



HAL
open science

Stochastic Online Convex Optimization; Application to probabilistic time series forecasting

Olivier Wintenberger

► **To cite this version:**

Olivier Wintenberger. Stochastic Online Convex Optimization; Application to probabilistic time series forecasting. 2021. hal-03125863v1

HAL Id: hal-03125863

<https://hal.science/hal-03125863v1>

Preprint submitted on 29 Jan 2021 (v1), last revised 11 Jan 2024 (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.

STOCHASTIC ONLINE CONVEX OPTIMIZATION; APPLICATION TO PROBABILISTIC TIME SERIES FORECASTING

A PREPRINT

Olivier Wintenberger

olivier.wintenberger@upmc.fr
Laboratoire de Probabilités, Statistique et Modélisation
Sorbonne Université, CNRS
4 place Jussieu, 75005 Paris, France

January 29, 2021

ABSTRACT

Stochastic regret bounds for online algorithms are usually derived from an "online to batch" conversion. Inverting the reasoning, we start our analyze by a "batch to online" conversion that applies in any Stochastic Online Convex Optimization problem under stochastic exp-concavity condition. We obtain fast rate stochastic regret bounds with high probability for non-convex loss functions. Based on this approach, we provide prediction and probabilistic forecasting methods for non-stationary unbounded time series.

Keywords Stochastic online learning, time series prediction, probabilistic forecasting.

1 Introduction

This paper considers a Stochastic version of the classical Online Convex Optimization (OCO) setting introduced in Zinkevich (2003). An OCO algorithm predicts $x_t \in \mathbb{R}^d$ and incurs $\ell_t(x_t)$ where ℓ_t is a convex loss function. Performance is measured by the regret $\sum_{t=1}^T \ell_t(x_t) - \sum_{t=1}^T \ell_t(x)$. The regret bounds are depending on the conditions satisfied by the losses; OGD algorithm satisfies a regret bound of order $O(\sqrt{T})$ for any convex loss function and ONS a regret bound of order $O(\log T)$ for exp-concave loss functions, see Hazan (2019). A Stochastic Online Convex Optimization (SOCO) algorithm predicts $x_t \in \mathbb{R}^d$ and incurs the risk $L_t(x_t) = \mathbb{E}[\ell_t(x_t) | \mathcal{F}_{t-1}]$ for some filtration \mathcal{F}_t , the risk function L_t being convex. Performance of a SOCO algorithm is measured by the stochastic regret $\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x)$.

Any regret bound turn into a stochastic regret bound in expectation. The usual "online to batch" conversion focuses on finding stochastic regret guarantees with high probability. In the independent identically distributed (iid) setting, applying exponential inequalities on the regret bound, Hazan (2019) and Mahdavi et al. (2015) proved that OGD and ONS satisfies a stochastic regret bound of order $O(d\sqrt{T})$ and $O(\log T)$, respectively. Note that the fast rate "online to batch" conversion is much more evolved and less general than the slow rate one, see Mehta (2017). In this paper we focus on fast rate stochastic regret bounds of the form

$$\mathbb{P}\left(\exists T \geq 1 : \sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) > O(\log T)\right) \leq \delta, \quad \delta > 0.$$

In the iid setting the risk function $L_t = L = \mathbb{E}[\ell_1]$ is deterministic. Applying Jensen's inequality, the averaging \bar{x}_T of $(x_t)_{1 \leq t \leq T}$ satisfies $\mathbb{P}(\exists T \geq 1 : L(\bar{x}_T) - L(x) > O(\log T)) \leq \delta, \delta > 0$.

Our approach differs from the usual "online to batch" conversion in many ways. First, we have in mind forecasting streaming data observed recursively in time such as time series. The iid assumption is not reasonable as \mathcal{F}_t refers to the information from the past at time t , ℓ_{t+1} is depending on \mathcal{F}_t and $L_{t+1}(x_{t+1})$ is the (random) risk of prediction.

Second, we aim at considering probabilistic forecasting where the prediction x_t parametrizes a probability measure P_{x_t} . Expected losses such as L_t are considered in order to measure the approximation of P_{x_t} and P_t , the conditional distribution of a variable of interest y_t given \mathcal{F}_{t-1} . We focus on the logarithmic loss that results into L_t being the KL divergence up to constants.

Related works. Online learning of time series is natural as recursive algorithms update their predictions at each time step when a new data is observed. However regret bounds are rare due to the dependence among the data that prevents to use standard exponential inequalities. For stationary dependent time series (geometric β and ϕ mixing), Agarwal and Duchi (2012) obtained fast rate regret bounds for strongly convex loss functions. Their risk function does not coincide with our L_t ; it is the (deterministic) expectation of the loss function with respect to the stationary distribution. Beyond the restrictive geometrically mixing time series setting, Anava et al. (2013) obtained fast rates of convergence for the ONS. Working in the well specified bounded setting, their notion of regret coincides with our stochastic regret notion. For exp-concave loss functions, Anava et al. (2013) proved that ONS achieves fast rates stochastic regret, robust to independent noise distributions chosen adversarially at each time step. Kuznetsov and Mohri (2015) obtained low rates of convergence on the future risk of prediction $L_{T+1}(x_{t+1})$. The task of bounding $L_{T+1}(x_{t+1})$ is more evolved than controlling the stochastic regret and requires the introduction of the crucial notion of discrepancy, controlling the differences between L_{T+1} and $\sum_{t=1}^T L_t$ uniformly over \mathcal{K} .

Our fast rate stochastic regret bounds hold for BOA aggregation of the ONS, an algorithm sharing some similarities with Metagrad developed by van Erven and Koolen (2016). Note that Metagrad enjoys optimal stochastic regret bounds in the iid setting as it adapts optimally to Bernstein conditions, see Koolen et al. (2016) for more details. These guarantees also hold in non iid settings under the assumption of the existence of a common minimizer x^* of all L_t , $t \geq 1$. The latter assumption is very restrictive in the non iid setting and impossible in the non-stationary setting. Even closer to our approach is KAO, an algorithm developed by Adjakossa et al. (2020) for aggregating Kalman recursions in an optimal way. Despite KAO stochastic regret bounds hold only in the well-specified setting, KAO performs better than other existing methods in applications on (misspecified) real-data such as electricity consumption prediction, see Adjakossa et al. (2020) for details.

Finally note that online aggregation algorithms for volatility prediction and probability forecasting have been recently developed by Werge and Wintenberger (2020) and Thorey et al. (2017), V'yugin and Trunov (2019), respectively.

Our contributions. Our first contribution is to define the SOCO problem in full generality, without any assumption on distributions, extending the OCO one of Zinkevich (2003). The main novelty of SOCO is that the convex assumption holds on L_t and not on ℓ_t . Distributions of the random losses can be chosen adversarially. Then we define the stochastic exp-concavity condition on L_t that is crucial for our "batch to online" conversion: we obtain the control of the risk function with a surrogate loss using a self bounded martingale arguments of Bercu and Touati (2008) and the stopping time technique of Freedman (1975). The surrogate loss coincides with the linearized loss plus a quadratic variation term that is naturally controlled by the ONS algorithm of Hazan and Kale (2011). Thus ONS solves SOCO problems with fast rate and high probability.

