# A new look at shifting regret

Nicolò Cesa-Bianchi, Pierre Gaillard, Gabor Lugosi, Gilles Stoltz

▶ **To cite this version:**

Nicolò Cesa-Bianchi, Pierre Gaillard, Gabor Lugosi, Gilles Stoltz. A new look at shifting regret. 2012. hal-00670514v1

**HAL Id: hal-00670514**
**https://hal.science/hal-00670514v1**

Preprint submitted on 15 Feb 2012 (v1), last revised 27 Sep 2012 (v2)

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

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

JMLR: Workshop and Conference Proceedings vol (2012) 1–18

# A New Look at Shifting Regret

**Nicolò Cesa-Bianchi** NICOLO.CESA-BIANCHI@UNIMI.IT
*DSI, Università degli Studi di Milano*

**Pierre Gaillard** PIERRE.GAILLARD@ENS.FR
*Ecole Normale Supérieure*[*]*, Paris*

**Gábor Lugosi** GABOR.LUGOSI@UPF.EDU
*ICREA & Universitat Pompeu Fabra*

**Gilles Stoltz** GILLES.STOLTZ@ENS.FR
*Ecole Normale Supérieure*[†]*, Paris & HEC Paris, Jouy-en-Josas, France*

## Abstract

We investigate extensions of well-known online learning algorithms such as fixed-share of Herbster and Warmuth (1998) or the methods proposed by Bousquet and Warmuth (2002). These algorithms use weight sharing schemes to perform as well as the best sequence of experts with a limited number of changes. Here we show, with a common, general, and simpler analysis, that weight sharing in fact achieves much more than what it was designed for. We use it to simultaneously prove new shifting regret bounds for online convex optimization on the simplex in terms of the total variation distance as well as new bounds for the related setting of adaptive regret. Finally, we exhibit the first logarithmic shifting bounds for exp-concave loss functions on the simplex.

**Keywords:** Prediction with expert advice, online convex optimization, tracking the best expert, shifting experts.

## 1. Introduction

Online convex optimization is a sequential prediction paradigm in which, at each time step, the learner chooses an element from a fixed convex set $\mathcal{S}$ and then is given access to a convex loss function defined on the same set. The value of the function on the chosen element is the learner's loss. Many problems such as prediction with expert advice, sequential investment, and online regression/classification can be viewed as special cases of this general framework. Online learning algorithms are designed to minimize the regret. The standard notion of regret is the difference between the learner's cumulative loss and the cumulative loss of the single best element in $\mathcal{S}$. A much harder criterion to minimize is shifting regret, which is defined as the difference between the learner's cumulative loss and the cumulative loss of an arbitrary sequence of elements in $\mathcal{S}$. Shifting regret bounds are typically expressed in terms of the *shift*, a notion of regularity measuring the length of the trajectory in $\mathcal{S}$ described by the comparison sequence (i.e., the sequence of elements against which the regret is evaluated).

In online convex optimization, shifting regret bounds for convex subsets $\mathcal{S} \subseteq \mathbb{R}^d$ are obtained for the online mirror descent (or follow-the-regularized-leader) algorithm. In this case the shift is

[†]. CNRS – Ecole normale supérieure, Paris – INRIA, within the project-team CLASSIC
[†]. CNRS – Ecole normale supérieure, Paris – INRIA, within the project-team CLASSIC

© 2012 .

typically computed in terms of the $p$-norm of the difference of consecutive elements in the comparison sequence —see Herbster and Warmuth (2001) and Cesa-Bianchi and Lugosi (2006). In this paper we focus on the important special case when $\mathcal{S}$ is the simplex, and investigate the online mirror descent with entropic regularizers. This family includes popular algorithms such as exponentially weighted average (EWA), Winnow, and exponentiated gradient. Proving general shifting bounds in this case is difficult due to the behavior of the regularizer at the boundary of the simplex. Herbster and Warmuth (2001) show shifting bounds for mirror descent with entropic regularizers using a 1-norm to measure the shift. In order to keep mirror descent from choosing points too close to the simplex boundary, they use a complex dynamic projection technique. When the comparison sequence is restricted to the corners of the simplex (which is the setting of prediction with expert advice), then the shift is naturally defined to be the number times the trajectory moves to a different corner. This problem is often called "tracking the best expert" —see, e.g., Herbster and Warmuth (1998); Vovk (1999); Herbster and Warmuth (2001); Bousquet and Warmuth (2002); György et al. (2005), and it is well known that EWA with weight sharing, which corresponds to the fixed-share algorithm of Herbster and Warmuth (1998), achieves a good shifting bound in this setting. Bousquet and Warmuth (2002) introduce a generalization of the fixed-share algorithm, and prove various shifting bounds for any trajectory in the simplex. However, their bounds are expressed using a quantity that corresponds to a proper shift only for trajectories on the simplex corners.

Our analysis unifies, generalizes (and simplifies) the previously quite different proof techniques and algorithms used in Herbster and Warmuth (1998) and Bousquet and Warmuth (2002). Our bounds are expressed in terms of a notion of shift based on the total variation distance. The generalization of the "small expert set" result in Bousquet and Warmuth (2002) leads us to obtain better bounds when the sequence against which the regret is measured is sparse. When the trajectory is restricted to the corners of the simplex, we recover, and occasionally improve, the known shifting bounds for prediction with expert advice. Besides, our analysis also captures the setting of adaptive regret, a related notion of regret introduced by Hazan and Seshadhri (2009). It was known that shifting regret and adaptive regret had some connections but this connection is now seen to be even tighter, as both regrets can be viewed as instances of the same *alma mater* regret, which we minimize. Finally, we also show how to dynamically tune the parameters of our algorithms and review briefly the special case of exp-concave loss functions, exhibiting the first logarithmic shifting bounds for exp-concave loss functions on the simplex.

## 2. Preliminaries

We first define the sequential learning framework we work with. Even though our results hold in the general setting of online convex optimization, we present them in the, somewhat simpler, *online linear optimization* setup. We point out in Section 6 how these results may be generalized. Online linear optimization may be cast as a repeated game between the *forecaster* and the *environment* as follows. We use $\Delta_d$ to denote the simplex $\left\{\boldsymbol{q} \in [0,1]^d \,:\, \|\boldsymbol{q}\|_1 = 1\right\}$.

**Online linear optimization.** For each round $t = 1, \ldots, T$,

1. Forecaster chooses $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \ldots, \widehat{p}_{d,t}) \in \Delta_d$ ;
2. Environment chooses a loss vector $\boldsymbol{\ell}_t = (\ell_{1,t}, \ldots, \ell_{d,t}) \in [0,1]^d$ ;
3. Forecaster suffers loss $\widehat{\boldsymbol{p}}_t^{\top} \boldsymbol{\ell}_t$ .

The goal of the forecaster is to minimize the accumulated loss $\widehat{L}_T = \sum_{t=1}^T \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t$. In the now classical problem of prediction with expert advice, the goal of the forecaster is to compete with the best fixed component (often called "expert") chosen in hindsight, that is, with $\min_{i=1,\dots,T} \sum_{t=1}^T \ell_{i,t}$. The focus of this paper is on more ambitious forecasters that compete with a richer class of *sequences* of components. Let $[d] = \{1, \dots, d\}$. We use $i_1^T = (i_1, \dots, i_T)$ to denote a sequence in $[d]^T$ and let $L_T(i_1^T) = \sum_{t=1}^T \ell_{i_t}$ be the cumulative linear loss of the sequence $i_1^T \in [d]^T$.

We start by introducing our main algorithmic tool, a generalized share algorithm. It is parametrized by the "mixing functions" $\psi_t : [0,1]^{(t+1)d} \to \Delta_d$ for $t = 1, \dots, T$ that assign probabilities to past "pre-weights" as defined below. In all examples discussed in this paper, these mixing functions are quite simple but working with such a general model makes the main ideas more transparent. We then provide a simple lemma that serves as the starting point for analyzing different instances of the generalized share algorithm.

