Electronic Journal of Statistics
Vol. 18 (2024) 429–464
ISSN: 1935-7524
https://doi.org/10.1214/23-EJS2208

Stochastic online convex optimization. Application to probabilistic time series forecasting

Olivier Wintenberger

Laboratoire de Probabilités, Statistique et Modélisation
Sorbonne Université, CNRS
4 place Jussieu, 75005 Paris, France, and,
Institut CNRS Pauli (ICP)
Vienna University, Oskar-Morgenstern-Platz 1, 1090 WIEN, Austria
e-mail: olivier.wintenberger@sorbonne-universite.fr

Abstract: In this paper, we propose a general framework for stochastic online convex optimization that allows for achieving fast-rate stochastic regret bounds. Specifically, we demonstrate that certain algorithms, including online Newton steps and a scale-free variant of Bernstein online aggregation, achieve the best-known rates in unbounded stochastic settings. To illustrate the usefulness of our approach, we apply it to calibrating parametric probabilistic forecasters of non-stationary sub-Gaussian time series. Importantly, our fast-rate stochastic regret bounds are valid at any time, providing a flexible and robust performance metric for sequential algorithms. Our proofs rely on combining self-bounded and Poissonian inequalities for martingales and sub-Gaussian random variables, respectively, under a stochastic exp-concavity assumption.

MSC2020 subject classifications: Primary 62L99, 62M10.
Keywords and phrases: Sequential learning, time series prediction, probabilistic forecasting.

Received April 2023.

Contents

1. Introduction

We present a stochastic version of the Online Convex Optimization (OCO) analysis proposed by Zinkevich (2003) to calibrate sequential parametric forecasters and evaluate their effectiveness in stochastic environments. Our approach, called Stochastic Online Convex Optimization (SOCO) analysis, deals with random loss functions ℓ_t, $t \geq 1$. While the SOCO analysis covers the deterministic OCO analysis when the distributions of the loss functions ℓ_t, $t \geq 1$, are Dirac masses, they differ from their definitions of regret. In OCO, the regret is defined as the cumulative loss, whereas in SOCO, it is the cumulative conditional risk, both measured relatively to their respective minima. Thus SOCO can also be viewed as a specific imperfect-information OCO problem miminizing the conditional risks that are not directly observable. The imperfect-information setting of SOCO has its unique characteristics. Firstly, the stochastic environment can enhance the convexity of the optimization problem since the conditional risk functions often exhibit better convex properties than the original loss functions. Secondly, it is worth noting that the deviations of the random loss functions from the conditional risks are likely to grow with the number of iterations, necessitating sequential algorithms to be robust against these deviations. Lastly, the regret bounds for stochastic and deterministic settings have different minimizers, with the former focusing on assessing the calibration of parametric probabilistic forecasters with the environment's conditional distributions.

In this study, we demonstrate that certain algorithms, such as online Newton steps and a scale-free variant of Bernstein online aggregation, are adaptable to the stochastic convex properties of the conditional risk functions and robust to the stochastic deviations. As a result, we can effectively use them to calibrate probabilistic forecasting.

In Section 3, our primary result is that the Online Newton Step (ONS) algorithm's calibration achieves an $O(\log T)$ stochastic regret bound for any conditionally sub-Gaussian sequence of random losses. The main assumption is a stochastic exp-concavity condition **(H2)**, which is valid for non-convex losses and unbounded gradients. We prove this result using a self-normalized martingale inequality and a Poissonian inequality applicable for conditional sub-Gaussian gradients, as specified in Condition **(H3)**. Our analysis provides insights into why second-order gradient algorithms like ONS produce a fast-rate calibration: ONS implicitly minimizes a surrogate loss involving second-order terms.

The learning with expert advice in Cesa-Bianchi and Lugosi (2006) consists on studying the regret of sequential aggregation algorithms. We propose in Section 4 a stochastic version called Stochastic Online Aggregation (SOA) analysis. In SOA, the experts are stochastic predictors and the aggregation algorithm competes with the best predictor. Denoting E the deviations of the predictors the best existing regret bounds achieve optimal rates $O(\log\log T + E)$. However, the value of E typically increases as $O(\sqrt{\log T})$ in sub-Gaussian environments. As a result, the rate is deteriorated by the ability to learn the unknown upper-bound E of the deviations at horizon T.

Our second finding presents a stochastic regret bound, as $O((\log\log T)^2 + \log E)$, achieved by a scale-free version of the Bernstein Online Aggregation (BOA) algorithm. We achieve this result by self-normalizing multiple learning rates such that the weights remain unaffected by the scaling of losses by a scalar. This feature is essential in dealing with stochastic losses, and the regret bound we obtain is an improvement over existing bounds in sub-Gaussian unbounded stochastic settings.

Section 5 presents our approach to calibrating parametric probabilistic forecasters using the SOCO analysis. Our focus is on Gaussian probabilistic forecasters of time series with logarithmic losses, where the conditional risk functions correspond to the Kullback-Leibler (KL) divergence. We interpret the stochastic regret bounds as cumulative KL bounds relative to a static optimal forecaster.

We demonstrate that the condition **(H2)** is satisfied by parametric Gaussian forecasters of a time series (y_t). We then apply the SOCO analysis to parametric forecasters using AR-ARCH modeling to predict the conditional expectations and variances. Although the corresponding logarithmic loss functions are not convex, the conditional risk functions are still locally stochastically exp-concave. Consequently, we combine the ONS and BOA algorithms to sequentially calibrate the parameters of the Gaussian probabilistic forecasters. Our study provides fast-rate non-asymptotic theoretical guarantees for such parametric probabilistic forecasters.

Our stochastic regret bounds are derived using Ville (1939)'s inequality and are applicable at any time. Recently, anytime-valid sequential inference has been successfully employed in several statistical problems, including testing, comparing forecasters, and designing confidence sequences, as demonstrated in Henzi and Ziegel (2022), Shafer et al. (2021), Waudby-Smith and Ramdas (2023). A comprehensive overview of this approach can be found in the textbook Shafer and Vovk (2019) and the survey paper Ramdas et al. (2022). Chapter 12 of Shafer and Vovk (2019) introduces sequentially calibrated non-parametric probabilistic forecasters with an $O(\sqrt{T})$ regret bound in any bounded stochastic environment. Faster regret bounds of $O(\log T)$ for parametric prediction of deterministic individual sequences are presented in Cesa-Bianchi and Lugosi (2006), Hazan (2016) under exp-concavity assumptions. We extend these fast-rate results to more general stochastic exp-concavity settings.

For independent and identically distributed (iid) loss functions ℓ_t, Hazan (2016), Mahdavi et al. (2015) proved that Online Gradient Descent (OGD) and ONS algorithms satisfy stochastic regret bounds of order $O(\sqrt{T})$ and $O(\log T)$, respectively. Learning with expert advice calibrated by Squint and BOA achieves a stochastic regret bound as $O(\log\log T)$ under the so-called Bernstein condition in the stationary bounded setting, as shown in Koolen et al. (2016) and Wintenberger (2017), respectively. The dependence on the deviation bound E in these results have been further improved by careful tuning the learning rate in Mhammedi et al. (2019), Orseau and Hutter (2021) to achieve the rate $O(\log\log T + E)$. All existing stochastic regret bounds have this linear dependence on the maximum deviations of the loss functions. They use a fast-rate "online to batch" conversion to convert deterministic regret bounds into stochastic ones, as explained in Mehta (2017). We propose a different approach, which uses surrogate losses to obtain results beyond the iid environment and improving the dependence on E.

Sequential learning is naturally applicable to time series, as recursive algorithms update their predictions when new data becomes available over time. However, obtaining high-probability regret bounds is challenging due to the temporal dependence of the data, which hinders the use of standard exponential inequalities. For stationary β- or ϕ-mixing time series, Agarwal and Duchi (2012) derived fast-rate regret bounds for the unconditional risk function $\mathbb{E}[\ell_t]$. Anava et al. (2013) obtained fast-rate regret bounds for the risk of the ONS algorithm for ARMA (Auto-Regressive Moving-Average) models, although their notion of stochastic regret differs from ours.

Our approach combines optimization (ONS) and aggregation (BOA) to achieve fast-rate sequential calibration. This strategy has similarities with existing algorithms developed by Giraud et al. (2015), van Erven et al. (2021), which achieve a fast rate of stochastic regret bound in some stationary environments (Koolen et al., 2016). The algorithm proposed by Adjakossa et al. (2023) aggregates Kalman recursions in non-stationary well-specified settings only. Finally, sequential algorithms for estimating volatilities or aggregating probability forecasters have recently been developed by Werge and Wintenberger (2022) and Thorey et al. (2017), V'yugin and Trunov (2019), respectively.

2. Preliminaries and assumptions

We use the notation $\mathbf{0} = (0,\ldots,0)^T$, $\mathbb{1} = (1,\ldots,1)^T$, and the operations implying vectors are thought componentwise. In the sequel $\|\cdot\|$ is the Euclidean norm $\|\cdot\|_2$. We consider a filtration $(\mathcal{F}_t)$, $t \geq 0$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$ by convention. The proofs of the main results are deferred to Appendix A.

Definition 2.1 (Stochastic online convex optimization)**.** *Consider a convex body (a convex, compact set with a non-empty interior) $\mathcal{K} \subset \mathbb{R}^d$ and an $\mathcal{F}_t$-adapted sequence of random loss functions (ℓ_t) defined over $\mathcal{K}$. An algorithm predicts $x_t \in \mathcal{K}$ that is $\mathcal{F}_{t-1}$-measurable and incurs the random conditional risk $L_t(x_t) = \mathbb{E}_{t-1}[\ell_t(x_t)] = \mathbb{E}[\ell_t(x_t) \mid \mathcal{F}_{t-1}]$ at each step $t \geq 1$. SOCO analyses the rate of the stochastic regret*

$$Regret_T = \sup_{x\in\mathcal{K}}\left\{\sum_{t=1}^{T} L_t(x_t) - \sum_{t=1}^{T} L_t(x)\right\}, \qquad T \geq 1, \tag{2.1}$$

as a function of $T \geq 1$ assuming the risk functions L_t being convex for all $t \geq 1$.

The main difference with the perfect-information OCO analysis is the use of the conditional risk functions L_t instead of the loss functions ℓ_t in the regret and the convex assumption. Another difference is that we do not consider randomized strategy in SOCO as it goes beyond the scope of this paper.

The SOCO regret depends on the stochastic environment via the minimizer of the cumulative conditional risk:

Example 2.1 (Regression with quadratic loss)**.** *Consider a sequence of random variables of interest $(y_t)_{t\geq 1}$ with finite variances and quadratic loss functions $\ell_t(x) = (x - y_t)^2$, $t \geq 1$. In OCO, the regret is the cumulative loss relative to its minimum achieved by the empirical mean $\frac{1}{T}\sum_{t=1}^{T} y_t$. The SOCO regret depends on the stochastic setting. If (y_t) is iid then the conditional risk is constant, $L_t(x) = \mathbb{E}[(x - y_1)^2]$ for every $t \geq 1$, and its minimum is achieved by $\mathbb{E}[y_1]$. On the opposite, if $y_t = y_1$, $t \geq 1$, then $L_t(x) = (x - y_1)^2$ as long as $y_1 \in \mathcal{F}_{t-1}$, $t \geq 2$, and is minimal at y_1.*

The SOCO setting extends the OCO setting.

Proposition 2.1. *Any OCO problem is a degenerate SOCO problem.*

Proof. We consider that ℓ_t has a degenerate distribution $\delta_{\{\ell_t\}}$, the Dirac mass at ℓ_t. It is a SOCO problem equipped with the natural filtration is $\mathcal{F}_t = \{\emptyset, \Omega\}$, $t \geq 1$, and $L_t = \ell_t$. □

The conditional distribution of the random loss function ℓ_t may depend adversarially on $x_t, \ldots, x_1 \in \mathcal{F}_{t-1}$, and, as in OCO, a boundedness assumption on $\mathcal{K}$ is necessary to obtain regret bounds.

(H1) The diameter of $\mathcal{K}$ is $D < \infty$ so that $\|x - y\| \leq D$, $x, y \in \mathcal{K}$, and the loss functions ℓ_t are continuously differentiable over $\mathcal{K}$ a.s. with integrable gradients.

Under **(H1)** and if the loss functions (ℓ_t) are convex the optimal rate is $O(\sqrt{T})$ for OCO and thus for SOCO problems by an application of Proposition 2.1. This optimal rate is satisfied in SOCO problems even if the loss functions (ℓ_t) are not convex but the risk functions (L_t) are. See Appendix B.2 for the case of the OGD when the gradients $\nabla \ell_t$ are a.s. bounded by $G > 0$. To obtain fast-rate $o(\sqrt{T})$ stochastic regret bounds, we assume stochastic exp-concavity.

(H2) The random loss functions ℓ_t, $t \geq 1$, are stochastically exp-concave if for some $\alpha \geq 0$:

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

Condition **(H2)** with $\alpha = 0$ coincides with the convexity assumption on L_t, $t \geq 1$. In the iid setting, stochastic exp-concavity has been studied by Koolen et al. (2016), making explicit a condition introduced in Rigollet et al. (2008). Condition **(H2)** was used by Gaillard and Wintenberger (2018) over the unit ℓ^1-ball and it implies the Bernstein condition of van Erven et al. (2021) introduced for convex losses. In the deterministic setting, an application of Lemma 4.3 of Hazan (2016) shows that Condition **(H2)** with $\alpha = 1/2(\mu \wedge 1/(GD))$ follows from the μ-exp-concavity of the loss functions. More generally, Condition **(H2)** follows from the exp-concavity of the loss functions (ℓ_t), as soon as the gradients are square integrable. Also **(H2)** with $\alpha \geq 0$ does not imply the convexity of ℓ_t and holds under strong convexity of the conditional risk L_t, $t \geq 1$.

Proposition 2.2. *Assume the loss functions are twice continuously differentiable. Then Condition **(H2)** implies*

$$\alpha \mathbb{E}_{t-1}\big[\nabla \ell_t(x) \nabla \ell_t(x)^T\big] \preceq \nabla^2 L_t(x), \qquad x \in \mathcal{K}, a.s., t \geq 1. \tag{2.2}$$

On the opposite, if L_t is μ-strongly convex and there exists $g > 0$ such that

$$\mathbb{E}_{t-1}\big[\nabla \ell_t(x) \nabla \ell_t(x)^T\big] \preceq g^2 I_d, \qquad x \in \mathcal{K}, a.s., t \geq 1, \tag{2.3}$$

*then Condition **(H2)** holds with $\alpha = \mu/g^2$.*

Proof. Inequalities (2.2) and (2.3) follow easily from a second-order Taylor expansion of L_t. □

We verify Condition **(H2)** when calibrating parametric Gaussian probabilistic forecasters in Section 5. Under exp-concavity assumptions, the optimal rate is $O(\log T)$ in OCO (Hazan, 2016) and thus in SOCO. In stochastic environments the constant $\alpha > 0$ depends on the conditional distributions of the losses and is unknown in practice.