Then we turn to the Stochastic Online Aggregation (SOA) problem. We remark that the stochastic exp-concavity condition is consistent in SOA as well and it coincides with the weak-exp concavity of Gaillard and Wintenberger (2018) with $\beta = 1$. The algorithm BOA of Wintenberger (2017) solves SOA in an optimal and parameter free way. Then BOA and ONS are combined for solving the SOCO problem of gaussian probabilistic forecast of time series.

Convexity (exp-concavity) of the loss function implies convexity (exp-concavity) of the risk function but not vice-versa. Thus we generalize the ARMA-based method and results of Anava et al. (2013) to unbounded non-stationary misspecified settings. We can also predict conditional variances, called volatilities, as the risk function is stochastically exp-concave despite the corresponding logarithmic loss function is not convex. Up to our knowledge it is the first time that a volatility prediction method achieves non-asymptotic theoretical guarantees. Combining both ARMA and volatility prediction methods, we design simple gaussian forecasters and a probabilistic forecasting algorithm mixing them.

2 Preliminaries and assumptions

We consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a filtration (\mathcal{F}_t) of non-decreasing σ -algebras.

Definition 1. A Stochastic OCO (SOCO) problem consists in a bounded convex set \mathcal{K} and a sequence of random loss functions (ℓ_t) defined over \mathcal{K} and adapted to the filtration (\mathcal{F}_t) so that the risk function $L_t = \mathbb{E}[\ell_t | \mathcal{F}_{t-1}] = \mathbb{E}_{t-1}[\ell_t]$ is convex.

A SOCO algorithm predicts a point $x_t \in \mathcal{K}$ from the information contained in \mathcal{F}_{t-1} and then incurs the cost $L_t(x_t) = \mathbb{E}_{t-1}[\ell_t(x_t)]$.

The natural filtration of a SOCO problem is

$$\mathcal{F}_t = \sigma(\ell_1, \dots, \ell_t), \quad t \geq 1.$$

The main difference with OCO is that the risk function is never observed. Instead one observes realizations of the gradients of the loss functions $\nabla \ell_1(x_1), \dots, \nabla \ell_t(x_t) \in \mathcal{F}_t$ at step t .

The main application we have in mind is the one of probabilistic forecasting.

Remark 1. A probabilistic forecasting algorithm predicts at each step x_t parametrizing a distribution function P_{x_t} . We consider a scoring rule, see Gneiting and Raftery (2007) for a formal definition, S and a random observation y_t so that the reward at step t is $S(P_{x_t}, y_t) = -\ell_t(x_t)$. Then the risk function is the inverse of the expected score

$$L_t(x_t) = \mathbb{E}_{t-1}[\ell(P_{x_t}, y_t)] = -\mathbb{E}_{t-1}[S(P_{x_t}, y_t)] = -S(P_{x_t}, P_{t-1})$$

where P_t denotes the distribution of y_t given \mathcal{F}_{t-1} . Considering the logarithmic score $S(P, y) = \log(p(y))$ where p is the density of P we get

$$L_t(x_t) = -S(P_{x_t}, P_{t-1}) = -\mathbb{E}_{t-1}[\log(p_{x_t}(y_t))] = KL(P_t, p_{x_t}) - \mathbb{E}_{t-1}[\log(p_t(y_t))]$$

where KL denotes the Kullback-Laibler divergence.

We note that in the SOCO problem the convex condition holds on the risk function and not on the loss function. All along the paper we will assume

(H1) The diameter of \mathcal{K} is $D < \infty$ so that $\|x - y\| \leq D$, $x, y \in \mathcal{K}$, ℓ_t are continuously differentiable over \mathcal{K} a.s. and the gradients are bounded by $G < \infty$: $\sup_{x \in \mathcal{K}} \|\nabla \ell_t(x)\| \leq G$ a.s., $t \geq 1$.

SOCO setting extend the OCO setting.

Proposition 1. Under **(H1)** any OCO problem is a SOCO problem and turns into a SOCO problem for any distribution on (ℓ_t) .

Proof. One can consider that ℓ_t is random with a degenerate distribution δ_{ℓ_t} , the Dirac mass at ℓ_t . Then the natural filtration is $\mathcal{F}_{t-1} = \emptyset$ and $L_t = \ell_t$ and the first point follows. Under **(H1)** the gradients are integrable so that one can differentiate under the integral and $\nabla L_t = \mathbb{E}_{t-1}[\nabla \ell_t]$. The expectation of a convex function is a convex function regardless of the distribution of ℓ_t . \square

We will focus on fast rates of convergence assuming the stochastic exp-concavity condition:

(H2) The random loss functions ℓ_t is stochastically exp-concave if it satisfies for some $\alpha > 0$

$$L_t(y) \leq L_t(x) + \nabla L_t(y)^T (y - x) - \frac{\alpha}{2} \mathbb{E}_{t-1}[(\nabla \ell_t(y)^T (y - x))^2], \quad x, y \in \mathcal{K}.$$

Stochastic exp-concavity has been introduced in an alternative way by Van Erven et al. (2015), making explicit a condition already used in Juditsky et al. (2008). Condition **(H2)** was also used by Gaillard and Wintenberger (2018) over the unit ℓ_1 -ball under the $\beta = 1$ -weak exp-concavity name. It implies the Bernstein condition of van Erven and Koolen (2016) with $\beta = 1$ in the iid setting with convex losses. Applying Lemma 4.3 of Hazan (2019), **(H2)** with $\alpha = 1/2(\mu \wedge 1/(4GD))$ is implied by the μ -exp-concavity of the loss function. It is also implied with $\alpha = \mu/G^2$ by the μ -strong convexity of the risk function but not by the exp-concavity of the risk function. Regarding probabilistic forecasting, it means that we restrict us to strictly proper scoring rules S as defined in Gneiting and Raftery (2007). Stochastic exp-concavity has strong links with classical statistical notions.

Proposition 2. If the loss function is twice continuously differentiable **(H2)** then we have

$$\alpha \mathbb{E}_{t-1}[\nabla \ell_t(x) \nabla \ell_t(x)^T] \preceq \mathbb{E}_{t-1}[\nabla^2 \ell_t(x)], \quad x \in \mathcal{K}. \quad (1)$$

If P_t is in the exponential family so that its density $p_t(y)$ is proportional to $e^{T(y)^T x_t^* - \ell_t(x_t^*)}$ with sufficient statistic $T(y)$ and some $x_t^* \in \mathcal{K}$ then for the logarithmic score

$$\mathbb{E}_{t-1}[\nabla \ell_t(x_t^*) \nabla \ell_t(x_t^*)^T] = \mathbb{E}_{t-1}[\nabla^2 \ell_t(x_t^*)],$$

and necessarily $\alpha \leq 1$. We refer to this setting as the well-specified one.

Proof. Inequality (1) follows easily from a second order expansion. The second assertion is the Fisher information identity. \square

In the iid setting, under (1) the asymptotic variance of the ERM given in Murata and Amari (1999) is bounded as

$$\text{Tr}(\mathbb{E}[\nabla^2 \ell_t(x^*)]^{-1} \mathbb{E}[\nabla \ell_t(x^*) \nabla \ell_t(x^*)^T]) \leq \frac{d}{\alpha} \quad (x^* = \arg \min_{x \in \mathcal{K}} \ell_t(x)).$$

The left hand side is also the optimal variance of any preconditioning online gradient algorithm as shown by Murata and Amari (1999).

