



HAL
open science

Optimal learning with Bernstein Online Aggregation

Olivier Wintenberger

► **To cite this version:**

| Olivier Wintenberger. Optimal learning with Bernstein Online Aggregation. 2015. hal-00973918v3

HAL Id: hal-00973918

<https://hal.science/hal-00973918v3>

Preprint submitted on 4 Feb 2015 (v3), 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 problem in the iid setting, see [Tsy03]; the excess of risk of its batch version achieves the fast rate of convergence $\log(M)/n$ in deviation when the loss is Lipschitz and strongly convex. 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 online to batch conversion that estimates the deviations in the stochastic environment with random second order terms. It converts a second order regret bound for BOA similar than in [GSVE14] to a bound on the cumulative predictive risk in a general stochastic context. The empirical 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 $\mathcal{F}_t = \{(X_1, Y_1), \dots, (X_t, Y_t)\}$ 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{F}_t . In this paper, we index with the subscript t any random element that is adapted with \mathcal{F}_t . A learner is a function $\mathcal{X} \mapsto \mathbb{R}$, denoted f_t , that depends only on the past observations \mathcal{F}_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

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

where $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a loss function. We define an online learner f as a recursive algorithm that produces at each time $t \geq 1$ a learner: $f = (f_0, f_1, f_2, \dots)$. The accuracy of an online learner is quantified by the cumulative predictive risk

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

Given a finite set $\mathcal{H} = \{f_1, \dots, f_M\}$ of online learners, we aim at finding optimal online aggregation procedures

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

with \mathcal{F}_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

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

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 set $\{1, \dots, M\}$. 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_π is denoted $\mathbb{E}_\pi[f_j]$ and, with some abuse, \hat{f} is denoted $\mathbb{E}_{\hat{\pi}}[f_j] = (\mathbb{E}_{\pi_0}[f_{j,0}], \mathbb{E}_{\pi_1}[f_{j,1}], \mathbb{E}_{\pi_2}[f_{j,2}], \dots)$. The predictive performance of an online aggregation procedure \hat{f} is compared with the best element of \mathcal{H} or with the best deterministic aggregation of \mathcal{H} . Following the pioneer works [Nem00, Tsy03], we refer to these two different objectives as, respectively, the convex Problem (C) or the model selection Problem (MS). The performance of online aggregation procedures is usually measured using the deterministic regret (in the context of individual sequences prediction, see the seminal book [CBL06]). In this paper, we use instead the excess of cumulative predictive risk, defined as

$$\text{Problem (C): } R_{n+1}(\hat{f}) - \inf_{\pi} R_{n+1}(f_\pi), \quad \text{Problem (MS): } R_{n+1}(\hat{f}) - \min_j R_{n+1}(f_j).$$

A new online to batch conversion is used to extend the second order regret bounds obtained for the deterministic regret to the cumulative predictive risk for any stochastic environment, see Theorem 4.1. Thanks to the use of the cumulative predictive risk, no assumption is required on the temporal dependence structure of the stochastic process (X_t, Y_t) . However, there is no warranty of the optimality of the procedure at that stage. To define properly the notion of optimality, we will 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. The batch version of an online aggregation procedure \hat{f} is defined as $\bar{f} = (n+1)^{-1} \sum_{t=0}^n \hat{f}_t$. The cumulative predictive risk $R_{n+1}(\hat{f})$ provides an upper bound

for the predictive risk $R(\bar{f})$ with $R(f) = \mathbb{E}[\ell(Y, f(X))]$: Applying Jensen's inequality, we have for any convex loss ℓ that

$$R(\bar{f}) \leq \frac{R_{n+1}(\hat{f})}{n+1} = \frac{1}{n+1} \sum_{t=1}^{n+1} \mathbb{E}[\ell(Y_t, \hat{f}_{t-1}(X_t)) \mid \mathcal{F}_{t-1}].$$

In the iid setting, lower bounds for the excesses of risk

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

are provided in [Nem00, Tsy03]. These lower bounds are called optimal rates of convergence. For Problem (C), the optimal rate is $\sqrt{\log M/n}$ when $M > \sqrt{n}$ and for Problem (MS) the optimal rate is $\log M/n$. The latter rate is called a fast rate of convergence. We are now ready to define the notion of optimality of an online aggregation procedure we will use:

Definition 1.1 (from [Tsy03] in the batch setting). *An online aggregation procedure is optimal for Problems (C) or (MS) if its batch version achieves, respectively, the optimal rates $\sqrt{\log M/n}$ or $\log M/n$ with high probability, i.e. there exists $C > 0$ such that with probability $1 - e^{-x}$, $x > 0$, it holds*

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

The notion of optimality for Problem (MS) is very restrictive. Very few known procedures achieve the fast rate $\log M/n$ with high probability and none of them are issued from online procedures. In this article, we provide the first online aggregation procedure, called Bernstein Online Aggregation (BOA), that is proved to be optimal. Before defining it properly, let us review the existing optimal procedures for Problem (MS).

The batch procedures in [Aud07, LM09, LR13] achieve the optimal rate in deviation. A priori, they are less explicit as they require to optimize a non regular criteria. This practical issue has been solved in the context of quadratic loss with gaussian noise in [DRXZ12]. The algorithm is a recursive greedy one and not an online one. On the opposite, before our work no online algorithm achieves the optimal fast rate in deviation. Most popular progressive aggregation rules are exponential weights algorithms (EWA) studied in [Vov90, HKW98]. Batch versions of EWA coincides with the Progressive Mixture Rules (PRMs). 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 in the iid context; see [Cat04, JRT08]. However, PRMs are suboptimal in deviation, i.e. the optimal rate cannot hold with high probability, see [Aud07, DRXZ12].

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. Other procedures already exist with different second 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

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

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

thanks to the new online to batch procedure that we describe below. Notice that the second order refinement $\ell_{j,t}$ 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 risk by following standard arguments. We first derive a second order bound on the regret (or excess of losses) in the deterministic setting: for any deterministic π

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

Then we extend it to an upper bound on the excess of risk $R_n(\hat{f}) - R_n(f_\pi)$ in any stochastic environment. In previous works, the online to batch conversion follows from an application of a Bernstein inequality for martingales. It provide a control of the deviations in the stochastic environment via the predictable quadratic variation, see for instance [Fre75, Zha05, KT08]. 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 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 $\mathbb{E}[\exp(M_n - [M]_n)] \leq 1$.*

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 as π_{t-1} on $\{1, \dots, M\}$ given \mathcal{F}_{n+1} ,
2. $M_{j,t} = \eta(R_t(\hat{f}) - R_t(f_j) - \mathcal{R}_t(\hat{f}) + \mathcal{R}_t(f_j))$ 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{F}_{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 2. will provide the online to batch conversion. In both cases, the second order term will be equal to $\eta^{-1} [M]_{n+1} = \eta \sum_{t=1}^{n+1} \ell_{j,t}^2$ after renormalization. It is the main motivation of BOA; as our notion of optimality requires

an online to batch conversion, the necessary cost is a second order term appearing in the empirical online to batch conversion. An online procedure will achieve good performances in the batch setting if the second order bound of the regret is similar than this necessary cost. We describe below why the BOA procedure achieves this aim.