In many stochastic environments it is unrealistic to work under a boundedness condition on the gradients of the loss that is independent of T. We introduce the conditional sub-Gaussian condition **(H3)** to control the growth of the gradients at a logarithmic rate. We consider unbounded sug-Gaussian gradients introducing the Orlicz function $\psi_2(x) = \exp(x^2) - 1$, $x \in \mathbb{R}$. Conditional sub-Gaussian random variables are such as the Orlicz norm

$$\|Y_t\|_{\psi_2,t} = \inf\{c > 0\,;\, \mathbb{E}_{t-1}\big[\psi_2(Y_t/c)\big] \le 1\, a.s.\}$$

is bounded by a constant for every $t \ge 1$. This norm is not precise enough for our purpose. We require a slightly more explicit condition involving two constants. Our assumption is a conditional version of the Bernstein condition, also related to the notion of Bernstein-Orlicz norm of van de Geer and Lederer (2013).

(H3) The gradients $\nabla\ell_t(x_t)$, $t \ge 1$, satisfy for G_{ψ_2}, $G_2 > 0$, and all $k \ge 1$, $t \ge 1$, $x \in \mathcal{K}$,

$$\begin{aligned}
\mathbb{E}_{t-1}\big[\big(\nabla\ell_t(x_t)^T(x_t - x)\big)^{2k}\big] &\le k!(G_{\psi_2}D)^{2(k-1)}\mathbb{E}_{t-1}\big[\big(\nabla\ell_t(x_t)^T(x_t - x)\big)^2\big] \quad a.s.,\\
\mathbb{E}_{t-1}\big[\|\nabla\ell_t(x_t)\|^{2k}\big] &\le k! G_{\psi_2}^{2(k-1)}\mathbb{E}_{t-1}\big[\|\nabla\ell_t(x_t)\|^2\big] \quad a.s.,\\
\mathbb{E}_{t-1}\big[\|\nabla\ell_t(x_t)\|^2\big] &\le G_2^2 \quad a.s.
\end{aligned}$$

If the gradients $\nabla\ell_t(x_t)$, $t \ge 1$, verify the condition **(H3)** then they are conditionally sub-Gaussian.

Proposition 2.3. *Assume that the gradient $\nabla\ell_t(x_t)$ satisfies Condition **(H3)**: Then $\|\nabla\ell_t(x_t)\|$ is conditionally sub-Gaussian with*

$$\max_{t\ge1} \|\nabla\ell_t(x_t)\|_{\psi_2,t} \le 2(G_{\psi_2} \vee G_2)^2, \qquad t \ge 1, \qquad a.s.$$

Proof. Denote $Y = \|\nabla\ell_t(x_t)\|$. We have

$$\mathbb{E}\big[\exp\big(Y^2/K\big)\big] \le 1 + \sum_{k=1}^{\infty} \frac{\mathbb{E}[Y^{2k}]}{k!K^k} \le 1 + \sum_{k=1}^{\infty} \frac{G_{\psi_2}^{2(k-1)}G_2^2}{K^k} \le 2$$

for $K = 2(G_{\psi_2} \vee G_2)^2$. We conclude by definition of the Orlicz' norm. □

Condition **(H3)** is satisfied in every bounded cases $\|\nabla\ell_t(x_t)\|^2 \le G^2$, $t \ge 1$, with $G_{\psi_2} = G_2 = G$, Thus our sub-Gaussian stochastic setting encompasses the classical bounded deterministic one. Condition **(H3)** is also verified for unbounded Gaussian gradients with second-order conditional moments bounded by the constant $G_2 > 0$. Condition **(H3)** is independent of the conditional risks $\nabla L_t(x_t) = \mathbb{E}_{t-1}[\nabla\ell_t(x_t)]$, $t \ge 1$, and it does not interfere with Condition **(H2)**.

Proposition 2.4. *Assume that the gradient $\nabla\ell_t(x_t)$ is normally distributed given $\mathcal{F}_{t-1}$. Then Condition **(H3)** is satisfied if $\mathbb{E}_{t-1}[\|\nabla\ell_t(x_t)\|^2] \le G_2^2$ a.s., $t \ge 1$, and then $G_{\psi_2} = 8.5\, G_2$.*

3. ONS achieves fast-rate stochastic regrets

3.1. Surrogate losses

We base our approach on an observed surrogate loss that upper-bounds the stochastic regret using an exponential inequality for martingales from Bercu and Touati (2008) on unbounded gradients $\nabla \ell_t$, $t \geq 1$.

Proposition 3.1. *Under **(H1)** and **(H2)**, for any predictable sequence* (x_t) *and deterministic* x *in* $\mathcal{K}$, *it holds with probability* $1-\delta$, $0<\delta\leq 1$,

$$\begin{aligned}
&\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \\
&\quad \leq \sum_{t=1}^T \nabla \ell_t(x_t)^T(x_t - x) + \frac{\lambda}{2}\sum_{t=1}^T \big(\nabla \ell_t(x_t)^T(x_t-x)\big)^2 \\
&\qquad + \frac{\lambda-\alpha}{2}\sum_{t=1}^T \mathbb{E}_{t-1}\big[\big(\nabla \ell_t(x_t)^T(x_t-x)\big)^2\big] + \frac{2}{\lambda}\log\big(\delta^{-1}\big), \quad \lambda>0,\, T\geq 1.
\end{aligned}$$

When the distributions of ℓ_t are degenerate the upper bound in Proposition 2.1 becomes

$$\sum_{t=1}^T \nabla \ell_t(x_t)^T(x_t-x) + \frac{2\lambda-\alpha}{2}\sum_{t=1}^T \big(\nabla \ell_t(x_t)^T(x_t-x)\big)^2 + \frac{2}{\lambda}\log\big(\delta^{-1}\big).$$

Since the result is valid with probability 1, the last term disappears letting $\delta \uparrow 1$. Forthcoming results, anytime-valid with a high probability in a stochastic environment, are surely valid in deterministic environments when suppressing the dependence in δ.

Following van Erven et al. (2021), we interpret

$$\widetilde{\ell}_t(x_t) = \nabla \ell_t(x_t)^T(x_t-x) + \frac{\lambda}{2}\big(\nabla \ell_t(x_t)^T(x_t-x)\big)^2, \qquad t\geq 1,$$

as a surrogate loss. The quadratic term in addition to the gradient term is necessary to upper-bound the unobserved conditional risk with high probability. In stochastic environments, algorithms should minimize the cumulative surrogate loss $\sum_{t=1}^T \widetilde{\ell}_t$ rather than the cumulative loss $\sum_{t=1}^T \ell_t$. Under Condition **(H2)** with $\alpha>0$, this additional quadratic term is counterbalanced by the compensator $\sum_{t=1}^T \mathbb{E}_{t-1}[(\cdots)^2]$ when $\lambda<\alpha/2$. The Poissonian inequality of Proposition 3.2 relates both quadratic terms.

3.2. The stochastic regret analysis of ONS

The cumulative surrogate losses $\sum_{t=1}^T \widetilde{\ell}_t$ is implicitly minimized in the ONS's regret analysis of Hazan (2016). Then the ONS algorithm achieves a fast stochastic regret bound.

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

$$\begin{aligned} A_t &= A_{t-1} + \nabla \ell_t(x_t) \nabla \ell_t(x_t)^T, \\ y_{t+1} &= x_t - \gamma^{-1} A_t^{-1} \nabla \ell_t(x_t), \\ x_{t+1} &= \arg\min_{x \in \mathcal{K}} (x - y_{t+1})^T A_t (x - y_{t+1}), \qquad \text{projection step.} \end{aligned}$$

Using the Sherman-Morrison formula, each step of ONS has a $O(d^2+P)$-cost, where P is the cost of the projection step. If the gradients $\nabla \ell_t(x_t)$, $t \geq 1$, verify the condition **(H3)** then the square of their Euclidean norm $\|\nabla \ell_t(x_t)\|^2$ satisfies a Poissonian exponential inequality.

Proposition 3.2. *Under Condition **(H3)** the gradients $\nabla \ell_t(x_t)$ satisfy*

$$\mathbb{E}_{t-1}\big[\exp(\eta\big(\|\nabla \ell_t(x_t)\|^2 - \mathbb{E}\big[\|\nabla \ell_t(x_t)\|^2\big]/\big(1 - \eta G_{\psi_2}^2\big)\big)\big] \leq 1, \qquad \forall \eta^2 < 1/G_{\psi_2},$$

for any time $t \geq 1$ a.s.

Proof. Expanding the exponential and using Condition (**H3**) we obtain

$$\begin{aligned} \mathbb{E}\big[\exp\big(\eta Y^2\big)\big] &= \sum_{k=0} \frac{\eta^k \mathbb{E}[Y^{2k}]}{k!} \leq 1 + \eta \mathbb{E}\big[Y^2\big]\Bigg(1 + \sum_{k \geq 2} \eta^{k-1} G_{\psi_2}^{2(k-1)}\Bigg) \\ &= 1 + \frac{\eta \mathbb{E}[Y^2]}{1 - \eta G_{\psi_2}^2} \leq \exp\big(\eta \mathbb{E}\big[Y^2\big]/\big(1 - \eta G_{\psi_2}^2\big)\big) \end{aligned}$$

for every $\eta G_{\psi_2}^2 < 1$ and the desired result follows. ☐

To control the second-order terms in Proposition 3.1, we combine the self-bounded martingale and Poissonian inequalities. We obtain a fast-rate stochastic regret bound for the ONS tuned choosing $\gamma = \alpha/3$.

Theorem 3.1. *Under **(H1)**, **(H2)** and **(H3)**, the ONS algorithm 1 for $\gamma = \alpha/3$ satisfies with probability $1 - 3\delta$ the stochastic regret bound*

$$\begin{aligned} Regret_T \leq \frac{3}{2\alpha}\Bigg(1 + d \log\Bigg(1 + \frac{2\alpha^2 D^2 (T G_2^2 + G_{\psi_2}^2 \log(\delta^{-1}))}{9}\Bigg)\Bigg) \\ + \Bigg(\frac{4\alpha (G_{\psi_2} D)^2}{9} + \frac{18}{\alpha}\Bigg) \log\big(\delta^{-1}\big) \end{aligned}$$

valid for every $T \geq 1$.

Our result extends fast-rate stochastic regret bounds for ONS far beyond existing results in the deterministic or iid bounded setting. It strengthens the result of Hazan (2016) only valid when ℓ_t is exp-concave for every $t \geq 1$.

4. BOA achieves fast-rate regret bounds in Stochastic Online Aggregation

4.1. Stochastic Online Aggregation

Cesa-Bianchi and Lugosi (2006) consider to learn expert advices $x_t^{(j)}$, for all $1 \le j \le K$, by aggregating them sequentially for $t \ge 1$. Here we consider $\mathbf{x}_t = [x_t^{(1)}, \dots, x_t^{(K)}]$ a $d \times K$ matrix whose columns are K different $\mathcal{F}_{t-1}$-adapted predictors $x_t^{(i)}$. We denote $\widehat{x}_t = \mathbf{x}_t \pi_t = \sum_{i=1}^K \pi_i x_t^{(i)}$ their aggregation, with π_t in the simplex $\Lambda_K = \{\pi \in \mathbb{R}^K; \pi > \mathbf{0}, \sum_{i=1}^K \pi_i = 1\}$. Aggregation algorithms combine the predictors with weights π_t minimizing the stochastic regret

$$Regret_T^{ag} = \max_{1 \le i \le K} \left\{ \sum_{t=1}^T L_t(\widehat{x}_t) - \sum_{t=1}^T L_t\big(x^{(i)}\big) \right\}, \qquad T \ge 1.$$

We have under Condition **(H2)** the relation

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

We consider the loss functions $\pi \to \ell_t(\mathbf{x}_t\pi)$ over $\mathcal{K} = \Lambda_K$ that is stochastically exp-concave with the same constant α as the original loss functions ℓ_t. Applying Proposition 3.1 under Condition **(H2)** we obtain

$$\begin{aligned}
&\sum_{t=1}^T L_t(\widehat{x}_t) - L_t(\mathbf{x}_t\pi) \\
&\quad \le \sum_{t=1}^T \nabla \ell_t(\mathbf{x}_t\pi_t)^T \mathbf{x}_t(\pi_t - \pi) + \frac{\lambda}{2}\sum_{t=1}^T \big(\nabla \ell_t(\mathbf{x}_t\pi_t)^T \mathbf{x}_t(\pi_t - \pi)\big)^2 \\
&\qquad + \frac{\lambda - \alpha}{2}\sum_{t=1}^T \mathbb{E}_{t-1}\big[\big(\nabla \ell_t(\mathbf{x}_t\pi_t)^T \mathbf{x}_t(\pi_t - \pi)\big)^2\big] + \frac{2}{\lambda}\log\big(\delta^{-1}\big). \qquad (4.1)
\end{aligned}$$

We identify the surrogate losses

$$(\pi_t - \pi)^T \boldsymbol{\ell}_t + \lambda/2\big((\pi_t - \pi)^T \boldsymbol{\ell}_t\big)^2$$

with gradients denoted by $\boldsymbol{\ell}_t = \mathbf{x}_t^T \nabla \ell_t(\mathbf{x}_t\pi_t)$. We analyze algorithms minimizing the sum of the surrogate losses in stochastic environments. We compare the aggregation strategy $\widehat{x}_t$ to $\pi \in \{e_i, 1 \le i \le K\}$, i.e., with the best predictor $x_t^{(i)}$, using the linear losses $(\pi_t - \pi)^T \boldsymbol{\ell}_t$ over $\mathcal{K} = \Lambda_K$. We call this problem, encompassing the learning with expert advice, the Stochastic Online Aggregation (SOA).

4.2. The stochastic regret for the scale-free version of BOA

The version of BOA described in Algorithm 2 is different than the original BOA algorithm in Wintenberger (2017), because of the specific tuning of the multiple

Algorithm 2: Bernstein Online Aggregation (Wintenberger, 2017), scale-free version

Initialization: Initial weights $\pi_1 \in \Lambda_K$ and $\eta_0^{-2} = \widetilde{L}_0 = \mathbf{0}$ ($\in \mathbb{R}^K$).
For each step $t \geq 1$: the predictors incur the losses $\boldsymbol{\ell}_t \in \mathbb{R}^K$.
Recursion: Update

$$\begin{aligned}
\eta_t^{-2} &= \eta_{t-1}^{-2} + 2.2\big(\boldsymbol{\ell}_t - \pi_t^T \boldsymbol{\ell}_t \mathbb{1}\big)^2, \\
\widetilde{L}_t &= \widetilde{L}_{t-1} + \big(\boldsymbol{\ell}_t - \pi_t^T \boldsymbol{\ell}_t \mathbb{1}\big) + \eta_t\big(\boldsymbol{\ell}_t - \pi_t^T \boldsymbol{\ell}_t \mathbb{1}\big)^2, \\
\pi_{t+1} &= \frac{\eta_t \exp(-\eta_t \widetilde{L}_t)\pi_1}{\pi_1^T(\eta_t \exp(-\eta_t \widetilde{L}_t))}.
\end{aligned}$$

learning rates η_t. The specific η_t provides a self-normalization and the algorithm is scale-free, i.e., insensitive to a multiplicative factor of the losses.