We give some examples from probabilistic forecasting literature where Condition **(H2)** is satisfied. Note that the loss function might not be convex.

Example 1. We provide examples where **(H2)** is satisfied in the probabilistic forecasting setting with logarithmic scoring rule.

1. If we consider $P_x = \mathcal{N}(x, \sigma^2)$ then $\ell_t(x) = (x - y_t)^2 / (2\sigma^2)$ plus constant. In the classical OCO problem, ℓ_t is σ^2/D^2 -exp-concave if $y_t \in \mathcal{K}$. Thus the distributions P_t are implicitly assumed to be \mathcal{K} supported. In the SOCO problem, cases $y_t \notin \mathcal{K}$ are tractable assuming that the conditional distribution P_t has mean m_t and variance σ_t^2 . The loss is not exp-concave but the risk still satisfies Condition **(H2)** with $\alpha = \sigma^2/D^2$ assuming that $m_t \in \mathcal{K}$ for all $t \geq 1$. Note that the well-specified case $P_t = \mathcal{N}(x, \sigma_t)$ is included in the SOCO problem, not in the OCO one.
2. If we consider $P_x = \mathcal{N}(m_t, x)$ then $\ell_t(x) = (\log(x) + (y_t - m_t)^2/x)/2$ plus constant is convex only if $0 < x \leq 2(y_t - m_t)^2$ which could be satisfied only for very unrealistic distributions P_t that do not put mass around their mean. On the other side, if the conditional distribution P_t has mean m_t and finite variance σ_t^2 then the risk is $L_t(x) = (\log(x) + \sigma_t^2/x)/2$. It is μ -strongly convex if $\sigma_t^2 \geq \bar{\sigma}^2/2 + \mu\bar{\sigma}^6$. Then condition (1) is satisfied with $\alpha = \mu/G^2$. It holds for instance if $\mathcal{K} = [c\bar{\sigma}^2/2, \bar{\sigma}^2]$, $1 < c < 2$, $\sigma_t^2 \in \mathcal{K}$ for all $t \geq 1$ and then $\alpha = (c-1)c^2/(2(2-c)^2)$ since $\mu = (c-1)/(2\bar{\sigma}^4)$ and $G = (2-c)/(c\bar{\sigma}^2)$.

3 ONS achieves fast rate in the SOCO problem

3.1 Surrogate loss from a "batch to online" conversion

As L_t is not observed, we base our approach on a "batch to online" conversion using a martingale inequality in order to exhibit a surrogate loss.

Proposition 3. With probability $1 - \delta$ for any $0 < \delta \leq 1$, it holds

$$\begin{aligned} \sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) &\leq \sum_{t=1}^T \nabla \ell_t(x_t)^T (x_t - x) + \frac{\lambda}{2} \sum_{t=1}^T (\nabla \ell_t(x_t)^T (x_t - x))^2 \\ &\quad + \frac{\lambda - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(x_t)^T (x_t - x))^2] + \frac{1}{\lambda} \log(\delta^{-1}). \end{aligned}$$

Proof. Denoting $Y_t = \nabla \ell_t(x_t)^T (x_t - x)$, we observe that

$$\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \leq \sum_{t=1}^T \mathbb{E}_{t-1} [Y_t] - \frac{\alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [Y_t^2].$$

Moreover, from Bercu and Touati (2008) for any random variable Y_t we have

$$\mathbb{E}_{t-1} \left[\exp(\lambda(Y_t - \mathbb{E}_{t-1}[Y_t]) - \frac{\lambda^2}{2}(\mathbb{E}_{t-1}[Y_t^2] + Y_{t-1}^2)) \right] \leq 1.$$

We obtain the desired result applying a classical martingale argument due to Freedman (1975) and recalled in Appendix A. \square

We interpret $\tilde{\ell}_t(x_t) = \nabla \ell_t(x_t)^T (x_t - x) + \frac{\lambda}{2} (\nabla \ell_t(x_t)^T (x_t - x))^2$ as a surrogate loss. Compared with the linearized regret appearing from the gradient trick in OCO, the additional second order term appears to control the quadratic variation of the martingale. This extra term seems necessary in a surrogate loss estimating the random risk function.

3.2 ONS in the SOCO problem

As the surrogate loss $\tilde{\ell}_t$ coincides with the one used in the proof of the regret bound of the ONS under exp-concavity, ONS described in Algorithm 1 achieves a fast stochastic regret bound. Using Sherman-Morrison formula, each step of ONS has a $O(d^2 + P)$ -cost where P is the cost of the projection, usually $P = O(d^3)$ as emphasized by Koren (2013).

Algorithm 1: Online Newton Step, Hazan and Kale (2011)

Parameter: $\gamma > 0$.

Initialization: Initial prediction $x_1 \in \mathcal{K}$ and $A_0 = \frac{1}{(\gamma D)^2} I_d$.

Predict: x_t

Incur: $\ell_t(x_t)$

Observe: $\nabla \ell_t(x_t) \in \mathbb{R}^d$

Recursion: Update

$$\begin{aligned} A_t &= A_{t-1} + \nabla \ell_t(x_t) \nabla f(x_t)^T, \\ y_{t+1} &= x_t - \frac{1}{\gamma} A_t^{-1} \nabla \ell_t(x_t), \\ x_{t+1} &= \arg \min_{x \in \mathcal{K}} \|x - y_{t+1}\|_{A_t}^2. \end{aligned}$$

Theorem 4. Under (H1) and (H2), ONS algorithm for $\gamma = \alpha/2$ satisfies with probability $1 - 2\delta$ the stochastic regret bound

$$\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \leq \frac{1}{\alpha} \left(1 + d \log \left(1 + T \frac{\alpha^2 (GD)^2}{4} \right) \right) + \left(\frac{\alpha (GD)^2}{3} + \frac{20}{\alpha} \right) \log(\delta^{-1})$$

valid for any $T \geq 1$ and any $x \in \mathcal{K}$.

In the iid setting we recover the stochastic regret bound of Mahdavi et al. (2015) and the rate $O(d \log T)$ is optimal.

