Mirror Descent Meets Fixed Share (and feels no regret)

Anonymous Author(s)
Affiliation
Address
email

Abstract

Mirror descent with an entropic regularizer is known to achieve shifting regret bounds that are logarithmic in the dimension. This is done using either a carefully designed projection or by a weight sharing technique. Via a novel unified analysis, we show that these two approaches deliver essentially equivalent bounds on a notion of regret generalizing shifting, adaptive, discounted, and other related regrets. Our analysis also captures and extends the generalized weight sharing technique of Bousquet and Warmuth, and can be refined in several ways, including improvements for small losses and adaptive tuning of parameters.

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 projected online mirror descent (or follow-the-regularized-leader) algorithm. In this case the shift is typically computed in terms of the p-norm of the difference of consecutive elements in the comparison sequence —see [1, 2] and [3].

We focus on the important special case when $\mathcal{S}$ is the simplex. In [1] shifting bounds are shown for projected mirror descent with entropic regularizers using a 1-norm to measure the shift.[1] 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 of times the trajectory moves to a different corner. This problem is often called "tracking the best expert" —see, e.g., [4, 5, 1, 6, 7], and it is well known that exponential weights with weight sharing, which corresponds to the fixed-share algorithm of [4], achieves a good shifting bound in this setting. In [6] the authors introduce a generalization of the fixed-share algorithm, and prove various shifting bounds for any trajectory in

[1] Similar 1-norm shifting bounds can also be proven using the analysis of [2]. However, without using entropic regularizers it is not clear how to achieve a logarithmic dependence on the dimension, which is one of the advantages of working in the simplex.

the simplex. However, their bounds are expressed using a quantity that corresponds to a proper shift only for trajectories on the simplex corners.

In this paper we offer a unified analysis of mirror descent, fixed share, and the generalized fixed share of [6] for the setting of online convex optimization in the simplex. Our bounds are expressed in terms of a notion of shift based on the total variation distance. Our analysis relies on a generalized notion of shifting regret which includes, as special cases, related notions of regret such as adaptive regret, discounted regret, and regret with time-selection functions. Perhaps surprisingly, we show that projected mirror descent and fixed share achieve essentially the same generalized regret bound. Finally, we show that widespread techniques in online learning, such as improvements for small losses and adaptive tuning of parameters, are all easily captured by our analysis.

2 Preliminaries

For simplicity, we derive our results in the setting of online linear optimization. As we show in the supplementary material, these results can be easily extended to the more general setting of online convex optimization through a standard linearization step.

Online linear optimization may be cast as a repeated game between the *forecaster* and the *environment* as follows. We use Δ_d to denote the simplex $\{\boldsymbol{q} \in [0,1]^d \,:\, \|\boldsymbol{q}\|_1 = 1\}$.

Online linear optimization in the simplex. For each round $t = 1, \dots, T$,

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

The goal of the forecaster is to minimize the accumulated loss, e.g., $\widehat{L}_T = \sum_{t=1}^T \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t$. 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 $\min_{i=1,\dots,T} \sum_{t=1}^T \ell_{i,t}$; or even to compete with a richer class of *sequences* of components. In Section 3 we state more specifically the goals considered in this paper.

We start by introducing our main algorithmic tool, described in Figure 1, a share algorithm whose formulation generalizes the seemingly unrelated formulations of the algorithms studied in [4, 1, 6]. It is parameterized 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[2] for analyzing different instances of this generalized share algorithm.

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

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

Proof. By Hoeffding's inequality (see, e.g., [3, Section A.1.1]),

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

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 implies $\displaystyle\widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t \leqslant \ell_{i,t} + \frac{1}{\eta} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8}$. The proof is concluded by taking a convex aggregation with respect to $\boldsymbol{q}_t$. ☐

[2]We only deal with linear losses in this paper. However, it is straightforward that for sequences of η–exp-concave loss functions, the additional term $\eta/8$ in the bound is no longer needed.

Parameters: learning rate $\eta > 0$ and mixing functions ψ_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

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

   $\underline{V}_{t+1} = \big[v_{i,s}\big]_{1 \leqslant i \leqslant d,\, 1 \leqslant s \leqslant t+1}$ the $d \times (t+1)$ matrix of all past and current pre-weights;
4. [shared update] Define $\widehat{\boldsymbol{p}}_{t+1} = \psi_{t+1}\big(\underline{V}_{t+1}\big)$.

Algorithm 1: The generalized share algorithm.

