality with $\mathbb{E}_\pi[\eta_j^{-1}] > 0$, we have

$$0 \le \inf_{\pi} \Big\{ \mathbb{E}_{\pi} \Big[\sum_{t=1}^{n+1} \ell_{J,t} + \eta_J \sum_{t=1}^{n+1} \ell_{J,t}^2 + \frac{\log(\pi'_J/\pi'_{J,0})}{\eta_J} \Big] \Big\},$$

The proof ends by identifying $\log(\pi'_j/\pi'_{j,0})$ and using the "gradient trick" as in the proof of Theorem 3.1. □

Notice that a simple corollary of the proof above is the simplified upper bound

$$\mathrm{Err}_{n+1}(\hat{f}) \le \min_{\pi} \Big\{ \mathrm{Err}_{n+1}(f_\pi) + \mathbb{E}_\pi \Big[\eta_J \sum_{t=1}^{n+1} \ell_{J,t+1}^2 + \frac{\log(1/\pi'_{J,0})}{\eta_J} \Big] \Big\},$$

where π'_0 is defined as $\pi'_{i,0} = \eta_i^{-1}\pi_{i,t}/\mathbb{E}_{\pi_t}[\eta_J^{-1}]$. The initial weights $\pi_{j,0}$ are modified and the upper bound favors the learners with small learning rates η_j. It constitutes a drawback of the multiple learning rates version of BOA as we will see that small learning rates will be associated with bad experts. One can solve this issue by choosing the initial weights differently than classically. For example, with no information on the learners f_j, the initial weights can be chosen equal to

$$\pi_{j,0} = \frac{\eta_j}{\sum_{j=1}^M \eta_j} \qquad \text{such that} \qquad \pi'_{j,0} = \frac{1}{M}, \qquad 1 \le j \le M.$$

In this case, $\log(1/\pi'_{j,0}) \le \log(M)$ and the weights have the expression

$$\pi_{j,t} = \frac{\eta_j \exp(-\eta_j \sum_{s=1}^t \ell_{j,s}(1+\eta_j \ell_{j,s})\pi_{j,0}}{\mathbb{E}_{\pi_0}[\eta_J \exp(-\eta_J \sum_{s=1}^t \ell_{J,s}(1+\eta_J \ell_{J,s})]}, \qquad 1 \le j \le M.$$

The form of the weights becomes similar than the one of the adaptive BOA introduced in Figure 2 and studied in the next section. The second order regret bounds becomes

$$\mathrm{Err}_{n+1}(\hat{f}) \le \min_{\pi} \left\{ \mathrm{Err}_{n+1}(f_\pi) + 2\sqrt{\log(M)}\mathbb{E}_\pi \left[\sqrt{\sum_{t=1}^{n+1} \ell_{J,t}^2} \right] \right\},$$

for learning rates tuned optimally

$$\eta_j = \sqrt{\frac{\log(M)}{\sum_{t=1}^{n+1} \ell_{j,t}^2}}, \qquad 1 \le j \le M.$$

However, the resulting procedure is not recursive. Such non recursive strategies are not convertible to the batch setting.

Second order regret bounds similar than the one of Theorem 3.2 have been obtained in [GSVE14, LS15, KVE15] in the context of individual sequences. In this context, studied in [CBL06], we consider that $Y_t = y_t$ for a deterministic sequence $y_0, \ldots, y_n$ ((X_t) is useless in this context). We have $\mathcal{D}_t = \{y_0, \ldots, y_t\}$, $0 \le t \le n$, and the online learners $f_j = (y_{j,1}, y_{j,2}, y_{j,3}, \ldots)$ of the dictionary are called the experts. The cumulative loss is now $\mathrm{Err}_{n+1}(\hat{f}) = \sum_{t=1}^{n+1} \ell(y_t, \hat{y}_t)$ for any aggregative strategy $\hat{y}_t = \hat{f}_{t-1} = \sum_{j=1}^M \pi_{j,t-1} y_{j,t}$ where $\pi_{j,t-1}$ are measurable functions of the past $\{y_0, \ldots, y_{t-1}\}$. We will compare the different second order regret bounds achieved by the different adaptative procedures at the end of the next Section.

3.2 A new adaptive method for exponential weights

We described in Figure 2 the adaptive version of the BOA algorithm. Notice that the adaptive version of the exponential weights

$$\pi_{j,t} = \frac{\eta_{j,t} \exp(-\eta_{j,t} L_{j,t})\pi_{j,0}}{\mathbb{E}_{\pi_0}[\eta_{J,t} \exp(-\eta_{J,t} L_{J,t})]},$$

is different from [CBMS07] as the multiple learning rates $\eta_{j,t}$ depend on j. Moreover, the multiple learning rates appear in the exponential and as a multiplicative factor to solve the issue concerning the modification of the initial weights described above. Adaptive procedures with such multiplicative forms have been studied in [GSVE14]. Another possibility consists in putting a prior on learning rates as in [KVE15]. Multiple learning rates versions can be investigated for other exponential weights than for those of BOA. Notice also that for the first time we apply a multiple version of the "doubling trick" of [CBMS07]; the ranges estimators $E_{j,t}$ also appear in the exponential weights as a penalization when the ranges estimators are exceeded. We obtain a second order regret bound for the BOA procedure similar than in [GSVE14]:

Theorem 3.3. *Assume that ℓ is convex with respect to its second argument and that E_j defined by*

$$\max_{1\le t\le n+1} |\ell_{j,t}| \le E_j, \qquad 1\le j\le M,$$

satisfies $2^{-c}\le E_j\le E$ for all $1\le j\le M$, we have

$$\mathrm{Err}_{n+1}(\hat f) \le \inf_{\pi}\Bigg\{\mathrm{Err}_{n+1}(f_\pi) + \mathbb{E}_\pi\Bigg[\sqrt{\sum_{t=1}^{n+1}\ell_{J,t}^2}\left(\frac{(2\sqrt{2}-1)\sqrt{\log(\pi_{J,0}^{-1})}}{\sqrt{2}-1} + \frac{2B_{n,E}+8E_J}{\sqrt{\log(\pi_{J,0}^{-1})}}\right) + 4E_J(2\log(\pi_{J,0}^{-1}) + 2B_{n,E}+3)\Bigg]\Bigg\},$$

where $B_{n,E} = \log(1+2^{-1}\log(n)+\log(E)+c\log(2))$ for all $n\ge 1$.