Proof. From the proof of the ONS regret bound in Hazan (2019), we get from the expression of the recursive steps (and not using the convexity of the loss)

$$\sum_{t=1}^T \nabla \ell_t(x_t)^T (x_t - x) \leq \frac{\gamma}{2} \sum_{t=1}^T (\nabla \ell_t(x_t)^T (x_t - x))^2 + \frac{1}{2\gamma} d \log(1 + T(\gamma GD)^2) + \frac{1}{2\gamma}.$$

Plugging this inequality into the previous bound we obtain

$$\begin{aligned} \sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) &\leq \frac{\lambda + \gamma}{2} \sum_{t=1}^T (\nabla \ell_t(x_t)^T (x_t - x))^2 \\ &\quad + \frac{\lambda - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(x_t)^T (x_t - x))^2] \\ &\quad + \frac{1}{2\gamma} d \log(1 + T(\gamma GD)^2) + \frac{1}{2\gamma} + \frac{1}{\lambda} \log(\delta^{-1}). \end{aligned}$$

Then we use a Poissonian exponential inequality from Gaillard and Wintenberger (2017) Appendix B1. More precisely, denoting $0 \leq Y_t = (\nabla \ell_t(x_t)^T (x_t - x))^2 / (GD)^2 \leq 1$, it holds

$$\mathbb{E}_{t-1} [\exp(Y_t - (e-1)\mathbb{E}_{t-1}[Y_t])] \leq 1.$$

We can apply the argument due to Freedman (1975) recalled in Appendix A yielding

$$\mathbb{P} \left(\exists t \geq 1 : \sum_{t=1}^T Y_t - (e-1) \sum_{t=1}^T \mathbb{E}_{t-1}[Y_t] > \log(\delta^{-1}) \right) \leq \delta, \quad 0 < \delta < 1. \quad (2)$$

Thus an union bound provides

$$\begin{aligned} \sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) &\leq \frac{e\lambda + (e-1)\gamma - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(x_t)^T (x_t - x))^2] \\ &\quad + \frac{1}{2\gamma} d \log(1 + T(\gamma GD)^2) + \frac{\gamma}{2} D^2 \epsilon + \left(\frac{\lambda + \gamma}{2} (GD)^2 + \frac{1}{\lambda} \right) \log(\delta^{-1}). \end{aligned}$$

Choosing $\lambda = (\alpha + (1 - e)\gamma)/e$ nonnegative since $\gamma \leq \alpha/(e - 1)$ we conclude

$$\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \leq \frac{1}{2\gamma} d \log(1 + T(\gamma GD)^2) + \frac{1}{2\gamma} + \left(\frac{\alpha + \gamma}{2e} (GD)^2 + \frac{e}{\alpha + (1 - e)\gamma} \right) \log(\delta^{-1}). \quad (3)$$

The desired result follows from the specific choice of γ . \square

4 BOA achieves fast rate in Stochastic Online Aggregation

4.1 Stochastic Online Aggregation

We consider $\mathbf{x}_t = [x_t^{(1)}, \dots, x_t^{(K)}]$ a $d \times K$ matrix whose columns are K different online predictors $x_t^{(i)}$ that are \mathcal{F}_{t-1} -adapted. We denote $\hat{x}_t = \mathbf{x}_t \pi = \sum_{i=1}^K \pi_i x_t^{(i)}$ their aggregation, π in the simplex $\Lambda_K = \{\pi \in \mathbb{R}^K; \pi > 0, \sum_{i=1}^K \pi_i = 1\}$. We are seeking an aggregation strategy that combines the predictors in an optimal way for the stochastic regret.

It is natural to aggregate online predictions issued from online procedures such as ONS. Indeed, since \mathbf{x}_t is \mathcal{F}_{t-1} -measurable by construction, we have under **(H2)** the relation

$$L_t(\mathbf{x}_t \pi_t) - L_t(\mathbf{x}_t \pi) \leq \nabla L_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi) - \frac{\alpha}{2} \mathbb{E}_{t-1} [(\nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi))^2].$$

Thus we are back to a SOCO problem with a loss $\pi \rightarrow L_t(\mathbf{x}_t \pi)$ over \mathcal{K} that is stochastically α -exp-concave.

Since the stochastic exp-concavity condition also hold for this problem, from Proposition 3 we have

$$\begin{aligned} \sum_{t=1}^T L_t(\mathbf{x}_t \pi_t) - L_t(\mathbf{x}_t \pi) &\leq \sum_{t=1}^T \nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi) + \frac{\lambda}{2} \sum_{t=1}^T (\nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi))^2 \\ &\quad + \frac{\lambda - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi))^2] + \frac{1}{\lambda} \log(\delta^{-1}). \end{aligned} \quad (4)$$

It would be tempting to use an aggregation strategy that minimizes the surrogate loss $(\pi_t - \pi)^T \tilde{\ell}_t - \alpha((\pi_t - \pi)^T \tilde{\ell}_t)^2$ with the notation

$$(\pi_t - \pi)^T \tilde{\ell}_t = \nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - \pi)$$

optimally. The trick is to solve an easier problem than the OCO one over \mathcal{K} ; it is enough to compare the aggregation strategy to $\pi \in \{e_i, 1 \leq i \leq K\}$, i.e. to compare the aggregation prediction \hat{x}_t with the best predictor $x_t^{(i)}$. We call this problem, which is not convex, the Stochastic Online Aggregation one. Exponential Weighted Aggregation (EWA) would be a natural candidate for obtaining an aggregation since the surrogate loss is exp-concave. However this procedure fails to provide fast rate with high probability as shown by Audibert (2008).

4.2 BOA in the SOA problem

The Bernstein Online Aggregation (BOA) described in Algorithm 2 has been designed as an alternative of EWA to get a second order regret bound that turns into a fast rate of convergence in a stochastic setting. For the sake of generality, we consider the setting of experts advices where the losses of the experts at time t are contained in a vector $\tilde{\ell}_t \in \mathbb{R}^K$ satisfying $|\tilde{\ell}_t| \leq GD$ for any $t \geq 1$. We denote $\mathbb{1} = (1, \dots, 1)^T$, the multiplications and inequalities implying vectors have to be thought componentwise.

Note that BOA is parameter-free and each step has a $O(d)$ -cost.

Theorem 5. *Under **(H1)** and **(H2)**, BOA algorithm satisfies with probability $1 - 2\delta$ the stochastic regret bound*

$$\sum_{t=1}^T L_t(\hat{x}_t) - \sum_{t=1}^T L_t(x_t^{(i)}) \leq 2 \left(\frac{2}{\alpha} + GD \right) \log(\pi_1^{-1} \log T) + 5GD + \left(\frac{\alpha(GD)^2}{2} + \frac{4}{\alpha} \right) \log(\delta^{-1}).$$

valid for any $T \geq 4$ and $1 \leq i \leq K$.

We recover the stochastic regret bound obtained by Wintenberger (2017) in the iid setting. A refined bound has been obtained in Gaillard and Wintenberger (2018) for a more complex variant of BOA.

Algorithm 2: Bernstein Online Aggregation, Wintenberger (2017)

Initialization: Initial weights $\pi_1 \in \Lambda_K$ and $\tilde{L}_0 = 0$.