**Algorithm 1**: The generalized share algorithm.

**Parameters:** learning rate $\eta > 0$ and mixing functions $\psi_t$ for $t = 1, \dots, T$
**Initialization:** $\widehat{\boldsymbol{p}}_1 = \boldsymbol{v}_1 = (1/d, \dots, 1/d)$

**For** each round $t = 1, \dots, T$,
1. Predict $\widehat{\boldsymbol{p}}_t$ ;
2. Observe loss $\boldsymbol{\ell}_t \in [0,1]^d$ ;
3. [loss update] **For** each $j = 1, \dots, d$ define

$$v_{j,t+1} = \frac{\widehat{p}_{j,t}\, e^{-\eta\, \ell_{j,t}}}{\sum_{i=1}^d \widehat{p}_{i,t}\, e^{-\eta\, \ell_{i,t}}} \qquad \text{the current pre-weights,}$$

$$\underline{V}_{t+1} = \big[v_{i,s}\big]_{i \in [d],\, 1 \leqslant s \leqslant t+1} \qquad \text{the } d \times (t+1) \text{ matrix of all past and current pre-weights;}$$

4. [shared update] Define $\widehat{\boldsymbol{p}}_{t+1} = \psi_{t+1}\big(\underline{V}_{t+1}\big)$.

**Lemma 1** *For all $t \geqslant 1$ and for all $\boldsymbol{q}_t \in \Delta_d$, Algorithm 1 satisfies*

$$\big(\widehat{\boldsymbol{p}}_t - \boldsymbol{q}_t\big)^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta} \sum_{i=1}^d q_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8}\,.$$

**Proof** By Hoeffding's inequality,

$$\sum_{j=1}^d \widehat{p}_{j,t}\, \ell_{j,t} \leqslant -\frac{1}{\eta} \ln \left( \sum_{j=1}^d \widehat{p}_{j,t}\, e^{-\eta\, \ell_{j,t}} \right) + \frac{\eta}{8}\,. \tag{1}$$

By definition of $v_{i,t+1}$, for all $i = 1, \dots, d$ we then have $\displaystyle \sum_{j=1}^d \widehat{p}_{j,t}\, e^{-\eta\, \ell_{j,t}} = \frac{\widehat{p}_{i,t}\, e^{-\eta\, \ell_{i,t}}}{v_{i,t+1}}$,

which entails $\displaystyle \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t \leqslant \ell_{i,t} + \frac{1}{\eta} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8}\,.$

The proof is concluded by taking a convex aggregation with respect to $\boldsymbol{q}_t$. ■

## 3. Shifting bounds

In this section we prove shifting regret bounds for the generalized share algorithm. We compare the cumulative loss $\sum_{t=1}^T \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t$ of the forecaster with the loss of an arbitrary sequence of vectors $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T$ in the simplex $\Delta_d$, that is, with $\sum_{t=1}^T \boldsymbol{q}_t^\top \boldsymbol{\ell}_t$. The bounds we obtain depend, of course, on the "regularity" of the comparison sequence. In the now classical results on tracking the best expert (as in Herbster and Warmuth 1998; Vovk 1999; Herbster and Warmuth 2001; Bousquet and Warmuth 2002), this regularity is measured as the number of times $\boldsymbol{q}_t \neq \boldsymbol{q}_{t+1}$ (henceforth referred to as "hard shifts"). The main results of this paper show not only that these results may be generalized to obtain bounds in terms of "softer" regularity measures but that the same algorithms that were proposed with hard shift tracking in mind achieve such, perhaps surprisingly good, performance. Building on the general formulation introduced in Section 2, we derive such regret bounds for the fixed-share algorithm of Herbster and Warmuth (1998) and for the algorithms of Bousquet and Warmuth (2002).

In fact, it is advantageous to extend our analysis so that we not only compare the performance of the forecaster with sequences $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T$ taking values in the simplex $\Delta_d$ of probability distributions but rather against arbitrary sequences $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T \in \mathbb{R}_+^d$ of vectors with non-negative components. The loss of such a sequence is defined by $\sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t$. For fair comparison, we measure the cumulative loss of the forecaster by $\sum_{t=1}^T \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t \|\boldsymbol{u}_t\|_1$. Of course, when $\boldsymbol{u}_t \in \Delta_d$, we recover the original notion of regret.

The norms $\|\boldsymbol{u}_1\|_1, \dots, \|\boldsymbol{u}_T\|_1$ may be viewed as a sequence of weights that give more or less importance to the instantaneous loss suffered at each step. Of particular interest is the case when $\|\boldsymbol{u}_t\|_1 \in [0,1]$ which is the setting of "time selection functions" (see Blum and Mansour 2007, Section 6). In particular, considering sequences $\|\boldsymbol{u}_t\|_1 \in \{0,1\}$ that include the zero vector will provide us a simple way of deriving "adaptive" regret bounds, a notion introduced by Hazan and Seshadhri (2009).

The first regret bounds derived below measure the regularity of the sequence $\boldsymbol{u}_1^T = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_T)$ in terms of the quantity

$$m(\boldsymbol{u}_1^T) = \sum_{t=1}^{T-1} D_{\mathrm{TV}}(\boldsymbol{u}_{t+1}, \boldsymbol{u}_t) \tag{2}$$

where for $\boldsymbol{x} = (x_1, \dots, x_d), \boldsymbol{y} = (y_1, \dots, y_d) \in \mathbb{R}_+^d$, we define $D_{\mathrm{TV}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{x_i \geqslant y_i} (x_i - y_i)$. Note that when $\boldsymbol{x}, \boldsymbol{y} \in \Delta_d$, we recover the total variation distance $D_{\mathrm{TV}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_1$, while for general $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^d$, the quantity $D_{\mathrm{TV}}(\boldsymbol{x}, \boldsymbol{y})$ is not necessarily symmetric and is always bounded by $\|\boldsymbol{x} - \boldsymbol{y}\|_1$. Note that when the vectors $\boldsymbol{u}_t$ are incidence vectors $(0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^d$ of elements $i_t \in [d]$, then $m(\boldsymbol{u}_1^T)$ corresponds to the number of shifts of the sequence $i_1^T \in [d]^T$, and we recover from the results stated below the classical bounds for tracking the best expert.

### 3.1. Fixed-share update

We now analyze a specific instance of the generalized share algorithm corresponding to the update

$$\widehat{p}_{j,t+1} = \sum_{i=1}^d \left( \frac{\alpha}{d} + (1-\alpha) \mathbb{1}_{i=j} \right) v_{i,t+1} = \frac{\alpha}{d} + (1-\alpha) v_{j,t+1} \,, \qquad 0 \leqslant \alpha \leqslant 1 \,. \tag{3}$$

Despite seemingly different statements, this update in Algorithm 1 can be seen to lead *exactly* to the fixed-share algorithm of Herbster and Warmuth (1998) for prediction with expert advice.

**Proposition 2** *With the above update, for all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{\|\boldsymbol{u}_1\|_1 \ln d}{\eta} + \frac{\eta}{8} \sum_{t=1}^T \|\boldsymbol{u}_t\|_1 + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{d}{\alpha} + \frac{\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 - m(\boldsymbol{u}_1^T) - 1}{\eta} \ln \frac{1}{1-\alpha} \, .$$

We emphasize that the fixed-share forecaster does not need to "know" anything about the sequence of the norms $\|\boldsymbol{u}_t\|$. Of course, in order to minimize the obtained upper bound, the tuning parameters $\alpha, \eta$ need to be optimized and their values will depend on the maximal value of $m(\boldsymbol{u}_i^T)$ for the sequences one wishes to compete against. In particular, we obtain the following corollary, in which $h(x) = -x \ln x - (1-x)\ln(1-x)$ denotes the binary entropy function for $x \in [0,1]$. We recall [1] that $h(x) \leqslant x \ln(e/x)$ for $x \in [0,1]$.

**Corollary 3** *Suppose Algorithm 1 is run with the update (3). Let $m_0 > 0$. For all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T \in \Delta_d$ with $m(\boldsymbol{q}_1^T) \leqslant m_0$,*

$$\sum_{t=1}^T \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{q}_t^\top \boldsymbol{\ell}_t \leqslant \sqrt{\frac{T}{2}\left( (m_0+1)\ln d + (T-1)\; h\left(\frac{m_0}{T-1}\right)\right)},$$

*whenever $\eta$ and $\alpha$ are optimally chosen in terms of $m_0$ and $T$.*

If we only consider vectors of the form $\boldsymbol{q}_t = (0, \dots, 0, 1, 0, \dots, 0)$ then $m(\boldsymbol{q}_1^T)$ corresponds to the number of times $\boldsymbol{q}_{t+1} \neq \boldsymbol{q}_t$ in the sequence $\boldsymbol{q}_1^T$. We thus recover Herbster and Warmuth (1998, Theorem 1) and Bousquet and Warmuth (2002, Lemma 6) from the much more general Proposition 2.