Proof. We adapt the reasoning of the proof of Theorem 3.2 for learning rates depending on t. Thus, the recursive argument holds only approximatively. For any $1\le t\le n$, let us consider the weights π'_t as

$$\pi'_{j,t} = \frac{\eta_{j,t}^{-1}\pi_{j,t}}{\mathbb{E}_{\pi_t}[\eta_{J,t}^{-1}]}.$$

We consider $\Delta M_{j,t+1} = -\eta_{j,t}\ell_{j,t+1}$ a centered random variable on $\{1,\dots,M\}$ when distributed as π'_t. As $\Delta M_{j,t+1}\ge -\eta_{j,t}E_{j,t+1}$, $j=1,\dots,M$, we apply the inequality (1):

$$\mathbb{E}_{\pi'_t}[\exp(-\eta_{J,t}2^{-1}(\ell_{J,t+1}(1+\eta_{J,t}\ell_{J,t+1}) + E_{J,t+1}1_{\eta_{J,t}\ell_{J,t+1}>1/2}))] \le 1.$$

By definition of the weights π'_t and π_t, we have for any $1\le t\le n$ the expression

$$\pi'_{j,t} = \frac{\exp(-\eta_{j,t}2^{-1}\sum_{s=1}^t(\ell_{j,s}(1+\eta_{j,s-1}\ell_{j,s}) + E_{j,s}1_{\eta_{j,s-1}\ell_{j,s}>1/2}))\pi_{j,0}}{\mathbb{E}_{\pi_0}[\exp(-\eta_{J,t}2^{-1}\sum_{s=1}^t(\ell_{J,s}(1+\eta_{J,s-1}\ell_{J,s}) + E_{J,s}1_{\eta_{J,s-1}\ell_{J,s}>1/2}))]}.$$

Using the expression of the weights in the preceding inequality provides

$$\mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta_{J,t}2^{-1}\sum_{s=1}^{t+1}(\ell_{J,s}(1+\eta_{J,s-1}\ell_{J,s}) + E_{J,s}1_{\eta_{J,s-1}\ell_{J,s}>1/2})\Big)\Big]$$
$$\le \mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta_{J,t}2^{-1}\sum_{s=1}^{t}(\ell_{J,s}(1+\eta_{J,s-1}\ell_{J,s}) + E_{J,s}1_{\eta_{J,s-1}\ell_{J,s}>1/2})\Big)\Big]. \quad (11)$$

Using the basic inequality $x \leq \alpha^{-1}x^{\alpha} + \alpha^{-1}(\alpha-1) \leq x^{\alpha} + \alpha^{-1}(\alpha-1)$ for

$$x = \exp\Big(- \eta_{j,t}2^{-1}\sum_{s=1}^{t}(\ell_{j,s}(1+\eta_{j,s-1}\ell_{j,s}) + E_{j,s}1_{\eta_{j,s-1}\ell_{j,s}>1/2})\Big) \geq 0$$

and $\alpha = \eta_{j,t-1}/\eta_{j,t} \geq 1$, we obtain for all $2 \leq t \leq n$

$$\begin{aligned}
&\mathbb{E}_{\pi_0}\Big[\exp\Big(- \eta_{J,t}2^{-1}\sum_{s=1}^{t}(\ell_{J,s}(1+\eta_{J,s-1}\ell_{J,s}) + E_{J,s}1_{\eta_{J,s-1}\ell_{J,s}>1/2})\Big)\Big] \\
&\leq \mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta_{J,t-1}2^{-1}\sum_{s=1}^{t}(\ell_{J,s}(1+\eta_{J,s-1}\ell_{J,s})+E_{J,s}1_{\eta_{J,s-1}\ell_{J,s}>1/2})\Big)\Big]+\mathbb{E}_{\pi_0}\Big[\frac{\eta_{J,t-1}-\eta_{J,t}}{\eta_{J,t-1}}\Big].
\end{aligned} \tag{12}$$

Then, combining the inequalities (11) and (12) recursively for $t = n, \ldots, 2$ and then (11) for $t = 1$ we obtain

$$\mathbb{E}_{\pi_0}\Big[\exp\Big(-\eta_{J,n}2^{-1}\sum_{t=1}^{n+1}(\ell_{J,t}(1+\eta_{J,t-1}\ell_{J,t})+E_{J,t}1_{\eta_{J,t-1}\ell_{J,t}>1/2})\Big)\Big] \leq 1+\sum_{t=2}^{n}\mathbb{E}_{\pi_0}\Big[\frac{\eta_{J,t-1}-\eta_{J,t}}{\eta_{J,t-1}}\Big].$$

We apply the variational form of the entropy (5) in order to derive that

$$\begin{aligned}
0 \leq \mathbb{E}_{\pi'}\Big[\eta_{J,n}2^{-1}\sum_{t=1}^{n+1}(\ell_{J,t}(1+\eta_{J,t-1}\ell_{J,t}) + E_{J,t}1_{\eta_{J,t-1}\ell_{J,t}>1/2})\Big] \\
+ \log\Big(1+\sum_{t=2}^{n}\mathbb{E}_{\pi_0}\Big[\frac{\eta_{J,t-1}-\eta_{J,t}}{\eta_{J,t-1}}\Big]\Big) + \mathcal{K}(\pi',\pi_0)
\end{aligned}$$

for any probability measure π' on $\{1,\ldots,M\}$. We bound the last term $\mathcal{K}(\pi',\pi_0) \leq \mathbb{E}_{\pi'}[\log(\pi_{J,0}^{-1})]$. By comparing with the integral of $1/x$ on the interval $[\eta_{j,n},\eta_{j,1}]$, we estimate

$$\sum_{t=2}^{n}\frac{\eta_{j,t-1}-\eta_{j,t}}{\eta_{j,t-1}} \leq \sum_{t=2}^{n}\int_{\eta_{j,t}}^{\eta_{j,t-1}}\frac{dx}{x} \leq \int_{\eta_{j,n}}^{\eta_{j,1}}\frac{dx}{x} \leq \log\Big(\frac{\eta_{j,1}}{\eta_{j,n}}\Big).$$