The factor 2.2 is not arbitrary and is chosen such as a small numeric constant satisfying

$$\exp\bigg(-\frac{y}{\sqrt{1+2.2y^2}} - \frac{y^2}{1+2.2y^2}\bigg) \leq 1 - \frac{y}{\sqrt{1+2.2y^2}}, \qquad y \in \mathbb{R}.$$

This relation is crucial in the proof of Theorem 4.2 to propagate the self-normalization in a recursive argument. The coordinate-wise learning rate $\eta_{t,i}$ is only well defined after the first non-null observation $\underline{m}_i := (\boldsymbol{\ell}_{t,i} - \pi_t^T \boldsymbol{\ell}_t) \neq 0$, $1 \leq i \leq K$. Before that time $\widetilde{L}_{t,i} = \widetilde{L}_{t-1,i} = \cdots = 0$ by convention. Contrary to the ONS, and thanks to the adaptive learning rates, the algorithm BOA is parameter-free as it does not require the knowledge of α, and each step has a $O(K)$-cost. We provide a deterministic regret bound valid for any deterministic sequence.

Theorem 4.1. *For every $1 \leq i \leq K$, the BOA algorithm 2 achieves the deterministic regret bound*

$$\begin{aligned}
&\sum_{t=1}^T \pi_t^T \boldsymbol{\ell}_t - \sum_{t=1}^T \pi_t^T \boldsymbol{\ell}_{t,i} \\
&\quad \leq \sqrt{2.2 \sum_{t=1}^T \big(\pi_t^T \boldsymbol{\ell}_t - \boldsymbol{\ell}_{t,i}\big)^2} \bigg(\frac{1}{1.1} + \log\big(\pi_{1,i}^{-1}\big) \\
&\qquad + \sum_{i=1}^K 1\Big\{\max_{2 \leq t \leq T} x_{t,i} > 1/4\Big\} \log\big(1 + (M_{T,i}/\underline{m}_i)^2\big) \\
&\qquad + \log\bigg(\mathrm{e} + \frac{1}{2}\pi_1^T \log\big(1 + (M_T/\underline{m})^2 T\big)\bigg)\bigg)
\end{aligned} \tag{4.2}$$

where $x_T = \eta_{T-1}(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1})$, $M_T = \max_{2 \leq t \leq T} |\boldsymbol{\ell}_t - \pi_t^T \boldsymbol{\ell}_t \mathbb{1}| \in \mathbb{R}^K$ and $\underline{m}_i$ is the first non null observation of $\boldsymbol{\ell}_{t,i} - \pi_t^T \boldsymbol{\ell}_t$.

The term

$$\sum_{i=1}^{K} 1\Big\{\max_{2\le t\le T} x_{t,i} > 1/4\Big\}\log\big(1+(M_{T,i}/\underline{m}_i)^2\big)$$

in the regret bound (4.2) replaces the term $\|M_T\|_\infty$ in the regret bounds of Mhammedi et al. (2019), Orseau and Hutter (2021). In some unbounded stochastic settings, our regret bound (4.2) is better for T large. For instance, if we consider that the first predictor is iid standard Gaussian and the other ones are bounded then $\|M_T\|_\infty \sim M_{T,1} \sim \sqrt{2\log T}$ is much larger than $\sum_{i=1}^K \log(1 + (M_{T,i}/\underline{m}_i)^2) \sim \log\log T$ for T large.

The deterministic regret bound in Theorem 4.1 is assumption-free. Its first term is the square root of the sum of the additional quadratic terms in the surrogate losses (4.1). It may increase at the rate $O(\sqrt{T})$ but, under condition **(H2)**, it becomes negligible. We provide a stochastic regret bound for sequential aggregation using BOA.

Theorem 4.2. *Assume Conditions **(H1)**, **(H2)** and **(H3)** hold on* $\mathbf{x}_t^T\nabla\ell_t(\mathbf{x}_t\pi_t)$ *a.s. for all* $t\ge 1$, $1\le i\le K$. *The scale-free BOA algorithm 2 with* $\pi_i \ge e^{-K}$ *for all* $1\le i\le K$ *satisfies, with probability* $1-3\delta$,

$$\begin{aligned} Regret_T^{ag} \le{}& \frac{3(K+1)^2}{\alpha}\left(\log\left(1+\frac{2G_{\psi_2}^2\log T}{\underline{m}^2}\right)\right)^2 \\ &+ O\big((\log\log\log T)^2\big) + \left(\frac{2\alpha}{3}(G_{\psi_2}D)^2 + \frac{6}{\alpha}\right)\log\big(\delta^{-1}\big), \end{aligned}$$

for every $T\ge 1$, *and* $\underline{m}>0$ *such that* $\mathbb{P}(\min_{1\le i\le K}\underline{m}_i \ge \underline{m}) \le 1-\delta$.

Aggregation problems are easier than optimization ones and BOA achieves a faster stochastic regret bound than ONS. This rate $O((\log\log T)^2)$ is suboptimal in the learning with expert advice setting. Condition **(H3)** implies that the deterministic gradients are bounded by a constant $G>0$, and Condition **(H2)** implies exp-concavity. Optimal strategies achieve $O(G\log K)$ deterministic regret (Cesa-Bianchi and Lugosi, 2006). Among them Exponentially Weighted Aggregation, but this algorithm achieves only a $O(\sqrt{T})$ stochastic regret as shown by Audibert (2007). Best-known aggregation algorithms in deterministic and unbounded stochastic settings are different. It is an open question to find an aggregation algorithm optimal in both settings whereas squint and the original version of BOA achieve optimal rates in bounded deterministic and stochastic settings. The choice of the initial weights π_1 being not crucial in the latter setting we choose implicitly uniform initial weights in the sequel.

4.3. The SOCO analysis to adapt to unknown stochastic exp-concavity constant $\boldsymbol{\alpha > 0}$

We study an example of BOA-ONS dealing with the adaptation to the best stochastic exp-concavity constant α. It is crucial for improving the ONS performances in any stochastic environment where, contrary to deterministic ones, there is no way to determine the optimal α as it depends on the conditional distributions of ℓ_t. Consider $\widehat{x}_t = \mathbf{x}_t\pi = \sum_{i=1}^K \pi_i x_t^{(i)}$ the BOA aggregation of $K \geq 1$ ONS predictions with different parameters $\gamma^{(i)}$ with $\gamma^{(i)} = \{2^{-1}, \dots, 2^{-K}\}$. The resulting BOA-ONS algorithm adapts to the optimal value of α that depends on the unknown stochastic environment. The algorithm Metagrad of van Erven et al. (2021) is also able to adapt to different rates of convergence at the price of an extra factor $\|M_T\|_\infty \log T$ that is sub-optimal in unbounded sub-Gaussian settings when $\|M_T\|_\infty$ increases as $\sqrt{\log T}$.

Corollary 4.1. *Under **(H1)**, **(H2)** and **(H3)** with $\alpha \geq 2^{-K-2}$, BOA-ONS algorithm satisfies with probability $1-4\delta$ the stochastic regret bound*

$$\sum_{t=1}^T L_t(\widehat{x}_t) - \sum_{t=1}^T L_t(x)$$
$$\leq \frac{1}{\alpha} O\big(d\log(T) + K^2 \log\log(T)^2\big) + O\bigg(\alpha(G_{\psi_2}D)^2 + \frac{1}{\alpha}\bigg)\log\big(\delta^{-1}\big).$$

Proof. We combine the stochastic regret bound of Theorem 4.2 with the inequality (A.4) choosing $-\log_2(\gamma)+1 \leq i \leq -\log_2(\gamma)+2$ so that $\alpha/4 \leq \gamma \leq \alpha/2$ for $\alpha \leq 1$. □

5. BOA-ONS for sequential prediction and probabilistic forecast of time series

5.1. Probabilistic forecasting

Observing a time series (y_t), we use the SOCO analysis to calibrate some parametric probabilistic forecasters in the sense of Chapter 12 of Shafer and Vovk (2019). In our setting sequential algorithms predict x_t and parametrize a probabilistic forecaster P_{x_t}. Scoring rules are real-valued functions of the forecast and the observation. They have been introduced by Gneiting and Raftery (2007) as summary measures for the evaluation of probabilistic forecasts. Given such a scoring rule S, we consider the loss at step t as $\ell_t(x_t) = S(P_{x_t}, y_t)$, $t \geq 1$. The expected score, also denoted by S in Gneiting and Raftery (2007), is a discrepancy measure between probabilities

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

where P_t denotes the distribution of y_t given $\mathcal{F}_{t-1}$. **(H2)** is a condition on the scoring rule S, the parametrization $x \mapsto P_x$ and the distribution P_t of the variable of interest y_t given $\mathcal{F}_{t-1}$. It writes as, for all $x, y \in \mathcal{K}$

$$S(P_y, P_t) \leq S(P_x, P_t) + \nabla_y S(P_y, P_t)^T(y-x) - \frac{\alpha}{2}\mathbb{E}_{t-1}\big[\big(\nabla_y S(P_y, y_t)^T(y-x)\big)^2\big].$$

If Condition **(H2)** is satisfied in well-specified settings $P_t = P_{x_t^*}$, for any $x_t^* \in \mathcal{K}$, then S is a proper scoring rule for the class $\{P_x; x \in \mathcal{K}\}$ in the sense of Gneiting and Raftery (2007); $S(P_y, P_t)$ is minimum when $P_y = P_t$ by convexity. The scoring rule is not necessarily strictly proper since this maximum is not unique when $\nabla_y S(P_y, P_t)$ is null in some directions y in the neighborhood of x_t^*.

We provide examples of time series probabilistic forecasting calibrated using the SOCO analysis by verifying Condition **(H2)**. We focus on the logarithmic score assuming that P_x, P_t admit densities p_x, p_t, $x \in \mathcal{K}$, $t \geq 1$. We have

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

where KL is the Kullback-Leibler divergence. This scoring rule is strictly proper because S is minimized when $P_y = P_t$ only. It is likely to satisfy the stochastic exp-concavity condition **(H2)** locally in well-specified settings.

Proposition 5.1. *If P_t is in the exponential family so that its conditional 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*

$$\begin{aligned}\mathbb{E}_{t-1}\big[\nabla \ell_t(x_t^*) \nabla \ell_t(x_t^*)^T\big] &= \mathbb{E}_{t-1}\big[\nabla_{x_t^*} S(P_{x_t^*}, y_t)^T \nabla_{x_t^*} S(P_{x_t^*}, y_t)^T\big] \\ &= \nabla^2_{x_t^*} S(P_{x_t^*}, P_{x_t^*}) = \nabla^2 L_t(x_t^*),\end{aligned}$$

*and necessarily $\alpha \leq 1$ if condition **(H2)** holds.*

Proof. We apply Proposition 2.2, noticing that the Fisher information identity holds in the well-specified setting. □

We use the logarithmic score for calibrating the first and second moments of Gaussian forecasters as recommended in Section 4.4 of Gneiting and Raftery (2007). Giraud et al. (2015) focus on the estimation of $m_t = \mathbb{E}_{t-1}[y_t]$, establishing fast-rate stochastic regret bounds in expectation.

Example 5.1 (Estimation of the conditional expectation)**.** *Let $P_x = \mathcal{N}(x, \sigma^2)$ so that $\ell_t(x) = (x - y_t)^2/(2\sigma^2)$ (plus constant). In the OCO analysis, ℓ_t is σ^2/D^2-exp-concave only if $y_t \in \mathcal{K}$ satisfying **(H1)**. This setting requires implicitly that the distributions P_t are $\mathcal{K}$ supported. Unbounded cases $y_t \notin \mathcal{K}$ are analyzed by SOCO assuming that the conditional distribution P_t has mean $m_t = \mathbb{E}_{t-1}[y_t] \in \mathcal{K}$ and finite conditional variance $\sigma_t^2 = \mathrm{Var}_{t-1}(y_t) \leq \overline{\sigma}^2$ a.s., for some $\overline{\sigma}^2 > 0$ and all $t \geq 1$. The losses ℓ_t are not exp-concave but still satisfy Condition **(H2)** with $\alpha = \sigma^2/(\overline{\sigma}^2 + D^2)$; See Proposition 5.2 for more details. The well-specified unbounded case $P_t = \mathcal{N}(x, \sigma_t^2)$ satisfied Condition **(H2)** with $\alpha = \sigma^2/(\overline{\sigma}^2 + D^2)$ when $m_t \in \mathcal{K}$ and $\sigma_t^2 \leq \overline{\sigma}^2$.*

We also focus on the estimation of the conditional variance or volatility $\sigma_t^2 = \mathrm{Var}_{t-1}(y_t)$ for Gaussian probabilistic forecasters. Up to our knowledge, stochastic regret bounds for sequential algorithms calibrating the volatility have not been established yet. However, the concept of volatility is important and required in many applications such as risk assessment and probabilistic forecasting in finance (McNeil et al., 2015, Shafer and Vovk, 2019). The logarithmic score is well-suited to measure the performances of volatility estimators as it is robust to extreme values (Patton, 2011).

Example 5.2 (Estimation of the volatility). *Let $P_x = \mathcal{N}(m_t, x)$ then $\ell_t(x) = (\log(x) + (y_t - m_t)^2/x)/2$ (plus constant) is convex only if $0 < x \leq 2(y_t - m_t)^2$. This assumption is unrealistic when y_t is concentrated around its conditional mean m_t. Using SOCO, if the conditional distribution P_t has mean m_t and volatility $\sigma_t^2 \in \mathcal{K} = [c\overline{\sigma}^2/2, \overline{\sigma}^2]$, $\overline{\sigma}^2 > 0$, $1 < c < 2$, then the risk $L_t(x) = (\log(x) + \sigma_t^2/x)/2$ is strongly convex with $\mu = (c-1)/(2\overline{\sigma}^4)$. Condition **(H2)** is satisfied with $\alpha = (c-1)c^4 2^{-6}$ if $\mathbb{E}_{t-1}[(y_t^2 - \sigma_t^2)^2] \leq 3\overline{\sigma}^4$; See Proposition 5.3 for more details.*

The stochastic exp-concavity condition is well-preserved for linear multivariate parametrization. Thus the conditional expectation and the volatility can be expressed as a linear combination of the past observations $y_{t-1}, \dots, y_1$ or their squares $y_{t-1}^2, \dots, y_1^2$. We obtain naturally AR and ARCH estimations for the conditional expectation and the volatility in Sections 5.2 and 5.3, respectively. Combining both, we obtain the AR-ARCH Gaussian forecaster studied in Section 5.4. The parametrization does not preserve the strictly proper property of the logarithmic loss function. Despite the logarithmic score being strictly proper overall probability measures, it is not for the AR-ARCH models because different linear combinations of past observations provide the same probabilistic forecaster. Stochastic exp-concavity condition **(H2)**, more general than strict properness, is crucial.