**Proof of Proposition 2** Applying Lemma 1 with $\boldsymbol{q}_t = \boldsymbol{u}_t / \|\boldsymbol{u}_t\|_1$, and multiplying by $\|\boldsymbol{u}_t\|_1$, we get for all $t \geqslant 1$ and $\boldsymbol{u}_t \in \mathbb{R}_+^d$

$$\|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta} \sum_{i=1}^d u_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8} \|\boldsymbol{u}_t\|_1 \, . \tag{4}$$

We now examine

$$\sum_{i=1}^d u_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} = \sum_{i=1}^d \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) + \sum_{i=1}^d \left( u_{i,t-1} \ln \frac{1}{v_{i,t}} - u_{i,t} \ln \frac{1}{v_{i,t+1}} \right) \, . \tag{5}$$

For the first term on the right-hand side, we have

$$\sum_{i=1}^d \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) = \sum_{i \,:\, u_{i,t} \geqslant u_{i,t-1}} \left( (u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \right)$$

1. As can be seen by noting that $\ln\big(1/(1-x)\big) < x/(1-x)$

$$+ \sum_{i\,:\,u_{i,t}<u_{i,t-1}} \Bigg( \underbrace{(u_{i,t}-u_{i,t-1}) \ln \frac{1}{v_{i,t}}}_{\leqslant 0} + u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \Bigg) . \qquad (6)$$

In view of the update (3), we have $1/\widehat{p}_{i,t} \leqslant d/\alpha$ and $v_{i,t}/\widehat{p}_{i,t} \leqslant 1/(1-\alpha)$. Substituting in (6), we get

$$\begin{aligned}
&\sum_{i=1}^{d} \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) \\
&\quad \leqslant \sum_{i\,:\,u_{i,t} \geqslant u_{i,t-1}} (u_{i,t}-u_{i,t-1}) \ln \frac{d}{\alpha} + \left( \sum_{i:\ u_{i,t} \geqslant u_{i,t-1}} u_{i,t-1} + \sum_{i:\ u_{i,t}<u_{i,t-1}} u_{i,t} \right) \ln \frac{1}{1-\alpha} \\
&\quad = D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \ln \frac{d}{\alpha} + \underbrace{\left( \sum_{i=1}^{d} u_{i,t} - \sum_{i\,:\,u_{i,t} \geqslant u_{i,t-1}} (u_{i,t}-u_{i,t-1}) \right)}_{= \|\boldsymbol{u}_t\|_1 - D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1})} \ln \frac{1}{1-\alpha} .
\end{aligned}$$

The sum of the second term in (5) telescopes. Substituting the obtained bounds in the first sum of the right-hand side in (5), and summing over $t = 2, \dots, T$, leads to

$$\begin{aligned}
\sum_{t=2}^{T} \sum_{i=1}^{d} u_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} \leqslant m(\boldsymbol{u}_1^T) \ln \frac{d}{\alpha} + \left( \sum_{t=2}^{T} \|\boldsymbol{u}_t\|_1 - 1 - m(\boldsymbol{u}_1^T) \right) \ln \frac{1}{1-\alpha} \\
+ \sum_{i=1}^{d} u_{i,1} \ln \frac{1}{v_{i,2}} - \underbrace{u_{i,T} \ln \frac{1}{v_{i,T+1}}}_{\leqslant 0} .
\end{aligned}$$

We hence get from (4), which we use in particular for $t = 1$,

$$\begin{aligned}
\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^{\top} \boldsymbol{\ell}_t - \boldsymbol{u}_t^{\top} \boldsymbol{\ell}_t \leqslant \frac{1}{\eta} \sum_{i=1}^{d} u_{i,1} \ln \frac{1}{\widehat{p}_{i,1}} + \frac{\eta}{8} \sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \\
+ \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{d}{\alpha} + \frac{\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 - 1 - m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{1}{1-\alpha} . \qquad \blacksquare
\end{aligned}$$

### 3.2. Sparse sequences: Bousquet-Warmuth updates

Bousquet and Warmuth (2002) proposed forecasters that are able to efficiently compete with the best sequence of experts among all those sequences that only switch a bounded number of times and also take a small number of different values. Such "sparse" sequences of experts appear naturally in many applications. In this section we show that their algorithms in fact work very well in comparison with a much larger class of sequences $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T$ that are "regular"—that is, $m(\boldsymbol{u}_1^T)$, defined in (2) is small—and "sparse" in the sense that the quantity

$$n(\boldsymbol{u}_1^T) = \sum_{i=1}^{d} \max_{t=1,\dots,T} u_{i,t}$$

is small. Note that when $\boldsymbol{q}_t \in \Delta_d$ for all $t$, then two interesting upper bounds can be provided. First, denoting the union of the supports of these convex combinations by $S \subseteq [d]$, we have $n(\boldsymbol{q}_1^T) \leqslant |S|$, the cardinality of $S$. Also,

$$n(\boldsymbol{q}_1^T) \leqslant \Big|\big\{\boldsymbol{q}_t,\ \ t = 1, \dots, T\big\}\Big|,$$

the cardinality of the pool of convex combinations. Thus, $n(\boldsymbol{u}_1^T)$ generalizes the notion of sparsity of Bousquet and Warmuth (2002).

Here we consider a family of shared updates of the form

$$\widehat{p}_{j,t} = (1-\alpha) v_{j,t} + \alpha \frac{w_{j,t}}{Z_t}\,, \qquad 0 \leqslant \alpha \leqslant 1\,, \tag{7}$$

where the $w_{j,t}$ are nonnegative weights that may depend on past and current pre-weights and $Z_t = \sum_{i=1}^d w_{i,t}$ is a normalization constant. Shared updates of this form were proposed by Bousquet and Warmuth (2002, Sections 3 and 5.2).

Apart from generalizing the regret bounds of Bousquet and Warmuth (2002), we believe that the analysis given below is significantly simpler and more transparent. We are also able to slightly improve their original bounds.

We focus on choices of the weights $w_{j,t}$ that satisfy the following conditions: there exists a constant $C \geqslant 1$ such that for all $j = 1, \dots, d$ and $t = 1, \dots, T$,

$$v_{j,t} \leqslant w_{j,t} \leqslant 1 \qquad \text{and} \qquad C\, w_{j,t+1} \geqslant w_{j,t}\,. \tag{8}$$

The next result improves on Proposition 2 when $T \ll d$ and $n(\boldsymbol{u}_1^T) \ll m(\boldsymbol{u}_1^T)$, that is, when the dimension (or number of experts) $d$ is large but the sequence $\boldsymbol{u}_1^T$ is sparse.