3 A generalized shifting regret for the simplex

We now introduce a generalized notion of shifting regret which unifies and generalizes the notions of discounted regret (see [3, Section 2.11]), adaptive regret (see [8]), and shifting regret (see [2]). For a fixed horizon T, a sequence of discount factors $\beta_{t,T} > 0$ for $t = 1, \dots, T$ assigns varying weights to the instantaneous losses suffered at each round. We compare the total loss of the forecaster with the loss of an arbitrary sequence of vectors $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T$ in the simplex Δ_d. Our goal is to bound the regret

$$\sum_{t=1}^T \beta_{t,T}\, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \beta_{t,T}\, \boldsymbol{q}_t^\top \boldsymbol{\ell}_t$$

in terms of the "regularity" of the comparison sequence $\boldsymbol{q}_1, \dots, \boldsymbol{q}_T$ and of the variations of the discounting weights $\beta_{t,T}$. By setting $\boldsymbol{u}_t = \beta_{t,T}\, \boldsymbol{q}_t^\top \in \mathbb{R}_+^d$, we can rephrase the above regret as

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1\, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t\,. \tag{2}$$

In the literature on tracking the best expert [4, 5, 1, 6], the regularity of the sequence $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T$ is measured as the number of times $\boldsymbol{u}_t \neq \boldsymbol{u}_{t+1}$. We introduce the following regularity measure

$$m(\boldsymbol{u}_1^T) = \sum_{t=2}^T D_{\mathrm{TV}}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \tag{3}$$

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 $\|\boldsymbol{x} - \boldsymbol{y}\|_1$. The traditional shifting regret of [4, 5, 1, 6] is obtained from (2) when all $\boldsymbol{u}_t$ are such that $\|\boldsymbol{u}_t\|_1 = 1$.

4 Projected update

The shifting variant of the EG algorithm analyzed in [1] is a special case of the generalized share algorithm in which the function ψ_{t+1} performs a projection of the pre-weights on the convex set $\Delta_d^\alpha = [\alpha/d, 1]^d \cap \Delta_d$. Here $\alpha \in (0,1)$ is a fixed parameter. We can prove (using techniques similar to the ones shown in the next section—see the supplementary material) the following bound which generalizes [1, Theorem 16].

Theorem 1. *For all $T \geqslant 1$, for all sequences $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_t \in [0,1]^d$ of loss vectors, and for all $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$, if Algorithm 1 is run with the above update, then*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{\|\boldsymbol{u}_1\|_1 \ln d}{\eta} + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{d}{\alpha} + \left(\frac{\eta}{8} + \alpha\right) \sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \,. \qquad (4)$$

This bound can be optimized by a proper tuning of α and η parameters. We show a similarly tuned (and slightly better) bound in Corollary 1.

5 Fixed-share update

Next, we consider a different instance of the generalized share algorithm corresponding to the update

$$\widehat{p}_{j,t+1} = \sum_{i=1}^d \left(\frac{\alpha}{d} + (1-\alpha)\mathbb{1}_{i=j}\right) v_{i,t+1} = \frac{\alpha}{d} + (1-\alpha) v_{j,t+1}\,, \qquad 0 \leqslant \alpha \leqslant 1 \qquad (5)$$

Despite seemingly different statements, this update in Algorithm 1 can be seen to lead *exactly* to the fixed-share algorithm of [4] for prediction with expert advice. We now show that this update delivers a bound on the regret almost equivalent to (though slightly better than) that achieved by projection on the subset Δ_d^α of the simplex.

Proposition 1. *With the above update, 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 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{\|\boldsymbol{u}_1\|_1 \ln d}{\eta} + \frac{\eta}{8} \sum_{t=1}^T \|\boldsymbol{u}_t\|_1 + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{d}{\alpha} + \frac{\sum_{t=2}^T \|\boldsymbol{u}_t\|_1 - m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{1}{1-\alpha} \,.$$

Note that if we only consider vectors of the form $\boldsymbol{u}_t = \boldsymbol{q}_t = (0, \ldots, 0, 1, 0, \ldots, 0)$ 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 [4, Theorem 1] and [6, Lemma 6] from the much more general Proposition 1.

The fixed-share forecaster does not need to "know" anything in advance about the sequence of the norms $\|\boldsymbol{u}_t\|$ for the bound above to be valid. Of course, in order to minimize the obtained upper bound, the tuning parameters α, η need to be optimized and their values will depend on the maximal values of $m(\boldsymbol{u}_1^T)$ and $\sum_{t=1}^T \|\boldsymbol{u}_t\|_1$ for the sequences one wishes to compete against. This is illustrated in the following corollary, whose proof is omitted. Therein, $h(x) = -x \ln x - (1-x)\ln(1-x)$ denotes the binary entropy function for $x \in [0,1]$. We recall[3] that $h(x) \leqslant x \ln(e/x)$ for $x \in [0,1]$.