5.2. Sequential ARMA prediction by BOA-ONS

AutoRegressive Moving Average (ARMA) modeling of the conditional mean is standard in time series analysis. See Brockwell and Davis (2009) for a reference textbook. We calibrate sequentially, using the SOCO analysis with the natural filtration $\mathcal{F}_t = \sigma(y_t, \dots, y_1)$, and the Gaussian forecaster $\mathcal{N}(\widehat{m}_t^{(p)}(x), \sigma^2)$ for arbitrary $\sigma^2 > 0$, with clipped mean

$$\widehat{m}_t^{(p)}(x) = x^T\big((y_{t-1} \wedge M/2) \vee (-M/2), \dots, (y_{t-p} \wedge M/2) \vee (-M/2)\big),\ M > 0,$$

and $\mathcal{K} = B_1(1)$, the ℓ^1 unit-ball of dimension p.

Proposition 5.2. *We assume that the distributions P_t of y_t given $y_{t-1}, \dots, y_1$, admit densities with means m_t satisfying $2|m_t| \leq M$, volatilities $\sigma_t^2 \leq \overline{\sigma}^2$ a.s., $M > 0$, $\overline{\sigma}^2 > 0$, for every $t \geq 1$, and satisfy **(H3)**. Then the Gaussian forecaster $\mathcal{N}(m_t^{(p)}(x), \sigma^2)$ calibrated by the ONS algorithm with $\gamma = \sigma^2/(3(\overline{\sigma}^2 + M^2))$ achieves the stochastic regret*

$$Regret_T \leq O\bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} p \log T + \bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} + \frac{\sigma^2}{\overline{\sigma}^2 + M^2} p\, G_{\psi_2}^2\bigg) \log\big(\delta^{-1}\big)\bigg),$$

for every $T \geq 1$, $x \in B_1(1)$, and with high probability $1 - \delta$.

To tackle the case of ARMA models with a moving average component, we consider increasing orders p since any invertible ARMA model admits an AR(∞)

representation. For the orders $p \in \{1, \dots, \sqrt{\log T}/\log\log T\}$, the ONS predictors $\widehat{m}_t^{(p)}(x_t)$ are aggregated with BOA in $\widehat{m}_t$. The obtained BOA-ONS algorithm achieves the cumulative KL-divergence bound

$$\begin{aligned}
&\sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(\widehat{m}_t, \sigma^2\big)\big) \\
&\quad \le \min_{1 \le p \le \sqrt{\log T}/\log\log T} \min_{x \in B_1(1)} \bigg\{ \sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(m_t^{(p)}(x), \sigma^2\big)\big) \\
&\qquad + O\bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} p \log T\bigg) + O\bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} + \frac{\sigma^2}{\overline{\sigma}^2 + M^2} p\, G_{\psi_2}^2\bigg) \log\big(\delta^{-1}\big)\bigg\},
\end{aligned} \tag{5.1}$$

refining the bound obtained by Anava et al. (2013) in the following ways. Our bound is valid in every sub-Gaussian stochastic adversarial setting where $2|m_t| \le D$, and the time series (y_t) does not have to be bounded as in Anava et al. (2013). Moreover, our bounds are anytime-valid with high probability.

The parameters (M, σ^2) should be tuned to find the best compromise in the regret bound (5.1). However, the task is not feasible using the SOA analysis because the loss functions depend on these parameters. The solution comes from the econometrics literature that provides better loss and risk functions introducing the concept of volatility.

5.3. Sequential ARCH prediction by BOA-ONS

In mathematical finance, the log-ratios (y_t) are commonly modeled using the Generalized AutoRegressive Conditionally Heteroscedastic (GARCH) model. Classical inference uses the Quasi-Likelihood approach (Francq and Zakoian, 2019) as if the conditional distributions were Gaussian. If the conditional means $m_t := \mathbb{E}[y_t \mid y_{t-1}, \dots, y_1]$ are null, the volatilities $\sigma_t^2 := \mathrm{Var}(y_t \mid y_{t-1}, \dots, y_1)$ are finite, $t \ge 1$, then the Quasi-Likelihood estimator $\widehat{\sigma}_t^2$ of the volatility minimizes the cumulative KL divergence $KL(P_t, \mathcal{N}(0, \widehat{\sigma}_t^2)) = (\log(2\pi\widehat{\sigma}_t^2) + \sigma_t^2/\widehat{\sigma}_t^2)/2$ (plus constant).

We assume that $\sigma_t^2 \in [c\overline{\sigma}^2/3, \overline{\sigma}^2]$, $1 < c < 2$, $\overline{\sigma}^2 > 0$, and we use a clipped-ARCH(q) model

$$\widehat{\sigma}_t^{2,(q)}(x) = c\overline{\sigma}^2/2 + x_1\big(y_{t-1}^2 \wedge \overline{\sigma}^2\big) + \cdots + x_q\big(y_{t-q}^2 \wedge \overline{\sigma}^2\big), \tag{5.2}$$

with $x \in \mathcal{K} = \{x \in \mathbb{R}^q : x \ge \mathbf{0} \text{ and } \|x\|_1 \le 1 - c/2\}$.

Proposition 5.3. *We assume that the distributions P_t of y_t given $y_{t-1}, \dots, y_1$, admit densities with means $m_t = 0$, volatilities $\sigma_t^2 \in [c\overline{\sigma}^2/2, \overline{\sigma}^2]$, $\mathbb{E}_{t-1}[(y_t^2 - \sigma_t)^2] \le 3\overline{\sigma}^4$, a.s., $1 < c < 2$, $\overline{\sigma}^2 > 0$ for every $t \ge 1$, and satisfy **(H3)**. Then the Gaussian forecaster $\mathcal{N}(0, \widehat{\sigma}_t^{2,(q)}(x))$ calibrated by the ONS algorithm with $\gamma = 2^6/(3(c-1)c^4)$ achieves the stochastic regret*

$$Regret_T \le O\big(q\big(\log T + G_{\psi_2}^2 \log\big(\delta^{-1}\big)\big)\big),$$

for every $T \geq 1$*,* $x \in \mathcal{K}$ *with high probability* $1-\delta$*.*

Any invertible GARCH model admits an ARCH(∞) representation. Thus we consider ARCH(q) models with increasing order q. We consider BOA-ONS $\widehat{\sigma}_t^2$ aggregating $\widehat{\sigma}_t^{2,(q)}(x_t)$, $q = 1, \dots, \sqrt{\log T}/\log\log T$ so that with high probability

$$\sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(0, \widehat{\sigma}_t^2\big)\big) \leq \min_{1 \leq q \leq \sqrt{\log T}/\log\log T} \min_{x \in \mathcal{K}} \Bigg\{ \sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(0, \sigma^{2,(q)}(x)\big)\big) + O\big(q\big(\log T + G_{\psi_2}^2 \log\big(\delta^{-1}\big)\big)\big) \Bigg\}.$$

We solve positively the question raised in the conclusion of Anava et al. (2013) about the optimization of GARCH forecasters. The main restriction of our approach is the small range of the volatilities $[c\overline{\sigma}^2/2, \overline{\sigma}^2]$, $1 < c < 2$. Otherwise, the risk functions arer not even convex when the volatility σ_t^2 can be over-estimated by a factor of 2. It is not surprising since Francq and Zakoïan (2010) showed that the Quasi-Likelihood approach is inconsistent with no lower boundedness assumption on the volatilities.

5.4. Online Gaussian probabilistic forecasting using BOA-ONS

We combine the ARMA and volatility prediction methods. We consider Gaussian probabilistic forecaster $\mathcal{N}(\widehat{m}_t^{(p)}(x_{1:p}), \widehat{\sigma}_t^{2,(q)}(x_{p+1:p+q}))$ with $M^2 = \overline{\sigma}^2$ and $x = (x_{1:p}, x_{p+1:p+q}) \in \mathcal{K}$ with

$$\mathcal{K} = \big\{ x \in \mathbb{R}^{p+q} : \|x_{1:p}\|_1 \leq 1,\ x_{p+1:p+q} \geq \mathbf{0}, \|x_{p+1:p+q}\|_1 \leq 1 - c/2 \big\}.$$

Proposition 5.4. *We assume that the distributions* P_t *of* y_t *given* $y_{t-1}, \dots, y_1$*, admit densities with means* $2|m_t| \leq \overline{\sigma}$*, volatilities* $\sigma_t^2 \in [c\overline{\sigma}^2/2, \overline{\sigma}^2]$*,* $\mathbb{E}_{t-1}[(y_t - m_t)^4] \leq 3\overline{\sigma}^4$*, a.s.,* $1 < c < 2$*,* $\overline{\sigma}^2 > 0$ *for every* $t \geq 1$*, and satisfy* ***(H3)****. Then the Gaussian forecaster* $\mathcal{N}(\widehat{m}_t^{(p)}(x), \widehat{\sigma}_t^{2,(q)}(x))$ *calibrated by the ONS algorithm with* $\gamma = 3 \times 2^5/((c-1)c^4)$ *achieves the stochastic regret*

$$Regret_T \leq O\big((p+q)\big(\log T + G_{\psi_2}^2\big) \log\big(\delta^{-1}\big)\big),$$

for every $T \geq 1$ *with high probability* $1-\delta$*.*

Aggregating such predictors for $1 \leq p, q \leq \sqrt{\log T}/\log\log T$ with BOA, we obtain a Gaussian probabilistic forecast $\mathcal{N}(\widehat{m}_t, \widehat{\sigma}_t^2)$ satisfying the cumulative KL-divergence bound

$$\sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(\widehat{m}_t, \widehat{\sigma}_t^2\big)\big)$$
$$\leq \min_{1 \leq q,p \leq \sqrt{\log T}/\log\log T} \min_{x \in \mathcal{K}} \Bigg\{ \sum_{t=1}^{T} KL\big(P_t, \mathcal{N}\big(\widehat{m}_t^{(p)}(x_{1:p}), \widehat{\sigma}_t^{2,(q)}(x_{p+1:p+q})\big)\big)$$

$$+ O\big((p+q)\big(\log T + G_{\psi_2}^2 \log\big(\delta^{-1}\big)\big)\big)\Bigg\},$$

with high probability. The sequential algorithms adapt to the random environment even in misspecified settings; It approximates the parametric Gaussian forecaster that is the closest to the unknown conditional distributions for the cumulative KL divergences and a penalty which increases such as $(p+q)\log(T)$. Thus the BOA-ONS forecaster regret minimizes automatically a Bayesian information type criterion at any-time and with high probability. It is comparable to a model selection procedure that would require to minimize a penalized log-likelihood at each step $1 \le t \le T$. The computational cost of our sequential method is $O(T((p+q)^2 + P))$ with explicit formulae except for the projection step of computational cost P, whereas the batch model selection has a computational cost $O(T(p+q)M)$ where M is the computational cost of the optimization of the likelihood in AR(p)-ARCH(q) models. This cost M is prohibitive when $p+q$ is large and the computational gain of our sequential procedure is important.

5.5. Sequential probabilistic forecasting using BOA-ONS

The main drawback of our BOA-ONS approach on Gaussian forecasters is the restriction $\sigma_t^2 \in [c\overline{\sigma}^2/2, \overline{\sigma}^2]$, $1 \le t \le T$. However, because the loss and risk functions depend on this hyperparameter, it is not possible to directly aggregate volatility estimators with different $\overline{\sigma}^2 > 0$ in a Gaussian forecaster to extend the range of the volatilities.

To circumvent the issue, we can aggregate the Gaussian probabilistic forecasters to obtain a probabilistic forecaster which is mixed Gaussian. Consider $\widehat{P}_t = (\widehat{P}_t^{(i)})_{1\le i\le K}$, K weak probabilistic forecasters with densities $\widehat{p}_t = (\widehat{p}_t^{(i)})_{1\le i\le K}$ such as $\widehat{P}_t^{(i)} = \mathcal{N}(\widehat{m}_t^{(j)}, \widehat{\sigma}_t^{2,\ell})$ with different localization, $jD + D/\sqrt{2} \le \widehat{m}_t^{(j)} \le (j+1)D + D/\sqrt{2}$, for $-K_1 \le j \le K_2$ and $\widehat{\sigma}_t^{2,\ell} \in [(c/2)^{\ell+1}\overline{\sigma}^2/2, (c/2)^{\ell}\overline{\sigma}^2]$ for $0 \le \ell \le K_3$. Consider the SOCO analysis of mixtures $x^T\widehat{P}$ with $\mathcal{K} = \Lambda_K$ and $\ell_t(x) = -\log(x^T\widehat{p}_t(y_t))$. We assume that $m \le \mathbb{E}_{t-1}[1/\widehat{p}_t^{(i)}(y_t)^2] \le M$ a.s. for $1 \le i \le K$, $t \ge 1$. The risk function is m-strongly convex and Condition **(H2)** is satisfied with $\alpha = m/M$.

Under Condition **(H3)**, we can use the ONS algorithm on the simplex $\mathcal{K} = \Lambda_K$, and we obtain

$$\sum_{t=1}^T KL\big(P_t, \pi_t^T\widehat{p}\big) \le \min_{\pi\in\Lambda_K} \sum_{t=1}^T KL\big(P_t, \pi^T\widehat{p}\big) + O\big(MK/m\big(\log T + \log\big(\delta^{-1}\big)\big)\big).$$

Similar fast-rate 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 constrain π_t to be in Λ_K. Contrary to our procedure, it is difficult to interpret their ensemble probabilistic forecast because they do not satisfy the axioms of a density function.