**Proposition 4** *Suppose Algorithm 1 is run with the shared update (7) with weights satisfying the conditions (8). Then for all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all sequences $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1\, \widehat{\boldsymbol{p}}_t^{\top} \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^{\top} \boldsymbol{\ell}_t \leqslant \frac{n(\boldsymbol{u}_1^T) \ln d}{\eta} + \frac{n(\boldsymbol{u}_1^T)\, T \ln C}{\eta} + \frac{\eta}{8} \sum_{t=1}^T \|\boldsymbol{u}_t\|_1 + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{\max_{t \leqslant T} Z_t}{\alpha} + \frac{\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 - m(\boldsymbol{u}_1^T) - 1}{\eta} \ln \frac{1}{1-\alpha}\,.$$

**Proof** The beginning and the end of the proof are similar to the one of Proposition 2, as they do not depend on the specific weight update. In particular, inequalities (4) and (5) remain the same. The proof is modified after (6), which this time we upper bound using the first condition in (8),

$$\sum_{i=1}^d \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) = \sum_{i\,:\,u_{i,t} \geqslant u_{i,t-1}} (u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} + \sum_{i\,:\,u_{i,t} < u_{i,t-1}} \underbrace{(u_{i,t} - u_{i,t-1})}_{\leqslant 0} \underbrace{\ln \frac{1}{v_{i,t}}}_{\geqslant \ln(1/w_{i,t})} + u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}}\,. \tag{9}$$

By definition of the shared update (7), we have $1/\widehat{p}_{i,t} \leqslant Z_t/(\alpha\, w_{i,t})$ and $v_{i,t}/\widehat{p}_{i,t} \leqslant 1/(1-\alpha)$. We then upper bound the quantity at hand in (9) by

$$\begin{aligned}
&\sum_{i\,:\,u_{i,t}\geqslant u_{i,t-1}} (u_{i,t}-u_{i,t-1}) \ln\left(\frac{Z_t}{\alpha\, w_{i,t}}\right) + \left(\sum_{i\,:\,u_{i,t}\geqslant u_{i,t-1}} u_{i,t-1} + \sum_{i\,:\,u_{i,t}<u_{i,t-1}} u_{i,t}\right) \ln\frac{1}{1-\alpha} \\
&\quad + \sum_{i\,:\,u_{i,t}<u_{i,t-1}} (u_{i,t}-u_{i,t-1}) \ln\frac{1}{w_{i,t}} \\
&= D_{\mathrm{TV}}(\boldsymbol{u}_t,\boldsymbol{u}_{t-1}) \ln\frac{Z_t}{\alpha} + \big(\|\boldsymbol{u}_t\|_1 - D_{\mathrm{TV}}(\boldsymbol{u}_t,\boldsymbol{u}_{t-1})\big) \ln\frac{1}{1-\alpha} + \sum_{i=1}^d (u_{i,t}-u_{i,t-1}) \ln\frac{1}{w_{i,t}}\,.
\end{aligned}$$

Proceeding as in the end of the proof of Proposition 2, we then get the claimed bound, provided that we can show that

$$\sum_{t=2}^T \sum_{i=1}^d (u_{i,t}-u_{i,t-1}) \ln\frac{1}{w_{i,t}} \leqslant n(\boldsymbol{u}_1^T)\,(\ln d + T\ln C) - \|\boldsymbol{u}_1\|_1 \ln d\,,$$

which we do next. Indeed, the left-hand side can be rewritten as

$$\begin{aligned}
&\sum_{t=2}^T \sum_{i=1}^d \left(u_{i,t}\ln\frac{1}{w_{i,t}} - u_{i,t}\ln\frac{1}{w_{i,t+1}}\right) + \sum_{t=2}^T \sum_{i=1}^d \left(u_{i,t}\ln\frac{1}{w_{i,t+1}} - u_{i,t-1}\ln\frac{1}{w_{i,t}}\right) \\
&\leqslant \left(\sum_{t=2}^T \sum_{i=1}^d u_{i,t}\ln\frac{C\,w_{i,t+1}}{w_{i,t}}\right) + \left(\sum_{i=1}^d u_{i,T}\ln\frac{1}{w_{i,T+1}} - \sum_{i=1}^d u_{i,1}\ln\frac{1}{w_{i,2}}\right) \\
&\leqslant \left(\sum_{i=1}^d \left(\max_{t=1,\dots,T} u_{i,t}\right) \sum_{t=2}^T \ln\frac{C\,w_{i,t+1}}{w_{i,t}}\right) + \left(\sum_{i=1}^d \left(\max_{t=1,\dots,T} u_{i,t}\right)\ln\frac{1}{w_{i,T+1}} - \sum_{i=1}^d u_{i,1}\ln\frac{1}{w_{i,2}}\right) \\
&= \sum_{i=1}^d \left(\max_{t=1,\dots,T} u_{i,t}\right)\left((T-1)\ln C + \ln\frac{1}{w_{i,2}}\right) - \sum_{i=1}^d u_{i,1}\ln\frac{1}{w_{i,2}}\,,
\end{aligned}$$

where we used $C \geqslant 1$ for the first inequality and the second condition in (8) for the second inequality. The proof is concluded by noting that (8) entails $w_{i,2} \geqslant (1/C) w_{i,1} \geqslant (1/C) v_{i,1} = 1/(dC)$ and that the coefficient $\max_{t=1,\dots,T} u_{i,t} - u_{i,1}$ in front of $\ln(1/w_{i,2})$ is nonnegative. ■

We now generalize Corollaries 8 and 9 of Bousquet and Warmuth (2002) by showing two specific instances of the generic update (7) that satisfy (8). The first update uses $w_{j,t} = \max_{s\leqslant t} v_{j,s}$. Then (8) is satisfied with $C = 1$. Moreover, since a sum of maxima of nonnegative elements is smaller than the sum of the sums, $Z_t \leqslant \min\{d,t\} \leqslant T$. This immediately gives the following result.

**Corollary 5** *Suppose Algorithm 1 is run with the update (7) with $w_{j,t} = \max_{s\leqslant t} v_{j,s}$. For all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1,\dots,\boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\boldsymbol{q}_1,\dots,\boldsymbol{q}_T \in \Delta_d$,*

$$\sum_{t=1}^T \widehat{\boldsymbol{p}}_t^{\top}\boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{q}_t^{\top}\boldsymbol{\ell}_t \leqslant \frac{n(\boldsymbol{q}_1^T)\ln d}{\eta} + \frac{\eta}{8}T + \frac{m(\boldsymbol{q}_1^T)}{\eta}\ln\frac{T}{\alpha} + \frac{T-m(\boldsymbol{q}_1^T)-1}{\eta}\ln\frac{1}{1-\alpha}\,.$$

The second update we discuss uses $w_{j,t} = \max_{s \leqslant t} e^{\gamma(s-t)} v_{j,s}$ in (7) for some $\gamma > 0$. Both conditions in (8) are satisfied with $C = e^{\gamma}$. One also has that

$$Z_t \leqslant d \qquad \text{and} \qquad Z_t \leqslant \sum_{\tau \geqslant 0} e^{-\gamma\tau} = \frac{1}{1 - e^{-\gamma}} \leqslant \frac{1}{\gamma}$$

as $e^x \geqslant 1 + x$ for all real $x$. The bound of Proposition 4 then instantiates as

$$\frac{n(\boldsymbol{q}_1^T) \ln d}{\eta} + \frac{n(\boldsymbol{q}_1^T)\, T\gamma}{\eta} + \frac{\eta}{8} T + \frac{m(\boldsymbol{q}_1^T)}{\eta} \ln \frac{\min\{d,\, 1/\gamma\}}{\alpha} + \frac{T - m(\boldsymbol{q}_1^T) - 1}{\eta} \ln \frac{1}{1-\alpha}$$

when sequences $\boldsymbol{u}_t = \boldsymbol{q}_t \in \Delta_d$ are considered. This bound is best understood when $\gamma$ is tuned optimally based on $T$ and on two bounds $m_0$ and $n_0$ over the quantities $m(\boldsymbol{q}_1^T)$ and $n(\boldsymbol{q}_1^T)$. Indeed, by optimizing $n_0 T\gamma + m_0 \ln(1/\gamma)$, i.e., by choosing $\gamma = m_0/(n_0\, T)$, one gets a bound that improves on the one of the previous corollary:

**Corollary 6** *Let $m_0,\, n_0 > 0$. Suppose Algorithm 1 is run with the update $w_{j,t} = \max_{s \leqslant t} e^{\gamma(s-t)} v_{j,s}$ where $\gamma = m_0/(n_0\, T)$. For all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_T \in \Delta_d$ such that $m(\boldsymbol{q}_1^T) \leqslant m_0$ and $n(\boldsymbol{q}_1^T) \leqslant n_0$, we have*

$$\begin{aligned} \sum_{t=1}^T \widehat{\boldsymbol{p}}_t^{\top} \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{q}_t^{\top} \boldsymbol{\ell}_t \;\leqslant\; & \frac{n_0 \ln d}{\eta} + \frac{m_0}{\eta} \left(1 + \ln \min\left\{d,\, \frac{n_0\, T}{m_0}\right\}\right) \\ & + \frac{\eta}{8} T + \frac{m_0}{\eta} \ln \frac{1}{\alpha} + \frac{T - m_0 - 1}{\eta} \ln \frac{1}{1-\alpha}\,. \end{aligned}$$

As the factors $e^{-\gamma t}$ cancel out in the numerator and denominator of the ratio in (7), there is a straightforward implementation of the algorithm (not requiring the knowledge of $T$) that needs to maintain only $d$ weights.