Corollary 1. *Suppose Algorithm 1 is run with the update (5). Let $m_0 > 0$ and $U_0 > 0$. 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 sequences $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$ with $\|\boldsymbol{u}_1\|_1 + m(\boldsymbol{u}_1^T) \leqslant m_0$ and $\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \leqslant U_0$,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \sqrt{\frac{U_0}{2}\left(m_0 \ln d + U_0\, h\!\left(\frac{m_0}{U_0}\right)\right)} \leqslant \sqrt{\frac{U_0\, m_0}{2}\left(\ln d + \ln\left(\frac{e\,U_0}{m_0}\right)\right)}$$

whenever η and α are optimally chosen in terms of m_0 and U_0.

Proof of Proposition 1. 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 $\boldsymbol{u}_t \in \mathbb{R}_+^d$

$$\|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta} \sum_{i=1}^d u_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8} \|\boldsymbol{u}_t\|_1 \,. \qquad (6)$$

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

We now examine

$$\sum_{i=1}^{d} u_{i,t} \ln \frac{v_{i,t+1}}{\widehat{p}_{i,t}} = \sum_{i=1}^{d} \left(u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) + \sum_{i=1}^{d} \left(u_{i,t-1} \ln \frac{1}{v_{i,t}} - u_{i,t} \ln \frac{1}{v_{i,t+1}} \right) . \tag{7}$$

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

$$\begin{aligned}\sum_{i=1}^{d} \left(u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) = & \sum_{i\,:\,u_{i,t} \geqslant u_{i,t-1}} \left((u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \right) \\ & + \sum_{i\,:\,u_{i,t} < u_{i,t-1}} \left(\underbrace{(u_{i,t} - u_{i,t-1}) \ln \frac{1}{v_{i,t}}}_{\leqslant 0} + u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \right). \end{aligned} \tag{8}$$

In view of the update (5), we have $1/\widehat{p}_{i,t} \leqslant d/\alpha$ and $v_{i,t}/\widehat{p}_{i,t} \leqslant 1/(1-\alpha)$. Substituting in (8), we get

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

The sum of the second term in (7) telescopes. Substituting the obtained bounds in the first sum of the right-hand side in (7), and summing over $t = 2, \ldots, T$, leads to

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

We hence get from (6), which we use in particular for $t = 1$,

$$\begin{aligned} \sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta} \sum_{i=1}^{d} u_{i,1} \ln \frac{1}{\widehat{p}_{i,1}} + \frac{\eta}{8} \sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \\ + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{d}{\alpha} + \frac{\sum_{t=2}^{T} \|\boldsymbol{u}_t\|_1 \, m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{1}{1-\alpha} . \end{aligned}$$

□

6 Applications

We now show how our regret bounds can be specialized to obtain bounds on adaptive and discounted regret, and on regret with time-selection functions. We show regret bounds just for the specific instance of the generalized share algorithm using update (5).

Adaptive regret, introduced by [8], can be viewed as a variant of discounted regret where the monotonicity assumption is dropped. For $\tau_0 \in \{1, \ldots, T\}$, the τ_0-*adaptive regret* of a forecaster is defined by

$$\mathcal{R}_T^{\tau_0 - \mathrm{adapt}} = \max_{\substack{[r,\,s] \subset [1,\,T] \\ s+1-r \leqslant \tau_0}} \left\{ \sum_{t=r}^{s} \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \min_{\boldsymbol{q} \in \Delta_d} \sum_{t=r}^{s} \boldsymbol{q}^\top \boldsymbol{\ell}_t \right\} . \tag{9}$$

The fact that this is a special case of (2) clearly emerges from the proof of Corollary 2 below here.

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). In this paper we note 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, i.e., Proposition 1. The work [8] essentially considered the case of online convex optimization with exp-concave loss function; in case of general convex functions, they also mentioned that the greedy projection forecaster of [2] enjoys adaptive regret guarantees. This is obtained in much the same way as we obtain an adaptive regret bound for the fixed-share forecaster in the next result.

Corollary 2. *Suppose that Algorithm 1 is run with the shared update (5). 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 $\tau_0 \in \{1, \dots, T\}$,*

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

whenever η and α are chosen optimally (depending on τ_0 and T).