We have the bounds $\sum_{t=1}^{n+1}E_{j,t}1_{\eta_{j,t-1}\ell_{j,t}>1/2} \leq 8E_j$ and $\log(\eta_{j,1}/\eta_{j,n}) \leq \log(\sqrt{n}E/2^{-c})$, $1 \leq j \leq M$. Then Theorem 3.3 is proved using similar arguments than in the proof of Theorem 6 in [CBMS07], choosing $\pi'_j = \eta_{j,n}^{-1}\pi_j/\mathbb{E}_{\pi}[\eta_{j,n}^{-1}]$ and using the "gradient trick" as in the proof of Theorem 3.1. □

The advantage of the adaptive BOA procedure compared with the procedures studied in [GSVE14, LS15, KVE15] is to be adaptive to unknown ranges. The price to pay is an additional logarithmic term $\log(E) + c\log(2)$ depending on the variability of the adaptive learning rates $\eta_{j,t}$ through time. Such losses are avoidable in the case of one single adaptive learning rate $\eta_{j,t} = \eta_t$, for all $1 \leq j \leq M$. Notice also that the relative entropy bound is

only achieved in the case of one single adaptive learning rate as then $\mathcal{K}(\pi', \pi_0) = \mathcal{K}(\pi, \pi_0)$. It is a drawback of the multiple learning rates procedures compared with the single ones of [LS15, KVE15] achieving such relative entropy bounds. Whether those extra terms can be avoided in multiple learning rates procedures is an open question.

4 Optimality of the BOA procedure in a stochastic environment

4.1 An empirical stochastic conversion

We now turn to a stochastic setting where (X_t, Y_t) are random elements observed recursively with $1 \leq t \leq n+1$ ($(X_0, Y_0) = (x_0, y_0)$ arbitrary are considered deterministic). Thanks to the empirical Bernstein inequality of Theorem 1.1, the cumulative predictive risk is bounded in term of the regret and a second order term. This new stochastic conversion is provided in Theorem 4.1 below. The main motivation of the introduction of the BOA procedure is the following reasoning: as a second order term appears necessarily in the stochastic conversion, a procedure that admits a similar second order term on the regret bound has nice properties in any stochastic environment. The BOA procedure achieves this strategy as the second order term of the regret bound is the same as the one appearing in the stochastic conversion. Let us go back for a moment to the most general case with no convex assumption on the loss and the notation:

$$\ell_{j,t} = \ell(X_t, f_{j,t}(X_{t-1})) - \mathbb{E}_{\pi_{t-1}}[\ell(X_t, f_{J,t}(X_t))]$$

for some online aggregation procedure $(\pi_t)_{0 \leq t \leq n}$, i.e. π_t is $\sigma(\mathcal{D}_t)$-measurable. Assume the existence of non increasing sequences $(\eta_{j,t})_t$ that are adapted to $(\mathcal{D}_t)$ for each $1 \leq j \leq M$ and that satisfy

$$\max_{1 \leq t \leq n+1} \max_{1 \leq j \leq M} \eta_{j,t-1} \ell_{j,t} \leq 1/2, \qquad a.s.. \tag{13}$$

We have the following general stochastic conversion that is also valid in the convex case with the associated linearized expression of $\ell_{j,t}$. It can be seen as an empirical counterpart of the online to batch conversion provided in [Zha05, KT08, GSVE14]. Thanks to the use of the cumulative predictive risk, the conversion holds in a completely general stochastic context; no assumption is done on the stochastic environment. Recall that $\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] = \sum_{t=1}^{n+1} \mathbb{E}_{\hat{\pi}_{t-1}}[\ell(Y_t f_{J,t}(X_t))]$.

Theorem 4.1. *Under* (13)*, the cumulative predictive risk of any aggregation procedure satisfies, with probability* $1 - e^{-x}$*,* $x > 0$*, for any* $1 \leq J \leq M$*:*

$$\begin{aligned}
&\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)] - R_{n+1}(f_j) \\
&\qquad \leq \mathbb{E}_{\hat{\pi}}[\mathrm{Err}_{n+1}(f_J)] - \mathrm{Err}_{n+1}(f_j) + \sum_{t=1}^{n+1} \eta_{j,t-1} \ell_{j,t}^2 + \frac{\log\left(1 + \mathbb{E}\left[\log\left(\frac{\eta_{j,1}}{\eta_{j,n}}\right)\right]\right) + x}{\eta_{j,n}}.
\end{aligned}$$

Proof. We first note that for each $1 \le j \le M$ the sequence $(M_{j,t})_t$ with $M_{j,t} = \eta(\mathbb{E}_{\hat{\pi}}[R_t(f_J)] - R_t(f_j) - (\mathbb{E}_{\hat{\pi}}[\mathrm{Err}_t(f_j)] - \mathrm{Err}_t(f_j)))$ is a martingale adapted to the filtration $(\mathcal{D}_t)$. Its difference is equal to $\Delta M_{j,t} = \eta(\mathbb{E}_{t-1}[\ell_{j,t}] - \ell_{j,t})$. Then the proof will follow from the classical recursive argument for supermartingales. However, as the learning rates $\eta_{j,t}$ are not necessarily constant, we adapt the recursive argument as in the proof of Theorem 3.3.

For any $1 \le j \le M$, $1 \le t \le n+1$, denoting $X = -\eta_{j,t-1}\ell_{j,t}$ we check that $X \ge -1/2$. We can apply (4) conditionally on $\mathcal{D}_{t-1}$ and we obtain

$$\mathbb{E}_{t-1}[\exp(-\eta_{j,t-1}(\ell_{j,t} - \mathbb{E}_{t-1}[\ell_{j,t}]) - \eta_{j,t-1}^2\ell_{j,t}^2)] \le 1.$$

Here we used the fact that $\eta_{j,t-1}$ is $\mathcal{D}_{t-1}$-measurable. Then we have

$$\mathbb{E}\Big[\exp\Big(-\eta_{j,t-1}\Big(\sum_{s=1}^{t}(\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big)\Big] \le \mathbb{E}\Big[\exp\Big(-\eta_{j,t-1}\Big(\sum_{s=1}^{t-1}(\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big)\Big].$$

To apply the recursive argument we use the basic inequality $x \le x^\alpha + (\alpha-1)/\alpha$ for $\alpha = \eta_{j,t-2}/\eta_{j,t-1} \ge 1$ and

$$x = \exp\Big(-\eta_{j,t-1}\Big(\sum_{s=1}^{t-1}(\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big).$$