Table 1
Statistics of the cumulative square losses based on 100 *Monte-Carlo experiments.*

Algorithm	Mean$\times 10^2$ (Sd$\times 10^2$) Unb. Experts			Mean$\times 10^3$ (Sd$\times 10^3$) Biased Experts		
	$\sigma = 1$	$\sigma = .1$	$\sigma = 10$	$\sigma = 1$	$\sigma = .1$	$\sigma = 10$
BOA	1.65 (.07)	.164 (.007)	16 (.7)	2.65 (.09)	.231 (.004)	31 (1)
S-F BOA	**1.17 (.05)**	**.117 (.006)**	**12 (.6)**	**1.37 (.04)**	.136 (.003)	**26 (1)**
Squint	**1.16 (.05)**	**.117 (.004)**	**12 (.5)**	**1.37 (.07)**	**.110 (.001)**	31 (2)

6. Numerical illustrations

6.1. Aggregations in stochastic environments

We study the impact of stochastic deviations on the aggregation of predictors for quadratic losses. We consider two sets of 100 predictors of $y_t = 0$, $t \geq 1$. As a baseline, the first set consists in $\sigma N_t^{(i)}$, with $N_t^{(i)}$ iid standard Gaussian random variables, independent for $1 \leq i \leq 100$. These 100 predictors are all unbiased. In this simple setting the uniform weights, which is the initial value of every implemented aggregation algorithms, is optimal. The second set consists in 100 predictors, the first one being negatively biased $-\sqrt{t} + \sigma N_t^{(1)}$, the other ones being positively biased $\sqrt{t} + \sigma N_t^{(i)}$, $1 \leq t \leq 1000$, $2 \leq i \leq 100$. Any aggregation half-weighting the first predictor does not suffer from the bias. We run 100 Monte-Carlo experiments of three different aggregation algorithms; the original version of BOA of Wintenberger (2017), the scale-free version of BOA described in Algorithm 2, and the squint algorithm of Koolen et al. (2016). The latter algorithm is not comparable to the others as it uses beforehand the maximum of the deviations for initializing the recursion.

In the baseline unbiased setting, the performances reported in Table 1 of scale-free BOA and squint algorithms are not distinguishable. Both algorithms are more stable than the original version of BOA which uses a doubling trick to adapt to the deviations of the predictors[1]. Note that the cumulative square losses are proportional to the standard deviations of the predictors.

In the biased setting, the comparison depends on the level of the stochastic deviations; see also the boxplots in Figure 1. The price for learning non-trivial weights to correct the bias of the predictors is approximately a factor 10 on the cumulative losses when compared with the baseline setting. The case $\sigma = 1$ is similar to the unbiased setting with scale-free BOA and squint algorithms both outperforming equally well the original version of BOA. When $\sigma = .1$ the scale-free version of BOA is outperformed by squint. Its regret bound suffers when the minimal observed loss $\underline{m}$ is small because of unstable self-normalization. Squint does not suffer from this drawback because the range of the deviations is provided directly in its initialization. On the opposite, the scale-free BOA algorithm outperforms the two other algorithms for large deviations when $\sigma = 10$. The performances, that deteriorate when stochastic deviations are of the

[1] The multiple tuning of the deviation bounds in the original version of BOA is flawed and replaced by the univariate doubling trick of Cesa-Bianchi et al. (2007).

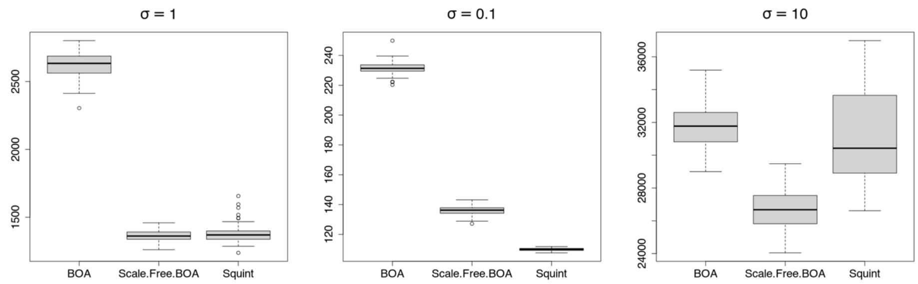

Fig 1. *Boxplots of the cumulative square losses based on* 100 *Monte-Carlo experiments of the biased setting.*

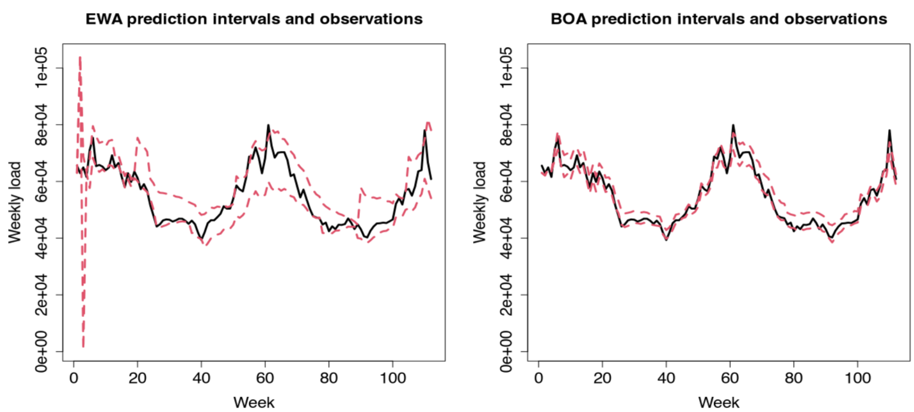

Fig 2. 90%*-prediction intervals of the electricity load based on EWA (left) and BOA (right) and the same 5 forecasters.*

same order as the bias, are in accordance with our regret bounds. Designing self-normalized sequential algorithms such as scale-free BOA is a robust alternative to the use of the range of the deviations as in BOA or squint algorithms.

6.2. Quantile prediction of electricity loads

We illustrate the impact of the SOCO anlysis on quantile predictions for weekly electricity load, data available in the Opera package developed by Gaillard et al. (2021). The 3 forecasters (GAM, AR, GBM) provided in Opera package plus 2 constant forecasters, 0 and 1.5 times the maximum of weekly loads, are aggregated to predict the upper and lower quantile of levels .5 and .95. We use the quantile loss funcion in 2 different sequential aggregation algorithms, Exponentially Weighted Algorithm (EWA) and BOA, and for the two levels .5 and .95. BOA aggregations provide accurate quantile predictions because it minimizes cumulative risks in the SOA analysis. It confirms the theoretical guarantees obtained in the paper since it is likely that the pinball risk is strongly convex

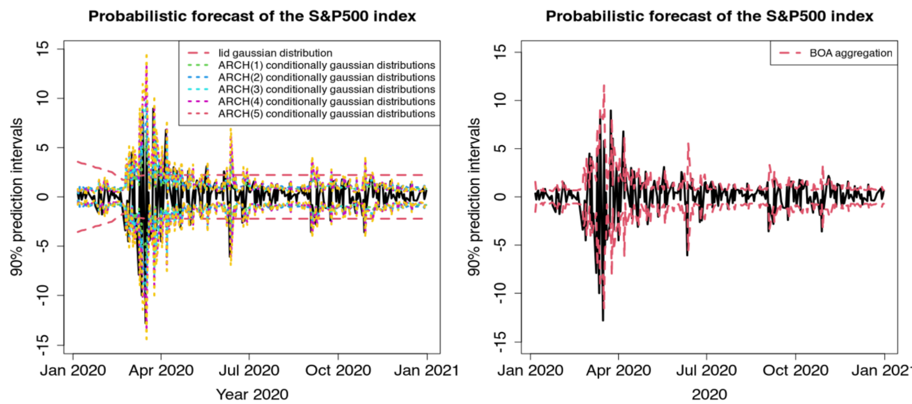

FIG 3. *90%-prediction intervals from 6 forecasters (left) and their BOA aggregation (right).*

(Steinwart and Christmann, 2011). On the contrary EWA aggregations fail to provide accurate quantile predictions because EWA algorithm minimizes the cumulative losses which are not exp-concave. Such visual validation of the predictions interval is enough to show the benefit of BOA but does not constitute any evidence of its good calibration. Biau and Patra (2011) analyze the asymptotic guarantees of a different sequential algorithm predicting quantiles.

6.3. Volatility estimation during the COVID crisis

We apply BOA-ONS for designing 90%-prediction intervals for the S&P500 index during 2020, including the COVID crisis in March. We use the iid $\mathcal{N}(0, x)$ and ARCH(p) Gaussian probabilistic forecasters for $p = 1, \dots, 5$. The forecasters are tuned sequentially with the ONS algorithm with $\gamma = 1$, and $\mathcal{K} = [c, \infty] \times B_1(1)$, $c = 0$ in the iid case, and $c = 10^{-16}$ in the ARCH cases. These 6 predictors of the volatility are then aggregated with BOA; See Figure 6.3. We notice that the iid forecast prediction interval is constant after some training period. The ARCH forecasts are required to predict intervals accurately during the crisis. BOA aggregations converge to weights $(0.01, 0.17, 0.09, 0.37, 0.21, 0.15)$ and improve the calibration of ARCH forecasters. A slightly more advanced sequentially calibrated volatility estimator developed by Werge and Wintenberger (2022) has been used in the forecast task of the M6 financial competition by de Vilmarest and Werge (2023). Its RPS performances rank 5th out of 163 competitors, showing that such sequential calibration o is competitive in probabilistic forecasting.

7. Conclusion and future works

In this paper, we derive fast-rate stochastic regret bounds for the ONS and BOA algorithms under stochastic exp-concavity. We alleviate the convexity assumption on the loss functions to calibrate sequentially parametric probabilistic forecasting using the logarithmic score. We achieve fast-rate stochastic regret bounds. Thus, BOA-ONS can adaptively and efficiently calibrate Gaussian probabilistic forecasters for any conditionally sub-Gaussian non-stationary time series. Our stochastic regret bounds are relative to a static prediction parametrized by $x \in \mathcal{K}$ for every $t \geq 1$. When forecasting non-stationary time series, we should also consider competitors that evolve through time. Key Propositions 3.1 and 3.2 extend readily to such settings called tracking optimization problems. Thus, one would like to develop SOCO and algorithms in more dynamic settings. A first step in that direction is made in Haddouche et al. (2023) using optimistic sequential algorithms.

Appendix A: Proofs of the main results

A.1. Proof of Proposition 2.4

We first show that

$$\left\| \frac{\nabla \ell_t(x_t)^T (x_t - x)}{\sqrt{\mathbb{E}_{t-1}[(\nabla \ell_t(x_t)^T (x_t - x))^2]}} \right\|_{\psi_2} \leq \sqrt{8/3 + (1/\log 2)^2} = K_{\psi_2} \approx 2.179.$$

Then we derive

$$\begin{aligned} &\frac{\mathbb{E}_{t-1}[(\nabla \ell_t(x_t)^T (x_t - x))^{2k}]}{\mathbb{E}_{t-1}[K_{\psi_2}^2 (\nabla \ell_t(x_t)^T (x_t - x))^2]^k} \\ &\quad \leq k! \mathbb{E}_{t-1}\left[\psi_2\left(\frac{\nabla \ell_t(x_t)^T (x_t - x)}{\sqrt{\mathbb{E}_{t-1}[K_{\psi_2}^2 (\nabla \ell_t(x_t)^T (x_t - x))^2]}}\right)\right] \\ &\quad \leq 2k!. \end{aligned}$$

Using Cauchy-Svhwarz inequality we derive that $\mathbb{E}_{t-1}[(\nabla \ell_t(x_t)^T (x_t - x))^2] \leq \mathbb{E}_{t-1}[\|\nabla \ell_t(x_t)\|^2] D^2 \leq G_2^2 D^2$ and

$$\begin{aligned} \mathbb{E}_{t-1}\big[\big(\nabla \ell_t(x_t)^T (x_t - x)\big)^{2k}\big] &\leq k! 2 K_{\psi_2}^{2k} \mathbb{E}_{t-1}\big[\big(\nabla \ell_t(x_t)^T (x_t - x)\big)^2\big]^k \\ &\leq k! 2 K_{\psi_2}^{2k} (G_2 D)^{2(k-1)} \mathbb{E}_{t-1}\big[\big(\nabla \ell_t(x_t)^T (x_t - x)\big)^2\big]. \end{aligned}$$

Then we fix $G_{\psi_2} = 2K_{\psi_2}^2 G_2$ so that Condition **(H3)** follows. Let us denote μ_t and σ_t the mean and the variance of the conditionally Gaussian random variable. Then, N being standard Gaussian distributed, we use the homogeneity and triangular inequality on the norm $\|\cdot\|_{\psi_2}$ to derive

$$\begin{aligned} \left\| \frac{\nabla \ell_t(x_t)^T (x_t - x)}{\sqrt{\mathbb{E}_{t-1}[(\nabla \ell_t(x_t)^T (x_t - x))^2]}} \right\|_{\psi_2} &= \left\| \frac{\sigma_t N + \mu_t}{\sqrt{\mathbb{E}_{t-1}[(\sigma_t N + \mu_t)^2]}} \right\|_{\psi_2} \\ &\leq \frac{\sigma_t \|N\|_{\psi_2} + \|\mu_t\|_{\psi_2}}{\sqrt{\sigma_t^2 + \mu_t^2}} \end{aligned}$$

$$= \frac{\sigma_t\sqrt{8/3} + \mu_t/\log 2}{\sqrt{\sigma_t^2 + \mu_t^2}}$$

and the desired results follows from Cauchy-Schwartz inequality.

A.2. Proof of Proposition 3.1

Denoting $Y_t = \nabla \ell_t(x_t)^T(x_t - x)$, we observe that under **(H2)** it holds

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

Moreover, from Lemma B.1 of Bercu and Touati (2008) for any random variable Y_t and any $\eta \in \mathbb{R}$ we have

$$\mathbb{E}_{t-1}\left[\exp\left(\eta\big(Y_t - \mathbb{E}_{t-1}[Y_t]\big) - \frac{\eta^2}{2}\big(\mathbb{E}_{t-1}\big[Y_t^2\big] - \mathbb{E}_{t-1}[Y_t]^2 + \big(Y_t - \mathbb{E}_{t-1}[Y_t]\big)^2\big)\right)\right] \le 1.$$

Developing the square, we obtain

$$\mathbb{E}_{t-1}\left[\exp(\eta\big(Y_t - \mathbb{E}_{t-1}[Y_t]\big) - \frac{\eta^2}{2}\big(\mathbb{E}_{t-1}\big[Y_t^2\big] + Y_t^2\big) + \eta^2\mathbb{E}_{t-1}[Y_t]Y_t\right] \le 1.$$

Using Young's inequality together with Jensen's one, we derive

$$\mathbb{E}_{t-1}[Y_t]Y_t \ge -\big(\mathbb{E}_{t-1}[Y_t]^2 + Y_t^2\big)/2 \ge -\big(\mathbb{E}_{t-1}\big[Y_t^2\big] + Y_t^2\big)/2$$

and the exponential inequality

$$\mathbb{E}_{t-1}\big[\exp(\eta\big(Y_t - \mathbb{E}_{t-1}[Y_t]\big) - \eta^2\big(\mathbb{E}_{t-1}\big[Y_t^2\big] + Y_t^2\big)\big] \le 1.$$

We obtain the desired result applying a classical martingale argument due to Ville (1939) and Freedman (1975) and recalled in Appendix B.1. Indeed, using the notation of Appendix B.1 with $Z_t = \eta(Y_t - \mathbb{E}_{t-1}[Y_t]) - \eta^2(\mathbb{E}_{t-1}[Y_t^2] + Y_t^2)$, we have

$$\mathbb{P}\big(\exists T \ge 1 :\ M_T > \delta^{-1}\big) \le \delta, \qquad 0 < \delta < 1,$$

where $M_T = \exp(\sum_{t=1}^T Z_t)$. Considering $\eta = -\lambda/2$ for any $\lambda > 0$, it holds with probability $1-\delta$ for any $T \ge 1$

$$\sum_{t=1}^T \left(-\frac{\lambda}{2}\big(Y_t - \mathbb{E}_{t-1}[Y_t]\big) - \frac{\lambda^2}{4}\big(\mathbb{E}_{t-1}\big[Y_t^2\big] + Y_t^2\big)\right) \le \log\big(\delta^{-1}\big)$$

$$\Leftrightarrow \sum_{t=1}^T \mathbb{E}_{t-1}[Y_t] \le \sum_{t=1}^T Y_t + \frac{\lambda}{2}\big(\mathbb{E}_{t-1}\big[Y_t^2\big] + Y_t^2\big) + \frac{2}{\lambda}\log\big(\delta^{-1}\big)$$

which, combines with (A.1), yields the desired result.