Proof. For $1 \leqslant r \leqslant s \leqslant T$ and $\boldsymbol{q} \in \Delta_d$, the regret in the right-hand side of (9) equals the regret considered in Proposition 1 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 $D_{\text{TV}}(\boldsymbol{u}_r, \boldsymbol{u}_{r-1}) = D_{\text{TV}}(\boldsymbol{q}, \mathbf{0}) = 1$ and $D_{\text{TV}}(\boldsymbol{u}_{s+1}, \boldsymbol{u}_s) = D_{\text{TV}}(\mathbf{0}, \boldsymbol{q}) = 0$ 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, $m(\boldsymbol{u}_1^T) + \|\boldsymbol{u}_1\|_1 = 1$, that is, $m_0 = 1$. Specializing the bound of Proposition 1 with the additional choice $U_0 = T$ gives the result. □

Discounted regret, introduced in [3, Section 2.11], is defined by

$$\max_{\boldsymbol{q} \in \Delta_d} \sum_{t=1}^{T} \beta_{t,T} \left(\widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{q}^\top \boldsymbol{\ell}_t \right) . \tag{10}$$

The discount factors $\beta_{t,T}$ measure the relative importance of more recent losses to older losses. For instance, for a given horizon T, the discounts $\beta_{t,T}$ may be larger as t is closer to T. On the contrary, in a game-theoretic setting, the earlier losses may matter more then the more recent ones (because of interest rates), in which case $\beta_{t,T}$ would be smaller as t gets closer to T. We mostly consider below monotonic sequences of discounts (both non-decreasing and non-increasing). Up to a normalization, we assume that all discounts $\beta_{t,T}$ are in $[0,1]$. As shown in [3], a minimal requirement to get non-trivial bounds is that the sum of the discounts satisfies $U_T = \sum_{t \leqslant T} \beta_{t,T} \to \infty$ as $T \to \infty$.

A natural objective is to show that the quantity in (10) is $o(U_T)$, for instance, by bounding it by something of the order of $\sqrt{U_T}$. We claim that Corollary 1 does so, at least whenever the sequences $(\beta_{t,T})$ are monotonic for all T. To support this claim, we only need to show that $m_0 = 1$ is a suitable value to deal with (10). Indeed, for all $T \geqslant 1$ and for all $\boldsymbol{q} \in \Delta_d$, the measure of regularity involved in the corollary satisfies

$$\|\beta_{1,T}\boldsymbol{q}\|_1 + m\big((\beta_{t,T}\boldsymbol{q})_{t \leqslant T}\big) \;=\; \beta_{1,T} + \sum_{t=2}^{T} \big(\beta_{t,T} - \beta_{t-1,T}\big)_+ \;=\; \max\big\{\beta_{1,T},\, \beta_{T,T}\big\} \;\leqslant\; 1\,,$$

where the second equality follows from the monotonicity assumption on the discounts.

The values of the discounts for all t and T are usually known in advance. However, the horizon T is not. Hence, a calibration issue may arise. The online tuning of the parameters α and η shown in Section 7.3 entails a forecaster that can get discounted regret bounds of the order $\sqrt{U_T}$ for all T. The fundamental reason for this is that the discounts only come in the definition of the fixed-share forecaster via their sums. In contrast, the forecaster discussed in [3, Section 2.11] weighs each instance t directly with $\beta_{t,T}$ (i.e., in the very definition of the forecaster) and enjoys therefore no regret guarantees for horizons other than T (neither before T nor after T). Therein, the knowledge of the horizon T is so crucial that it cannot be dealt with easily, not even with online calibration of the parameters or with a doubling trick. We insist that for the fixed-share forecaster, much flexibility is gained as some of the discounts $\beta_{t,T}$ can change in a drastic manner for a round T to values $\beta_{t,T+1}$ for the next round. As for the comparison to the setting of discounted losses of [9], we note that the latter can be cast as a special case of our setting (since the discounting weights take the special form $\beta_{t,T} = \gamma_t \dots \gamma_{T-1}$ therein, for some sequence γ_s of positive numbers). In particular, the fixed-share forecaster can satisfy the bound stated in [9, Theorem 2], for instance, by using the online tuning techniques of Section 7.3.

A final reference to mention is the setting of time-selection functions of [10, Section 6], which basically corresponds to knowing in advance the weights $\|\boldsymbol{u}_t\|_1$ of the comparison sequence $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T$ the forecaster will be evaluated against. We thus generalize their results as well.

7 Refinements and extensions

We now show that a set of techniques for refining the standard online analysis can be easily applied to our framework. In particular, we focus on the following: improvement for small losses, sparse target sequences, and dynamic tuning of parameters. Not all of them where within reach of previous analyses.