We obtain

$$\mathbb{E}\Big[\exp\Big(-\eta_{j,t-1}\Big(\sum_{s=1}^{t}(\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big)\Big] \le \mathbb{E}\Big[\exp\Big(-\eta_{j,t-2}\Big(\sum_{s=1}^{t-1}(\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big)\Big] + \mathbb{E}\Big[\frac{\eta_{j,t-2} - \eta_{j,t-1}}{\eta_{j,t-1}}\Big].$$

The same recursive argument than in the proof of Theorem 3.3 is applied; we get

$$\mathbb{E}\Big[\exp\Big(-\eta_{j,n}\Big(\sum_{t=1}^{n}(\ell_{j,t} - \mathbb{E}_{t-1}[\ell_{j,t}]) - \eta_{j,s-1}\ell_{j,s}^2\Big)\Big)\Big] \le 1 + \mathbb{E}\Big[\log\Big(\frac{\eta_{j,1}}{\eta_{j,n}}\Big)\Big].$$

We end the proof by an application of the Chernoff device. □

4.2 Second order bounds on the excess of the cumulative predictive risk

Based on the result of Theorem 4.1, we derive from the main results of Section 3 second order bounds on the cumulative predictive risk of the BOA procedure. As an example, using the second order regret bound of Theorem 3.1 and the stochastic conversion of Theorem 4.1 we obtain

Theorem 4.2. *Assume that the non adaptive BOA procedure described in Figure 1 is such that $\eta = \eta_{j,t}$ for $1 \le j \le M$, $0 \le t \le n$ satisfies condition (6). The BOA procedure with $\ell_{j,t} = \ell(Y_t, f_{j,t-1}(X_t)) - \ell(Y_t, \hat{f}_{t-1}(X_t))$ has its cumulative predictive risk that satisfies, with probability $1 - e^{-x}$, $x > 0$:*

$$\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)] \le \inf_{\pi} \Big\{ \mathbb{E}_{\pi}\Big[R_{n+1}(f_j) + 2\eta \sum_{t=1}^{n+1} \ell_{J,t}^2\Big] + \frac{\mathcal{K}(\pi, \pi_0) + x}{\eta} \Big\}.$$

Moreover, if ℓ is convex with respect to its second argument, the BOA procedure with $\ell_{j,t} = \ell'(Y_t, \hat{f}_{t-1}(X_t))(f_{j,t-1}(X_t) - \hat{f}_{t-1}(X_t))$ has its cumulative predictive risk that also satisfies, with probability $1 - e^{-x}$, $x > 0$:

$$R_{n+1}(\hat{f}) \le \inf_{\pi} \Big\{ R_{n+1}(f_\pi) + 2\eta \sum_{t=1}^{n+1} \mathbb{E}_{\pi}[\ell_{J,t}^2] + \frac{\mathcal{K}(\pi, \pi_0) + x}{\eta} \Big\}.$$

Proof. We prove the result by integrating the result of Theorem 4.1 with respect to any deterministic π and noticing that, as the learning rates $\eta_{j,t} = \eta$ are constant in the BOA procedure described in figure 1, $\log(\eta_{j,1}/\eta_{j,n}) = 0$. □

The main advantage of the new stochastic conversion compared with the one of [GSVE14] is that the empirical second order bound can be chosen to be the same than in the regret bound. Similarly, we can extend Theorem 3.3; as the boundedness condition (13) is no longer satisfied for any $1 \le t \le n+1$, Theorem 4.1 does not apply directly. We still have

Theorem 4.3. *Under the hypothesis of Theorem 3.3, the cumulative predictive risk of the adaptive BOA procedure described in Figure 2 satisfies with probability $1 - e^{-x}$, $x > 0$,*

$$R_{n+1}(\hat{f}) \le \inf_{\pi} \Bigg\{ R_{n+1}(f_\pi) + 2\mathbb{E}_{\pi} \Bigg[\sqrt{\sum_{t=1}^{n+1} \ell_{J,t}^2} \left(\frac{\sqrt{2}\sqrt{\log(\pi_{J,0}^{-1})}}{\sqrt{2}-1} + \frac{2B_{n,E} + 8E_J + x}{\sqrt{\log(\pi_{J,0}^{-1})}} \right) + 8E_J(\log(\pi_{J,0}^{-1}) + 2B_{n,E} + 3 + x) \Bigg] \Bigg\},$$

where $B_{n,E} = \log(1 + 2^{-1}\log(n) + \log(E) + c\log(2))$ for all $n \ge 1$.

Proof. As the boundedness condition (13) is not satisfied for all $1 \le t \le n+1$, we cannot apply directly the result of Theorem 4.1. However, one can adapt the proof of the Theorem 4.1 as we adapted the proof of Theorem 4.1 for proving Theorem 3.3. As $-\eta_{j,t}(\ell_{j,t+1} - \mathbb{E}_{t-1}[\ell_{j,t+1}]) \le \eta_{j,t}E_{j,t+1}$, $1 \le j \le M$, we can apply the inequality (1)

$$\mathbb{E}_{t-1}\Big[\exp\Big(-\eta_{j,t}2^{-1}(\ell_{j,t+1} - \mathbb{E}_{t-1}[\ell_{j,t+1}]) - \eta_{j,t}E_{j,t+1}1_{\eta_{j,t}\ell_{j,t+1} > 1/2}\Big)\Big] \le 1.$$

Thus the proof ends by an application of the same recursive argument as in the proof of Theorem 3.3. □

Let us discuss the constants for uniform initial weights $\pi_{j,0} = M^{-1}$. First, one can use nE^2 as a crude estimate of $\sum_{t=1}^n \ell_{j,t}^2$. Second, if $n \leq e^M$ then $B_n \leq \log(M)$. Thus, if M is sufficiently large (for instance $M > \sqrt{n}$), $B_{n,E}$ is comparable with $\log(M)$. We continue the discussion for M sufficiently large satisfying $n \leq e^M$ and $M \geq c$. The second order bound becomes

$$R_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ R_{n+1}(f_\pi) + 2\mathbb{E}_\pi \left[\sqrt{\max_{1\leq j\leq M} \sum_{t=1}^{n+1} \ell_{j,t}^2} \right] \frac{8(\log(M)+E)+x}{\sqrt{\log(M)}} \right\} + 24E(\log M + 1 + x). \quad (14)$$

The second order term $\max_{1\leq j\leq M} \sum_{t=1}^n \ell_{j,t}^2$ is a natural candidate to assert the complexity of Problem (C); the more the $\sum_{t=1}^n \ell_{j,t}^2$ are uniformly small and the more one can aggregate the elements of the dictionary optimally. Moreover, this complexity term is observable and it would be interesting to develop a parsimonious strategy that would only aggregate the elements of the dictionary with small complexity terms $\sum_{t=1}^n \ell_{j,t}^2$. Reducing also the size M of the dictionary, the second order bound (14) can be reduced at the price to decrease the generality of Problem (C), i.e. the number of learners in $\mathcal{H}$.