In the first application 1. of Theorem 1.1, we have $\mathbb{E}_{\pi_{t-1}}[\Delta M_t] = 0$ and we verify that ΔM_t is centered. An application of Theorem 1.1 yields the regret bound of Theorem 3.1:

$$\mathbb{E}_{\hat{\pi}}[\mathcal{R}_{n+1}(f_j)] \leq \inf_{\pi} \left\{ \mathbb{E}_{\pi} \left[\mathcal{R}_{n+1}(f_j) + \eta \sum_{t=1}^{n+1} \ell_{j,t}^2 + \frac{\log(\pi_j/\pi_{j,0})}{\eta} \right] \right\}.$$

Such second order bounds also hold for the regret of other algorithms, see [CBMS07]. The procedure introduced in [HK10] achieves a second order regret bound that is not comparable with the one of BOA. In [GSVE14], the authors introduced the ML-prod and ML-pol procedures that satisfies the same regret bound. Using our new online to batch conversion, this second order regret bound is converted to the stochastic setting. Application 2. of Theorem 1.1 yields the bound on the excess of cumulative risk of Theorem 4.2. With probability $1 - e^{-x}$, $x > 0$, we have

$$\mathbb{E}_{\hat{\pi}}[R_{n+1}(f)] \leq \inf_{\pi} \left\{ \mathbb{E}_{\pi} \left[R_{n+1}(f) + 2\eta \sum_{t=1}^{n+1} \ell_{j,t}^2 + \frac{\log(\pi_j/\pi_{j,0}) + x}{\eta} \right] \right\}.$$

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 sub-gradient trick to bound the excess of 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).

In order to solve Problem (C), we use the sub gradient trick, 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

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

Then we consider the second order refinement

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

With some abuse of notation, we will still denote it as $\ell_{j,t}$. 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 on the regret for Problem (C)

$$\mathcal{R}_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \eta \sum_{t=1}^{n+1} \mathbb{E}_{\pi}[\ell_{j,t}^2] + \frac{\mathbb{E}_{\pi}[\log(\pi_j/\pi_{j,0})]}{\eta} \right\}.$$

The bounds ale holds on the excess of risk with high probability from an application of the new online to batch conversion, see Theorem 4.2.

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 also solve this issue by introducing the multiple learning rates version of BOA in Figure 2. Here we are

Parameters: Learning rates $\eta_j > 0$.

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 $\pi_t = (\pi_{j,t})_{1 \leq j \leq M}$:

$$\pi_{j,t} = \frac{\exp(-\eta_j \ell_{j,t}(1 + \eta_j \ell_{j,t})) \pi_{j,t-1}}{\mathbb{E}_{\pi_{t-1}}[\exp(-\eta_j \ell_{j,t}(1 + \eta_j \ell_{j,t}))]}. \quad (2)$$

Figure 2: The multiple learning rates BOA algorithm

able to extend the second order regret bounds for Problem (C) to the multiple learning rates BOA procedure in Theorem 3.2. The multiple learning rates η_j can be tuned as $\sqrt{\log(\pi_{j,0}^{-1})}/\sqrt{\max_j \sum_{t=1}^{n+1} \ell_{j,t}^2}$. The resulting procedure satisfies a second order regret bound but is not recursive. The random variable $M_t = R_t(\hat{f}) - R_t(f_\pi) - (\mathcal{R}_t(\hat{f}) - \mathcal{R}_t(f_\pi))$ is not adapted to \mathcal{F}_t but to \mathcal{F}_{n+1} . One cannot apply the online to batch conversion and one cannot provide bounds in deviation on the excess of risk.

Thus, we introduce a recursive procedure with adaptive multiple learning rates $(\eta_{j,t})$ in Figure 3. The novelty, compared with classical adaptive procedures developed in

Parameter: a rule to sequentially pick the learning rates $(\eta_{j,t})$ for $1 \leq j \leq M$ and $1 \leq t \leq n$.

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

$$L_{j,t} = L_{j,t-1} + \ell_{j,t}(1 + \eta_{j,t-1} \ell_{j,t}),$$

2. 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 3: The adaptive BOA procedure

[CBMS07], is the dependence of the learning rates with respect to j and that the learning rates appear in the exponential and as a factor. We consider the learning rates

$$\eta_{j,t} = \min \left\{ \frac{1}{2E_j}, \sqrt{\frac{\log(\pi_{j,0}^{-1})}{\sum_{s=1}^t \ell_{j,s}^2}} \right\}, \quad t \geq 0,$$

where E_j is a known estimate of the range of the linearized loss (1) of the learner f_j , $1 \leq j \leq M$. We also give a fully adaptive version of the algorithm for cases when the ranges E_j are unknown. For these adaptive BOA procedures, we obtain regret bounds such as

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

for some "constant" $C > 0$ that grows as $\log \log(n)$, see Theorems 3.3 and 3.4 for details. Such second order bounds involving excess losses terms as the $\ell_{j,t}$ s have been proved for other algorithms in [GSVE14], and we refer to this article for nice consequences of such bounds. Here again, the online to batch conversion holds in any stochastic environment and we obtain with probability $1 - e^{-x}$, $x > 0$

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

It is remarkable to obtain a result for the excess of risk with no assumption on the dependence of the stochastic observations. Formerly, such inequalities were derived under very restrictive dependent assumptions, see [ALW13, MR10, AD13]. The generality of our results is due to the use of the cumulative predictive risk. It is the correct criteria to assert the prediction accuracy of online algorithms as it coincides with the cumulative loss for deterministic observations and with the classical risk 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 optimality for Problem (C) is proved in the very general setting: under a boundedness assumption, the rate of convergence of $R_{n+1}(\hat{f})/(n+1)$ is smaller than $\sqrt{\log(M)/n}$.