A.3. Proof of Theorem 3.1

From the proof of the ONS regret bound in Hazan (2016), we obtain from the expression of the recursive steps (and not using the convexity of the loss)

$$\sum_{t=1}^T \nabla\ell_t(x_t)^T(x_t-x)\le\frac{\gamma}{2}\sum_{t=1}^T\big(\nabla\ell_t(x_t)^T(x_t-x)\big)^2+\frac{1}{2\gamma}\log\big(\det(A_T)/\det(A_0)\big)+\frac{1}{2\gamma}.$$

Plugging this inequality into the previous bound we obtain

$$\begin{aligned}\sum_{t=1}^T L_t(x_t)-\sum_{t=1}^T L_t(x)\le{}&\frac{\lambda+\gamma}{2}\sum_{t=1}^T\big(\nabla\ell_t(x_t)^T(x_t-x)\big)^2\\&+\frac{\lambda-\alpha}{2}\sum_{t=1}^T\mathbb{E}_{t-1}\big[\big(\nabla\ell_t(x_t)^T(x_t-x)\big)^2\big]\\&+\frac{1}{2\gamma}\log\big(\det(A_T)/\det(A_0)\big)+\frac{1}{2\gamma}+\frac{2}{\lambda}\log\big(\delta^{-1}\big).\end{aligned}$$

Then we apply the Poissonian exponential inequality from Proposition 3.2 on the second-order terms. More precisely, denoting $0\le Y_t=(\nabla\ell_t(x_t)^T(x_t-x))^2/(2(G_{\psi_2}D)^2)$, we obtain

$$\mathbb{E}_{t-1}\big[\exp\big(Y_t-2\mathbb{E}_{t-1}[Y_t]\big)\big]\le 1. \tag{A.2}$$

Combined with the argument due to Freedman (1975) recalled in Appendix B.1 we derive

$$\mathbb{P}\left(\exists T\ge 1:\sum_{t=1}^T Y_t-2\sum_{t=1}^T\mathbb{E}_{t-1}[Y_t]>\log\big(\delta^{-1}\big)\right)\le\delta,\qquad 0<\delta<1. \tag{A.3}$$

Thus an union bound provides

$$\begin{aligned}\sum_{t=1}^T L_t(x_t)-\sum_{t=1}^T L_t(x)\le{}&\frac{3\lambda+2\gamma-\alpha}{2}\sum_{t=1}^T\mathbb{E}_{t-1}\big[\big(\nabla\ell_t(x_t)^T(x_t-x)\big)^2\big]\\&+\frac{1}{2\gamma}\log\big(\det(A_T)/\det(A_0)\big)+\frac{1}{2\gamma}\\&+\left((\lambda+\gamma)(G_{\psi_2}D)^2+\frac{2}{\lambda}\right)\log\big(\delta^{-1}\big).\end{aligned}$$

Choosing $3\lambda=\alpha-2\gamma>0$ since $\gamma<\alpha/2$ we conclude

$$\begin{aligned}\sum_{t=1}^T L_t(x_t)-\sum_{t=1}^T L_t(x)\le{}&\frac{1}{2\gamma}\log\big(\det(A_T)/\det(A_0)\big)+\frac{1}{2\gamma}\\&+\left(\frac{\alpha+\gamma}{3}(G_{\psi_2}D)^2+\frac{6}{\alpha-2\gamma}\right)\log\big(\delta^{-1}\big).\end{aligned}\tag{A.4}$$

From the initialization $A_0 = \frac{1}{(\gamma D)^2} I_d$, we obtain bound

$$\log\bigl(\det(A_T)/\det(A_0)\bigr) \le d\log\left(1 + (\gamma D)^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^2\right).$$

We apply the Poissonian exponential inequality from Propostion 3.2 on the second-order terms $0 \le Y_t = \|\nabla \ell_t(x_t)\|^2/(2G_{\psi_2}^2)$ and, combined with the argument due to Freedman (1975) and Condition **(H3)** ensuring $\mathbb{E}_{t-1}[Y_t] \le G_2^2/(2G_{\psi_2}^2)$, we obtain

$$\mathbb{P}\left(\exists T \ge 1 : \sum_{t=1}^T Y_t - TG_2^2/G_{\psi_2}^2 > \log\bigl(\delta^{-1}\bigr)\right) \le \delta, \qquad 0 < \delta < 1.$$

We derive that, with probability $1-\delta$, it holds

$$\log\bigl(\det(A_T)/\det(A_0)\bigr) \le d\log\bigl(1 + 2(\gamma D)^2\bigl(TG_2^2 + G_{\psi_2}\log\bigl(\delta^{-1}\bigr)\bigr)\bigr), \qquad T \ge 1.$$

The desired result follows from the specific choice of γ and a union bound.

A.4. Proof of Theorem 4.1

We keep the same notation and convention as in Section 4. In particular, inequalities involving vectors are coordinate-wise. With no loss of generality we assume that $\eta_{1,i} \neq 0$ for all $1 \le i \le K$. To prove the regret bound (4.2) we will show that

$$\begin{aligned}&\pi_1^T \exp(-\eta_T \widetilde{L}_T)\\ &\le \underbrace{\exp\left(\sum_{i=1}^K 1\bigl\{\max_{2\le t\le T} x_{t,i} > 1/4\bigr\}\log\bigl(1+(\eta_{1,i}M_{T,i})^2\bigr)\right)\left(\mathrm{e}+\frac{1}{2}\pi_1^T\log\bigl(\mathbb{1}+(\eta_1 M_T)^2 T\bigr)\right)}_{=:A_T}.\end{aligned} \tag{A.5}$$

From (A.5) we derive

$$-\eta_T\widetilde{L}_T = \eta_T\left(\sum_{t=1}^T\bigl(\pi_t^T\boldsymbol{\ell}_t\mathbb{1} - \boldsymbol{\ell}_t\bigr) - \sum_{t=1}^T \eta_{t-1}\bigl(\pi_t^T\boldsymbol{\ell}_t\mathbb{1} - \boldsymbol{\ell}_t\bigr)^2\right) \le \log\bigl(\pi_1^{-1}A_T\bigr)$$

so that

$$\sum_{t=1}^T \pi_t^T\boldsymbol{\ell}_t\mathbb{1} \le \sum_{t=1}^T \boldsymbol{\ell}_t + \sum_{t=1}^T \eta_{t-1}\bigl(\pi_t^T\boldsymbol{\ell}_t\mathbb{1} - \boldsymbol{\ell}_t\bigr)^2 + \frac{\log(\pi_1^{-1})}{\eta_T} + \frac{\log(A_T)}{\eta_T}.$$

Since $\eta_t^{-2} = \eta_{t-1}^{-2} + 2.2(\boldsymbol{\ell}_t - \pi_t^T\boldsymbol{\ell}_t\mathbb{1})^2$ we obtain by rearranging the sum

$$\sum_{t=1}^T \eta_{t-1}\bigl(\pi_t^T\boldsymbol{\ell}_t\mathbb{1} - \boldsymbol{\ell}_t\bigr)^2 = \frac{1}{2.2}\sum_{t=1}^T \eta_{t-1}\bigl(\eta_t^{-2} - \eta_{t-1}^{-2}\bigr) \le \frac{1}{2.2}\left(\sum_{t=1}^T \frac{\eta_{t-1}-\eta_t}{\eta_t^2} + \frac{1}{\eta_T}\right).$$

Thus we derive from a comparison sum-integral

$$\sum_{t=1}^{T} \frac{\eta_{t-1}-\eta_t}{\eta_t^2} \le \frac{1}{\eta_T} \qquad \Longrightarrow \qquad \sum_{t=1}^{T} \eta_{t-1}\big(\pi_t^T \boldsymbol{\ell}_t \mathbb{1} - \boldsymbol{\ell}_t\big)^2 \le \frac{1}{1.1\eta_T}.$$

The learning rate satisfying the relation

$$\begin{aligned} &\frac{1/1.1+\log(\pi_1^{-1})+\log(A_T)}{\eta_T} \\ &\qquad \le \big(1/1.1+\log\big(\pi_1^{-1}\big)+\log(A_T)\big)\sqrt{2.2\sum_{t=1}^{T}\big(\pi_t^T\boldsymbol{\ell}_t\mathbb{1}-\boldsymbol{\ell}_t\big)^2}, \end{aligned}$$

and the regret bound (4.2) follows from the expression of $\log(A_T)$.

It remains to prove the exponential inequality (A.5). We use the identity

$$\exp(-\eta_T\widetilde{L}_T) = \exp\big(\eta_T\big(\pi_T^T\boldsymbol{\ell}_T\mathbb{1}-\boldsymbol{\ell}_T\big) - \eta_T^2\big(\boldsymbol{\ell}_T-\pi_T^T\boldsymbol{\ell}_T\mathbb{1}\big)^2\big)\exp(-\eta_T\widetilde{L}_{T-1}).$$

To initiate the recursion, we use the basic inequality $x \le x^\alpha + e^{-1}(\alpha-1)/\alpha$ for $x \ge 0$ and $\alpha \ge 1$ with $x = \exp(-\eta_T\widetilde{L}_{T-1})$ and $\alpha = \eta_{T-1}/\eta_T$ so that

$$\exp(-\eta_T\widetilde{L}_{T-1}) \le \exp(-\eta_{T-1}\widetilde{L}_{T-1}) + e^{-1}\frac{\eta_{T-1}-\eta_T}{\eta_{T-1}}.$$

We obtain

$$\begin{aligned} \exp(-\eta_T\widetilde{L}_T) &\le \exp\big(\eta_T\big(\pi_T^T\boldsymbol{\ell}_T\mathbb{1}-\boldsymbol{\ell}_T\big) - \eta_T^2\big(\boldsymbol{\ell}_T-\pi_T^T\boldsymbol{\ell}_T\mathbb{1}\big)^2\big) \\ &\qquad \times\left(\exp(-\eta_{T-1}\widetilde{L}_{T-1}) + e^{-1}\frac{\eta_{T-1}-\eta_T}{\eta_{T-1}}\right). \end{aligned}$$

Then we use the expression

$$\eta_T = \frac{\eta_{T-1}}{\sqrt{1+2.2\eta_{T-1}^2(\boldsymbol{\ell}_T-\pi_T^T\boldsymbol{\ell}_T\mathbb{1})^2}},$$

and the notation $x_T = \eta_{T-1}(\boldsymbol{\ell}_T - \pi_T^T\boldsymbol{\ell}_T\mathbb{1})$ to derive

$$\begin{aligned} &\exp(-\eta_T\widetilde{L}_T) \\ &\quad \le \exp\left(-\frac{x_T}{\sqrt{1+2.2x_T^2}} - \frac{x_T^2}{1+2.2x_T^2}\right)\left(\exp(-\eta_{T-1}\widetilde{L}_{T-1}) + \frac{\eta_{T-1}-\eta_T}{\eta_{T-1}}\right). \end{aligned}$$

We use different bounds over the function $\varphi :\ y \in \mathbb{R} \mapsto \exp(-\frac{y}{\sqrt{1+2.2y^2}} - \frac{y^2}{1+2.2y^2})$: $\varphi(y) \le e/2$, $\varphi(y) \le 1 - \frac{y}{\sqrt{1+2.2y^2}}$ for any $y \in \mathbb{R}$ and $\varphi(y) \le 1-y$ if $y \le 1/4$. Distinguishing whether x_T is larger or not than $1/4$, we deduce

$$\exp(-\eta_T\widetilde{L}_T) \le \big(\mathbb{1} - \eta_{T-1}\big(\boldsymbol{\ell}_T - \pi_T^T\boldsymbol{\ell}_T\mathbb{1}\big)\big)\exp(-\eta_{T-1}\widetilde{L}_{T-1})\mathbb{1}\{x_T \le 1/4\}$$

$$+ \big(\mathbb{1} - \eta_T\big(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}\big)\big) \exp(-\eta_{T-1}\widetilde{L}_{T-1})\mathbb{1}\{x_T > 1/4\}$$
$$+ 1/2 \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}}.$$

Using the relations $\eta_{T-1}/\eta_T \geq \mathbb{1}$ and $1 - \frac{y}{\sqrt{1+y^2}} > 0$, $y \in \mathbb{R}$ we upper bound the second term by

$$\begin{aligned}
&\frac{\eta_{T-1}}{\eta_T}\big(\mathbb{1} - \eta_T\big(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}\big)\big) \exp(-\eta_{T-1}\widetilde{L}_{T-1})\mathbb{1}\{x_T > 1/4\} \\
&\quad = \left(\frac{\eta_{T-1}}{\eta_T} - \eta_{T-1}\big(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}\big)\right) \exp(-\eta_{T-1}\widetilde{L}_{T-1})\mathbb{1}\{x_T > 1/4\}.
\end{aligned}$$

Combining it with the previous bound we achieve

$$\begin{aligned}
\exp(-\eta_T\widetilde{L}_T) \leq{}& \left(\frac{\eta_{T-1}}{\eta_T}\right)^{\mathbb{1}\{x_T > 1/4\}} \exp(-\eta_{T-1}\widetilde{L}_{T-1}) \\
&- \eta_{T-1}\big(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}\big) \exp(-\eta_{T-1}\widetilde{L}_{T-1}) + 1/2 \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}}.
\end{aligned}$$

