

# Damped Online Newton Step for Portfolio Selection

**Zakaria Mhammedi**

*Massachusetts Institute of Technology*

MHAMMEDI@MIT.EDU

**Alexander Rakhlin**

*Massachusetts Institute of Technology*

RAKHLIN@MIT.EDU

**Editors:** Po-Ling Loh and Maxim Raginsky

## Abstract

We revisit the classic online portfolio selection problem, where at each round a learner selects a distribution over a set of portfolios to allocate its wealth. It is known that for this problem a logarithmic regret with respect to Cover’s loss is achievable using the Universal Portfolio Selection algorithm, for example. However, all existing algorithms that achieve a logarithmic regret for this problem have per-round time and space complexities that scale polynomially with the total number of rounds, making them impractical. In this paper, we build on the recent work by Luo et al. (2018) and present the first practical online portfolio selection algorithm with a logarithmic regret and whose per-round time and space complexities depend only logarithmically on the horizon. Behind our approach are two key technical novelties of independent interest. We first show that the Damped Online Newton steps can approximate mirror descent iterates well, even when dealing with time-varying regularizers. Second, we present a new meta-algorithm that achieves an adaptive logarithmic regret (i.e. a logarithmic regret on any sub-interval) for mixable losses.

## 1. Introduction

In this paper, we consider the problem of online portfolio selection where, at each round  $t$ , a learner chooses a distribution  $\mathbf{p}_t \in \Delta_d$  over a fixed set of  $d$  portfolios. Then, the environment reveals a return vector  $\mathbf{r}_t \in \mathbb{R}_{\geq 0}^d$ , and the learner suffers a loss  $\ell_t(\mathbf{p}_t) := -\ln\langle \mathbf{p}_t, \mathbf{r}_t \rangle$ . The goal of the learner is to minimize the regret  $\mathcal{R}_T(\mathbf{u}) := \sum_{t=1}^T (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}))$  after  $T \geq 1$  rounds, which is the difference between the cumulative loss of the learner minus that of any distribution  $\mathbf{u}$  over portfolios. For this problem, it is known that Cover’s Universal Portfolio Algorithm (UPA) (Cover, 1991) guarantees the optimal  $O(d \ln T)$  regret bound. One implication of this is that if a distribution  $\mathbf{u}$  has an exponential return growth rate with constant  $\lambda > 0$ , i.e.  $\prod_{t \in [T]} \langle \mathbf{u}, \mathbf{r}_t \rangle \propto e^{\lambda T}$ , then the total return of UPA also has an exponential growth rate with constant at least  $\lambda - O(d \ln(T)/T)$ .

The main shortcoming of the UPA is that the expression of its outputs involves multi-variate integrals that make its implementation impractical. One way of approximating these integrals is via log-concave sampling as done by Kalai and Vempala (2002). The algorithm of the latter has a computational complexity of order  $O(d^4 T^{15})$ , measured after  $T$  rounds. Even though this computational complexity can be reduced using more modern log-concave sampling methods (see e.g. (Narayanan and Rakhlin, 2017; Bubeck et al., 2018)), it remains a large polynomial of  $T$ , making these approaches impractical.

It is possible to use other more efficient online learning algorithms for the portfolio selection problem. Algorithms such as Online Gradient Descent (Zinkevich, 2003), Online Newton Step (Hazan et al., 2007), and Exponentiated Gradients (Helmbold et al., 1998) all have regret bounds that scale with the largest observed gradient norm  $G$ . One way to ensure that the gradients are bounded in

Algorithm	Regret	Run-Time	Space Comp.	References
Universal Portfolio	$d \ln T$	$d^4 T^{15}$	$dT$	(Cover, 1991) (Kalai and Vempala, 2002)
ONS	$Gd \ln T$	$d^{3.5} T$	$d^2$	(Hazan et al., 2007)
FTRL	$G^2 d \ln(dT)$	$d^{2.5} T^2$	$dT$	(Agarwal and Hazan, 2005)
EG	$G\sqrt{T \ln d}$	$dT$	$d$	(Helmbold et al., 1998)
Soft-Bayes	$\sqrt{dT \ln d}$	$dT$	$d$	(Orseau et al., 2017)
Ada-BARRONS	$d^2 \ln^4 T$	$d^{2.5} T^2 + d^{3.5} T$	$d^2 T$	(Luo et al., 2018)
PAE + DONS	$d^2 \ln^5 T$	$d^3 T \ln^2 T$	$d^2 \ln^2 T$	(this work—Thm. 17)