The fast rate of convergence $\log(M)/n$ of Problem (MS) comes from a careful study of the second order terms $\sum_{t=1}^{n+1} \ell_{j,t}^2$. It is known that it also requires more conditions on the loss, see [Lec07, 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. Inspired by the work of [LR13], we modify the loss in the BOA procedure and we use the mixture of the original and the linearized loss. The second order refinement in the BOA procedure becomes:

$$\ell_{j,t} = \ell'(Y_t, \hat{f}_{t-1}(X_t))(f_{j,t-1}(X_t) - \hat{f}_{t-1}(X_t)) + \ell(Y_t, f_{j,t-1}(X_t)) - \mathbb{E}_{\pi_{t-1}}[\ell(Y_t, f_{j,t-1}(X_t))],$$

for all $1 \leq j \leq M$, $1 \leq t \leq n+1$. We also fix the initial weights uniformly $\pi_{j,0} = M^{-1}$. We obtain in Theorem 4.4 the fast rate of convergence for the BOA procedure; with probability $1 - e^{-x}$, $x > 0$,

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

for η smaller than a constant depending on the range. The result follows from our online to batch conversion. The second order term is estimated with the excess of risk using the strong convexity assumption on the loss. Such optimality should also hold for the other algorithms provided in [GSVE14]. We conclude by providing the version of the fast rate bound with the best possible η for 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 online to batch conversion and the excess of 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 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

$$X - X^2 \leq \log(1+X) \Leftrightarrow \exp(X - X^2) \leq 1 + X \Leftrightarrow \mathbb{E}_{t-1}[\exp(X - X^2)] \leq 1. \quad (3)$$

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 (3) 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 $t = 1, \dots, n$ provides the desired result.

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 $h, g : \mathcal{X} \rightarrow \mathbb{R}$ we have:*

$$\begin{aligned} \mathbb{E}_P[\exp(h - \mathbb{E}_P[h] - g)] &\leq 1 \\ \iff \mathbb{E}_Q[h] - \mathbb{E}_P[h] &\leq \mathbb{E}_Q[g] + \mathcal{K}(Q, P), \quad \text{for any probability measure } Q. \end{aligned} \quad (4)$$

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{F}_{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 $\mathcal{R}_{n+1}(f)$ quantify the prediction of $f = (f_0, f_1, f_2, \dots)$. We have

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

$$\eta \max_{1 \leq t \leq n+1} \max_{1 \leq j \leq M} \ell_{j,t} \leq 1/2, \quad (5)$$

then the cumulative loss of the BOA procedure satisfies

$$\mathbb{E}_{\hat{\pi}}[\mathcal{R}_{n+1}(f_j)] \leq \inf_{\pi} \left\{ \mathbb{E}_{\pi}[\mathcal{R}_{n+1}(f_j)] + \eta \sum_{t=1}^{n+1} \ell_{j,t}^2 \right\} + \frac{\mathcal{K}(\pi, \pi_0)}{\eta}.$$

Moreover, if ℓ is convex with respect to its second argument, we have

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

Proof. We consider $\Delta M_{j,t+1} = -\eta \ell_{j,t+1}$ that is a centered random variable on $\{1, \dots, M\}$ when distributed as π_t . Under the assumption (5), $\Delta M_{j,t+1} \geq -1/2$ for any $0 \leq t \leq n$ a.s.. An application of the inequality (3) provides the inequality

$$\mathbb{E}_{\pi_t}[\exp(-\eta \ell_{j,t+1}(1 + \eta \ell_{j,t+1}))] \leq 1. \quad (6)$$

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 (6) provides

$$\mathbb{E}_{\pi_0} \left[\exp \left(-\eta \sum_{s=1}^{t+1} \ell_{j,s}(1 + \eta \ell_{j,s}) \right) \right] \leq \mathbb{E}_{\pi_0} \left[\exp \left(-\eta \sum_{s=1}^t \ell_{j,s}(1 + \eta \ell_{j,s}) \right) \right].$$

By the recursive argument for supermartingales we obtain

$$\mathbb{E}_{\pi_0} \left[\exp \left(-\eta \sum_{t=1}^{n+1} \ell_{j,t}(1 + \eta \ell_{j,t}) \right) \right] \leq 1.$$

Equivalently, using the variational form of the entropy (4),

$$0 \leq \inf_{\pi} \left\{ \mathbb{E}_{\pi} \left[\eta \sum_{t=1}^{n+1} \ell_{j,t} + \eta^2 \sum_{t=1}^{n+1} \ell_{j,t}^2 \right] + \mathcal{K}(\pi, \pi_0) \right\}, \quad (7)$$

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

$$\begin{aligned} \mathcal{R}_{n+1}(\hat{f}) - \mathcal{R}_{n+1}(f_{\pi}) &= \sum_{t=1}^{n+1} \ell(Y_t, \hat{f}_{t-1}(X_t)) - \ell(Y_t, \mathbb{E}_{\pi}[f_{j,t-1}](X_t)) \\ &\leq \sum_{t=1}^{n+1} \ell'(Y_t, \hat{f}_{t-1}(X_t)) (\hat{f}_{t-1}(X_t) - \mathbb{E}_{\pi}[f_{j,t-1}](X_t)) \\ &= \mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \ell'(Y_t, \hat{f}_{t-1}(X_t)) (\hat{f}_{t-1}(X_t) - f_{j,t-1}(X_t)) \right] \\ &= -\mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \ell_{j,t} \right]. \end{aligned}$$

□

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 behavior 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 described in Figure 2. 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 , $1 \leq j \leq M$, that are positive. If*

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

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

$$\mathcal{R}_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\eta_j \sum_{t=1}^{n+1} \ell_{j,t}^2 + \frac{\log(\pi_j/\pi_{j,0}) + \log(\mathbb{E}_{\pi_0}[\eta_j^{-1}]/\mathbb{E}_{\pi}[\eta_j^{-1}])}{\eta_j} \right] \right\}.$$

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

$$\mathbb{E}_{\pi'_t}[\eta_j h_j] = \mathbb{E}_{\pi_t}[h_j] / \mathbb{E}_{\pi_t}[\eta_j^{-1}], \quad 1 \leq t \leq n+1. \quad (8)$$