For each step $t \geq 1$:

Recursion: Update

$$\eta_t = \sqrt{\frac{\log(\pi_1^{-1})}{\sum_{i=1}^T (\pi_i^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2}} \wedge \frac{1}{2GD}$$

$$\tilde{L}_t = \tilde{L}_{t-1} + (\tilde{\ell}_t - \pi_t^T \tilde{\ell}_t \mathbf{1}) + \eta_{t-1} (\tilde{\ell}_t - \pi_t^T \tilde{\ell}_t \mathbf{1})^2,$$

$$\pi_{t+1} = \frac{\eta_t \exp(-\eta_t \tilde{L}_t) \pi_1}{\pi_1^T (\eta_t \exp(-\eta_t \tilde{L}_{t+1}))}.$$

Proof. The proof is a direct consequence of the regret bound (proved in Appendix B)

$$\sum_{t=1}^T \pi_t^T \tilde{\ell}_t - \sum_{t=1}^T \pi_t^T \tilde{\ell}_{t,i} \leq \sqrt{\log(K \log T) \sum_{t=1}^T (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2} + GD(5 + 2 \log(\pi_1^{-1} \log T)) \quad (5)$$

under Condition **(H2)**. That such slow rate second order regret bound turns into a fast rate stochastic regret bound since the variance term is adaptive has already been noticed by van Erven and Koolen (2016). From the regret bound (5) and Young's inequality we infer that for any $\eta > 0$

$$\sum_{t=1}^T \pi_t^T \tilde{\ell}_t - \sum_{t=1}^T \tilde{\ell}_{t,i} \leq \frac{\eta}{2} \sum_{t=1}^T (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2 + \frac{\log(\pi_1^{-1} \log T)}{2\eta} + GD(5 + 2 \log(\pi_1^{-1} \log T)).$$

Plugging this bound into (4) and identifying \hat{x}_t we get

$$\begin{aligned} \sum_{t=1}^T L_t(\hat{x}_t) - \sum_{t=1}^T L_t(x_t^{(i)}) &\leq \frac{\lambda + \eta}{2} \sum_{t=1}^T (\nabla \ell_t(\hat{x}_t)^T (\hat{x}_t - x_t^{(i)})^2 \\ &\quad + \frac{\lambda - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(\hat{x}_t)^T (\hat{x}_t - x_t^{(i)}))^2] \\ &\quad + \frac{1}{\lambda} \log(\delta^{-1}) + \frac{\log(\pi_1^{-1} \log T)}{2\eta} + GD(5 + 2 \log(\pi_1^{-1} \log T)). \end{aligned}$$

Applying once again the Poissonian inequality (2) we derive that

$$\sum_{t=1}^T (\nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - e_i))^2 \leq (e-1) \sum_{t=1}^T \mathbb{E}_{t-1} [(\nabla \ell_t(\mathbf{x}_t \pi_t)^T \mathbf{x}_t (\pi_t - e_i))^2] + (GD)^2 \log(\delta^{-1}).$$

Thus choosing $2\eta = \alpha/4$ and $\lambda = \alpha/4$ so that $e\lambda + (e-1)\eta - \alpha \leq 0$ we get the desired result. \square

5 Examples

5.1 BOA-ONS adapts to unknown stochastic exp-concavity constant $\alpha > 0$

Consider $\hat{x}_t = \mathbf{x}_t \pi = \sum_{i=1}^K \pi_i x_t^{(i)}$ the aggregation of $K \geq 1$ ONS with different parameters $\gamma^{(i)}$ with $\gamma^{(i)} = \{2^{-1}, \dots, 2^{-K}\}$. With uniform prior $\pi_1 = K^{-1} \mathbf{1}$ the algorithm adapts to α .

Corollary 6. Under **(H1)** and **(H2)** with $\alpha \geq 2^{-K-2}$, BOA-ONS algorithm satisfies with probability $1 - 4\delta$ the stochastic regret bound

$$\begin{aligned} \sum_{t=1}^T L_t(\hat{x}_t) - \sum_{t=1}^T L_t(x) &\leq \frac{2}{\alpha} \vee 1 \left(1 + d \log \left(1 + T \frac{(\alpha GD)^2}{4} \right) \right) \\ &\quad + 2 \left(\frac{2}{\alpha} \vee 1 + GD \right) \log(K \log T) \\ &\quad + 5GD + \left(\alpha (GD)^2 + 22 \left(\frac{2}{\alpha} \vee 1 \right) \right) \log(\delta^{-1}). \end{aligned}$$

A reasonable choice for K is $d \vee 100$ so that the stochastic regret bound remains of the order $O((\alpha^{-1} \vee 1)d \log T)$. In the iid setting, De Vilmares and Wintenberger (2020) show that the Extended Kalman Filter is a recursive algorithm that tunes automatically the constant of exp-concavity without any aggregation procedure and with steps of $O(d^2)$ -cost.

Proof. Thanks to an union bound we combine the stochastic regret bound of Theorem 5 together with the inequality (3) choosing $-\log_2(\gamma) + 1 \leq i \leq -\log_2(\gamma) + 2$ so that $\alpha/4 \leq \gamma \leq \alpha/2$ for $\alpha \leq 1$. \square

5.2 BOA-ONS for online prediction and probabilistic forecast of time series

We observe an unbounded non-stationary time series (y_t) and we denote $\mathcal{F}_t = \sigma(y_t, \dots, y_1)$. The aim of the probabilistic forecaster is to predict the distribution P_{t+1} of the next observation y_{t+1} given \mathcal{F}_t . We combine the algorithms ONS and BOA in order to achieve fast rate stochastic regret guarantees in various SOCO problems. We focus on the logarithmic score associated with the KL-divergence risk function.

5.2.1 Online ARMA prediction by BOA-ONS

AutoRegressive Moving Average (ARMA) modeling is standard in time series prediction, see Brockwell and Davis (1991) for a reference textbook. Using the AR(p) model we consider the gaussian forecast $\mathcal{N}(\widehat{m}_t^{(p)}(x), \sigma^2)$ for any arbitrary $\sigma^2 > 0$ (note that σ^2 cancels in the ONS algorithm),

$$\widehat{m}_t^{(p)}(x) = x^T((y_{t-1} \wedge D/\sqrt{2}) \vee (-D/\sqrt{2}), \dots, (y_{t-p} \wedge D/\sqrt{2}) \vee (-D/\sqrt{2}))$$

and $\mathcal{K} = B_1(1)$, the ℓ^1 unit-ball of dimension p . We assume that the mean m_t of P_t satisfies $2m_t^2 \leq D^2$ a.s. and finite variance σ_t^2 . The KL-divergence

$$KL(P_t, \mathcal{N}(\widehat{m}_t^{(p)}(x), \sigma^2)) = \frac{1}{2} \left(\log(2\pi) + \log(\sigma^2) + \frac{(m_t - \widehat{m}_t^{(p)}(x))^2 + \sigma_t^2}{\sigma^2} \right) + \mathbb{E}_t[\log(p_t(y_t))]$$

is, up to a constant, the quadratic loss of the conditional expectations $(m_t - \widehat{m}_t^{(p)}(x))^2 / (2\sigma^2)$ which is quadratic in x . Thus it satisfies condition **(H2)** with $\alpha = \sigma^2/D^2$ and $\mathcal{K} = [-D/\sqrt{2}, D/\sqrt{2}]$.

Applying Theorem 4, the ONS achieves the stochastic regret

$$\sum_{t=1}^T KL(P_t, \mathcal{N}(\widehat{m}_t^{(p)}(x_t), \sigma^2)) - \sum_{t=1}^T KL(P_t, \mathcal{N}(m_t^{(p)}(x), \sigma^2)) \leq O(D^2(p \log T + \log(\delta^{-1})))$$

with high probability. We extend the regret bound of Anava et al. (2013) for adversarial noise to unbounded observation y_t and fixed order p .