Table 1: Result comparison.

the online portfolio setting is to mix the outputs of such algorithms with a small amount of uniform distribution. This approach leads to a regret bound of order  $d\sqrt{T \ln d}$  in the best case (which is not logarithmic in  $T$ ), even after optimizing for the amount of uniform distribution used. The Soft-Bayes algorithm (Orseau et al., 2017) provides a  $\sqrt{d}$  improvement over this regret bound. Finally, Agarwal and Hazan (2005) showed that the Follow-the-Regularized-Leader (FTRL) algorithm achieves a regret bound of order  $O(G^2 d \ln(dT))$ ,  $G$  (the largest gradient norm) may be merely an artifact of the analysis. In the concurrent work by Zimmert et al. (2022) it was shown that the FTRL algorithm cannot guarantee a  $O(d \ln(dT))$  regret (i.e. without the dependence on the largest gradient norm  $G$ ), refuting a conjecture by Van Erven et al. (2020). Table 1 compares the regret bounds and computational complexities of the different algorithms mentioned here.

Among known algorithms that achieve a logarithmic regret in the online portfolio setting, Ada-BARRONS (Luo et al., 2018) is the best in terms of computational cost (see Tab. 1). Ada-BARRONS consists of I) a base algorithm that is essentially mirror descent with a log-barrier plus a quadratic regularizer with a parameter  $\beta$ ; and II) a meta-algorithm that implements a clever restart scheme to learn the parameter  $\beta$  and achieve a logarithmic regret. The algorithmic idea behind Ada-BARRONS can be traced back to the problem of combining bandit algorithms (Agarwal et al., 2017; Wei and Luo, 2018), where the use of a non-decreasing learning rate schedule is used to extract crucial negative terms in the regret analysis of mirror descent (see §3).

The main drawback of Ada-BARRONS is that its time [resp. space] complexity is quadratic [resp. linear] in the total number of rounds (see Table 1). Though Ada-BARRONS has a substantially better computational complexity compared to approaches based on log-concave sampling, it is still not a practical algorithm when the horizon is large. The main reason for the quadratic time complexity is the restarts of Ada-BARRONS, which require computing the regularized leader at each round. Luo et al. (2018) posed the question of whether there exists an algorithm that improves on either the regret or the computational complexity of Ada-BARRONS without hurting the other.

**Contributions.** We answer the above question in the positive by presenting an online algorithm for portfolio selection with a logarithmic regret and that has near constant per-round time and space complexities. Behind our solution are two techniques of independent interest in online learning. We first show that one can use the *damped Newton steps* (Nesterov et al., 2018) to approximate the mirror descent iterates in Ada-BARRONS without sacrificing the logarithmic regret. Here, existing results due to Abernethy et al. (2012) do not apply (due to time-varying regularizers in the mirror descent objective—see §3.1). Even if they did, they would lead a suboptimal  $O(\sqrt{T})$  regret, and so

a new analysis is needed, which we provide. Using online damped Newton steps confers a  $O(\sqrt{d})$  improvement in the computational cost.

The second and crucial tool we use is a meta-algorithm that achieves an adaptive logarithmic regret for exp-concave losses (Hazan and Seshadhri, 2007; Zhang et al., 2019); that is, an algorithm that achieves a logarithmic regret on any interval  $I \subseteq [T]$ , whenever the losses are exp-concave (we also generalize the algorithm to work with mixable losses). Thanks to a novel analysis, we show that using such a meta-algorithm removes the need for computing the regularized leader, which is required by Ada-BARRONS. This further improves the time and space complexities by an  $\tilde{O}(T)$  factor, leading to our final algorithm that has  $O(d^3 T \ln^2 T)$  and  $O(d^2 \ln^2 T)$  total time and space complexities, respectively.

The techniques we develop are transferable to another prominent online learning problem; that of learning linear models with the log-loss (Rakhlin and Sridharan, 2015, Section 6).