We consider $\Delta M_t = -\eta_j \ell_{j,t+1}$. Thanks to relation (8), 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 (2). Thus, one can apply the same reasoning than in the proof of Theorem 3.1. We obtain, equivalently than the inequality (7), that

$$0 \leq \inf_{\pi'} \left\{ \mathbb{E}_{\pi'} \left[\eta_j \sum_{t=1}^{n+1} \ell_{j,t} + \eta_j^2 \sum_{t=1}^{n+1} \ell_{j,t}^2 \right] + \mathcal{K}(\pi', \pi'_0) \right\},$$

for π' denoting any probability measure on $\{1, \dots, M\}$. Using the identity (8) to define π from π' , and multiplying the above inequality with $\mathbb{E}_{\pi}[\eta_j^{-1}] > 0$, we have

$$0 \leq \inf_{\pi} \left\{ \mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \ell_{j,t} + \eta_j \sum_{t=1}^{n+1} \ell_{j,t}^2 + \frac{\log(\pi'_j/\pi'_{j,0})}{\eta_j} \right] \right\},$$

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

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

$$\mathcal{R}_{n+1}(\hat{f}) \leq \min_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\eta_j \sum_{t=1}^{n+1} \ell_{j,t+1}^2 + \frac{\log(1/\pi'_{j,0})}{\eta_j} \right] \right\},$$

where π'_0 is defined as $\pi'_{i,0} = \eta_i^{-1} \pi_{i,0} / \mathbb{E}_{\pi_0}[\eta_j^{-1}]$. Then the initial weights $\pi_{j,0}$ are modified. They favor 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 strategies. 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

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

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

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

The form of the weights becomes similar than the one of the adaptive BOA introduced in Figure 3 and studied in the next section. The second order regret bounds becomes

$$\mathcal{R}_{n+1}(\hat{f}) \leq \min_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + 2\sqrt{\log(M)} \mathbb{E}_{\pi} \left[\sqrt{\sum_{t=1}^{n+1} \ell_{j,t}^2} \right] \right\},$$

for learning rates tuned optimally

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

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] 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, \dots, y_n ((X_t) is useless in this context). We have $\mathcal{F}_t = \{y_0, \dots, y_t\}$, $0 \leq t \leq n$, and the online learners $f_j = (y_{j,1}, y_{j,2}, y_{j,3}, \dots)$ of the dictionary are called the experts. The cumulative loss is now $\mathcal{R}_{n+1}(\hat{f}) = \sum_{t=1}^{n+1} \ell(y_t, \hat{y}_t)$ 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 $\{y_0, \dots, y_{t-1}\}$.

3.2 A new adaptive method for exponential weights

We described in Figure 3 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]. Multiple learning rates versions can be investigated for other exponential weights than for those of BOA. 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 the learning rates $(\eta_{j,t})_t$ are non increasing and satisfy

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

then the cumulative loss of the adaptive BOA procedure satisfies

$$\mathcal{R}_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \eta_{j,t-1} \ell_{j,t}^2 + \frac{\log(\pi_{j,0}^{-1}) + \log \left(1 + \mathbb{E}_{\pi_0} \left[\log \left(\frac{\eta_{j,1}}{\eta_{j,n}} \right) \right] \right)}{\eta_{j,n}} \right] \right\}.$$

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 $1 \leq t \leq n$, 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 $\Delta M_{j,t+1} \geq -1/2$, $j = 1, \dots, M$, we apply the inequality (3):

$$\mathbb{E}_{\pi'_t}[\exp(-\eta_{j,t} \ell_{j,t+1} (1 + \eta_{j,t} \ell_{j,t+1}))] \leq 1.$$

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

$$\pi'_{j,t} = \frac{\exp(-\eta_{j,t} \sum_{s=1}^t \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s})) \pi_{j,0}}{\mathbb{E}_{\pi_0}[\exp(-\eta_{j,t} \sum_{s=1}^t \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s}))]}. \quad (9)$$

Using the expression of the weights in the preceding inequality provides

$$\mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t} \sum_{s=1}^{t+1} \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s}) \right) \right] \leq \mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t} \sum_{s=1}^t \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s}) \right) \right]. \quad (10)$$

Using the basic inequality $x \leq \alpha^{-1} x^{\alpha} + \alpha^{-1} (\alpha - 1) \leq x^{\alpha} + \alpha^{-1} (\alpha - 1)$ for $x = \exp(-\eta_{j,t} \sum_{s=0}^{t-1} \ell_{j,s+1} (1 + \eta_{j,s} \ell_{j,s+1})) \geq 0$ and $\alpha = \eta_{j,t-1} / \eta_{j,t} \geq 1$, we obtain for all $2 \leq t \leq n$

$$\begin{aligned} & \mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t} \sum_{s=1}^t \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s}) \right) \right] \\ & \leq \mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t-1} \sum_{s=1}^t \ell_{j,s} (1 + \eta_{j,s-1} \ell_{j,s}) \right) \right] + \mathbb{E}_{\pi_0} \left[\frac{\eta_{j,t-1} - \eta_{j,t}}{\eta_{j,t-1}} \right]. \quad (11) \end{aligned}$$

Then, combining the inequalities (10) and (11) recursively for $t = n, \dots, 2$ and then (10) for $t = 1$ we obtain

$$\mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,n} \sum_{t=1}^{n+1} \ell_{j,t} (1 + \eta_{j,t-1} \ell_{j,t}) \right) \right] \leq 1 + \sum_{t=2}^n \mathbb{E}_{\pi_0} \left[\frac{\eta_{j,t-1} - \eta_{j,t}}{\eta_{j,t-1}} \right].$$

We apply the variational form of the entropy (4) in order to derive that

$$0 \leq \mathbb{E}_{\pi'} \left[\eta_{j,n} \sum_{t=1}^{n+1} \ell_{j,t} (1 + \eta_{j,t-1} \ell_{j,t}) \right] + \log \left(1 + \sum_{t=2}^n \mathbb{E}_{\pi_0} \left[\frac{\eta_{j,t-1} - \eta_{j,t}}{\eta_{j,t-1}} \right] \right) + \mathcal{K}(\pi', \pi_0)$$

for any probability measure π' on $\{1, \dots, M\}$. 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

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

The proof ends by choosing $\pi'_j = \eta_{j,n}^{-1} \pi_j / \mathbb{E}_{\pi}[\eta_{j,n}^{-1}]$ and using the sub-gradient trick as in the proof of Theorem 3.1. \square

3.3 The adaptive BOA procedure when the range is known

We now consider the case where the effective ranges $E_j > 0$ of the linearized errors are known:

$$\max_{1 \leq t \leq n+1} |\ell_{j,t}| \leq E_j, \quad 1 \leq j \leq M. \quad (12)$$