The upper bound in (14) is an observable bound for Problem (C) similar than those arising from the PAC bayesian approach, see [Cat04] for a detailed study of such empirical bounds in the iid context. It would be interesting to know whether the complexity term $\max_{1\leq j\leq M} \sum_{t=1}^n \ell_{j,t}^2$ is the optimal one or not. We are not aware of empirical lower bounds for Problem (C). The bounds developed by [Nem00, Tsy03, Rig12] are deterministic. To assert the optimality of BOA, it is easy to turn from an empirical bound to a deterministic one. In the iid context, (X_t, Y_t) are iid copies of (X, Y) and the learners are assumed to be constant $f_j = f_{j,t}$, $t \geq 0$, $1 \leq j \leq M$. We then have $(n+1)^{-1} R_{n+1}(f_j) = \bar{R}(f_j) = \mathbb{E}[\ell(Y, f_j(X))]$. It is always preferable to convert any online learner $\hat{f}$ to a batch learner by averaging

$$\bar{f} = \frac{1}{n+1} \sum_{t=0}^{n} \hat{f}_t$$

as an application of the Jensen inequality gives $\bar{R}(\bar{f}) \leq (n+1)^{-1} R_{n+1}(\hat{f})$.As $\ell_{j,t}^2 \leq E^2$, Equation (14) implies that the batch version of BOA satisfies, with high probability in the iid setting,

$$\bar{R}(\bar{f}) \leq \inf_{\pi} \bar{R}(f_\pi) + 16E\sqrt{\frac{\log(M)}{n+1}} + \frac{16E^2 + 2x}{\sqrt{(n+1)\log(M)}} + 24E\frac{\log M + 1 + x}{n+1}.$$

Then the BOA procedure is optimal for Problem (C) in the sense of the Definition 1.1: in the iid context, the excess of risk of the batch version of the BOA procedure is of order $\sqrt{\log(M)/n}$. Notice that the generality of this optimal rate is remarkable. With no assumption on the dependence structure of the stochastic environment, this rate is also the one of the mean predictive risk for Problem (C). However, as the mean predictive risk is not deterministic, it should be more natural to have an empirical lower bound in order than a deterministic one in order to define the notion of optimality in non iid settings.

4.3 Optimal learning for Problem (MS)

The BOA procedure is optimal for Problem (C) and the optimal rate of convergence is also valid in any stochastic environment on the excess of mean predictive risk. To turn to Problem (MS), we restrict our study to the context of Lipschitz strongly convex losses with iid observations. Remind that from [Tsy03, Rig12] the optimal rate for Problem (MS) is a fast rate of convergence $\log(M)/n$. Such fast rates cannot be obtained without regularity assumption on the loss ℓ that makes it comparable to the square loss locally, see for instance [Aud09]. In the sequel $\ell : \mathbb{R}^2 \to \mathbb{R}$ is a loss function satisfying the assumption called **(LIST)** after [KT08]

(LIST) the loss function ℓ is C_ℓ-strongly convex and C_b-Lipschitz continuous in its second coordinate on a convex set $\mathcal{C} \subset \mathbb{R}$.

Recall that a function g is c strongly convex on $\mathcal{C} \subset \mathbb{R}$ if there exists a constant $c > 0$ such that

$$g(\alpha a + (1-\alpha)a') \leq \alpha g(a) + (1-\alpha)g(a') - \frac{c}{2}\alpha(1-\alpha)(a-a')^2,$$

for any $a, a' \in \mathcal{C}$, $0 < \alpha < 1$. Under the condition **(LIST)**, few algorithms are known to be optimal in deviation, see [Aud07, LM09, LR14].

Notice that Assumption **(LIST)** is restrictive and can hold only locally; on a compact set $\mathcal{C}$, the minimizer $f(y)^*$ of $f(y) \in \mathbb{R} \to \ell(y, f(y))$ exists and verifies, by strong convexity,

$$\ell(y, f(y)) \geq \ell(y, f(y)^*) + \frac{C_\ell}{2}(f(y) - f(y)^*)^2.$$

Moreover, by Lipschitz continuity, $\ell(y, f(y)) \leq \ell(y, f(y)^*) + C_b|f(y) - f(y)^*|$. Thus, necessarily the diameter D of $\mathcal{C}$ is finite and satisfies $C_\ell D \leq 2C_b$. Then we deduce that $|\ell_{j,t}| \leq C_b D$, $1 \leq t \leq n+1$, $1 \leq j \leq M$, and under **(LIST)** the ranges are estimated by $E = C_b D$.

We obtain the optimality of the BOA procedure for Problem (MS). The result extends easily (with different constants) to any online procedures achieving second order regret bounds on the linearized loss similar than BOA such as the procedures described in [GSVE14, LS15, KVE15].

Theorem 4.4. *In the iid setting, under the condition* **(LIST)**, *for the uniform initial weights $\pi_{j,0} = M^{-1}$, $1 \leq j \leq M$, and any learning rate satisfying*

$$C_b^2(16(e-1)/C_\ell \vee D^2) \leq \eta^{-1}, \tag{15}$$

then the cumulative predictive risk of the BOA procedure described in Figure 1 and the risk of its batch version satisfy, with probability $1 - e^{-x}$,

$$\bar{R}(\bar{f}) + \frac{C_\ell}{2(n+1)} \sum_{t=0}^{n} \mathbb{E}[(\hat{f}_t(X) - \bar{f}(X))^2] \leq \frac{R_{n+1}(\hat{f})}{n+1} \leq \min_{1 \leq j \leq M} \bar{R}(f_j) + \frac{2\log(M) + 3x}{\eta(n+1)}.$$