**Limitation.** Our new analysis of the damped Newton Step requires a small learning rate which leads to a large constant in the regret bound of our final algorithm. This, of course, affects the practicality of our proposed method. It is unclear whether this is a true limitation of the damped Newton Step or merely an artifact of our current analysis. We leave such investigation for future work. Nevertheless, we emphasize that our method for avoiding the computation of the regularized leader in Ada-BARRONS, which costs  $O(T)$  arithmetic operations per round, need not be used together with the damped Newton Step specifically. In fact, when used with BARRONS—the base algorithm of Ada-BARRONS (see §3.1)—it already confers a substantial computational improvement over Ada-BARRONS (by shaving off an  $\tilde{O}(T)$  factor in the computational complexity) without sacrificing the regret bound of Ada-BARRONS by much (the regret bound is at most a  $3 \ln T$  factor worse).

Finally, we acknowledge the concurrent work by Zimmert et al. (2022) that achieve similar results to ours for the Portfolio Selection problem, albeit with a completely different method.

**Outline.** In §2, we introduce the notation and definitions we need. We also include some results on self-concordance that we require in our analysis. In §3, we describe the Ada-BARRONS algorithm in more detail and highlight the challenges involved in the design of an efficient alternative. There, we also outline our solution and give a sketch of why it works. Finally, in §4, we present the full details of our algorithm and its guarantee. The proofs are deferred to the appendix.

## 2. Preliminaries

We define the set  $\mathcal{C}_{d-1} := \{\mathbf{u} \in \mathbb{R}_{\geq 0}^{d-1} : \langle \mathbf{1}, \mathbf{u} \rangle \leq 1\}$ . Throughout, for any  $\mathbf{v} \in \mathcal{C}_{d-1}$ , we denote

$$\mathbf{v}' := (1 - 1/T)\mathbf{v} + \mathbf{1}/(dT), \quad \text{and} \quad \bar{\mathbf{v}} := \mathbf{e}_d + J^T \mathbf{v}, \quad \text{where} \quad J := \begin{bmatrix} I & -\mathbf{1} \end{bmatrix}.$$

We may combine the notation and write  $\bar{\mathbf{v}}' := (1 - 1/T)\bar{\mathbf{v}} + \mathbf{1}/(dT)$  and  $\mathbf{v}'' := (1 - 1/T)\mathbf{v}' + \mathbf{1}/(dT)$ . We will be working with Cover's loss  $\ell_t$ , which for a return vector  $\mathbf{r}_t \in \mathbb{R}_{\geq 0}^d$ , is given by

$$\forall \mathbf{u} \in \mathcal{C}_{d-1}, \quad \ell_t(\mathbf{u}) := -\ln \langle \mathbf{r}_t, \bar{\mathbf{u}} \rangle. \tag{1}$$

Our goal is to design an efficient algorithm whose outputs  $(\mathbf{u}_t)$  are such that the regret

$$\text{Regret}_T(\mathbf{u}) := \sum_{t=1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u})) = \ln \frac{\prod_{t=1}^T \langle \mathbf{r}_t, \bar{\mathbf{u}} \rangle}{\prod_{t=1}^T \langle \mathbf{r}_t, \bar{\mathbf{u}}_t \rangle}$$

against any comparator  $\mathbf{u} \in \mathcal{C}_{d-1}$  is bounded by a poly-logarithmic factor in  $T$ . Since the regret is invariant to the scale of  $(\mathbf{r}_t)$ , we may assume without loss of generality that  $(\mathbf{r}_t) \subset [0, 1]^d$ . The next lemma (taken from (Luo et al., 2018, Lemma 10)), implies that a regret against a comparator  $\mathbf{u} \in \mathcal{C}_{d-1}$  is bounded by the regret against  $\mathbf{u}'$  up to an additive factor—this will be useful throughout:

**Lemma 1** *For any  $\mathbf{u} \in \mathcal{C}_{d-1}$  and  $\mathbf{u}' = (1 - \frac{1}{T})\mathbf{u} + \frac{1}{dT}\mathbf{1}$ , we have  $\sum_{t=1}^T \ell_t(\mathbf{u}) \leq \sum_{t=1}^T \ell_t(\mathbf{u}') + 2$ .*

**Self-Concordant Functions.** We now present some results on self-concordant functions that we will make heavy use of in the proofs of our results. We start by the definition of a self-concordant function. For the rest of this section, we let  $\mathcal{K}$  be convex compact set with non-empty interior  $\text{int } \mathcal{K}$ . For a twice [resp. thrice] differentiable function, we let  $\nabla^2 f(\mathbf{u})$  [resp.  $\nabla^3 f(\mathbf{u})$ ] be the Hessian [resp. third derivative tensor] of  $f$  at  $\mathbf{u}$ .