The second order regret bound provided in Theorem 3.3 can be easily optimized in $\eta_{j,t}$ and we tune the learning rates as

$$\eta_{j,t} = \min \left\{ \frac{1}{2E_j}, \sqrt{\frac{\log(\pi_{j,0}^{-1})}{\sum_{s=1}^t \ell_{j,s}^2}} \right\}, \quad 1 \leq t \leq n, 1 \leq j \leq M. \quad (13)$$

The learning rates are similar than those of Section 4.1 in [CBMS07] except that they depend on j through E_j , $\sum_{s=1}^t \ell_{j,s}^2$ and $\log(\pi_{j,0}^{-1})$; see [GSVE14] for similar multiple learning rates. We restrict ourself to cases where $\pi_{j,0} < 1$ to consider only positive learning rates $\eta_{j,t} > 0$ for all $1 \leq j \leq M$. We provide below a second order regret bound for the adaptive BOA procedure:

Theorem 3.4. *If ℓ is convex with respect to its second argument, if (12) holds and if the learning rates are tuned as in (13) then the cumulative loss of the adaptive BOA procedure satisfies*

$$\mathcal{R}_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\sqrt{\sum_{t=1}^{n+1} \ell_{j,t}^2} \left(\frac{\sqrt{2 \log(\pi_{j,0}^{-1})}}{\sqrt{2} - 1} + \frac{B_n}{\sqrt{\log(\pi_{j,0}^{-1})}} \right) + E_j (2 \log(\pi_{j,0}^{-1}) + 2B_n + 1) \right] \right\},$$

where $B_n = \log(1 + 2^{-1} \log(n))$ for all $n \geq 1$.

Proof. The proof is an application of the result of Theorem 3.3 using similar arguments than the proof of Theorem 5 in [CBMS07]. An exception is the additional term $\log(1 + \mathbb{E}_{\pi_0}[\log(\eta_{j,1}/\eta_{j,n})])$ that is easily bounded using the estimates

$$\log\left(\frac{\eta_{j,1}}{\eta_{j,n}}\right) = \log\left(\frac{1}{2E_j\eta_{j,n}}\right)_+ \leq \frac{1}{2} \log\left(\frac{\sum_{t=1}^n \ell_{j,t}^2}{4E_j^2 \log(M)}\right)_+ \leq \frac{\log(n)}{2}.$$

□

The second order regret bound for the adaptive BOA procedure is very similar to the ones obtained in [GSVE14]. Thus, the regret bounds in case of small excess losses and against iid sequences developed in [GSVE14] are also satisfied by the BOA procedure.

3.4 The adaptive BOA procedure when the ranges are unknown

When the effective ranges of the linearized error are not known, we have to estimate it. Let $c > 0$ be some constant chosen arbitrarily big enough to consider that 2^{-c} is negligible. To adapt the reasoning of [CBMS07], we consider an estimator $E_{j,t}$ of the range at time t : $E_{j,t} = 2^{k+1}$ where $k \geq -c$ is the smallest integer such that $\max_{1 \leq s \leq t} |\ell_{j,t}| \leq 2^k$, $1 \leq j \leq M$. Then we define the learning rates as

$$\eta_{j,t} = \min\left\{\frac{1}{E_{j,t}}, \sqrt{\frac{\log(\pi_{j,0}^{-1})}{\sum_{s=1}^t \ell_{j,s}^2}}\right\}, \quad 1 \leq t \leq n, 1 \leq j \leq M. \quad (14)$$

This rule for updating learning rates is similar than the one in [CBMS07] except the dependence on j . We also restrict the range of $E_{j,t}$ to define correctly $\eta_{j,1}$ when $\ell_{j,1} = 0$. Assuming that the ranges E_j defined in (12) satisfies $2^{-c} \leq E_j \leq E$ for all $1 \leq j \leq M$, we have

Theorem 3.5. *If ℓ is convex with respect to its second argument, if (12) holds and if the learning rates are tuned as in (14) then the cumulative loss of the adaptive BOA procedure satisfies,*

$$\mathcal{R}_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ \mathcal{R}_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\sqrt{\sum_{t=1}^{n+1} \ell_{j,t}^2} \left(\frac{\sqrt{2 \log(\pi_{j,0}^{-1})}}{\sqrt{2} - 1} + \frac{B_{n,E} + 8E_j}{\sqrt{\log(\pi_{j,0}^{-1})}} \right) + 4E_j(\log(\pi_{j,0}^{-1}) + B_{n,E} + 3) \right] \right\},$$

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

Proof. We apply the same recursive argument than in the proof of Theorem 3.3. But we have to distinguish two cases, depending whether $\eta_{j,t} \ell_{j,t+1} \leq 1/2$ or not.

We consider the set of indices that \mathcal{T}_j such that $E_{j,t} < E_{j,t+1}$. For such indices, possibly $\eta_{j,t} \ell_{j,t+1} > 1/2$ but as $-\eta_{j,t} \ell_{j,t+1} \leq \eta_{j,t} E_{j,t+1}$ we have

$$\mathbb{E}_{\pi_t'}[\exp(-\eta_{j,t}(\ell_{j,t+1}(1 + \eta_{j,t} \ell_{j,t+1}) + E_{j,t+1} 1_{\eta_{j,t} \ell_{j,t+1} > 1/2}))] \leq 1.$$

Then, from the expression of the weights π'_t in (9) we obtain

$$\begin{aligned} \mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t} \left(\sum_{s=0}^t \ell_{j,s+1} (1 + \eta_{j,s} \ell_{j,s+1}) + E_{j,t+1} 1_{\eta_{j,t} \ell_{j,t+1} > 1/2} \right) \right) \right] \\ \leq \mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,t} \sum_{s=0}^{t-1} \ell_{j,s+1} (1 + \eta_{j,s} \ell_{j,s+1}) \right) \right]. \quad (15) \end{aligned}$$

Second, we consider the indices t that do not belong to \mathcal{T}_j . Then $\eta_{j,t} \ell_{j,t+1} \leq 1/2$ and the same reasoning than in the proof of Theorem 3.3 applies. The recursive formulas (10) and (11) hold. To conclude, we apply recursive formulas (10), (11) and (15) and we obtain the upper bound

$$\mathbb{E}_{\pi_0} \left[\exp \left(-\eta_{j,n} \sum_{t=0}^n \left(\ell_{j,t+1} (1 + \eta_{j,t} \ell_{j,t+1}) + E_{j,t+1} 1_{\eta_{j,t} \ell_{j,t+1} > 1/2} \right) \right) \right] \leq 1 + \mathbb{E}_{\pi_0} \left[\log \left(\frac{\eta_{j,1}}{\eta_{j,n}} \right) \right].$$

We have the elementary bounds $\sum_{t \in \mathcal{T}_j} E_{j,t+1} \leq 8E_j$ and $\log(\eta_{j,1}/\eta_{j,n}) \leq \log(\sqrt{n}E/2^{-c})$. Then Theorem 3.5 is proved using similar arguments than in the proof of Theorem 6 in [CBMS07]. \square

The advantage of the adaptive BOA procedure compared with the procedures studied in [GSVE14] is to be adaptive to unknown ranges. The price to pay is an additional term 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 $1 \leq j \leq M$. Whether this extra term can be avoided in the multiple learning rates case is an open question.