The second inequality is obtained. We have

$$\begin{aligned}
\pi_1^T \exp(-\eta_T\widetilde{L}_T) \leq{}& \left\| \left(\frac{\eta_{T-1}}{\eta_T}\right)^{\mathbb{1}\{x_T > 1/4\}} \right\|_\infty \pi_1^T \exp(-\eta_{T-1}\widetilde{L}_{T-1}) \\
&- \big(\pi_1 \eta_{T-1} \exp(-\eta_{T-1}\widetilde{L}_{T-1})\big)^T \big(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}\big) \\
&+ 1/2 \pi_1^T \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}}.
\end{aligned}$$

We recognize the weights

$$\pi_1 \eta_{T-1} \exp(-\eta_{T-1}\widetilde{L}_{T-1}) = \pi_T \big(\pi_1^T \eta_{T-1} \exp(-\eta_{T-1}\widetilde{L}_{T-1})\big)$$

and the second term in the upper bound is proportional to $\pi_T^T(\boldsymbol{\ell}_T - \pi_T^T \boldsymbol{\ell}_T \mathbb{1}) = 0$ and thus vanishes. We obtain

$$\begin{aligned}
&\pi_1^T \exp(-\eta_T\widetilde{L}_T) \\
&\quad \leq \left\| \left(\frac{\eta_{T-1}}{\eta_T}\right)^{\mathbb{1}\{x_T > 1/4\}} \right\|_\infty \pi_1^T \exp(-\eta_{T-1}\widetilde{L}_{T-1}) + 1/2 \pi_1^T \frac{\eta_{T-1} - \eta_T}{\eta_T},
\end{aligned}$$

and a recursive argument yields

$$\begin{aligned}
\pi_1^T \exp(-\eta_T\widetilde{L}_T) \leq \exp\left(\sum_{t=2}^T \left\| \log\left(\frac{\eta_{t-1}}{\eta_t}\right) \mathbb{1}\{x_t > 1/4\} \right\|_\infty \right) \Bigg(& \pi_1^T \exp(-\eta_1\widetilde{L}_1) \\
&+ 1/2 \sum_{t=2}^T \pi_1^T \frac{\eta_{T-1} - \eta_T}{\eta_{T-1}} \Bigg).
\end{aligned}$$

We bound the exponent term such as

$$
\begin{aligned}
&\sum_{t=2}^{T}\Big\|\log\Big(\frac{\eta_{t-1}}{\eta_t}\Big)\mathbb{1}\{x_t>1/4\}\Big\|_\infty \\
&\le \sum_{i=1}^{K}\sum_{t=2}^{T}\log\Big(\frac{\eta_{t-1,i}}{\eta_{t,i}}\Big)\mathbb{1}\{x_{t,i}>1/4\} \\
&\le \sum_{i=1}^{K}1\Big\{\max_{2\le t\le T}x_{t,i}>1/4\Big\}\Big(\sum_{t=2}^{T}\log\Big(\frac{\eta_{t-1,i}}{\eta_{t,i}}\Big)\mathbb{1}\{x_{t,i}>1/4\}\Big) \\
&\le \sum_{i=1}^{K}1\Big\{\max_{2\le t\le T}x_{t,i}>1/4\Big\}\Big(\log\Big(\frac{\eta_{1,i}}{\eta_{T-1,i}}\Big)+\log\Big(\frac{\eta_{T-1,i}}{\eta_{T,i}}\Big)\Big),
\end{aligned}
$$

assuming with no loss of generality that if $\max_{2\le t\le T}x_{t,i}>1/4$ then it happens for the last iterate $x_{T,i}=\eta_{T-1,i}\ell_{T,i}>1/4$. Notice also that $x_{T,i}>1/4$ implies that $\eta_{T-1,i}^{-1}\le M_{T,i}/4$. Combined with

$$
\frac{\eta_{T-1,i}}{\eta_{T,i}}=\sqrt{1+2.2\eta_{T-1,i}^2\big(\boldsymbol{\ell}_{T,i}-\pi_T^T\boldsymbol{\ell}_T\big)^2}\le\sqrt{1+2.2\eta_{1,i}^2M_{T,i}^2},
$$

we obtain

$$
\begin{aligned}
&\sum_{t=2}^{T}\Big\|\log\Big(\frac{\eta_{t-1}}{\eta_t}\Big)\mathbb{1}\{x_t>1/4\}\Big\|_\infty \\
&\le \sum_{i=1}^{K}1\Big\{\max_{2\le t\le T}x_{t,i}>1/4\Big\}\Big(\log(\eta_{1,i}M_{T,i}/4)+\frac{1}{2}\log\big(1+2.2\eta_{1,i}^2M_{T,i}^2\big)\Big) \\
&\le \sum_{i=1}^{K}1\Big\{\max_{2\le t\le T}x_{t,i}>1/4\Big\}\log\big(1+\eta_{1,i}^2M_{T,i}^2\big).
\end{aligned}
$$

We have $\exp(-\eta_1\widetilde{L}_1)\le\exp(\mathbb{1})$ using the relation $|\eta_1\widetilde{L}_1|=\mathbb{1}$ and the comparison sum-integral

$$
\sum_{t=2}^{T}\frac{\eta_{t-1}-\eta_t}{\eta_{t-1}}\le\log(\eta_1/\eta_T)=\frac{1}{2}\log\big(\mathbb{1}+(\eta_1M_T)^2T\big),
$$

we achieve (A.5).

A.5. Proof of Theorem 4.2

From the regret bound (4.2), keeping the notation of (A.5) and applying Young's inequality, we infer that for any $\eta>0$

$$
\sum_{t=1}^{T}\pi_t^T\boldsymbol{\ell}_t-\sum_{t=1}^{T}\boldsymbol{\ell}_{t,i}\le\frac{\eta}{2}\sum_{t=1}^{T}\big(\pi_t^T\boldsymbol{\ell}_t-\boldsymbol{\ell}_{t,i}\big)^2+\frac{(1/1.1+\log(\pi_1^{-1})+\log(A_T))^2}{2\eta}.
$$

Plugging this bound into (4.1) and identifying $\boldsymbol{\ell}_t = \mathbf{x}_t^T \nabla \ell_t(\mathbf{x}_t \pi_t)$ and $\widehat{x}_t = \mathbf{x}_t \pi_t$ we obtain

$$\sum_{t=1}^T L_t(\widehat{x}_t) - \sum_{t=1}^T L_t\big(x_t^{(i)}\big) \le \frac{\lambda+\eta}{2} \sum_{t=1}^T \nabla \ell_t(\widehat{x}_t)^T \big(\widehat{x}_t - x_t^{(i)}\big)^2 + \frac{\lambda-\alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1}\big[(\nabla \ell_t(\widehat{x}_t)^T \big(\widehat{x}_t - x_t^{(i)}\big)^2\big] + \frac{(1/1.1 + \log(\pi_1^{-1}) + \log(A_T))^2}{2\eta} + \frac{2}{\lambda}\log\big(\delta^{-1}\big).$$

Applying once again the Poissonian inequality (A.3), using that the diameter of the simplex satisfies is less than 1, we derive that with probability $1-\delta$

$$\sum_{t=1}^T (\nabla \ell_t(\widehat{x}_t)^T \big(\widehat{x}_t - x_t^{(i)}\big)^2 \le 2 \sum_{t=1}^T \mathbb{E}_{t-1}\big[(\nabla \ell_t(\widehat{x}_t)^T \big(\widehat{x}_t - x_t^{(i)}\big)^2\big] + 2(G_{\psi_2} D)^2 \log\big(\delta^{-1}\big).$$

Then we obtain

$$\sum_{t=1}^T L_t(\widehat{x}_t) - \sum_{t=1}^T L_t\big(x_t^{(i)}\big) \le \frac{3\lambda + 2\eta - \alpha}{2} \sum_{t=1}^T \mathbb{E}_{t-1}\big[(\nabla \ell_t(\widehat{x}_t)^T \big(\widehat{x}_t - x_t^{(i)}\big)^2\big] + \frac{(1/1.1 + \log(\pi_1^{-1}) + \log(A_T))^2}{2\eta} + \left((\lambda+\eta)(G_{\psi_2} D)^2 + \frac{2}{\lambda}\right)\log\big(\delta^{-1}\big).$$

Thus choosing $\lambda = \eta = \alpha/3$ and introducing $\nabla \ell_t(\hat{x}_t)$ for bounding roughly $\log(A_T)$, we obtain

$$\begin{aligned}&\sum_{t=1}^T L_t(\widehat{x}_t) - \sum_{t=1}^T L_t\big(x_t^{(i)}\big)\\ &\quad\le \frac{3}{\alpha}\Bigg(K \log\left(1 + \frac{\max_{1\le t\le T} \|\nabla \ell_t(\hat{x}_t)\|^2}{\underline{m}^2}\right)\\ &\qquad + \log\left(\mathrm{e} + \log\left(1 + \frac{\max_{1\le t\le T} \|\nabla \ell_t(\hat{x}_t)\|^2}{\underline{m}^2} T\right)\right) + 1/1.1 + \log\big(\pi_1^{-1}\big)\Bigg)^2\\ &\qquad + \left(\frac{2\alpha}{3}(G_{\psi_2} D)^2 + \frac{6}{\alpha}\right)\log\big(\delta^{-1}\big).\end{aligned}$$

From the proof Proposition 2.3 on the second-order terms $0 \le Y_t = \|\nabla \ell_t(\widehat{x}_t)\|^2 / (2G_{\psi_2}^2)$ we obtain

$$\mathbb{E}_{t-1}\big[\exp(Y_t)\big] \le 1 + 2\mathbb{E}_{t-1}\big[Y_t^2\big] \le 1 + (G_2/G_{\psi_2})^2.$$

Thus, for any $x > 0$ we have

$$\mathbb{P}\Big(\max_{1\le t\le T} Y_t > x\Big) \le \mathbb{E}\big[\exp(\max Y_t)\big]\exp(-x)$$

$$\leq \sum_{t=1}^{T} \mathbb{E}\big[\exp(Y_t)\big] \exp(-x) \leq T\big(1 + (G_2/G_{\psi_2})^2\big) \exp(-x),$$

and with probability $1-\delta$ it holds

$$\max_{1\leq t\leq T} Y_t \leq \log(T) + \log\big(1 + (G_2/G_{\psi_2})^2\big) + \log\big(\delta^{-1}\big).$$

Finally, we obtain the desired result using a union bound.

A.6. Proof of Proposition 5.2

We denote

$$\overline{y}^M_{t-1,t-p} = \big((y_{t-1} \wedge M/2) \vee (-M/2), \dots, (y_{t-p} \wedge M/2) \vee (-M/2)\big) \in \mathbb{R}^p.$$

Let $P_x = \mathcal{N}(\widehat{m}^{(p)}_t(x), \sigma^2)$ then $\ell_t(x) = (y_t - \widehat{m}^{(p)}_t(x))^2/(2\sigma^2)$ (plus constant) and

$$\begin{aligned}
\mathbb{E}_{t-1}\big[\nabla \ell_t(x) \nabla \ell_t(x)^T\big] &= \frac{\mathbb{E}_{t-1}[(y_t - \widehat{m}^{(p)}_t(x))^2]}{\sigma^4} \overline{y}^M_{t-1,t-p}\big(\overline{y}^M_{t-1,t-p}\big)^T,\\
\mathbb{E}_{t-1}\big[\nabla^2 \ell_t(x)\big] &= \frac{1}{\sigma^2} \overline{y}^M_{t-1,t-p}\big(\overline{y}^M_{t-1,t-p}\big)^T.
\end{aligned}$$

Because the second derivatives do not depend on x a Taylor expansion provides

$$\begin{aligned}
L_t(y) &= L_t(x) + \nabla L_t(y)^T(y-x) - \frac{1}{\sigma^2}(y-x)^T \overline{y}^M_{t-1,t-p}\big(\overline{y}^M_{t-1,t-p}\big)^T (y-x)\\
&= L_t(x) + \nabla L_t(y)^T(y-x) - \frac{1}{\sigma^2}\big(\widehat{m}^{(p)}_t(y) - \widehat{m}^{(p)}_t(x)\big)^2\\
&\leq L_t(x) + \nabla L_t(y)^T(y-x) - \frac{\mathbb{E}_{t-1}[(y_t - \widehat{m}^{(p)}_t(x))^2]}{\sigma^2(\overline{\sigma}^2 + M^2)}\big(\widehat{m}^{(p)}_t(y) - \widehat{m}^{(p)}_t(x)\big)^2\\
&\leq L_t(x) + \nabla L_t(y)^T(y-x) - \frac{\sigma^2}{\overline{\sigma}^2 + M^2}(y-x)^T \mathbb{E}_{t-1}\big[\nabla \ell_t(x) \nabla \ell_t(x)^T\big](y-x).
\end{aligned}$$

The first inequality comes from the relations

$$\mathbb{E}_{t-1}\big[\big(y_t - \widehat{m}^{(p)}_t(x)\big)^2\big] = \mathbb{E}_{t-1}\big[(y_t - m_t)^2\big] + \big(m_t - \widehat{m}^{(p)}_t(x)\big)^2 \leq \overline{\sigma}^2 + M^2.$$

Thus Condition **(H2)** is satisfied with $\alpha = \sigma^2/(\overline{\sigma}^2 + M^2)$.

Applying Theorem 3.1, the ONS achieves the stochastic regret against every $x \in B_1(1)$ (satisfying $\|x\| \leq \sqrt{p}$)

$$\begin{aligned}
&\sum_{t=1}^{T} L_t(x_t) - \sum_{t=1}^{T} L_t(x)\\
&\quad \leq O\bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} p \log T + \bigg(\frac{\overline{\sigma}^2 + M^2}{\sigma^2} + \frac{\sigma^2}{\overline{\sigma}^2 + M^2} p\, G^2_{\psi_2}\bigg) \log\big(\delta^{-1}\big)\bigg)
\end{aligned}$$

with high probability. Since the risk satisfies the relation

$$L_t(x) = \frac{1}{2}\left(\log(2\pi) + \log\bigl(\sigma^2\bigr) + \frac{(m_t - \widehat{m}_t^{(p)}(x))^2 + \sigma_t^2}{\sigma^2}\right)$$
$$= KL\bigl(P_t, \mathcal{N}\bigl(m_t^{(p)}(x), \sigma^2\bigr)\bigr) + cst.,$$

we obtain the desired result.