Proof. Inspired by the Q-aggregation procedures of [LR14], we start the proof by adding the two second order empirical bounds obtained in Theorem 4.2 (using that $\mathcal{K}(\pi, \pi_0) \leq \log(M)$):

$$R_{n+1}(\hat{f}) + \mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)] \leq \inf_{\pi} \Big\{ R_{n+1}(f_\pi) + \mathbb{E}_\pi[R_{n+1}(f_J)] + 4\eta \sum_{t=1}^{n+1} \mathbb{E}_\pi[\ell_{J,t}^2] + 2\frac{x + \log(M)}{\eta} \Big\}. \tag{16}$$

Then we convert the empirical second order term into a deterministic one. From the "poissonnian" inequality of Lemma A3 of [CBL06], as under (15) we have $0 \leq \eta \ell_{j,t}^2 \leq \eta C_b^2 D^2 \leq 1$, we have

$$\mathbb{E}_{t-1}[\exp(\eta \ell_{j,t}^2 - (e-1)\eta \mathbb{E}_{t-1}[\ell_{j,t}^2])] \leq 1$$

Applying the recursive supermartingales argument and an union bound, we obtain the deterministic version of the upper bound (16): with probability $1 - e^{-x}$, $x > 0$,

$$\begin{aligned} R_{n+1}(\hat{f}) + \mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)] \leq \inf_{\pi} \Big\{ & R_{n+1}(f_\pi) + \mathbb{E}_\pi[R_{n+1}(f_J)] \\ & + 4(e-1)\eta \sum_{t=1}^{n+1} \mathbb{E}_\pi[\mathbb{E}_{t-1}[\ell_{J,t}^2]] + \frac{3x + 2\log(M)}{\eta} \Big\}. \end{aligned}$$

The optimal fast rate comes from a careful analysis of the second order deterministic bound. From the Lipschitz property, the sub-gradient ℓ' is bounded by C_b and thus

$$\begin{aligned} \mathbb{E}_\pi[\mathbb{E}_{t-1}[\ell_{J,t}^2]] &\leq C_b^2 \mathbb{E}_\pi[\mathbb{E}_{t-1}[(f_J(X_{t-1}) - \hat{f}_{t-1}(X_{t-1}))^2]] \\ &\leq C_b^2 (V(\pi) + \mathbb{E}[(f_\pi(X) - \hat{f}_{t-1}(X))^2]). \end{aligned}$$

where $V(\pi) = \mathbb{E}_\pi[\mathbb{E}[(f_J(X) - f_\pi(X))^2]]$. As $R_{n+1}(f_\pi) = (n+1)\bar{R}(f_\pi)$, $\mathbb{E}_\pi[R_{n+1}(f_J)] = (n+1)\mathbb{E}_\pi[\bar{R}(f_J)]$ combining those bounds we obtain

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)]}{n+1} \leq \inf_{\pi} \Big\{ & \bar{R}(f_\pi) + \mathbb{E}_\pi[\bar{R}(f_J)] \\ & + \gamma \Big(V(\pi) + \frac{1}{n+1} \sum_{t=0}^{n} \mathbb{E}[(f_\pi(X) - \hat{f}_t(X))^2] \Big) + \frac{3x + 2\log(M)}{\eta(n+1)} \Big\} \end{aligned} \tag{17}$$

with $\gamma = 4C_b^2(e-1)\eta$. The rest of the proof is inspired by the reasoning of [LR14]. First, one can check that

$$V(\pi) - V(\pi') = < \nabla V(\pi'), (\pi - \pi') > - \mathbb{E}[(f_\pi(X) - f_{\pi'}(X))^2]$$

where π and π' are any weights vectors and $< \cdot, \cdot >$ denotes the scalar product on $\mathbb{R}^M$. Moreover, by C_ℓ-strong convexity, one can also check that

$$\bar{R}(f_\pi) - \bar{R}(f_{\pi'}) \geq < \nabla \bar{R}(f_{\pi'}), (\pi - \pi') > + \frac{C_\ell}{2} \mathbb{E}[(f_\pi(X) - f_{\pi'}(X))^2].$$

Thus the function $H\colon \pi \to \bar{R}(f_\pi) + \mathbb{E}_\pi[\bar{R}(f_j)] + \gamma V(\pi)$ is convex as $0 \le \gamma \le C_\ell/2$ under (15). Moreover, if one denotes π^* a minimizer of H, we have for any weights π

$$H(\pi) - H(\pi^*) \ge \Big(\frac{C_\ell}{2} - \gamma\Big)\mathbb{E}[(f_\pi(X) - f_{\pi^*}(X))^2].$$

Thus, applying this inequality to $\hat{\pi}$ we have

$$\frac{C_\ell/2 - \gamma}{n+1}\sum_{t=0}^{n}\mathbb{E}[(f_{\pi^*}(X) - \hat{f}_t(X))^2] \le \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)]}{n+1} - \Big(\bar{R}(f_{\pi^*}) + \mathbb{E}_{\pi^*}[\bar{R}(f_J)] + \gamma V(\pi^*)\Big) + \frac{\gamma}{n+1}\sum_{t=0}^{n} V(\pi_t).$$

Then, combining this last inequality with the inequality (17), we derive that

$$\frac{C_\ell/2 - \gamma}{n+1}\sum_{t=0}^{n}\mathbb{E}[(f_{\pi^*}(X) - \hat{f}_t(X))^2] \le \frac{3x + 2\log(M)}{\eta(n+1)} + \frac{\gamma}{n+1}\sum_{t=0}^{n} V(\pi_t).$$

Plugging in this new estimate into (17) we obtain

$$\frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_J)]}{n+1} - \frac{2\gamma^2}{C_\ell - 2\gamma}\frac{1}{n+1}\sum_{t=0}^{n} V(\pi_t) \le \bar{R}(f_{\pi^*}) + \mathbb{E}_{\pi^*}[\bar{R}(f_J)] + \gamma V(\pi^*) + \frac{C_\ell}{C_\ell - 2\gamma}\frac{3x + 2\log(M)}{\eta(n+1)}.$$

Now, using C_ℓ-strong convexity as in Proposition 2 of [LR14], we have for any probability measure π

$$\bar{R}(f_\pi) \le \mathbb{E}_\pi[\bar{R}(f_J)] - \frac{C_\ell V(\pi)}{2}.$$

As under condition (15) it holds

$$\frac{2\gamma^2}{C_\ell - 2\gamma} \le \frac{C_\ell}{2} \qquad \text{and} \qquad \frac{C_\ell}{C_\ell - 2\gamma} \le 2,$$

we can use the strong convexity argument for any π_t, $0 \le t \le n$. We obtain

$$2\frac{R_{n+1}(\hat{f})}{n+1} \le \bar{R}(f_{\pi^*}) + \mathbb{E}_{\pi^*}[\bar{R}(f_J)] + \gamma V(\pi^*) + 2\frac{3x + 2\log(M)}{\eta(n+1)}.$$