Following the approach of Anava et al. (2013), one considers increasing orders p since any contracting ARMA model admits an AR(∞) representation. We consider $p \in \{1, \dots, \log T\}$ different ONS predictors $\widehat{m}_t^{(p)}(x_t)$ and their BOA aggregation \widehat{m}_t with initial weights proportional to $T^{-1}, \dots, T^{-\log T}$. The aggregation allows us to penalize the KL-divergence in order to prevent overfitting and BOA-ONS cumulative KL-divergence is bounded by

$$\begin{aligned} \sum_{t=1}^T KL(P_t, \mathcal{N}(\widehat{m}_t(x_t), \sigma^2)) &\leq \min_{1 \leq p \leq \log T} \min_{x \in B_\infty(1)} \left\{ \sum_{t=1}^T KL(P_t, \mathcal{N}(m_t^{(p)}(x), \sigma^2)) + O(p \log T) \right\} \\ &\quad + O(D^2(\log \log T + \log(\delta^{-1}))) \end{aligned}$$

improving the bound $\min_{x \in B_\infty(1)} \sum_{t=1}^T KL(P_t, \mathcal{N}(m_t^{(\log T)}(x), \sigma^2)) + O(\log^2 T)$ obtained by Anava et al. (2013).

5.2.2 Online volatility prediction by BOA-ONS

In mathematical finance the log ratios y_t are usually considered as centered and modeled by Generalized AutoRegressive Conditionally Heteroscedastic (GARCH). The inference uses the likelihood approach as if the conditional distributions were gaussian, see Francq and Zakoian (2019) for a reference textbook. Thus P_t has mean $m_t = 0$ and conditional variance σ_t^2 called the volatility. The aim is to provide an online estimator $\widehat{\sigma}_t^2$ of the volatility minimizing the KL-divergence $KL(P_t, \mathcal{N}(0, \widehat{\sigma}_t^2)) = (\log(2\pi\widehat{\sigma}_t^2) + \sigma_t^2/\widehat{\sigma}_t^2)/2$ which is known as the QLik risk and widely used in econometrics since it is robust to extreme values, see Patton (2011).

We consider $\bar{\sigma}_t^2, \hat{\sigma}_t^2 \in [c\bar{\sigma}^2/2, \bar{\sigma}^2]$ with $1 < c < 2$ such that $\hat{\sigma}_t^2 \rightarrow KL(P_t, \mathcal{N}(0, \hat{\sigma}_t^2))$ satisfies Condition **(H2)** with α as in Example 1. Using the ARCH(q) model we parametrize

$$\hat{\sigma}_t^{2,(q)}(x) = c\bar{\sigma}^2/2 + x_1(y_{t-1}^2 \wedge \bar{\sigma}^2) + \dots + x_q(y_{t-q}^2 \wedge \bar{\sigma}^2) \quad (6)$$

and $\mathcal{K} = \{x \in \mathbb{R}^q : x_i \geq 0 \text{ and } \|x\|_1 \leq 1 - c/2\}$ so that $\hat{\sigma}_t^{2,(q)} \in [c\bar{\sigma}^2/2, \bar{\sigma}^2]$. Then Condition **(H2)** also holds on the function $x \rightarrow KL(P_t, \mathcal{N}(0, \hat{\sigma}_t^{2,(q)}(x)))$ and applying Theorem 4, the ONS achieves the stochastic regret with probability $1 - 2\delta$

$$\sum_{t=1}^T KL(P_t, \mathcal{N}(0, \hat{\sigma}_t^{2,(q)}(x_t))) - \sum_{t=1}^T KL(P_t, \mathcal{N}(0, \hat{\sigma}_t^{2,(q)}(x))) \leq O(\bar{\sigma}^2(q \log T + \log(\delta^{-1}))).$$

Similar as for ARMA, any contracting GARCH model admits an ARCH(∞) representation. We consider BOA-ONS $\hat{\sigma}_t^2$ aggregating $\hat{\sigma}_t^{2,(q)}(x_t)$, $q = 1, \dots, \log T$ with initial weights proportional to $T^{-1}, \dots, T^{-\log T}$ so that with probability $1 - 4\delta$

$$\begin{aligned} \sum_{t=1}^T KL(P_t, \mathcal{N}(0, \hat{\sigma}_t^2)) &\leq \min_{1 \leq q \leq \log T} \min_{x \in \mathcal{K}} \left\{ \sum_{t=1}^T KL(P_t, \mathcal{N}(0, \sigma^{2,(q)}(x))) + O(p \log T) \right\} \\ &\quad + O(\bar{\sigma}^2(\log \log T + \log(\delta^{-1}))). \end{aligned}$$

5.2.3 Online gaussian prediction using BOA-ONS

We combine the ARMA and volatility prediction methods. We consider the simple probabilistic forecast $\mathcal{N}(\hat{m}_t^{(p)}(x), \hat{\sigma}_t^{2,(q)}(x))$ where $x \in \mathbb{R}^{p+q} \in \mathcal{K} = \{x \in \mathbb{R}^{p+q} : \|x_{1:p}\| \leq 1, x_{p+1:q} \geq 0 \text{ and } \|x_{p+1:q}\|_1 \leq 1 - c/2\}$. Under the condition $D^2 < (c-1)\bar{\sigma}^2/2$ then L_t is strongly convex and thus Condition **(H2)** holds with $\alpha = \frac{c^2((c-1)\bar{\sigma}^2/2 - D^2)}{4D^2(\bar{\sigma}^2 + 1/2)}$. Aggregating such predictors for $1 \leq p, q \leq \log T$ with BOA and initial weights proportional to $T^{-(p+q)}$, $1 \leq p, q \leq \log T$, we obtain a gaussian probabilistic forecast $\mathcal{N}(\hat{m}_t, \hat{\sigma}_t^2)$ satisfying the cumulative KL -divergence bound

$$\begin{aligned} \sum_{t=1}^T KL(P_t, \mathcal{N}(\hat{m}_t, \hat{\sigma}_t^2)) &\leq \min_{1 \leq p, q \leq \log T} \min_{x \in \mathcal{K}} \left\{ \sum_{t=1}^T KL(P_t, \mathcal{N}(m_t^{(p)}(x), \sigma^{2,(q)}(x))) + O((p+q) \log T) \right\} \\ &\quad + O(D^2 \bar{\sigma}^2(\log \log T + \log(\delta^{-1}))) \end{aligned}$$

with probability $1 - 4\delta$.

Note that the value of α provided above is lousy. One could use BOA to aggregate several ONS predictors with different orders (p, q) but also different exp-concavity constants α as in Section 5.1. We emphasize that this stochastic regret bound holds for any unbounded non-stationary time series (y_t) such that the conditional distributions P_t admits a density and satisfies the local restrictions $2m_t^2 \leq D^2$ and $\sigma_t^2 \in [c\bar{\sigma}^2/2, \bar{\sigma}^2]$, $t \geq 1$.