**Definition 2** *A convex function  $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$  is called self-concordant with constant  $M_f \geq 0$ , if  $f$  is  $C^3$  and satisfies **I**  $f(\mathbf{x}_k) \rightarrow +\infty$  for  $\mathbf{x}_k \rightarrow \mathbf{x} \in \partial\mathcal{K}$ ; and **II***

$$\forall \mathbf{x} \in \text{int } \mathcal{K}, \forall \mathbf{u} \in \mathbb{R}^d, \quad |\nabla^3 f(\mathbf{x})[\mathbf{u}, \mathbf{u}, \mathbf{u}]| \leq 2M_f \|\mathbf{u}\|_{\nabla^2 f(\mathbf{x})}^3.$$

Note that by definition, if  $f$  is self-concordant with constant  $M_f \geq 0$  it is also self-concordant with any constant  $M \geq M_f$ . Another property that we will use is that if  $f_1$  and  $f_2$  are self-concordant functions with constants  $M_1$  and  $M_2$ , respectively, then for any  $\alpha, \beta > 0$ , the function  $\alpha f_1 + \beta f_2$  is self-concordant with constant  $\frac{M_1}{\sqrt{\alpha}} \wedge \frac{M_2}{\sqrt{\beta}}$  (Nesterov et al., 2018, Theorem 5.1.1).

For a self-concordant function  $f$  and  $\mathbf{x} \in \text{dom } f$ , the quantity  $\lambda(\mathbf{x}, f) := \|\nabla f(\mathbf{x})\|_{\nabla^{-2} f(\mathbf{x})}$ , known as the *Newton decrement*, will be instrumental in our proofs. The following two lemmas contain properties of the Newton decrement and Hessians of self-concordant functions, which we will use repeatedly throughout (see e.g. (Nemirovski and Todd, 2008; Nesterov et al., 2018)).

**Lemma 3** *Let  $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$  be a self-concordant function with constant  $M_f \geq 1$ . Further, let  $\mathbf{x} \in \text{int } \mathcal{K}$  and  $\mathbf{x}_f \in \text{argmin}_{\mathbf{x} \in \mathcal{K}} f(\mathbf{x})$ . Then, **I** whenever  $\lambda(\mathbf{x}, f) < 1/M_f$ , we have*

$$\|\mathbf{x} - \mathbf{x}_f\|_{\nabla^2 f(\mathbf{x}_f)} \vee \|\mathbf{x} - \mathbf{x}_f\|_{\nabla^2 f(\mathbf{x})} \leq \lambda(\mathbf{x}, f)/(1 - M_f \lambda(\mathbf{x}, f));$$

*and **II** for any  $M \geq M_f$ , the damped Newton step  $\mathbf{x}_+ := \mathbf{x} - \frac{1}{1 + M\lambda(\mathbf{x}, f)} \nabla^{-2} f(\mathbf{x}) \nabla f(\mathbf{x})$  satisfies  $\mathbf{x}_+ \in \text{int } \mathcal{K}$  and  $\lambda(\mathbf{x}_+, f) \leq M\lambda(\mathbf{x}, f)^2(1 + (1 + M\lambda(\mathbf{x}, f))^{-1})$ .*

**Lemma 4** *Let  $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$  be a self-concordant function with constant  $M_f$  and  $\mathbf{x} \in \text{int } \mathcal{K}$ . Then, for any  $\mathbf{y}$  such that  $r := \|\mathbf{y} - \mathbf{x}\|_{\nabla^2 f(\mathbf{x})} < 1/M_f$ , we have*

$$(1 - M_f r)^2 \nabla^2 f(\mathbf{y}) \preceq \nabla^2 f(\mathbf{x}) \preceq (1 - M_f r)^{-2} \nabla^2 f(\mathbf{x}).$$

A consequence of the latter lemma is the following useful result whose proof is in Appendix B:

**Lemma 5** *Let  $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$  be a self-concordant function with constant  $M_f > 0$ . Then, for any  $\mathbf{x}, \mathbf{y} \in \text{int } \mathcal{K}$  such that  $r := \|\mathbf{x} - \mathbf{y}\|_{\nabla^2 f(\mathbf{x})} < 1/M_f$ , we have*

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_{\nabla^{-2} f(\mathbf{x})}^2 \leq \frac{1}{(1 - M_f r)^2} \|\mathbf{y} - \mathbf{x}\|_{\nabla^2 f(\mathbf{x})}^2.$$

The result of the lemma is reminiscent of the relationship between the Bregman divergence with respect to a function  $f$  and the one with respect to its Fenchel dual  $f^*$ ; that is,  $D_f(\mathbf{u}, \mathbf{v}) = D_{f^*}(\nabla f(\mathbf{v}), \nabla f(\mathbf{u}))$  (Cesa-Bianchi and Lugosi, 2006, Proposition 11.1).

**Mixability.** As a by-product of our efficient solution to the portfolio problem, we present an algorithm that guarantees an adaptive logarithmic regret for mixable losses.

**Definition 6** A sequence  $(f_t) \subset \{f: \mathcal{K} \rightarrow \mathbb{R}\}$  of convex functions is said to be  $\eta$ -mixable for  $\eta > 0$  if for any distribution  $P$  on  $\mathcal{K}$ , there exists  $\mathbf{u}_* \in \mathcal{K}$  such that

$$\forall t \geq 1, \quad f_t(\mathbf{u}_*) \leq -\eta^{-1} \ln \mathbb{E}_{\mathbf{u} \sim P} e^{-\eta f_t(\mathbf{u})}.$$

Formally, given an algorithm  $A$  whose outputs  $(\mathbf{u}_t)$  achieve a logarithmic regret against any sequence of  $\eta$ -mixable losses, i.e.  $\sum_{t=1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u})) \leq O(\ln T)$  for any  $\mathbf{u} \in \mathcal{K}$ , we design a meta-algorithm that aggregates instances of  $A$  and generates outputs  $(\mathbf{w}_t)$  that satisfy  $\sum_{t \in I} (\ell_t(\mathbf{w}_t) - \ell_t(\mathbf{u})) \leq O(\ln^2 T)$ , for any interval  $I \subset [T]$  and  $\mathbf{u} \in \mathcal{K}$ .

**Additional Notation.** For a differentiable convex function  $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$ , we denote by  $D_f(\mathbf{u}, \mathbf{v}) := f(\mathbf{u}) - f(\mathbf{v}) - \langle \nabla f(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle$  the *Bregman divergence* between  $\mathbf{u}, \mathbf{v} \in \text{int } \mathcal{K}$  with respect to  $f$ . We use the notation  $\tilde{O}(\cdot)$  to hide poly-log factors in  $T$  and  $d$ .

### 3. Background, Challenges, and Solution Sketch

In this section, we start by describing the algorithm Ada-BARRONS (Luo et al., 2018) that we build on. We then point out key challenges we tackle to design our efficient portfolio selection algorithm. The analysis we sketch in §3.2 and §3.3 is of independent interest as we discuss below.

#### 3.1. The Ada-BARRONS Algorithm

The Ada-BARRONS algorithm consists of a base algorithm, BARRONS, and a meta-algorithm that restarts the former under a certain condition on the sequence of returns and iterates of the algorithm.

**Base Algorithm.** BARRONS is simply mirror descent with a barrier regularizer. In particular, if we let  $\bar{\Delta}_d := \{\mathbf{x} \in \Delta_d: x_i \geq 1/T, \forall i \in [d]\}$ , the outputs  $(\mathbf{p}_t)$  of BARRONS are such that  $\mathbf{p}_1 := \mathbf{1}/d$  and  $\mathbf{p}_{t+1} = \text{argmin}_{\mathbf{p} \in \bar{\Delta}_d} \langle \mathbf{p}, \mathbf{g}_t \rangle + D_{\Phi_t}(\mathbf{p}, \mathbf{p}_t)$ , where