In contrast, the corresponding algorithm of Bousquet and Warmuth (2002), using the updates $\widehat{p}_{j,t} = (1-\alpha) v_{j,t} + \alpha S_t^{-1} \sum_{s \leqslant t-1} (s-t)^{-1} v_{j,s}$ or $\widehat{p}_{j,t} = (1-\alpha) v_{j,t} + \alpha S_t^{-1} \max_{s \leqslant t-1} (s-t)^{-1} v_{j,s}$, where $S_t$ denote normalization factors, needs to maintain $O(dT)$ weights with a naive implementation, and $O(d \ln T)$ weights with a more sophisticated one. In addition, the obtained bounds are slightly worse than the one stated above in Corollary 6 as an additional factor of $m_0 \ln(1 + \ln T)$ is present in Bousquet and Warmuth (2002, Corollary 9).

## 4. Adaptive regret

Next we show how the results of the previous section, e.g., Proposition 2, imply guarantees in terms of adaptive regret —a notion introduced by Hazan and Seshadhri (2009) as follows. For $\tau_0 \in \{1, \ldots, T\}$, the *$\tau_0$-adaptive regret* of a forecaster is defined by

$$\mathcal{R}_T^{\tau_0 - \text{adapt}} = \max_{\substack{[r,\, s] \,\subset\, [1,\, T] \\ s + 1 - r \,\leqslant\, \tau_0}} \left\{ \sum_{t=r}^{s} \widehat{\boldsymbol{p}}_t^{\top} \boldsymbol{\ell}_t - \min_{\boldsymbol{q} \in \Delta_d} \sum_{t=r}^{s} \boldsymbol{q}^{\top} \boldsymbol{\ell}_t \right\}. \tag{10}$$

Adaptive regret is an alternative way to measure the performance of a forecaster against a changing environment. It is a straightforward observation that adaptive regret bounds also lead to shifting

regret bounds (in terms of hard shifts). Here we show that these two notions of regret share an even tighter connection, as they can be both viewed as instances of the same *alma mater* bound, e.g., Proposition 2.

Hazan and Seshadhri (2009) essentially considered the case of online convex optimization with exp-concave loss function (see Section 6 below). In case of general convex functions, they also mentioned that the greedy projection forecaster of Zinkevich (2003) —i.e., mirror descent with a quadratic regularizer— enjoys adaptive regret guarantees. This forecaster can be implemented on the simplex in time $O(d)$ —see, e.g., Duchi et al. (2008). We now show that the simpler fixed-share algorithm has a similar adaptive regret bound.

**Proposition 7** *Suppose that Algorithm 1 is run with the shared update (3). Then for all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\tau_0 \in \{1, \dots, T\}$,*

$$\mathcal{R}_T^{\tau_0-\mathrm{adapt}} \leqslant \frac{1}{\eta} \ln \frac{d}{\alpha} + \frac{\tau_0 - 1}{\eta} \ln \frac{1}{1-\alpha} + \frac{\eta}{8} \tau_0 \,.$$

*In particular, when $\eta$ and $\alpha$ are chosen optimally (depending on $\tau_0$ and $T$)*

$$\mathcal{R}_T^{\tau_0-\mathrm{adapt}} \leqslant \sqrt{\frac{\tau_0}{2} \left( \tau_0 \, h\!\left(\frac{1}{\tau_0}\right) + \ln d \right)} \leqslant \sqrt{\frac{\tau_0}{2} \, \ln(ed\tau_0)} \,.$$

**Proof** For $1 \leqslant r \leqslant s \leqslant T$ and $\boldsymbol{q} \in \Delta_d$, the regret in the right-hand side of (10) equals the regret considered in Proposition 2 against the sequence $\boldsymbol{u}_1^T$ defined as $\boldsymbol{u}_t = \boldsymbol{q}$ for $t = r, \dots, s$ and $\mathbf{0} = (0, \dots, 0)$ for the remaining $t$. When $r \geqslant 2$, this sequence is such that $D_{TV}(\boldsymbol{u}_r, \boldsymbol{u}_{r-1}) = D_{TV}(\boldsymbol{q}, \mathbf{0}) = 1$ and $D_{TV}(\boldsymbol{u}_{s+1}, \boldsymbol{u}_s) = D_{TV}(\mathbf{0}, \boldsymbol{q}) = 0$ so that $m(\boldsymbol{u}_1^T) = 1$, while $\|\boldsymbol{u}_1\|_1 = 0$. When $r = 1$, we have $\|\boldsymbol{u}_1\|_1 = 1$ and $m(\boldsymbol{u}_1^T) = 0$. In all cases, $m(\boldsymbol{u}_1^T) + \|\boldsymbol{u}_1\|_1 = 1$. Specializing the bound of Proposition 2 to the thus defined sequence $\boldsymbol{u}_1^T$ gives the result. ■

## 5. Online tuning of the parameters

The forecasters studied above need their parameters $\eta$ and $\alpha$ to be tuned according to various quantities, including the time horizon $T$. We show here how the trick of Auer et al. (2002) of having these parameters vary over time can be extended to our setting. For the sake of concreteness we focus on the fixed-share update, i.e., Algorithm 1 run with the update (3). We respectively replace steps 3 and 4 of its description by the loss and shared updates

$$v_{j,t+1} = \frac{\widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}}}{\sum_{i=1}^d \widehat{p}_{i,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{i,t}}} \qquad \text{and} \qquad p_{j,t+1} = \frac{\alpha_t}{d} + (1 - \alpha_t) \, v_{j,t+1} \,, \tag{11}$$

for all $t \geqslant 1$ and all $j \in [d]$, where $(\eta_\tau)$ and $(\alpha_\tau)$ are two sequences of positive numbers, indexed by $\tau \geqslant 1$. We also conventionally define $\eta_0 = \eta_1$. Proposition 2 is then adapted in the following way (when $\eta_t \equiv \eta$ and $\alpha_t \equiv \alpha$, Proposition 2 is exactly recovered).

**Proposition 8** *The forecaster based on the above updates* (11) *is such that whenever* $\eta_t \leqslant \eta_{t-1}$ *and* $\alpha_t \leqslant \alpha_{t-1}$ *for all* $t \geqslant 1$*, the following performance bound is achieved. For all* $T \geqslant 1$*, for all sequences* $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ *of loss vectors* $\boldsymbol{\ell}_t \in [0,1]^d$*, and for all* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T \in \mathbb{R}_+^d$*,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \left( \frac{\|\boldsymbol{u}_t\|_1}{\eta_1} + \sum_{t=2}^T \|\boldsymbol{u}_t\|_1 \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \right) \ln d$$
$$+ \frac{m(\boldsymbol{u}_1^T)}{\eta_T} \ln \frac{d(1-\alpha_T)}{\alpha_T} + \sum_{t=2}^T \frac{\|\boldsymbol{u}_t\|_1}{\eta_{t-1}} \ln \frac{1}{1-\alpha_t} + \sum_{t=1}^T \frac{\eta_{t-1}}{8} \|\boldsymbol{u}_t\|_1 \, .$$

Due to space constraints, we only instantiate the obtained bound to the case of $T$-adaptive regret guarantees, when $T$ is unknown and/or can increase without bounds.