5.2.4 Online probabilistic forecasting using BOA-ONS

Following the approach of Thorey et al. (2017), we can combine simple probabilistic forecasters for obtaining a probabilistic forecast algorithm as a mixture distribution. Consider $\hat{P}_t = (\hat{P}_t^{(i)})_{1 \leq i \leq K}$, K probabilistic forecasters with densities $\hat{p}_t = (\hat{p}_t^{(i)})_{1 \leq i \leq K}$. For instance $\hat{P}_t^{(i)} = \mathcal{N}(\hat{m}_t^{(j)}, \hat{\sigma}_t^{2,\ell})$ is the simple gaussian probabilistic forecaster as above but with different local assumptions, $jD + D/\sqrt{2} \leq \hat{m}_t^{(j)} \leq (j+1)D + D/\sqrt{2}$, for $-K_1 \leq j \leq K_2$ and $\hat{\sigma}_t^{2,\ell} \in [(c/2)^{\ell+1}\bar{\sigma}^2/2, (c/2)^\ell\bar{\sigma}^2]$ for $0 \leq \ell \leq K_3$. Consider the SOCO problem of finding the mixture $x^T \hat{P}$ with $\mathcal{K} = \Lambda_K$ and $\ell_t(x) = -\log(x^T \hat{p}_t(y_t))$. We assume that $m \leq \hat{p}_t^{(i)}(y_t) \leq M$ a.s. for $1 \leq i \leq K$, $t \geq 1$ so that the gradients are bounded by $G = \sqrt{KM}/m$. Since the loss function is $\mu = 1$ -exp-concave, Condition **(H2)** is satisfied with $\alpha = 1/2(1 \wedge 1/(8G))$ by an application of Lemma 4.3 of Hazan (2019) because $\|x - y\| \leq \|x - y\|_1 \leq 2$ for $x, y \in \mathcal{K}$.

We use ONS to solve this SOCO problem and we get

$$\sum_{t=1}^T KL(P_t, \pi_t^T \hat{p}) \leq \min_{\pi \in \Lambda_K} \sum_{t=1}^T KL(P_t, \pi^T \hat{p}) + O(G(K \log T + \log(\delta^{-1}))).$$

Similar fast rates regret bounds were obtained by Thorey et al. (2017) for the CRPS score instead of the KL-divergence. They used the Recursive Least Square algorithm without projection that does not constrained π_t to be in Λ_K . On the contrary to our procedure, it is difficult to interpret their probabilistic forecast as it does not satisfy the axioms of a density function. Note that our method may be applied on $\hat{p}_t^{(i)}$ being the non-parametric ensemble forecasters considered in Thorey et al. (2017). Applied to simple gaussian probabilistic $\mathcal{N}(\hat{m}_t^{(j)}, \hat{\sigma}_t^{2,\ell})$, it provides a way to get around their local restrictions imposed on each simple gaussian probabilistic forecasters.

6 Conclusion and future work

In this paper, we derive fast rate stochastic regret bounds for the ONS and BOA algorithms under stochastic exp-concavity of the risk function (see Sections 3 and 4.2 respectively). Free of the convexity assumption on the loss function, we investigate online probabilistic forecasting for the logarithmic score by localizing simple gaussian forecasters and then mixing them (in Section 5).

Our approach is valid in any adversarial stochastic setting. Thus one can predict non-stationary time series. However our stochastic regret bounds are relative to a static prediction. One would like to develop a SOCO tracking problems and algorithms, in the spirit of Fixed Share of Herbster and Warmuth (1998) and FLH of Hazan (2019), in order to establish stochastic adaptive regret bounds.

References

- E. Adjakossa, Y. Goude, and O. Wintenberger. Kalman recursions aggregated online. *arXiv preprint arXiv:2002.12173*, 2020.
- A. Agarwal and J. C. Duchi. The generalization ability of online algorithms for dependent data. *IEEE Transactions on Information Theory*, 59(1):573–587, 2012.
- O. Anava, E. Hazan, S. Mannor, and O. Shamir. Online learning for time series prediction. In *Conference on learning theory*, pages 172–184, 2013.
- J.-Y. Audibert. Progressive mixture rules are deviation suboptimal. In *Advances in Neural Information Processing Systems*, pages 41–48, 2008.
- B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *The Annals of Applied Probability*, 18(5):1848–1869, 2008.
- P. J. Brockwell and R. A. Davis. *Time series : theory and methods*. Springer Series in Statistics. Springer, New York, 1991.
- J. De Vilmorest and O. Wintenberger. Stochastic online optimization using kalman recursion. *arXiv preprint arXiv:2002.03636*, 2020.
- C. Francq and J.-M. Zakoian. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons, 2019.
- D. Freedman. On tail probabilities for martingales. *Annals of Probability*, 3:100–118, 1975.
- P. Gaillard and O. Wintenberger. Sparse accelerated exponential weights. In *Artificial Intelligence and Statistics*, pages 75–82, 2017.
- P. Gaillard and O. Wintenberger. Efficient online algorithms for fast-rate regret bounds under sparsity. In *Advances in Neural Information Processing Systems*, pages 7026–7036, 2018.
- T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.
- E. Hazan. Introduction to online convex optimization. *arXiv preprint arXiv:1909.05207*, 2019.
- E. Hazan and S. Kale. Beyond the regret minimization barrier: an optimal algorithm for stochastic strongly-convex optimization. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 421–436, 2011.
- M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine learning*, 32(2):151–178, 1998.
- A. Juditsky, P. Rigollet, A. B. Tsybakov, et al. Learning by mirror averaging. *The Annals of Statistics*, 36(5):2183–2206, 2008.
- W. M. Koolen, P. Grünwald, and T. van Erven. Combining adversarial guarantees and stochastic fast rates in online learning. In *Advances in Neural Information Processing Systems*, pages 4457–4465, 2016.

- T. Koren. Open problem: Fast stochastic exp-concave optimization. In *Conference on Learning Theory*, pages 1073–1075. PMLR, 2013.
- V. Kuznetsov and M. Mohri. Learning theory and algorithms for forecasting non-stationary time series. In *NIPS*, pages 541–549. Citeseer, 2015.
- M. Mahdavi, L. Zhang, and R. Jin. Lower and upper bounds on the generalization of stochastic exponentially concave optimization. In *Conference on Learning Theory*, pages 1305–1320. PMLR, 2015.
- N. Mehta. Fast rates with high probability in exp-concave statistical learning. In *Artificial Intelligence and Statistics*, pages 1085–1093. PMLR, 2017.
- N. Murata and S.-i. Amari. Statistical analysis of learning dynamics. *Signal Processing*, 74(1):3–28, 1999.
- A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, 2011.
- J. Thorey, V. Mallet, and P. Baudin. Online learning with the continuous ranked probability score for ensemble forecasting. *Quarterly Journal of the Royal Meteorological Society*, 143(702):521–529, 2017.
- T. van Erven and W. M. Koolen. Metagrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems*, pages 3666–3674, 2016.
- T. Van Erven, P. Grunwald, N. A. Mehta, M. Reid, R. Williamson, et al. Fast rates in statistical and online learning. *Journal of Machine Learning Research*, 16, 2015.
- V. V. V’yugin and V. G. Trunov. Online learning with continuous ranked probability score. In A. Gammerman, V. Vovk, Z. Luo, and E. Smirnov, editors, *Proceedings of the Eighth Symposium on Conformal and Probabilistic Prediction and Applications*, volume 105 of *Proceedings of Machine Learning Research*, pages 163–177, Golden Sands, Bulgaria, 09–11 Sep 2019. PMLR. URL <http://proceedings.mlr.press/v105/v-yugin19a.html>.
- N. Werge and O. Wintenberger. Adavol: An adaptive recursive volatility prediction method. 2020.
- O. Wintenberger. Optimal learning with bernstein online aggregation. *Machine Learning*, 106(1):119–141, 2017.
- M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning, ICML 2003*, 2003.