$$\Phi_t(\mathbf{p}) := \sum_{i=1}^d \frac{-\ln p_i}{\eta_{t,i}} + \frac{d \|\mathbf{p}\|^2}{2} + \frac{\beta}{2} \sum_{s=1}^t \langle \nabla_s, \mathbf{p} \rangle^2, \quad \eta_{t,i} := \eta \cdot \max_{s \in [t]} e^{-\log_T(dp_{s,i})}, \quad (2)$$

and  $\nabla_t := \mathbf{r}_t / \langle \mathbf{r}_t, \mathbf{p}_t \rangle$ . Using the standard analysis of mirror descent and the fact that Cover's loss is exp-concave, Luo et al. (2018) show that the regret  $R_T(\mathbf{u}) = \sum_{t=1}^T (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}))$  against a comparator  $\mathbf{u} \in \bar{\Delta}_d$  (competing against comparators in  $\bar{\Delta}_d$  is sufficient—see Lem. 1) is bounded as

$$R_T(\mathbf{u}) \leq \sum_{t=1}^T \langle \nabla_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + \sum_{t=1}^T (D_{\Phi_t}(\mathbf{u}, \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}, \mathbf{p}_{t+1}) - \beta \langle \nabla_t, \mathbf{p}_t - \mathbf{u} \rangle / 2), \quad (3)$$

as long as  $\beta$ , the parameter in the regularizer in (2), is less than  $\alpha_T(\mathbf{u}) := \frac{1}{2} \wedge \min_{t \in [T]} \frac{1}{8 \langle \mathbf{u} - \mathbf{p}_t, \nabla_t \rangle}$ . This condition seems strong since the algorithm does not have access to the sequence of returns  $(\mathbf{r}_t)$  or the comparator  $\mathbf{u}$  up-front to ensure that  $\beta \leq \alpha_T(\mathbf{u})$ . However, this issue is resolved via a clever restart scheme as we describe further below.

The fact that the regularizers  $(\Phi_t)$  have a quadratic term and a log-barrier ensures that the iterates  $(\mathbf{p}_t)$  are stable. In particular, the first sum on the RHS of (3) can be bounded by  $O(\beta^{-1} d \ln T)$ .

However, this term can still be problematic since  $\beta^{-1}$  may be large; after all, (3) only holds when  $\beta \leq \alpha_T(\mathbf{u})$  and  $\alpha_T(\mathbf{u})$  may be as small as  $1/(dT)$ .

Fortunately, terms of the form  $O(\beta^{-1}d \ln T)$  can be canceled by the second sum on the RHS of (3), thanks to the log-barrier regularizer  $\Psi_t(\mathbf{p}) := \sum_{i=1}^d -\eta_{t,i}^{-1} \ln p_i$  in the definition of  $\Phi_t$  and the non-decreasing nature of the learning rates  $(\eta_{t,i})$ . In particular, Luo et al. (2018) show that the second sum in (3) is bounded from above by  $O(\eta^{-1}d \ln T)$  plus

$$\sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}, \mathbf{p}_t) - D_{\Psi_{t-1}}(\mathbf{u}, \mathbf{p}_t)) \leq -\frac{1}{8\eta \ln T} \sum_{i \in [d]} \max_{t \in [T]} \frac{u_i}{p_{t,i}}. \quad (4)$$

where the inequality follows by (Luo et al., 2018, proof of Lem. 6). Now the RHS of (4) can cancel the bound  $O(\beta^{-1}d \ln T)$  on the stability term as long as  $\beta \geq \alpha_T(\mathbf{u})/2$  (see App. F for details).

**The Meta-Algorithm.** Since the sequence of returns  $(r_t)$  is not known up-front, it is not possible for any algorithm to pick  $\beta$  so that the condition  $\alpha_T(\mathbf{u})/2 \leq \beta \leq \alpha_T(\mathbf{u})$  is always satisfied. Aggregating multiple instances of BARRONS with different  $\beta$ 's also fails since  $\alpha_T(\mathbf{u})$  depends on the outputs of the algorithm; and so, changing  $\beta$  changes the target  $\alpha_T(\mathbf{u})$  for the base algorithm (see also discussion in (Luo et al., 2018)). Instead of aggregating base algorithms, the approach taken by Luo et al. (2018) consists of restarting the base algorithm on round  $t$  if the current estimate for  $\beta$  satisfies  $\beta > \alpha_t(\mathbf{u}_t)$ , where  $\mathbf{u}_t$  is the regularized leader:

$$\mathbf{u}_t \in \operatorname{argmin}_{\mathbf{u} \in \bar{\Delta}_d} \sum_{i=1}^d \frac{-\ln u_i}{\eta_{t,i}} + \sum_{s=\tau}^t \ell_s(\mathbf{u}) \quad (5)$$

and  $\tau$  is the round where the current instance of the base algorithm was initialized. The technical reason for why this works is sketched in Appendix F.