7.1 Improvement for small losses

The regret bounds of the fixed-share forecaster can be significantly improved when the cumulative loss of the best sequence of experts is small. The next result improves on Corollary 1 whenever $L_0 \ll U_0$. For concreteness, we focus on the fixed-share update (5).

Corollary 3. *Suppose Algorithm 1 is run with the update (5). Let $m_0 > 0$, $U_0 > 0$, and $L_0 > 0$. 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 sequences $\boldsymbol{u}_1, \dots, \boldsymbol{u}_T \in \mathbb{R}_+^d$ with $\|\boldsymbol{u}_1\|_1 + m(\boldsymbol{u}_1^T) \leqslant m_0$, $\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \leqslant U_0$, and $\sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant L_0$,*

$$\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^{T} \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \sqrt{L_0 \, m_0 \left(\ln d + \ln\left(\frac{e\,U_0}{m_0}\right)\right)} + \ln d + \ln\left(\frac{e\,U_0}{m_0}\right)$$

whenever η and α are optimally chosen in terms of m_0, U_0, and L_0.

Here again, the parameters α and η may be tuned online using the techniques shown in Section 7.3. The above refinement is obtained by mimicking the analysis of Hedge forecasters for small losses (see, e.g., [3, Section 2.4]). In particular, one should substitute Lemma 1 with the following lemma in the analysis carried out in Section 5; its proof follows from the mere replacement of Hoeffding's inequality by [3, Lemma A.3], which states that for all $\eta \in \mathbb{R}$ and for all random variable X taking values in $[0,1]$, one has $\ln \mathbb{E}[e^{-\eta X}] \leqslant (e^{-\eta} - 1)\mathbb{E}X$.

Lemma 2. *For all $t \geqslant 1$ and $\boldsymbol{q}_t \in \Delta_d$, Algorithm 1 satisfies*

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

7.2 Sparse target sequences

The work [6] introduced 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 (3) 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 $S \subseteq \{1, \dots, d\}$, we have $n(\boldsymbol{q}_1^T) \leqslant |S|$, the cardinality of S. Also, $n(\boldsymbol{q}_1^T) \leqslant \bigl|\{\boldsymbol{q}_t, \;\; t = 1, \dots, T\}\bigr|$, the cardinality of the pool of convex combinations. Thus, $n(\boldsymbol{u}_1^T)$ generalizes the notion of sparsity of [6].

Here we consider a family of shared updates of the form

$$\widehat{p}_{j,t} = (1-\alpha) v_{j,t} + \alpha \frac{w_{j,t}}{Z_t}\,, \qquad 0 \leqslant \alpha \leqslant 1\,, \tag{11}$$

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 [6, Sections 3 and 5.2]. Apart from generalizing the regret bounds of [6], 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 $j = 1, \ldots, d$ and $t = 1, \ldots, T$,

$$v_{j,t} \leqslant w_{j,t} \leqslant 1 \qquad \text{and} \qquad C\, w_{j,t+1} \geqslant w_{j,t}\,. \tag{12}$$

The next result improves on Proposition 1 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. Its proof can be found in the supplementary material; it is a variation on the proof of Proposition 1.

Proposition 2. *Suppose Algorithm 1 is run with the shared update (11) with weights satisfying the conditions* (12)*. Then 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 sequences $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

$$\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 \, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \sum_{t=1}^T \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{n(\boldsymbol{u}_1^T)\ln d}{\eta} + \frac{n(\boldsymbol{u}_1^T)\, T \ln C}{\eta} + \frac{\eta}{8}\sum_{t=1}^T \|\boldsymbol{u}_t\|_1 + \frac{m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{\max_{t \leqslant T} Z_t}{\alpha} + \frac{\sum_{t=2}^T \|\boldsymbol{u}_t\|_1 - m(\boldsymbol{u}_1^T)}{\eta} \ln \frac{1}{1-\alpha}\,.$$

Corollaries 8 and 9 of [6] can now be generalized (and even improved); we do so—in the supplementary material—by showing two specific instances of the generic update (11) that satisfy (12).