4 Optimality of the BOA procedure in a stochastic environment

4.1 An empirical online to batch conversion

We now turn to a stochastic setting where (X_t, Y_t) are random elements observed recursively with $1 \leq t \leq n$. 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 online to batch 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 online to batch conversion, a procedure that admits a similar second order term on the regret bound has nice properties in the stochastic environment. The BOA procedure achieves this strategy as the second order term of the regret bound is the same than the one appearing in the online to batch conversion. Let us go back for a moment to the most general case:

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

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

$$\max_{1 \leq t \leq n+1} \max_{1 \leq j \leq M} \eta_{j,t-1} \ell_{j,t} \leq 1/2, \quad a.s.. \quad (16)$$

We have the following general online to batch conversion. It can be seen as an empirical counterpart of the conversion provided in [Zha05, KT08]. 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.

Theorem 4.1. *The cumulative predictive risk of any aggregation procedure satisfies, with probability $1 - e^{-x}$, $x > 0$:*

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

Proof. We first note that for each $1 \leq j \leq M$ the sequence $(M_{j,t})_t$ with $M_{j,t} = \eta(R_t(\hat{f}) - R_t(f_j) - (\mathcal{R}_t(f_\pi) - \mathcal{R}_t(f_j)))$ is a martingale adapted with the filtration (\mathcal{F}_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 constant, we adapt this recursive argument as in the proof of Theorem 3.3.

For any $1 \leq j \leq M$, $1 \leq t \leq n+1$, denoting $X = -\eta_{j,t-1} \ell_{j,t}$ we check that $X \geq -1/2$. We can apply (3) conditionally on \mathcal{F}_{t-1} and we obtain

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

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

$$\begin{aligned} & \mathbb{E} \left[\exp \left(-\eta_{j,t-1} \left(\sum_{s=1}^t (\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1} \ell_{j,s}^2 \right) \right) \right] \\ & \leq \mathbb{E} \left[\exp \left(-\eta_{j,t-1} \left(\sum_{s=1}^{t-1} (\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1} \ell_{j,s}^2 \right) \right) \right]. \end{aligned}$$

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

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

We obtain

$$\begin{aligned} & \mathbb{E} \left[\exp \left(- \eta_{j,t-1} \left(\sum_{s=1}^t (\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1} \ell_{j,s}^2 \right) \right) \right] \\ & \leq \mathbb{E} \left[\exp \left(- \eta_{j,t-2} \left(\sum_{s=1}^{t-1} (\ell_{j,s} - \mathbb{E}_{s-1}[\ell_{j,s}]) - \eta_{j,s-1} \ell_{j,s}^2 \right) \right) \right] + \mathbb{E} \left[\frac{\eta_{j,t-2} - \eta_{j,t-1}}{\eta_{j,t-1}} \right]. \end{aligned}$$

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

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

We end the proof by the Chernoff device. \square

4.2 Second order bounds on the excess of cumulative predictive risk

Based on the result of Theorem 4.1, we derive batch versions of the main results of Section 3 for the BOA procedure. As an example, using the second order regret bound of Theorem 3.1 and the online to batch conversion of Theorem 4.1 we obtain

Theorem 4.2. *Assume that $\eta > 0$ satisfies condition (5). The cumulative risk of the BOA procedure satisfies, with probability $1 - e^{-x}$, $x > 0$:*

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

Moreover, if ℓ is convex with respect to its second argument, we have, with probability $1 - e^{-x}$, $x > 0$:

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

Proof. We prove the result by integrating the result of Theorem 4.1 with respect to any deterministic π and noticing that, as $\eta_{j,t} = \eta$ is constant, $\log(\eta_{j,1}/\eta_{j,n}) = 0$. \square

Similarly, we can extend Theorems 3.2 and 3.4. Below we provide the extension of the adaptive case of Theorem 3.5 as the proof is more involved; the boundedness condition (16) is no longer satisfied for any $1 \leq t \leq n+1$ and Theorem 4.1 does not apply directly. We have

Theorem 4.3. *Under the hypothesis of Theorem 3.5, the cumulative risk of the adaptive BOA procedure satisfies with probability $1 - e^{-x}$, $x > 0$,*

$$\begin{aligned} R_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ R_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\sqrt{\sum_{t=1}^{n+1} \ell_{j,t}^2} \left((\sqrt{2} + 1)^2 \sqrt{\log(\pi_{j,0}^{-1})} + \frac{2B_{n,E} + 16E_j + x}{\sqrt{\log(\pi_{j,0}^{-1})}} \right) \right. \right. \\ \left. \left. + 4E_j(\log M + 2B_{n,E} + 6 + x) \right] \right\}, \end{aligned}$$

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

Proof. As the boundedness condition (16) is not satisfied for all $1 \leq t \leq n + 1$, 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 3.4 for proving Theorem 3.5. Let us denote \mathcal{T}_j the set of indices such that $E_{j,t} < E_{j,t+1}$. For $t \in \mathcal{T}_j$ possibly $\eta_{j,t} \ell_{j,t+1} > 1/2$ and the recursive argument for supermartingales do not apply directly. But as $-\eta_{j,t}(\ell_{j,t+1} - \mathbb{E}_{t-1}[\ell_{j,t+1}]) \leq \eta_{j,t} E_{j,t+1}$ we apply instead

$$\mathbb{E}_{t-1} \left[\exp \left(-\eta_{j,t}(\ell_{j,t+1} - \mathbb{E}_{t-1}[\ell_{j,t+1}]) - \eta_{j,t} E_{j,t+1} \mathbf{1}_{\eta_{j,t} \ell_{j,t+1} > 1/2} \right) \right] \leq 1.$$

Thus the proof ends by an application of the recursive argument for supermartingales as in the proof of Theorem 3.5. \square