**Computational Considerations.** The computational complexity is dominated by the computation of the mirror descent iterates for the base algorithm and the FTRL computation (5) for the meta-algorithm. Both problems can be solved using an interior point method leading to a computational cost of  $\tilde{O}(d^{3.5}T + d^{2.5}T^2)$  after  $T$  rounds. We will reduce the computational complexity to  $\tilde{O}(d^3T)$  (where the  $O(d^3)$  is due to the computation of a matrix inverse) by **I**) avoiding the expensive FTRL computation in (5) thanks to an adaptive algorithm for exp-concave losses; and **II**) providing a new analysis for the damped Newton step to approximate mirror descent iterates. These techniques, which we describe next, are of independent interest.

### 3.2. Avoiding the FTRL Computation

To avoid computing the regularized leader in (5) that is needed to trigger restarts, we will use a meta-algorithm that aggregates base algorithms initialized at different rounds (one may think of these as “restarted” instances of the base algorithm). If the meta-algorithm has a small regret against any of the base algorithms, then this would emulate the effect of performing restarts, without the expensive cost of FTRL computations. More formally, if we denote by  $(\mathbf{u}_i^\tau)$  the outputs of an instance of the base algorithm  $A^\tau$  that is initialized at round  $\tau$ , we can emulate the effect of restarts if the outputs  $(\mathbf{u}_t)$  of the meta-algorithm satisfy

$$\sum_{s=\tau}^t (\ell_s(\mathbf{u}_s) - \ell_s(\mathbf{u}_s^\tau)) \leq O(\text{poly-log}(T)), \quad (6)$$

for all  $\tau \in [T]$  and  $t > \tau$ . A regret bound of this type may be achieved using sleeping experts algorithms, where in our case the instance  $A^\tau$  is considered “asleep” during the rounds  $s < \tau$  and “awake” for  $s \geq \tau$ . There are two challenges that come with using standard sleeping experts algorithms such as those in (Adamskiy et al., 2012; Gaillard et al., 2014). First, such techniques operate on linearized losses, which is sufficient when seeking a  $O(\sqrt{T})$  regret. This is not the case in our setting as we are aiming for a logarithmic regret. Second, if we want a regret bound of the form (6) to hold for all  $\tau \in [T]$ , a naive sleeping experts strategy would require keeping track of  $T$  experts. This would imply a  $O(T)$ -per-round computational complexity in the worst case, which would defeat the purpose of seeking an efficient alternative to the FTRL computation in (5).

We manage to circumvent these issues by using a meta-algorithm due to Zhang et al. (2019) that enjoys a logarithmic regret on any subinterval for exp-concave losses. To reduce the computational complexity, we show that it is sufficient to ensure a low regret against base algorithms indexed by a small set of geometric intervals (Daniely et al., 2015), reducing the number of experts at any round to at most  $O(\ln T)$  (this is discussed in §4.2).

### 3.3. Damped Newton Step for Mirror Descent

Now that we have a way of avoiding the expensive FTRL computations of Ada-BARRONS, it remains to find a more efficient alternative to the Mirror Descent (MD) computations of its base algorithm BARRONS. Before describing how we use Damped Newton Steps (DNS) for this purpose, we first describe the shortcomings of existing approaches.