A The stopping time argument of Freedman (1975)

We recall the argument of Freedman (1975) as we apply it several times in the paper. Consider $M_T = \exp(\sum_{t=1}^T Z_t)$ for any Z_t adapted to a filtration \mathcal{F}_t and satisfying the exponential inequality $\mathbb{E}[\exp(Z_t) | \mathcal{F}_{t-1}] \leq 1$. Then we have

$$\mathbb{P}\left(\exists T \geq 1 : \sum_{t=1}^T Z_t > \log(\delta^{-1})\right) \leq \delta$$

for any $0 < \delta < 1$ by an application of the following lemma.

Lemma 7. *Let (\mathcal{F}_t) be a filtration and (M_t) a super-martingale satisfying $M_0 = 1$ a.s., M_t is adapted to \mathcal{F}_t and $\mathbb{E}[M_t | \mathcal{F}_{t-1}] \leq M_{t-1}$ a.s., $t \geq 1$. Then for any $0 < \delta < 1$ it holds*

$$\mathbb{P}(\exists T \geq 1 : M_T > \delta^{-1}) \leq \delta.$$

Proof. We apply the optional stopping theorem with Markov’s inequality defining the stopping time $\tau = \inf\{t > 1 : M_t > \delta^{-1}\}$ so that

$$\mathbb{P}(\exists t \geq 1 : M_t > \delta^{-1}) = \mathbb{P}(M_\tau > \delta^{-1}) \leq \mathbb{E}[M_\tau] \delta \leq \mathbb{E}[M_0] \delta \leq \delta.$$

□

B Proof of the regret bound (5)

We keep the same notation as in Section 4.2. In order to prove the regret bound (5) we will show that

$$\pi_1^T \exp(-\eta_T \tilde{L}_T) \leq \log T. \tag{7}$$

We derive that for any $1 \leq i \leq K$ it holds

$$-\eta_{T,i} \tilde{L}_{T,i} = \eta_{T,i} \left(\sum_{t=1}^T (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i}) - \sum_{t=1}^T \eta_{t-1,i} (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2 \right) \leq \log(\pi_1^{-1} \log T)$$

so that

$$\sum_{t=1}^T \pi_t^T \tilde{\ell}_t \leq \tilde{\ell}_{t,i} + \sum_{t=1}^T \eta_{t-1,i} (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2 + \frac{\log(\pi_1^{-1} \log T)}{\eta_{T,i}}.$$

Since $\eta_{t,i}^{-2} \leq \eta_{t-1,i}^{-2} + (\pi_t^T \tilde{\ell}_t \mathbf{1} - \tilde{\ell}_t)^2$ we get by Abel's summation

$$\sum_{t=1}^T \eta_{t-1,i} (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2 = \sum_{t=1}^T \eta_{t-1,i} (\eta_{t,i}^{-2} - \eta_{t-1,i}^{-2}) = \sum_{t=1}^T \frac{\eta_{t-1,i} - \eta_{t,i}}{\eta_{t,i}^2} - \frac{1}{\eta_{0,i}} + \frac{1}{\eta_{T,i}}.$$

Since $(\pi_t^T \tilde{\ell}_t \mathbf{1} - \tilde{\ell}_t)^2 \leq (GD)^2$ we have

$$\begin{aligned} \sum_{t=1}^T \eta_{t-1,i} (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2 &\leq \sum_{t=1}^T \frac{\eta_{t-1,i} - \eta_{t,i}}{\eta_{t-1,i}^2} + \eta_{0,i} (GD)^2 + \frac{1}{\eta_{T,i}} \\ &\leq \frac{2}{\eta_{T,i}} + \frac{GD}{2} \end{aligned}$$

where the last inequality follows from a comparison sum-integral. The expression of the learning rates

$$\frac{1}{\eta_{T,i}} \leq \sqrt{\frac{\log \pi_1^{-1}}{\sum_{t=1}^T (\pi_t^T \tilde{\ell}_t - \tilde{\ell}_{t,i})^2}} + 2GD$$

yields the regret bound (5).

It remains to prove the exponential inequality (7). One has to consider weights $\tilde{\pi}_{t+1} = \eta_t^{-1} \pi_{t+1} / \pi_{t+1}^T \eta_t^{-1} \in \Lambda_K$ so that

$$\tilde{\pi}_{t+1}^T (\eta_t (\pi_{t+1}^T \tilde{\ell}_{t+1} \mathbf{1} - \tilde{\ell}_{t+1})) = \pi_{t+1}^T (\pi_{t+1}^T \tilde{\ell}_{t+1} \mathbf{1} - \tilde{\ell}_{t+1}) = 0 \text{ and } \tilde{\pi}_{t+1} = \frac{\exp(-\eta_t \tilde{L}_t) \pi_1}{\pi_1^T \exp(-\eta_t \tilde{L}_t)}.$$

Using the basic inequality $x \leq x^\alpha + \alpha^{-1}(\alpha - 1)$ for $x \geq 0$ and $\alpha \geq 1$ with $x = \exp(-\eta_T \tilde{L}_T)$ and $\alpha = \eta_{t-1} / \eta_t$ we get

$$\begin{aligned} \pi_1^T \exp(-\eta_T \tilde{L}_T) &\leq \pi_1^T \exp(-\eta_{T-1} \tilde{L}_T) + \pi_1^T \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}} \\ &\leq \tilde{\pi}_T^T \exp(\eta_{T-1} (\pi_T^T \tilde{\ell}_T \mathbf{1} - \tilde{\ell}_T) - \eta_{T-1}^2 (\tilde{\ell}_T - \pi_T^T \tilde{\ell}_T \mathbf{1})^2) \pi_1^T \exp(-\eta_{T-1} \tilde{L}_{T-1}) \\ &\quad + \pi_1^T \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}}. \end{aligned}$$

Using the basic inequality $\exp(x - x^2) \leq 1 + x$ for any $|x| \leq 1/2$ using that $|\eta_{T-1} \tilde{\ell}_T| \leq 1/2$ we obtain

$$\tilde{\pi}_T^T \exp(\eta_{T-1} (\pi_T^T \tilde{\ell}_T \mathbf{1} - \tilde{\ell}_T) - \eta_{T-1}^2 (\tilde{\ell}_T - \pi_T^T \tilde{\ell}_T \mathbf{1})^2) \leq 1 + \pi_T^T (\tilde{\ell}_T - \pi_T^T \tilde{\ell}_T \mathbf{1}) \leq 1$$

The desired result (7) follows by a recursive argument using $\pi_1^T \exp(-\eta_1 \tilde{L}_1) \leq \exp(1/2)$ and the comparison sum-integral

$$\sum_{t=2}^T \frac{\eta_{t-1} - \eta_t}{\eta_{t-1}} \leq \log(\eta_1 / \eta_T) = \frac{1}{2} \log \left(1 + \sum_{t=1}^T (\pi_t^T \tilde{\ell}_t \mathbf{1} - \tilde{\ell}_t)^2 / (4GD)^2 \right) \leq \frac{1}{2} \log(1 + T/4).$$

Then we use that $\exp(-\exp(1/2)) + \frac{1}{2} \log(1 + T/4) \leq \log T$ for $T \geq 4$.