Let us discuss the constants. First, nE^2 is 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, $B_{n,E}$ is comparable with $\log(M)$. We continue the discussion for M sufficiently large satisfying $n \leq e^M$ and $M \geq c$. Choosing π_0 uniform on $\{1, \dots, M\}$, the second order bound becomes

$$R_{n+1}(\hat{f}) \leq \inf_{\pi} \left\{ R_{n+1}(f_{\pi}) + \mathbb{E}_{\pi} \left[\sqrt{\max_{1 \leq j \leq M} \sum_{t=1}^{n+1} \ell_{j,t}^2} \right] \frac{8(\log(M) + 2E) + x}{\sqrt{\log(M)}} \right\} + 12E(2 \log M + 2 + x). \quad (17)$$

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 the problem of aggregation; 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 (17) 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 (17) is an empirical bound for Problem (C), see also [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] are deterministic. To assert the optimality of BOA, it is easy to turn from an empirical bound to a deterministic one. As $\ell_{j,t}^2 \leq E^2$, Equation (17) implies

$$\frac{R_{n+1}(\hat{f})}{n+1} \leq \frac{\inf_{\pi} R_{n+1}(f_{\pi})}{n+1} + 8E \sqrt{\frac{\log(M)}{n+1}} + \frac{16E^2 + x}{\sqrt{(n+1) \log(M)}} + 12E \frac{\log M + 2 + x}{n+1}.$$

Then the BOA procedure is optimal for Problem (C) in the sense of the Definition (17): 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 optimal rate.

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 the general stochastic context. To turn to Problem (MS), we restrict our study to the context of Lipschitz strongly convex losses with iid observations. 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 $(n+1)^{-1}R_{n+1}(f_j) = R(f_j) = \mathbb{E}[\ell(Y, f_j(X))]$. 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 Jensen inequality gives $R(\bar{f}) \leq (n+1)^{-1}R_{n+1}(\hat{f})$. Remind that from [Tsy03] 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 , see [Lec07, Aud09]. In the sequel $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a loss function satisfying the following 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, LR13]. All of them are batch procedures.

Notice that Assumption **(LIST)** is restrictive and can hold only locally; on a compact set \mathcal{C} , the minimizer $f(y)^*$ of $f(y) \in \mathbb{R} \rightarrow \ell(y, f(y))$ 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 $E = C_b D$.

Inspired by the Q-aggregation procedures of [LR13], we consider under **(LIST)** a mixture of the original and the linearized loss. The second order refinement in the BOA procedure becomes:

$$\ell_{j,t} = \ell'(Y_t, \hat{f}_{t-1}(X_t))(f_{j,t-1}(X_t) - \hat{f}_{t-1}(X_t)) + \ell(Y_t, f_{j,t-1}(X_t)) - \mathbb{E}_{\pi_{t-1}}[\ell(Y_t, f_{j,t-1}(X_t))],$$

for all $1 \leq j \leq M$, $1 \leq t \leq n+1$. We obtain the optimality of the BOA procedure for Problem (MS).

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*

$$48C_b^2(1 + 3C_b D/100)\eta \leq C_\ell, \quad (18)$$

then the cumulative predictive risk of the BOA procedure satisfies, with probability $1 - e^{-x}$,

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

Proof. The proof starts from the second order empirical bound provided in Theorem 4.2. As the optimal rate is deterministic, we first convert the empirical second order term into a deterministic one. From the classical Bennett's inequality, as $\eta \ell_{j,t} \leq (48C_b^2)^{-1} C_\ell C_b D \leq (24)^{-1}$ under (18), then $\eta \ell_{j,t}^2 \leq C_b D/24$ and we have

$$\mathbb{E}_{t-1}[\exp(\eta \ell_{j,t}^2 - (1 + C_b D(e-2)/24)\eta \mathbb{E}_{t-1}[\ell_{j,t}^2])] \leq 1$$

We estimate $(e-2)/24 \leq 3/100$. Applying the recursive supermartingales argument and an union bound, we obtain the deterministic version of the result of Theorem 4.2: with probability $1 - e^{-x}$, $x > 0$,

$$\begin{aligned} R_{n+1}(\hat{f}) + \mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)] &\leq \inf_{\pi} \left\{ R_{n+1}(f_\pi) + \mathbb{E}_{\pi}[R_{n+1}(f_j)] \right. \\ &\quad \left. + 2(1 + 3C_b D/100)\eta \sum_{t=1}^{n+1} \mathbb{E}_{\pi}[\mathbb{E}_{t-1}[\ell_{j,t}^2]] + \frac{2x + \mathcal{K}(\pi, \pi_0)}{\eta} \right\}. \end{aligned}$$

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

$$\mathbb{E}_{\pi}[\mathbb{E}_{t-1}[\ell_{j,t}^2]] \leq 4C_b^2 \mathbb{E}_{\pi}[\mathbb{E}_{t-1}[(f_j(X_{t-1}) - \hat{f}_{t-1}(X_{t-1}))^2]] \leq 4C_b^2 (V(\pi) + \mathbb{E}[(f_\pi(X) - \hat{f}_{t-1}(X))^2])$$

where $V(\pi) = \mathbb{E}_{\pi}[\mathbb{E}[(f_j(X) - f_\pi(X))^2]]$. As $R_{n+1}(f_\pi) = R(f_\pi)$, $\mathbb{E}_{\pi}[R_{n+1}(f_j)] = \mathbb{E}_{\pi}[R(f_j)]$ and $\mathcal{K}(\pi, \pi_0) \leq \log(M)$, combining those bounds we obtain

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} &\leq \inf_{\pi} \left\{ R(f_\pi) + \mathbb{E}_{\pi}[R(f_j)] \right. \\ &\quad \left. + \gamma \left(V(\pi) + \frac{1}{n+1} \sum_{t=0}^n \mathbb{E}[(f_\pi(X) - \hat{f}_t(X))^2] \right) + \frac{2x + \log(M)}{\eta(n+1)} \right\} \quad (19) \end{aligned}$$

with $\gamma = 8C_b^2(1 + C_b D(e-2))\eta$. The rest of the proof is inspired by the reasoning of [LR13]. First, one can check that

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

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

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

Thus the function $H: \pi \rightarrow R(f_\pi) + \mathbb{E}_\pi[R(f_j)] + \gamma V(\pi)$ is convex as $0 \leq \gamma \leq C_\ell/2$ under (18). Moreover, if one denotes π^* a minimizer of H , we have for any weights π