**Corollary 9** *The forecaster based on the above updates with* $\eta_t = \sqrt{\big(\ln(dt)\big)/t}$ *for* $t \geqslant 3$ *and* $\eta_0 = \eta_1 = \eta_2 = \eta_3$ *on the one hand,* $\alpha_t = 1/t$ *on the other hand, is such that for all* $T \geqslant 3$ *and for all sequences* $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ *of loss vectors* $\boldsymbol{\ell}_t \in [0,1]^d$*,*

$$\max_{[r,s] \subset [1,T]} \left\{ \sum_{t=r}^s \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \min_{\boldsymbol{q} \in \Delta_d} \sum_{t=r}^s \boldsymbol{q}^\top \boldsymbol{\ell}_t \right\} \leqslant \sqrt{2T \ln(dT)} + \sqrt{3 \ln(3d)} \, .$$

**Proof** The sequence $n \mapsto \ln(n)/n$ is only non-increasing after round $n \geqslant 3$, so that the defined sequences of $(\alpha_t)$ and $(\eta_t)$ are non-increasing, as desired. For a given pair $(r,s)$ and a given $\boldsymbol{q} \in \Delta_d$, we consider the sequence $\nu_1^T$ defined in the proof of Proposition 7; it satisfies that $m(\boldsymbol{u}_1^T) \leqslant 1$ and $\|\boldsymbol{u}_t\|_1 \leqslant 1$ for all $t \geqslant 1$. Therefore, Proposition 8 ensures that

$$\sum_{t=r}^s \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \min_{\boldsymbol{q} \in \Delta_d} \sum_{t=r}^s \boldsymbol{q}^\top \boldsymbol{\ell}_t \leqslant \frac{\ln d}{\eta_T} + \frac{1}{\eta_T} \ln \underbrace{\frac{d(1-\alpha_T)}{\alpha_T}}_{\leqslant dT} + \underbrace{\sum_{t=2}^T \frac{1}{\eta_{t-1}} \ln \frac{1}{1-\alpha_t}}_{\leqslant (1/\eta_T) \sum_{t=2}^T \ln(t/(t-1)) = (\ln T)/\eta_T} + \sum_{t=1}^T \frac{\eta_{t-1}}{8} \, .$$

It only remains to substitute the proposed values of $\eta_t$ and to note that

$$\sum_{t=1}^T \eta_{t-1} \leqslant 3\eta_3 + \sum_{t=3}^{T-1} \frac{1}{\sqrt{t}} \sqrt{\ln(dT)} \leqslant 3\sqrt{\frac{\ln(3d)}{3}} + 2\sqrt{T} \sqrt{\ln(dT)} \, . \qquad \blacksquare$$

## 6. Online convex optimization and exp-concave loss functions

By using a standard reduction, the results of the previous sections can be applied to online convex optimization on the simplex. In this setting, at each step $t$ the forecaster chooses $\widehat{\boldsymbol{p}}_t \in \Delta_d$ and then is given access to a convex loss $\ell_t : \Delta_d \to [0,1]$. Now, using Algorithm 1 with the loss vector $\boldsymbol{\ell}_t \in \partial \ell_t(\widehat{\boldsymbol{p}}_t)$ given by a subgradient of $\ell_t$ leads to the desired bounds. Indeed, by the convexity of $\ell_t$, the regret at each time $t$ with respect to any vector $\boldsymbol{u}_t \in \mathbb{R}_+^d$ with $\|\boldsymbol{u}_t\|_1 > 0$ is then bounded as

$$\|\boldsymbol{u}_t\|_1 \left( \ell_t(\widehat{\boldsymbol{p}}_t) - \ell_t\!\left( \frac{\boldsymbol{u}_t}{\|\boldsymbol{u}_t\|_1} \right) \right) \leqslant \big( \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t - \boldsymbol{u}_t \big)^\top \boldsymbol{\ell}_t \, .$$

### 6.1. Exp-concave loss functions

Recall that a loss function $\ell_t$ is called $\eta_0$-exp-concave if $e^{-\eta_0 \ell_t}$ is concave. (In particular, exp-concavity implies convexity.) Bousquet and Warmuth (2002) study shifting regret for exp-concave loss functions. However, they define the regret of an element $\boldsymbol{q}_1^T$ of the comparison class (a sequence of elements in $\Delta_d$) by

$$\sum_{t=1}^{T} \Big( \ell_t(\widehat{\boldsymbol{p}}_t) - \boldsymbol{q}_t^\top \boldsymbol{\ell}_t \Big) \tag{12}$$

where $\boldsymbol{\ell}_t = \big(\ell_t(\boldsymbol{e}_1), \ldots, \ell_t(\boldsymbol{e}_d)\big)$ and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$ are the elements of the canonical basis of $\mathbb{R}^d$. This corresponds to the linear optimization case studied in the previous sections. However, due to exp-concavity, (1) can be replaced by an application of Jensen's inequality, namely,

$$\ell_t(\widehat{\boldsymbol{p}}_t) \leqslant -\frac{1}{\eta_0} \ln \left( \sum_{j=1}^{d} \widehat{p}_{j,t}\, e^{-\eta_0\, \ell_t(\boldsymbol{e}_j)} \right) .$$

Hence the various propositions and corollaries of Sections 3 and 4 still hold true for the regret (12) up to some modifications (deletion of the terms linear in $\eta$, assumption of exp-concavity, boundedness no longer needed). For the sake of concreteness, we illustrate the required modifications on Proposition 4.

**Proposition 10** *Suppose Algorithm 1 is run with the shared update (7) with weights satisfying the conditions* (8) *and for the choice $\eta = \eta_0$. Then for all $T \geqslant 1$, for all sequences $\ell_1, \ldots, \ell_T$ of $\eta_0$–exp-concave loss functions, and for all sequences $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

$$\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1\, \ell_t(\widehat{\boldsymbol{p}}_t) - \sum_{t=1}^{T} \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{(\boldsymbol{u}_1^T) \ln d}{\eta_0} + \frac{n(\boldsymbol{u}_1^T)\, T \ln C}{\eta_0} + \frac{m(\boldsymbol{u}_1^T)}{\eta_0} \ln \frac{\max_{t \leqslant T} Z_t}{\alpha} + \frac{\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 - m(\boldsymbol{u}_1^T) - 1}{\eta_0} \ln \frac{1}{1-\alpha} .$$

We now turn to the more ambitious goal of controlling regrets of the form $\sum_{t=1}^{T} \big(\ell_t(\widehat{\boldsymbol{p}}_t) - \ell_t(\boldsymbol{q}_t)\big)$ where losses $\ell_t$ are exp-concave. Hazan and Seshadhri (2009) constructed algorithms with $T$–adaptive regret of the order of $O(\ln^2 T)$ and running in time $\mathrm{poly}(d, \log T)$. They also constructed different algorithms with $T$–adaptive regret bounded by $O(\ln T)$) and running time $\mathrm{poly}(d, T)$.

Next, we show the first logarithmic shifting bounds for exp-concave loss functions. However, we only do so against sequences $\boldsymbol{q}_1^T$ of elements in $\Delta_d$, i.e., we offer here no general bound in terms of linear vectors $\boldsymbol{u}_1^T$ that would unify here as well the view between tracking bounds and adaptive regret bounds. Besides, we get shifting bounds only in terms of hard shifts

$$s(\boldsymbol{q}_1^T) = \big|\big\{t = 2, \ldots, T \,:\, \boldsymbol{q}_t \neq \boldsymbol{q}_{t-1}\big\}\big| .$$

Obviously, getting unifying bounds in terms of soft shifts of sequences $\boldsymbol{u}_1^T$ of linear vectors is an important open question, which we leave for future research. To get our bound, we mix ideas of Herbster and Warmuth (1998) and Blum and Kalai (1997). We define a prior over the sequences of convex weight vectors as the distribution of the following homogeneous Markov chain $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots$: The starting vector $\boldsymbol{Q}_1$ is drawn at random according to the uniform distribution $\mu$ over $\Delta_d$. Then,

given $\boldsymbol{Q}_{t-1}$, the next element $\boldsymbol{Q}_t$ is equal to $\boldsymbol{Q}_{t-1}$ with probability $1-\alpha$ and with probability $\alpha$ is drawn at random according to $\mu$. In the sequel, all probabilities $\mathbb{P}$ and expectations $\mathbb{E}$ will be with respect to this Markov chain. Now, the convex weight vector used at time $t \geqslant 1$ by the forecaster is

$$\widehat{\boldsymbol{p}}_t = \frac{\mathbb{E}\Big[\boldsymbol{Q}_t\, e^{-\eta_0 L_{t-1}(\boldsymbol{Q}_1^{t-1})}\Big]}{\mathbb{E}\Big[e^{-\eta_0 L_{t-1}(\boldsymbol{Q}_1^{t-1})}\Big]}\,, \qquad \text{where} \quad L_{t-1}(\boldsymbol{Q}_1^{t-1}) = \sum_{s=1}^{t-1} \ell_s(\boldsymbol{Q}_s) \tag{13}$$

(with the convention that an empty sum is null). For this forecaster, we get the following performance bound, whose proof can be found in appendix.