The proof ends by noticing that by definition of π^* we have

$$\bar{R}(f_{\pi^*}) + \mathbb{E}_{\pi^*}[\bar{R}(f_J)] + \gamma V(\pi^*) \le 2\min_{1\le j\le M}\bar{R}(f_j).$$

The lower bound on $R_{n+1}(\hat{f})/(n+1)$ follows by an application of the strong convexity argument applied to $\bar{f} = (n+1)^{-1}\sum_{t=0}^{n}\hat{f}_t$. □

Theorem 4.4 provides the optimality of the BOA procedure for Problem (MS) because

$$\bar{R}(\bar{f}) \leq \min_{1\leq j\leq M} \bar{R}(f_j) + \frac{2\log(M)+3x}{\eta(n+1)}.$$

The additional term

$$\frac{C_\ell}{2(n+1)} \sum_{t=0}^{n} \mathbb{E}[(\hat{f}_t(X) - \bar{f}(X))^2]$$

is the benefit of considering the batch version of BOA under the strong convexity assumptions **(LIST)**. As the fast rate is optimal, the partial sums $\sum_{t=0}^{n} \mathbb{E}[(\hat{f}_t(X) - \bar{f}(X))^2]$ might converge to a small constant. Assuming that $\bar{f}$ is converging with n, the convergence of the partial sums implies that $\mathbb{E}[(\hat{f}_n(X) - \bar{f}(X))^2] = o(n^{-1})$. Thanks to the Lipschitz assumption on the loss, it implies that the difference $|\bar{R}(\bar{f}) - \bar{R}(\hat{f}_n)| \leq C_b \mathbb{E}[(\hat{f}_n(X) - \bar{f}(X))^2]$ is small. The difference $|\bar{R}(\bar{f}) - \bar{R}(\hat{f}_n)|$ is then negligible compared with the fast rate $\log(M)/n$. Then, at the price of some constant $C > 1$, we also have

$$\bar{R}(\hat{f}_n) \leq \min_{1\leq j\leq M} \bar{R}(f_j) + C\frac{2\log(M)+3x}{\eta(n+1)}.$$

It means that in the iid setting, the predictive risk of the online procedure $\bar{R}(\hat{f}_n) = \mathbb{E}[\ell(Y_{n+1}, \hat{f}_n(X_{n+1}) \mid \mathcal{D}_n]$ might be optimal. It would be interesting to verify when it is the case and to extend this result to non iid settings following the reasoning developed in [MR10].

The tuning parameter η can be considered as the inverse of the temperature β of the Q-aggregation procedure studied in [LR14]. In the Q-aggregation, the tuning parameter β is required to be larger than $60C_b^2/C_\ell$. It is a condition similar than our restriction (15) on η. The larger is η satisfying the condition (15) and the best is the rate of convergence. The choice $\eta^* C_b^2(16(e-1)/C_\ell \vee D^2) = 1$ is optimal. The resulting BOA procedure is non adaptive in the sense that it depends on the constants appearing in the condition **(LIST)** and on the range $C_b D$ that can be unknown. On the contrary, the multiple learning rates BOA procedure achieves to tune automatically the learning rates. At the price of larger "constants" that grow as $\log\log(n)$, we extend the preceding optimal rate of convergence to the adaptive BOA procedure:

Theorem 4.5. *In the iid setting, under the condition* **(LIST)**, *for the uniform initial weights $\pi_{j,0} = M^{-1}$, $1 \leq j \leq M$ with $M \geq 3$, the mean predictive risk of the adaptive BOA procedure described in Figure 2 and the risk of its batch version satisfy, with probability $1 - e^{-x}$,*

$$\begin{aligned}
&\bar{R}(\bar{f}) + \frac{C_\ell}{2(n+1)} \sum_{t=0}^{n} \mathbb{E}[(\hat{f}_t(X) - \bar{f}(X))^2] \\
&\quad \leq \frac{R_{n+1}(\hat{f})}{n+1} \leq \min_{1\leq j\leq M} \bar{R}(f_j) + \frac{48\log(M) + 16B_{n,C_bD}^2 + 256(C_bD)^2 + 4x^2 + x}{\eta^*(n+1)} \\
&\qquad\qquad + \frac{16C_bD(\log(M) + 2B_{n,C_bD} + 3 + x)}{n+1}.
\end{aligned}$$

where $B_{n,C_bD} = \log(1 + 2^{-1}\log(n) + \log(C_bD) + c\log(2))$ *for all* $n \geq 1$.

Proof. The proof starts from the second order empirical bound provided in Theorem 3.3 in the iid context under (**LIST**), where $|\ell_{j,t}| \leq C_bD$; using the Young inequality, we have for any $\eta > 0$

$$R_{n+1}(\hat{f}) + \mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)] \leq \inf_{\pi}\Big\{\bar{R}(f_\pi) + \mathbb{E}_\pi[R(\bar{f}_j)] + 4\eta\mathbb{E}_\pi\Big[\sum_{t=1}^{n+1}\ell_{j,t}^2\Big]\Big\}$$
$$+ 4\frac{B(n, C_bD, M) + x^2}{\eta} + 16C_bD(\log(M) + x + 2B_{n,C_bD} + 3),$$

where, using that $(a+b+c+d)^2 \leq 4(a^2+b^2+c^2+d^2)$, a, b, c and $d > 0$, $(\sqrt{2}/(\sqrt{2}-1))^2 \leq 12$, $B_{n,E_j} \leq B_{n,C_bD}$ and $\log(M) \geq 1$,

$$B(n, C_bD, M) \leq 12\log(M) + 4B_{n,C_bD}^2 + 64(C_bD)^2.$$

Then we can use the "poissonnian" inequality as in the proof of Theorem 4.4 to obtain the deterministic second order bound, with $\gamma = 4C_b^2(e-1)\eta$,

$$\frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} \leq \inf_{\pi}\Big\{\bar{R}(f_\pi) + \mathbb{E}_\pi[\bar{R}(f_j)] + \frac{\gamma}{n+1}\mathbb{E}_\pi\Big[\sum_{t=1}^{n+1}\mathbb{E}_{t-1}[\ell_{j,t}^2]\Big]\Big\}$$
$$+ \frac{4B(n, C_bD, M) + 4x^2 + x}{\eta(n+1)} + \frac{16C_bD(\log(M) + x + 2B_{n,C_bD} + 3)}{n+1}.$$