A.7. Proof of Proposition 5.3

We denote

$$\overline{y}^{2,\overline{\sigma}}_{t-1,t-q} = \bigl(y_{t-1}^2 \wedge \overline{\sigma}^2, \dots, y_{t-q}^2 \wedge \overline{\sigma}^2\bigr) \in \mathbb{R}^q.$$

Let $P_x = \mathcal{N}(0, \widehat{\sigma}_t^{2,(q)}(x))$ then $\ell_t(x) = (\log(\widehat{\sigma}_t^{2,(q)}(x)) + y_t^2/\widehat{\sigma}_t^{2,(q)}(x))/2$ and

$$\mathbb{E}_{t-1}\bigl[\nabla\ell_t(x)\nabla\ell_t(x)^T\bigr] = \frac{\mathbb{E}_{t-1}[(y_t^2 - \widehat{\sigma}_t^{2,(q)}(x))^2]}{2(\widehat{\sigma}_t^{2,(q)}(x))^4}\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigl(\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigr)^T,$$
$$\mathbb{E}_{t-1}\bigl[\nabla^2\ell_t(x)\bigr] = \frac{2\sigma_t^2 - \widehat{\sigma}_t^{2,(q)}(x)}{2(\widehat{\sigma}_t^{2,(q)}(x))^3}\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigl(\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigr)^T.$$

Because $1/2 \le \sigma_t^2/\widehat{\sigma}_t^{2,(q)}(x) \le 2$ under our assumptions, the second derivatives are decreasing in $\widehat{\sigma}_t^{2,(q)}(x)$ and thus

$$\mathbb{E}_{t-1}\bigl[\nabla^2\ell_t(x)\bigr] \succeq \frac{2\sigma_t^2 - \overline{\sigma}^2}{2\overline{\sigma}^6}\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigl(\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigr)^T \succeq \frac{c-1}{2\overline{\sigma}^4}\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigl(\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigr)^T.$$

Combining this lower bound with a Taylor expansion, we obtain

$$\begin{aligned}
L_t(y) &\le L_t(x) + \nabla L_t(y)^T(y-x) - \frac{c-1}{2\overline{\sigma}^4}(y-x)^T\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigl(\overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\bigr)^T(y-x)\\
&\le L_t(x) + \nabla L_t(y)^T(y-x)\\
&\quad - \frac{(c-1)\mathbb{E}_{t-1}[(y_t^2 - \widehat{\sigma}_t^{2,(q)}(x))^2]}{2\overline{\sigma}^4(3\overline{\sigma}^4 + \overline{\sigma}^4)}\bigl(\widehat{\sigma}_t^{2,(q)}(y) - \widehat{\sigma}_t^{2,(q)}(x)\bigr)^2\\
&\le L_t(x) + \nabla L_t(y)^T(y-x)\\
&\quad - \frac{(c-1)(c\overline{\sigma}^2/2)^4}{4\overline{\sigma}^8}(y-x)^T\mathbb{E}_{t-1}\bigl[\nabla\ell_t(x)\nabla\ell_t(x)^T\bigr](y-x).
\end{aligned}$$

The second inequality comes from the relations

$$\mathbb{E}_{t-1}\bigl[\bigl(y_t^2 - \widehat{\sigma}_t^{2,(q)}(x)\bigr)^2\bigr] = \mathbb{E}_{t-1}\bigl[\bigl(y_t^2 - \sigma_t^2\bigr)^2\bigr] + \bigl(\sigma_t^2 - \widehat{\sigma}_t^{2,(q)}(x)\bigr)^2 \le 3\overline{\sigma}^4 + \overline{\sigma}^2.$$

Thus Condition **(H2)** is satisfied with $\alpha = (c-1)c^4 2^{-6}$.

Applying Theorem 3.1, the ONS achieves the stochastic regret with high probability against every $x \in \mathcal{K}$ (satisfying $\|x\| \le \sqrt{q}$)

$$\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \le O\bigl(q\log T + \bigl(1 + qG_{\psi_2}^2\bigr)\log\bigl(\delta^{-1}\bigr)\bigr).$$

We conclude the proof by identifying the KL divergence with L_t up to additive constants.

Proof of Proposition 5.3

We use similar arguments than in the proofs of Propositions 5.2 and 5.3, keeping the same notation with

$$\ell_t(x) = \frac{1}{2}\left(\log\big(\widehat{\sigma}_t^{2,(q)}(x_{p+1:p+q})\big) + \frac{(y_t - \widehat{m}_t^{(p)}(x_{1:p}))^2}{\widehat{\sigma}_t^{2,(q)}(x_{p+1:p+q})}\right).$$

Adapting previous computations, we similarly obtain a lower bound on the second derivatives

$$\mathbb{E}_{t-1}\big[\nabla^2 \ell_t(x)\big] \succeq \frac{c-1}{2\overline{\sigma}^4}\big(\overline{y}^M_{t-1,t-p}, \overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\big)\big(\overline{y}^M_{t-1,t-p}, \overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\big)^T.$$

We can also upper bound the first derivatives to obtain

$$\begin{aligned}\mathbb{E}_{t-1}\big[\nabla \ell_t(x)\nabla \ell_t(x)^T\big] \preceq \frac{2^4}{2(c\overline{\sigma}^2)^4}&\mathbb{E}_{t-1}\Big[\Big(\widehat{\sigma}_t^{2,(q)} - \big(y_t - \widehat{m}_t^{(p)}(x_{1:p})\big)^2\Big)^2\Big] \\ &\times \big(\overline{y}^M_{t-1,t-p}, \overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\big)\big(\overline{y}^M_{t-1,t-p}, \overline{y}^{2,\overline{\sigma}}_{t-1,t-q}\big)^T.\end{aligned}$$

Under our assumptions, we roughly estimate

$$\begin{aligned}&\mathbb{E}_{t-1}\Big[\Big(\widehat{\sigma}_t^{2,(q)} - \big(y_t - \widehat{m}_t^{(p)}(x_{1:p})\big)^2\Big)^2\Big] \\ &\le 2\Big(\big(\widehat{\sigma}_t^{2,(q)}\big)^2 + \mathbb{E}_{t-1}\Big[\big(y_t - \widehat{m}_t^{(p)}(x_{1:p})\big)^4\Big]\Big) \\ &\le 2\Big(\overline{\sigma}^4 + 2\Big(\mathbb{E}_{t-1}\big[(y_t - m_t)^4\big] + \mathbb{E}_{t-1}\Big[\big(m_t - \widehat{m}_t^{(p)}(x_{1:p})\big)^4\Big]\Big)\Big) \\ &\le 18\overline{\sigma}^4.\end{aligned}$$

Thus Condition **(H2)** is satisfied with $\alpha = (c-1)c^4 3^{-2} 2^{-5}$.

Applying Theorem 3.1, the ONS achieves the stochastic regret with high probability against every $x \in \mathcal{K}$ (satisfying $\|x\| \le D = \sqrt{p+q}$)

$$\sum_{t=1}^T L_t(x_t) - \sum_{t=1}^T L_t(x) \le O\Big((p+q)\log T + \big(1 + (p+q)G^2_{\psi_2}\big)\log\big(\delta^{-1}\big)\Big).$$

We conclude the proof by identifying the KL divergence with L_t up to additive constants.

Appendix B: Auxiliary results

B.1. The stopping time argument of Ville (1939) and Freedman (1975)

We recall the argument of Ville (1939) and Freedman (1975) as we apply it several times in the proofs of 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 $\mathbb{E}[\exp(Z_t) \mid \mathcal{F}_{t-1}] \le 1$. Then we have

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

for any $0 < \delta < 1$ by applying the following lemma.

Lemma B.1. *If M_t is adapted to $\mathcal{F}_t$, $M_0 = 1$ a.s. and $\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] \le M_{t-1}$ a.s., $t \ge 1$, then, for any $0 < \delta < 1$, it holds*

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

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

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

B.2. SOCO analysis of the OGD algorithm

In this section we work under **(H1)** and **(H2)** with $\alpha = 0$. Proposition 3.1 holds, $\lambda > 0 = \alpha$ and the compensator term in Proposition 3.1 is positive. In this section we assume that the gradients are bounded by $G < \infty$. A slow rate stochastic regret bound $O(GD\sqrt{T})$ is expected and the surrogate loss in Proposition 3.1 is useless. The classical Online Gradient Descent (OGD) of Zinkevich (2003)

$$x_{t+1} = \arg\min_{x \in \mathcal{K}} \left\| x - \frac{D}{G\sqrt{t}} \nabla \ell_t(x_t) \right\| \qquad \text{starting from } x_0 \in \mathcal{K},$$

satisfies the following linearized regret bound in any SOCO problem, see the proof in Hazan (2016) that does not use any convex assumption,

$$\sum_{t=1}^{T} \nabla \ell_t(x_t)^T (x_t - x) \leq \frac{3}{2} DG\sqrt{T}.$$

Under **(H1)** we easily bound a.s. both extra quadratic terms in Proposition 3.1 with the same quantity $\lambda/2\, G^2 D^2 T$. Choosing $\lambda = \sqrt{2\log(\delta^{-1})}/(GD\sqrt{T})$ we immediately obtain a new slow rate stochastic regret bound for the OGD valid in any SOCO problem:

Theorem B.1. *Assume that **(H1)** holds and that $\sup_{x\in\mathcal{K}} \|\nabla \ell_t(x)\| \le G$ a.s., $t \ge 1$. The OGD algorithm satisfies with probability $1 - \delta$ the stochastic regret bound*

$$\sum_{t=1}^{T} L_t(x_t) - \sum_{t=1}^{T} L_t(x) \leq \left(\frac{3}{2} + 2\sqrt{2\log\bigl(\delta^{-1}\bigr)}\right) DG\sqrt{T}$$

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

This simple extension of the usual iid setting to any stochastic adversarial setting could be obtained by classical arguments such as Azuma's inequality used in Chapter 9 of Hazan (2016). It relies on the martingale $\sum_{t=1}^{T}(\nabla L_t(x_t) - \nabla \ell_t(x_t))^T(x_t - x^*)$ and the gradient trick on L_t to remove the assumption of convexity on the losses ℓ_t.

Acknowledgments

The author would like to thank the Institut für Statistik und Operations Research at Vienna University for its kind hospitality. He is also grateful to an anonymous Referee for helpful comments and suggestions on an earlier version of the paper.

References

E. Adjakossa, Y. Goude, and O. Wintenberger. Kalman recursions aggregated online. *Statistical Papers*, pages 1–36, 2023.

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. MR3008166

O. Anava, E. Hazan, S. Mannor, and O. Shamir. Online learning for time series prediction. In *Conference on Learning Theory*, pages 172–184. PMLR, 2013.

J.-Y. Audibert. Progressive mixture rules are deviation suboptimal. *Advances in Neural Information Processing Systems*, 20, 2007.

B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *Annals of Applied Probability*, 18:1848–1869, 2008. MR2462551

G. Biau and B. Patra. Sequential quantile prediction of time series. *IEEE Transactions on Information Theory*, 57(3):1664–1674, 2011. MR2815841

P. J. Brockwell and R. A. Davis. *Time series: theory and methods*. Springer Science & Business Media, 2009. MR1093459

N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006. MR2409394

N. Cesa-Bianchi, Y. Mansour, and G. Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66:321–352, 2007. MR2203264

J. de Vilmarest and N. Werge. An adaptive volatility method for probabilistic forecasting and its application to the m6 financial forecasting competition. *arXiv preprint arXiv:2303.01855*, 2023.

C. Francq and J.-M. Zakoïan. Inconsistency of the mle and inference based on weighted ls for larch models. *Journal of Econometrics*, 159(1):151–165, 2010. MR2720849

C. Francq and J.-M. Zakoian. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons, 2019. MR3186556

D. A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, pages 100–118, 1975. MR0380971

P. Gaillard and O. Wintenberger. Efficient online algorithms for fast-rate regret bounds under sparsity. *Advances in Neural Information Processing Systems*, 31, 2018.

P. Gaillard, Y. Goude, L. Plagne, T. Dubois, and B. Thieurmel. *opera: Online Prediction by Expert Aggregation*, 2021. URL `https://CRAN.R-project.org/package=opera`. R package version 1.2.0.

C. Giraud, F. Roueff, and A. Sanchez-Perez. Aggregation of predictors for non stationary sub-linear processes and online adaptive forecasting of time varying autoregressive processes. *Annals of Statistics*, 43(6):2412–2450, 2015. MR3405599

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007. MR2345548

M. Haddouche, B. Guedj, and O. Wintenberger. Optimistic dynamic regret bounds. *arXiv preprint arXiv:2301.07530*, 2023.

E. Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

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. JMLR Workshop and Conference Proceedings, 2011.

A. Henzi and J. F. Ziegel. Valid sequential inference on probability forecast performance. *Biometrika*, 109(3):647–663, 2022. MR4472840

W. M. Koolen, P. Grünwald, and T. Van Erven. Combining adversarial guarantees and stochastic fast rates in online learning. *Advances in Neural Information Processing Systems*, 29, 2016.

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.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton University Press, 2015. MR3445371

N. Mehta. Fast rates with high probability in exp-concave statistical learning. In *Artificial Intelligence and Statistics*, pages 1085–1093. PMLR, 2017.

Z. Mhammedi, W. M. Koolen, and T. Van Erven. Lipschitz adaptivity with multiple learning rates in online learning. In *Conference on Learning Theory*, pages 2490–2511. PMLR, 2019.

L. Orseau and M. Hutter. Isotuning with applications to scale-free online learning. *arXiv preprint arXiv:2112.14586*, 2021.

A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, 2011. MR2745881

A. Ramdas, P. Grünwald, V. Vovk, and G. Shafer. Game-theoretic statistics and safe anytime-valid inference. *arXiv preprint arXiv:2210.01948*, 2022. MR4665027

P. Rigollet, A. Tsybakov, and A. Juditsky. Learning by mirror averaging. *Annals of Statistics*, 36(6):2183, 2008. MR2458184

G. Shafer and V. Vovk. *Game-theoretic foundations for probability and finance*, volume 455. John Wiley & Sons, 2019. MR1852450

G. Shafer et al. Testing by betting: A strategy for statistical and scientific communication. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 184(2):407–431, 2021. MR4255905

I. Steinwart and A. Christmann. Estimating conditional quantiles with the help of the pinball loss. *Bernoulli*, 17(1):211–225, 2011. MR2797989

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.

S. van de Geer and J. Lederer. The bernstein–orlicz norm and deviation inequalities. *Probability Theory and Related Fields*, 157(1-2):225–250, 2013. MR3101846

T. van Erven, W. M. Koolen, and D. van der Hoeven. Metagrad: Adaptation using multiple learning rates in online learning. *Journal of Machine Learning Research*, 22(161):1–61, 2021. MR4318517

J. Ville. *Etude critique de la notion de collectif*. Gauthier-Villars, 1939. MR3533075

V. V. V'yugin and V. G. Trunov. Online learning with continuous ranked probability score. In *Conformal and Probabilistic Prediction and Applications*, pages 163–177. PMLR, 2019.

I. Waudby-Smith and A. Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2023.

N. Werge and O. Wintenberger. Adavol: An adaptive recursive volatility prediction method. *Econometrics and Statistics*, 23:19–35, 2022.

O. Wintenberger. Optimal learning with bernstein online aggregation. *Machine Learning*, 106:119–141, 2017. MR3592152

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 928–936, 2003.