7.3 Online tuning of the parameters

The forecasters studied above need their parameters η and α to be tuned according to various quantities, including the time horizon T. We show here how the trick of [11] 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 (5). We respectively replace steps 3 and 4 of its description by the loss and shared updates

$$v_{j,t+1} = \frac{\widehat{p}_{j,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{j,t}}}{\sum_{i=1}^d \widehat{p}_{i,t}^{\frac{\eta_t}{\eta_{t-1}}} e^{-\eta_t \ell_{i,t}}} \qquad \text{and} \qquad p_{j,t+1} = \frac{\alpha_t}{d} + (1-\alpha_t)\, v_{j,t+1}\,, \tag{13}$$

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

Proposition 3. *The forecaster based on the updates* (13) *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 $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$ of loss vectors $\boldsymbol{\ell}_t \in [0,1]^d$, and for all $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T \in \mathbb{R}_+^d$,*

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

Due to space constraints, we provide an illustration of this bound only in the supplementary material.

References

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

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

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

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

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

[6] 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.

[7] 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.

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

[9] A. Chernov and F. Zhdanov. Prediction with expert advice under discounted loss. In *Proceedings of the 21st International Conference on Algorithmic Learning Theory, ALT 2010*, pages 255–269. Springer, 2008.

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

[11] 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 Online convex optimization on the simplex

By using a standard reduction, the results of the main body of the paper (for linear optimization on the simplex) 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 ℓ_t leads to the desired bounds. Indeed, by the convexity of ℓ_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

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

B Proof of Proposition 2; application of the bound to two different updates

Proof. The beginning and the end of the proof are similar to the one of Proposition 1, as they do not depend on the specific weight update. In particular, inequalities (6) and (7) remain the same. The proof is modified after (8), which this time we upper bound using the first condition in (12),

$$\begin{aligned}\sum_{i=1}^d \left(u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) = & \sum_{i\,:\,u_{i,t} \geqslant u_{i,t-1}} (u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \\ & + \sum_{i\,:\,u_{i,t} < u_{i,t-1}} \underbrace{(u_{i,t} - u_{i,t-1})}_{\leqslant 0} \underbrace{\ln \frac{1}{v_{i,t}}}_{\geqslant \ln(1/w_{i,t})} + u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \,. \qquad (14)\end{aligned}$$

By definition of the shared update (11), 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 (14) by

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

Proceeding as in the end of the proof of Proposition 1, 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

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

where we used $C \geqslant 1$ for the first inequality and the second condition in (12) for the second inequality. The proof is concluded by noting that (12) entails $w_{i,2} \geqslant (1/C) w_{i,1} \geqslant (1/C) v_{i,1} = 1/(dC)$ and that the coefficient $\max_{t=1,\ldots,T} u_{i,t} - u_{i,1}$ in front of $\ln(1/w_{i,2})$ is nonnegative. □

The first update uses $w_{j,t} = \max_{s \leqslant t} v_{j,s}$. Then (12) 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 4. *Suppose Algorithm 1 is run with the update* (11) *with* $w_{j,t} = \max_{s \leqslant t} v_{j,s}$. *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* $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_T \in \Delta_d$,

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

The second update we discuss uses $w_{j,t} = \max_{s \leqslant t} e^{\gamma(s-t)} v_{j,s}$ in (11) for some $\gamma > 0$. Both conditions in (12) 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 2 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 γ 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 5. *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*

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

As the factors $e^{-\gamma t}$ cancel out in the numerator and denominator of the ratio in (11), 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 [6], 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 5 as an additional factor of $m_0 \ln(1 + \ln T)$ is present in [6, Corollary 9].

C Proof of Proposition 3; illustration of the obtained bound

We first adapt Lemma 1.

Lemma 3. *The forecaster based on the loss and shared updates* (13) *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 (13), for all $i \in \{1, \ldots, d\}$,

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

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 $i \in \{1, \ldots, d\}$. The proof is concluded by taking a convex aggregation over i with respect to $\boldsymbol{q}_t$. □

The proof of Proposition 3 follows the steps of the one of Proposition 1; we sketch it below.

Proof of Proposition 3. Applying Lemma 3 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 $\boldsymbol{u}_t \in \mathbb{R}_+^d$,

$$\begin{aligned} \|\boldsymbol{u}_t\|_1\, \widehat{\boldsymbol{p}}_t^\top \boldsymbol{\ell}_t - \boldsymbol{u}_t^\top \boldsymbol{\ell}_t \leqslant \frac{1}{\eta_{t-1}} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{v_{i,t+1}} \\ + \|\boldsymbol{u}_t\|_1 \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \ln d + \frac{\eta_{t-1}}{8}\, \|\boldsymbol{u}_t\|_1\,. \end{aligned} \tag{15}$$