**Proposition 11** *For all $T \geqslant 1$, for all sequences $\ell_1, \dots, \ell_T$ of $\eta_0$–exp-concave loss functions taking values in $[0, L]$, the cumulative loss of the above forecaster is bounded for all sequences $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T \in \Delta_d$ by*

$$\sum_{t=1}^{T} \ell_t(\widehat{\boldsymbol{p}}_t) - \sum_{t=1}^{T} \ell_t(\boldsymbol{q}_t) \leqslant \frac{\big(s(\boldsymbol{q}_1^T)+1\big)(d-1)}{\eta} \max\left\{1,\ \ln \frac{e\,\eta\,L\,T}{\big(s(\boldsymbol{q}_1^T)+1\big)(d-1)}\right\} + \frac{s(\boldsymbol{q}_1^T)}{\eta} \ln \frac{1}{\alpha} + \frac{T - s(\boldsymbol{q}_1^T) - 1}{\eta} \ln \frac{1}{1-\alpha}\,.$$

Under the imposition of a bound $s_0$ on the numbers of hard shifts $s(\boldsymbol{q}_1^T)$ and up to a tuning of $\alpha$ in terms of $s_0$ and $T$, the last two terms of the bound are smaller than $T\,h(s_0/T) \leqslant s_0\,\ln(es_0/T)$ and therefore, the whole regret bound is $O\big((ds_0/\eta_0)\,\ln T\big)$.

## References


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

A. Blum and A. Kalai. Universal portfolios with and without transaction costs. In *Proceedings of the 10th Annual Conference on Learning Theory (COLT)*, pages 309–313. Springer, 1997.

A. Blum and Y. Mansour. From extermal to internal regret. *Journal of Machine Learning Research*, 8:1307–1324, 2007.

O. Bousquet and M.K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3:363–396, 2002.

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

J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra. Efficient projections onto the $\ell_1$–ball for learning in high dimensions. In *Proceedings of the 25th International Conference on Machine Learning, ICML 2008*, 2008.

A. György, T. Linder, and G. Lugosi. Tracking the best of many experts. In *Proceedings of the 18th Annual Conference on Learning Theory (COLT)*, pages 204–216, Bertinoro, Italy, Jun. 2005. Springer.

E. Hazan and C. Seshadhri. Efficient learning algorithms for changing environment. *Proceedings of the 26th International Conference of Machine Learning (ICML)*, 2009.

M. Herbster and M. Warmuth. Tracking the best expert. *Machine Learning*, 32:151–178, 1998.

M. Herbster and M. Warmuth. Tracking the best linear predictor. *Journal of Machine Learning Research*, 1:281–309, 2001.

V. Vovk. Derandomizing stochastic prediction strategies. *Machine Learning*, 35(3):247–282, Jun. 1999.

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning, ICML 2003*, 2003.

## Appendix A. Proof of Proposition 8

We first adapt Lemma 1.

**Lemma 12** *The forecaster based on the loss and shared updates* (11) *satisfies, for all* $t \geqslant 1$ *and for all* $\boldsymbol{q}_t \in \Delta_d$*,*

$$\left(\widehat{\boldsymbol{p}}_t - \boldsymbol{q}_t\right)^\top \boldsymbol{\ell}_t \leqslant \sum_{i=1}^d q_{i,t} \left( \frac{1}{\eta_{t-1}} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \ln \frac{1}{v_{i,t+1}} \right) + \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \ln d + \frac{\eta_{t-1}}{8} \,,$$

*whenever* $\eta_t \leqslant \eta_{t-1}$.

**Proof** By Hoeffding's inequality,

$$\sum_{j=1}^d \widehat{p}_{j,t}\, \ell_{j,t} \leqslant -\frac{1}{\eta_{t-1}} \ln \left( \sum_{j=1}^d \widehat{p}_{j,t}\, e^{-\eta_{t-1}\, \ell_{j,t}} \right) + \frac{\eta_{t-1}}{8} \,.$$

By Jensen's inequality, since $\eta_t \leqslant \eta_{t-1}$ and thus $x \mapsto x^{\frac{\eta_{t-1}}{\eta_t}}$ is convex,

$$\frac{1}{d} \sum_{j=1}^d \widehat{p}_{j,t}\, e^{-\eta_{t-1} \ell_{j,t}} = \frac{1}{d} \sum_{j=1}^d \left( \widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}} \right)^{\frac{\eta_{t-1}}{\eta_t}} \geqslant \left( \frac{1}{d} \sum_{j=1}^d \widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}} \right)^{\frac{\eta_{t-1}}{\eta_t}} .$$

Substituting in Hoeffding's bound we get

$$\widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t \leqslant -\frac{1}{\eta_t} \ln \left( \sum_{j=1}^d \widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}} \right) + \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \ln d + \frac{\eta_{t-1}}{8} \,.$$

Now, by definition of the loss update in (11), for all $i \in [d]$,

$$\sum_{j=1}^d \widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}} = \frac{1}{v_{i,t+1}}\, \widehat{p}_{i,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{i,t}} \,,$$

which, after substitution in the previous bound leads to the inequality

$$\widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t \leqslant \ell_{i,t} + \frac{1}{\eta_{t-1}} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \ln \frac{1}{v_{i,t+1}} + \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \ln d + \frac{\eta_{t-1}}{8} \,,$$

valid for all $i \in [d]$. The proof is concluded by taking a convex aggregation over $i$ with respect to $\boldsymbol{q}_t$. ∎

The proof of Proposition 8 follows the steps of the one of Proposition 2; we sketch it below.

**Proof of Proposition 8** Applying Lemma 12 with $\boldsymbol{q}_t = \boldsymbol{u}_t / \left\|\boldsymbol{u}_t\right\|_1$, and multiplying by $\left\|\boldsymbol{u}_t\right\|_1$, we get for all $t \geqslant 1$ and $\boldsymbol{u}_t \in \mathbb{R}_+^d$,

$$\|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta_{t-1}} \sum_{i=1}^d u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^d u_{i,t} \ln \frac{1}{v_{i,t+1}} + \|\boldsymbol{u}_t\|_1 \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \ln d + \frac{\eta_{t-1}}{8} \|\boldsymbol{u}_t\|_1 \,. \quad (14)$$

We will sum these bounds over $t \geqslant 1$ to get the desired result but need to perform first some additional boundings for $t \geqslant 2$; in particular, we examine

$$\frac{1}{\eta_{t-1}} \sum_{i=1}^d u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^d u_{i,t} \ln \frac{1}{v_{i,t+1}} = \frac{1}{\eta_{t-1}} \sum_{i=1}^d \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) + \sum_{i=1}^d \left( \frac{u_{i,t-1}}{\eta_{t-1}} \ln \frac{1}{v_{i,t}} - \frac{u_{i,t}}{\eta_t} \ln \frac{1}{v_{i,t+1}} \right) , \quad (15)$$

where the first difference in the right-hand side can be bounded as in (6) by

$$\begin{aligned}
&\sum_{i=1}^d \left( u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) \\
&\quad \leqslant \sum_{i \,:\, u_{i,t} \geqslant u_{i,t-1}} \left( (u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \right) + \sum_{i \,:\, u_{i,t} < u_{i,t-1}} u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \\
&\quad \leqslant D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \ln \frac{d}{\alpha_t} + \big( \|\boldsymbol{u}_t\|_1 - D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \big) \ln \frac{1}{1-\alpha_t} \\
&\quad \leqslant D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \ln \frac{d(1-\alpha_T)}{\alpha_T} + \|\boldsymbol{u}_t\|_1 \ln \frac{1}{1-\alpha_t} \,, \qquad (16)
\end{aligned}$$

where we used for the second inequality that the shared update in (11) is such that $1/\widehat{p}_{i,t} \leqslant d/\alpha_t$ and $v_{i,t}/\widehat{p}_{i,t} \leqslant 1/(1-\alpha_t)$, and for the third inequality, that $\alpha_t \geqslant \alpha_T$ and $x \mapsto (1-x)/x$ is increasing on $(0,1]$. Summing (15) over $t = 2, \dots, T$ using (16) and the fact that $\eta_t \geqslant \eta_T$, we get

$$\begin{aligned}
&\sum_{t=2}^T \left( \frac{1}{\eta_{t-1}} \sum_{i=1}^d u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^d u_{i,t} \ln \frac{1}{v_{i,t+1}} \right) \\
&\quad \leqslant \frac{m(\boldsymbol{u}_1^T)}{\eta_T} \ln \frac{d(1-\alpha_T)}{\alpha_T} + \sum_{t=2}^T \frac{\|\boldsymbol{u}_t\|_1}{\eta_{t-1}} \ln \frac{1}{1-\alpha_t} + \sum_{i=1}^d \Bigg( \frac{u_{i,1}}{\eta_1} \ln \frac{1}{v_{i,2}} - \underbrace{\frac{u_{i,T}}{\eta_T} \ln \frac{1}{v_{i,T+1}}}_{\geqslant 0} \Bigg).
\end{aligned}$$