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

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

$$\begin{aligned} \frac{C_\ell/2 - 2\gamma}{n+1} \sum_{t=0}^n \mathbb{E}[(f_{\pi^*}(X) - \hat{f}_t(X))^2] &\leq \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} + \\ &\quad - \left(R(f_{\pi^*}) + \mathbb{E}_{\pi^*}[R(f_j)] + \gamma V(\pi^*)\right) + \frac{\gamma}{n+1} \sum_{t=0}^n V(\pi_t). \end{aligned}$$

Then, combining this last inequality with the inequality (19), we derive that

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

Plugging in this new estimate into (19) we obtain

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} - \frac{2\gamma^2}{C_\ell - 4\gamma} \frac{1}{n+1} \sum_{t=0}^n V(\pi_t) &\leq R(f_{\pi^*}) + \mathbb{E}_{\pi^*}[R(f_j)] + \gamma V(\pi^*) \\ &\quad + \frac{C_\ell - 2\gamma}{C_\ell - 4\gamma} \frac{2x + \log(M)}{\eta(n+1)}. \end{aligned}$$

Now, using C_ℓ -strong convexity as in Proposition 2 of [LR13], we have for any probability measure π

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

As under condition (18) it holds

$$\frac{2\gamma^2}{C_\ell - 4\gamma} \leq \frac{C_\ell}{2} \quad \text{and} \quad \frac{C_\ell - 2\gamma}{C_\ell - 4\gamma} \leq 2,$$

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

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

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

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

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

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

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

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

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

The tuning parameter η can be considered as the inverse of the temperature β of the Q -aggregation procedure studied in [LR13]. In the Q -aggregation, the tuning parameter β is required to be larger than $60C_b^2/C_\ell$. It is a condition similar to our restriction (18) on η . The larger is η satisfying the condition (18) and the best is the rate of convergence. The choice $48C_b^2(1 + 3C_bD/100)\eta^* = C_\ell$ is optimal. The resulting BOA procedure is non adaptive in the sense that it depends on the constants appearing in the condition (**LIST**). As $C_\ell/C_b^2 \leq (C_bD)^{-1}$, it also depends on the range C_bD 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 grows 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 cumulative predictive risk of the adaptive BOA procedure satisfies, with probability $1 - e^{-x}$,*

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} &\leq \min_{1 \leq j \leq M} R(f_j) + \frac{34 \log(M) + 2B_{n,C_bD} + 16C_bD + 2x}{\eta^*(n+1)} \\ &\quad + \frac{C_bD(\log(M) + 2B_{n,C_bD} + 6 + x)}{n+1}. \end{aligned}$$

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 3.5 in the iid context under **(LIST)**, where $|\ell_{j,t}| \leq C_b D$; using the Young inequality, we have for any $\eta > 0$

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} &\leq \inf_{\pi} \left\{ R(f_{\pi}) + \mathbb{E}_{\pi}[R(f_j)] + 2\eta \mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \ell_{j,t}^2 \right] \right\} \\ &\quad + \frac{B(n, C_b D, M) + x}{\eta} + C_b D (\log(M) + x + 2B_{n, C_b D} + 6), \end{aligned}$$

where, using that $(a+b)^2 \leq 2(a^2 + b^2)$, $(\sqrt{2} + 1)^4 \leq 34$, $B_{n, E_j} \leq B_{n, C_b D}$ and $\log(M) \geq 1$,

$$B(n, C_b D, M) \leq 34 \log(M) + 2B_{n, C_b D} + 16C_b D.$$

Then we can use the Bennett inequality as in the proof of Theorem 4.4 to obtain the deterministic second order bound

$$\begin{aligned} \frac{R_{n+1}(\hat{f})}{n+1} + \frac{\mathbb{E}_{\hat{\pi}}[R_{n+1}(f_j)]}{n+1} &\leq \inf_{\pi} \left\{ R(f_{\pi}) + \mathbb{E}_{\pi}[R(f_j)] + \frac{2(1 + 3C_b D/100)\eta}{n+1} \mathbb{E}_{\pi} \left[\sum_{t=1}^{n+1} \mathbb{E}_{t-1}[\ell_{j,t}^2] \right] \right\} \\ &\quad + \frac{B(n, C_b D, M) + 2x}{\eta(n+1)} + \frac{E(\log(M) + x + 2B_{n, C_b D} + 6)}{n+1}. \end{aligned}$$

The proof ends similarly than the one of Theorem 4.4. For η^* satisfying the equality in the condition (18), we obtain

$$\frac{R_{n+1}(\hat{f})}{n+1} \leq \min_{1 \leq j \leq M} R(f_j) + \frac{B(n, C_b D, M) + 2x}{\eta^*(n+1)} + \frac{E(\log(M) + x + 2B_{n, C_b D} + 6)}{n+1}.$$

The result follows from the expression of $B(n, C_b D, M)$. \square

The BOA procedure is explicitly computed with complexity $O(Mn)$. It is a practical advantage compared with the batch procedure studied in [Aud07, LM09, LR13] that require an optimization routine. This issue has been solved in [DRXZ12] for the square loss using greedy sequential algorithms that approximate the Q -aggregation procedure. Note that in the case of the square loss, the second order refinement is equal to

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

The quality of a learner is asserted by the prediction accuracy of the average of its prediction and the aggregative prediction.

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. Stat. **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. Stat. **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*, The Journal of Machine Learning Research **14** (2013), 729–769.

- [GSVE14] P. Gaillard, G. Stoltz, and T. Van Erven, *A second-order bound with excess losses*, arXiv preprint arXiv:1402.2044, 2014.
- [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. Stat. **36** (2008), 2183–2206.
- [KT08] S. M. Kakade and A. Tewari, *On the generalization ability of online strongly convex programming algorithms*, NIPS, 2008.
- [Lec07] G. Lecué, *Optimal rates of aggregation in classification under low noise assumption*, Bernoulli **13** (2007), 1000–1022.
- [LM09] G. Lecué and S. Mendelson, *Aggregation via empirical risk minimization*, Probab. theory and related fields **145** (2009), 591–613.
- [LR13] G. Lecué and P. Rigollet, *Optimal learning with Q -aggregation*, arXiv preprint arXiv:1301.6080., 2013.
- [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.
- [Tsy03] A. B. Tsybakov, *Optimal rates of aggregation*, Learning Theory and Kernel Machines, Springer Berlin Heidelberg, 2003.
- [Vov90] V.G. Vovk, *Aggregating strategies*, Proc. Third Workshop on Computational Learning Theory, 1990.
- [Zha05] T. Zhang, *Data dependent concentration bounds for sequential prediction algorithms*, Learning Theory, Springer Berlin Heidelberg, 2005.