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

$$\begin{aligned} & \frac{1}{\eta_{t-1}} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{v_{i,t+1}} \\ & \quad = \frac{1}{\eta_{t-1}} \sum_{i=1}^{d} \left(u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) + \sum_{i=1}^{d} \left(\frac{u_{i,t-1}}{\eta_{t-1}} \ln \frac{1}{v_{i,t}} - \frac{u_{i,t}}{\eta_t} \ln \frac{1}{v_{i,t+1}} \right)\,, \end{aligned} \tag{16}$$

where the first difference in the right-hand side can be bounded as in (8) by

$$\begin{aligned} & \sum_{i=1}^{d} \left(u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - u_{i,t-1} \ln \frac{1}{v_{i,t}} \right) \\ & \quad \leqslant \sum_{i\,:\, u_{i,t} \geqslant u_{i,t-1}} \left((u_{i,t} - u_{i,t-1}) \ln \frac{1}{\widehat{p}_{i,t}} + u_{i,t-1} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \right) + \sum_{i\,:\, u_{i,t} < u_{i,t-1}} u_{i,t} \ln \frac{v_{i,t}}{\widehat{p}_{i,t}} \end{aligned}$$

$$\leqslant \quad D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \ln \frac{d}{\alpha_t} + \big(\|\boldsymbol{u}_t\|_1 - D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1})\big) \ln \frac{1}{1-\alpha_t}$$

$$\leqslant \quad D_{TV}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1}) \ln \frac{d(1-\alpha_T)}{\alpha_T} + \|\boldsymbol{u}_t\|_1 \ln \frac{1}{1-\alpha_t}\,, \tag{17}$$

where we used for the second inequality that the shared update in (13) 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 (16) over $t = 2, \ldots, T$ using (17) and the fact that $\eta_t \geqslant \eta_T$, we get

$$\sum_{t=2}^{T} \left(\frac{1}{\eta_{t-1}} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{\widehat{p}_{i,t}} - \frac{1}{\eta_t} \sum_{i=1}^{d} u_{i,t} \ln \frac{1}{v_{i,t+1}} \right)$$

$$\leqslant \quad \frac{m(\boldsymbol{u}_1^T)}{\eta_T} \ln \frac{d(1-\alpha_T)}{\alpha_T} + \sum_{t=2}^{T} \frac{\|\boldsymbol{u}_t\|_1}{\eta_{t-1}} \ln \frac{1}{1-\alpha_t} + \sum_{i=1}^{d} \bigg(\frac{u_{i,1}}{\eta_1} \ln \frac{1}{v_{i,2}} - \underbrace{\frac{u_{i,T}}{\eta_T} \ln \frac{1}{v_{i,T+1}}}_{\geqslant 0} \bigg).$$

An application of (15) —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. $\square$

We now instantiate the obtained bound to the case of, e.g., T–adaptive regret guarantees, when T is unknown and/or can increase without bounds.

Corollary 6. *The forecaster based on the updates discussed above 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 $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$ 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 (α_t) and (η_t) are non-increasing, as desired. For a given pair (r,s) and a given $\boldsymbol{q} \in \Delta_d$, we consider the sequence ν_1^T defined in the proof of Corollary 2; it satisfies that $m(\boldsymbol{u}_1^T) \leqslant 1$ and $\|\boldsymbol{u}_t\|_1 \leqslant 1$ for all $t \geqslant 1$. Therefore, Proposition 3 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 η_t and to note that

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

D Proof of Theorem 1

We recall that the forecaster at hand is the one described in Algorithm 1, with the shared update $\widehat{\boldsymbol{p}}_{t+1} = \psi_{t+1}\big(\underline{V}_{t+1}\big)$ for

$$\psi_{t+1}\big(\underline{V}_{t+1}\big) \in \operatorname*{argmin}_{\boldsymbol{x} \in \Delta_d^\alpha} \mathcal{K}(\boldsymbol{x}, \boldsymbol{v}_{t+1})\,, \qquad \text{where} \quad \mathcal{K}(\boldsymbol{x}, \boldsymbol{v}_{t+1}) = \sum_{i=1}^{d} x_i \ln \frac{x_i}{v_{i,t+1}} \tag{18}$$

is the Kullback-Leibler divergence and $\Delta_d^\alpha = [\alpha/d, 1]^d \cap \Delta_d$ is the simplex of convex vectors with the constraint that each component be larger than α/d.

The proof of the performance bound starts with an extension of Lemma 1.