The proof ends similarly than the one of Theorem 4.4. For η^* satisfying the equality in the condition (15), we obtain

$$\frac{R_{n+1}(\hat{f})}{n+1} \leq \min_{1\leq j\leq M}\bar{R}(f_j) + \frac{4B(n, C_bD, M) + 4x^2 + x}{\eta^*(n+1)} + \frac{16C_bD(\log(M) + x + 2B_{n,C_bD} + 3)}{n+1}.$$

The result follows from the expression of $B(n, C_bD, M)$ and the strong convexity of the risk. □

The BOA procedure is explicitly computed with complexity $O(Mn)$. It is a practical advantage compared with the batch procedures studied in [Aud07, LM09, LR14] that require a computational optimization technique. This issue has been solved in [DRXZ12] for the square loss using greedy iterative algorithms that approximate the Q-aggregation procedure.

Acknowledgments

I am grateful to two anonymous referees for their helpful comments. I would also like to thank Pierre Gaillard and Gilles Stoltz for valuable comments on a preliminary version.

References

[AABR09] J. Abernethy, A. Agarwal, P.L. Bartlett, and A. Rakhlin, *A stochastic view of optimal regret through minimax duality*, COLT, 2009.

[AD13] A. Agarwal and J. C. Duchi, *The generalization ability of online algorithms for dependent data*, Information Theory, IEEE Trans. **59** (2013), 573–587.

[ALW13] P. Alquier, X. Li, and O. Wintenberger, *Prediction of time series by statistical learning: General losses and fast rates*, Dependence Modeling **1** (2013), 65–93.

[AMS06] J. Y. Audibert, R. Munos, and C. Szepesvari, *Use of variance estimation in the multi-armed bandit problem*, NIPS, 2006.

[Aud07] J.-Y. Audibert, *Progressive mixture rules are deviation suboptimal*, Advances in Neural Information Processing Systems, 2007, pp. 41–48.

[Aud09] ———, *Fast learning rates in statistical inference through aggregation*, Ann. Statist. **37** (2009), 1591–1646.

[BM05] A. Blum and Y. Mansour, *From external to internal regret*, Learning theory, Springer, 2005, pp. 621–636.

[Cat04] O. Catoni, *Statistical learning theory and stochastic optimization*, Notes in Mathematics (Saint-Flour Summer School on Probability Theory 2001), Springer, 2004.

[Cat07] ———, *Pac-bayesian supervised classification: the thermodynamics of statistical learning*, Institute of Mathematical Statistics, Beachwood, OH, 2007.

[CBL06] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*, Cambridge University Press, NY, USA, 2006.

[CBMS07] N. Cesa-Bianchi, Y. Mansour, and G. Stoltz, *Improved second-order bounds for prediction with expert advice*, Mach. Learn. **66** (2007), 321–352.

[DRXZ12] D. Dai, P. Rigollet, L. Xia, and T. Zhang, *Deviation optimal learning using greedy Q-aggregation.*, Ann. Statist. **40** (2012), 1878–1905.

[DV75] M. D. Donsker and S. S. Varadhan, *Asymptotic evaluation of certain markov process expectations for large time, I*, Commun. Pure Appl. Math. **28** (1975), 1–47.

[Fre75] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probab. **3** (1975), 100–118.

[Ger13] S. Gerchinovitz, *Sparsity regret bounds for individual sequences in online linear regression*, JMLR **14** (2013), 729–769.

[GSVE14] P. Gaillard, G. Stoltz, and T. Van Erven, *A second-order bound with excess losses*, COLT, 2014, arXiv preprint arXiv:1402.2044.

[HK10] E. Hazan and S. Kale, *Extracting certainty from uncertainty: Regret bounded by variation in costs*, Mach. Learn. **80** (2010), 165–188.

[HKW98] D. Haussler, J. Kivinen, and M. K. Warmuth, *Sequential prediction of individual sequences under general loss functions*, Information Theory, IEEE Transactions **44** (1998), 1906–1925.

[JRT08] A. Juditsky, P. Rigollet, and A. B. Tsybakov, *Learning by mirror averaging*, Ann. Statist. **36** (2008), 2183–2206.

[KT08] S. M. Kakade and A. Tewari, *On the generalization ability of online strongly convex programming algorithms*, NIPS, 2008.

[KVE15] W. Koolen and T. Van Erven, *Second-order quantile methods for experts and combinatorial games*, COLT, 2015, pp. 1155–1175.

[LM09] G. Lecué and S. Mendelson, *Aggregation via empirical risk minimization*, PTRF **145** (2009), 591–613.

[LR14] G. Lecué and P. Rigollet, *Optimal learning with Q-aggregation*, Ann. Statist. **42** (2014), no. 1, 211–224.

[LS15] H. Luo and R.E. Schapire, *Achieving all with no parameters: Adanormalhedge*, Proceedings of The 28th Conference on Learning Theory, 2015, pp. 1286–1304.

[MP09] A. Maurer and M. Pontil, *Empirical bernstein bounds and sample variance penalization*, COLT, 2009.

[MR10] M. Mohri and A. Rostamizadeh, *Stability bounds and for ϕ-mixing and β-mixing processes*, JMLR **4** (2010), 1–26.

[Nem00] A. Nemirovski, *Topics in non-parametric statistics*, Lectures on Probability Theory and Statistics: Ecole d'Ete de Probabilites de Saint-Flour XXVIII-1998 **28** (2000), 85.

[Rig12] P. Rigollet, *Kullback-Leibler aggregation and misspecified generalized linear models*, Ann. Statist. **40(2)** (2012), 639–665.

[Tsy03] A. B. Tsybakov, *Optimal rates of aggregation*, COLT, Springer Berlin Heidelberg, 2003.

[Vov90] V.G. Vovk, *Aggregating strategies*, COLT, 1990.

[Zha05] T. Zhang, *Data dependent concentration bounds for sequential prediction algorithms*, Learning Theory, Springer Berlin Heidelberg, 2005.