An application of (14) —including for $t = 1$, for which we recall that $\widehat{p}_{i,1} = 1/d$ and $\eta_1 = \eta_0$ by convention— concludes the proof. ■

## Appendix B. Proof of Proposition 11

**Proof** By the definition of exp-concavity and by application of Jensen's inequality to the distribution $\mathbb{P}_t$ over $(\Delta_d)^t$ with density

$$\boldsymbol{r}_1^t \longmapsto \frac{1}{\mathbb{E}\Big[e^{-\eta_0 L_{t-1}(\boldsymbol{r}_1^{t-1})}\Big]}\, e^{-\eta_0 L_{t-1}(\boldsymbol{r}_1^{t-1})} \times 1$$

with respect to the marginal distribution of $\mathbb{P}$ over $(\Delta_d)^t$, we have that

$$\exp\big(-\eta_0\, \ell_t(\widehat{\boldsymbol{p}}_t)\big) = \exp\Big(-\eta_0\, \ell_t\big(\mathbb{E}_t[\boldsymbol{Q}_t]\big)\Big) \geqslant \mathbb{E}_t\Big[\exp\big(-\eta_0\, \ell_t(\boldsymbol{Q}_t)\big)\Big] = \frac{\mathbb{E}\Big[e^{-\eta_0 L_t(\boldsymbol{Q}_1^t)}\Big]}{\mathbb{E}\Big[e^{-\eta_0 L_{t-1}(\boldsymbol{Q}_1^{t-1})}\Big]}\,.$$

Thus, a telescoping sum appears,

$$\sum_{t=1}^T \ell_t(\widehat{\boldsymbol{p}}_t) = \sum_{t=1}^T -\frac{1}{\eta_0} \ln e^{-\eta_0 \ell_t(\widehat{\boldsymbol{p}}_t)} \leqslant -\frac{1}{\eta_0} \ln \mathbb{E}\Big[e^{-\eta_0 L_T(\boldsymbol{Q}_1^T)}\Big]\,.$$

It suffices to lower bound the expectation. To do so, we define for all sequences $\boldsymbol{r}_1^k$ the set of the sequences of $k$ weight vectors that only shift when $\boldsymbol{r}_1^k$ does and that at each such shift are $\varepsilon$–close to the corresponding values of the $\boldsymbol{r}_t$:

$$S_{\varepsilon,\boldsymbol{r}_1^k} = \Big\{\boldsymbol{s}_1^k \in \mathcal{X}^k : \quad \forall t \in \{2,\dots,k\},\ \boldsymbol{s}_t \neq \boldsymbol{s}_{t-1} \ \Rightarrow\ \boldsymbol{r}_t \neq \boldsymbol{r}_{t-1} \\ \text{and} \qquad \forall t \in \{1,\dots,k\}, \quad \boldsymbol{s}_t = (1-\varepsilon)\boldsymbol{r}_t + \varepsilon \boldsymbol{w}_t \text{ for some } \boldsymbol{w}_t \in \mathcal{X}\Big\}\,.$$

Note that the second defining constraint is equivalent to the same constraint only at the shifting times of $\boldsymbol{r}_1^k$, in view of the first constraint. Since exp-concave loss functions are in particular convex, we get that for all $\boldsymbol{s}_1^T \in S_{\varepsilon,\boldsymbol{q}_1^T}$,

$$\sum_{t=1}^T \ell_t(\boldsymbol{s}_t) \leqslant (1-\varepsilon)\sum_{t=1}^T \ell_t(\boldsymbol{q}_t) + \varepsilon \sum_{t=1}^T \ell_t(\boldsymbol{w}_t) \leqslant \sum_{t=1}^T \ell_t(\boldsymbol{q}_t) + \varepsilon L T\,.$$

Thus,

$$-\frac{1}{\eta_0} \ln \mathbb{E}\Big[e^{-\eta_0 L_T(\boldsymbol{Q}_1^T)}\Big] \leqslant -\frac{1}{\eta_0} \ln \mathbb{E}\left[e^{-\eta_0 L_T(\boldsymbol{Q}_1^T)}\ \mathbb{I}_{\{\boldsymbol{Q}_1^T \in S_{\varepsilon,\boldsymbol{q}_1^T}\}}\right] \\ \leqslant \sum_{t=1}^T \ell_t(\boldsymbol{q}_t) + \varepsilon L T - \frac{1}{\eta_0} \ln \mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^T}\big)\,.$$

Furthermore, we show by induction on $t$ that for all $t \geqslant 1$,

$$\mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^t}\big) \geqslant \varepsilon^{d-1}(1-\alpha)^{t-s(\boldsymbol{q}_1^T)-1}\left(\alpha\varepsilon^{d-1}\right)^{s(\boldsymbol{q}_1^T)}\,.$$

This is true for $t = 1$ as $S_{\varepsilon,\boldsymbol{q}_1} = (1-\varepsilon)\boldsymbol{q}_1 + \varepsilon\mathcal{X}$ has a $\mathbb{P}$–probability given by its $\mu$–probability, which is equal to $\varepsilon^{d-1}\,\mu(\mathcal{X}) = \varepsilon^{d-1}$, and as by convention, $s(\boldsymbol{q}_1) = 0$. Besides, when $t \geqslant 2$, we have by definition of $\mathbb{P}$ (cf. its defining transition probability distributions) and $S_{\varepsilon,\boldsymbol{q}_1^t}$ (cf. the $\boldsymbol{s}_1^t$ can only shift when the $\boldsymbol{q}_1^t$ do) that

$$\mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^t}\big) \geqslant \begin{cases} (1-\alpha)\ \mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^{t-1}}\big) & \text{when } \ \boldsymbol{q}_t = \boldsymbol{q}_{t-1} \\ \alpha\ \mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^{t-1}}\big)\ \mu(S_{\varepsilon,\boldsymbol{r}_t}) = \alpha\,\varepsilon^{d-1}\,\mathbb{P}\big(S_{\varepsilon,\boldsymbol{q}_1^{t-1}}\big) & \text{when } \ \boldsymbol{q}_t \neq \boldsymbol{q}_{t-1}\,, \end{cases}$$

which concludes the induction.

Substituting the obtained bound, we have proved so far that

$$\sum_{t=1}^{T} \ell_t(\widehat{\boldsymbol{p}}_t) - \sum_{t=1}^{T} \ell_t(\boldsymbol{q}_t) \leqslant \varepsilon LT - \frac{1}{\eta_0} \ln\Big(\varepsilon^{d-1}(1-\alpha)^{t-s(\boldsymbol{q}_1^T)-1}\,\Big(\alpha\varepsilon^{d-1}\Big)^{s(\boldsymbol{q}_1^T)}\Big)\,.$$

$\varepsilon \in [0,1]$ is a parameter of the analysis, it can be optimized to minimize

$$\varepsilon LT + \frac{\big(s(\boldsymbol{q}_1^T)+1\big)(d-1)}{\eta_0} \ln\frac{1}{\varepsilon}$$

and get the claimed bound. This is achieved by choosing

$$\varepsilon = \min\left\{1,\ \frac{\big(s(\boldsymbol{q}_1^T)+1\big)(d-1)}{\eta_0 LT}\right\}.$$

■