Lemma 4. *For all $t \geqslant 1$ and for all $\boldsymbol{q}_t \in \Delta_d^\alpha$, the generalized forecaster with the shared update* (18) *satisfies*

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

Proof. We rewrite the bound of Lemma 1 in terms of Kullback-Leibler divergences,

$$
\begin{aligned}
(\widehat{\boldsymbol{p}}_t - \boldsymbol{q}_t)^\top \boldsymbol{\ell}_t \quad &\leqslant \quad \frac{1}{\eta}\sum_{i=1}^{d} q_{i,t} \ln \frac{v_{i,t+1}}{p_{i,t}} + \frac{\eta}{8} = \frac{\mathcal{K}(\boldsymbol{q}_t, \widehat{\boldsymbol{p}}_t) - \mathcal{K}(\boldsymbol{q}_t, \boldsymbol{v}_{t+1})}{\eta} + \frac{\eta}{8} \\
&\leqslant \quad \frac{\mathcal{K}(\boldsymbol{q}_t, \widehat{\boldsymbol{p}}_t) - \mathcal{K}(\boldsymbol{q}_t, \widehat{\boldsymbol{p}}_{t+1})}{\eta} + \frac{\eta}{8} = \frac{1}{\eta}\sum_{i=1}^{d} q_{i,t} \ln \frac{\widehat{p}_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8} \,,
\end{aligned}
$$

where the last inequality holds by applying a generalized Pythagorean theorem for Bregman divergences (here, the Kullback-Leibler divergence) —see, e.g., [3, Lemma 11.3]. □

Proof. Let $\boldsymbol{q}_t = \frac{\alpha}{d} + (1-\alpha)\frac{\boldsymbol{u}_t}{\|\boldsymbol{u}_t\|_1} \in \Delta_d^\alpha$. We have by rearranging the terms for all t,

$$
\begin{aligned}
\big(\|\boldsymbol{u}_t\|_1 \widehat{\boldsymbol{p}}_t - \boldsymbol{u}_t\big)^\top \boldsymbol{\ell}_t &= \|\boldsymbol{u}_t\|_1 (\widehat{\boldsymbol{p}}_t - \boldsymbol{q}_t)^\top \boldsymbol{\ell}_t + \left(\frac{\alpha}{d}\|\boldsymbol{u}_t\|_1 - \alpha \boldsymbol{u}_t\right)^\top \boldsymbol{\ell}_t \\
&\leqslant \|\boldsymbol{u}_t\|_1 (\widehat{\boldsymbol{p}}_t - \boldsymbol{q}_t)^\top \boldsymbol{\ell}_t + \alpha \|\boldsymbol{u}_t\|_1 \;.
\end{aligned}
$$

Therefore, by applying Lemma 4 with $\boldsymbol{q}_t \in \Delta_d^\alpha$, we further upper bound the quantity of interest as

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

The upper bound is rewritten by summing over t and applying an Abel transform to its first term,

$$
\begin{aligned}
&\sum_{t=1}^{T} \frac{\|\boldsymbol{u}_t\|_1}{\eta}\sum_{i=1}^{d} q_{i,t} \ln \frac{\widehat{p}_{i,t+1}}{\widehat{p}_{i,t}} + \frac{\eta}{8}\|\boldsymbol{u}_t\|_1 + \alpha \|\boldsymbol{u}_t\|_1 \\
= \quad & \frac{\|\boldsymbol{u}_1\|_1 \ln d}{\eta} + \frac{\|\boldsymbol{u}_T\|_1}{\eta} \underbrace{\sum_{i=1}^{d} q_{i,T} \ln \widehat{p}_{i,T+1}}_{\leqslant 0} + \frac{1}{\eta}\sum_{t=2}^{T}\sum_{i=1}^{d} \underbrace{\big(\|\boldsymbol{u}_t\|_1 q_{i,t} - \|\boldsymbol{u}_{t-1}\| q_{i,t-1}\big)}_{=(1-\alpha)(u_{i,t}-u_{i,t-1})} \underbrace{\ln \frac{1}{\widehat{p}_{i,t}}}_{0 \leqslant \,\cdot\, \leqslant \ln \frac{d}{\alpha}} \\
&+ \left(\frac{\eta}{8} + \alpha\right)\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \\
\leqslant \quad & \frac{\|\boldsymbol{u}_1\|_1 \ln d}{\eta} + \frac{1-\alpha}{\eta}\left(\sum_{t=2}^{T} D_{\mathrm{TV}}(\boldsymbol{u}_t, \boldsymbol{u}_{t-1})\right) \ln \frac{d}{\alpha} + \left(\frac{\eta}{8} + \alpha\right)\sum_{t=1}^{T} \|\boldsymbol{u}_t\|_1 \;.
\end{aligned}
$$

□