**Shortcomings of previous DNS results.** Abernethy et al. (2012) showed how one can use damped Newton steps to approximate the iterates  $(\mathbf{p}_t)$  of FTRL given by  $\mathbf{p}_{t+1} \in \operatorname{argmin}_{\mathbf{p} \in \mathcal{C}} f_{t+1}(\mathbf{p}) := \sum_{s=1}^t \langle \mathbf{p}, \mathbf{g}_s \rangle + \Phi(\mathbf{p})$ , where  $\Phi$  is a self-concordant barrier for some set of interest  $\mathcal{C}$ . In particular, Abernethy et al. (2012) showed that for an appropriate scaling of  $\Phi$  the damped Newton steps  $(\mathbf{w}_t)$  defined by  $\mathbf{w}_{t+1} = \mathbf{w}_t - \frac{1}{1+\lambda(\mathbf{w}_t, f_{t+1})} \nabla^{-2} f_{t+1}(\mathbf{w}_t) \nabla f_{t+1}(\mathbf{w}_t)$  are close enough to the FTRL iterates  $(\mathbf{p}_t)$  so that the regret w.r.t.  $(\mathbf{w}_t)$  is bounded by the regret w.r.t.  $(\mathbf{p}_t)$  up to an additive  $O(\sqrt{T})$ .

We note that for a fixed regularizer  $\Phi$  that is a self-concordant barrier for  $\mathcal{C}$ , it is known that the FTRL iterates in the previous paragraph match the MD ones; that is,  $\mathbf{p}_{t+1} \in \operatorname{argmin}_{\mathbf{p} \in \mathcal{C}} \langle \mathbf{p}, \mathbf{g}_t \rangle + D_\Phi(\mathbf{p}, \mathbf{p}_t)$ , for all  $t$ . However, this is no longer the case when dealing with time-varying regularizers  $(\Phi_t)$ ; that is, when  $f_{t+1}(\cdot) = \langle \cdot, \mathbf{g}_t \rangle + D_{\Phi_t}(\cdot, \mathbf{p}_t)$  (as in the case of BARRONS). This means that we cannot directly use the analysis of Abernethy et al. (2012) to show that damped Newton steps are good approximations of MD iterates with varying regularizers. What is more, the damped Newton steps with respect to  $(f_t)$  can no longer be computed directly in this case since the gradient  $\nabla f_{t+1}(\mathbf{w}_t) = \mathbf{g}_t + \nabla \Phi_t(\mathbf{w}_t) - \nabla \Phi_t(\mathbf{p}_t)$  in the expression of the DNS depends on the iterate  $\mathbf{p}_t$ , which is what we seek to efficiently approximate in the first place. Using  $\mathbf{w}_t$  as an estimator for  $\mathbf{p}_t$  does not work since the approximation errors accumulate across rounds in an unfavorable way (breaking the analysis of Abernethy et al. (2012)).

One tempting approach around these issues is to target the FTRL iterates  $(\mathbf{p}_t)$  given by  $\mathbf{p}_{t+1} = \operatorname{argmin}_{\mathbf{p} \in \mathcal{C}} \sum_{s=1}^t \langle \mathbf{p}, \mathbf{g}_s \rangle + \Phi_t(\mathbf{p})$  instead of the MD ones. However, we are not aware of an existing analysis of FTRL that yields negative terms from Bregman divergences in the regret bound as in (4) (negative terms were needed to cancel the problematic  $O(\beta^{-1} d \ln T)$  term in the regret bound).

**Our approach.** Our solution consists of approximating FTRL iterates w.r.t. modified gradients that are chosen in a way to still allow us to use the MD analysis to derive our regret bound (similar in

spirit to the approach by Foster et al. (2020)). In particular, we consider the objective

$$f_{t+1}(\mathbf{p}) := \sum_{s=1}^t \mathbf{p}^\top (\mathbf{g}_s - \nabla \Phi_s(\mathbf{w}_s) + \nabla \Phi_{s-1}(\mathbf{w}_s)) + \Phi_t(\mathbf{p}),$$

with  $(\mathbf{w}_s)$  being the damped Newton iterates w.r.t.  $(f_t)$ . Under mild conditions on  $(\Phi_t)$ , the FTRL iterates  $(\mathbf{p}_t \in \operatorname{argmin}_{\mathbf{p} \in \mathcal{C}} f_t(\mathbf{p}))$  can be efficiently approximated by  $(\mathbf{w}_t)$ . Despite the fact that the objective  $f_{t+1}$  does not contain any “unknown” MD iterates, we are still able to take advantage of the MD analysis and bring back negative terms from Bregman divergences (as in (4)) in the regret bound. The key fact that enables this is that the FTRL iterates  $(\mathbf{p}_t)$  with respect to  $(f_t)$  match the MD iterates with modified gradients  $(\tilde{\mathbf{g}}_t)$ . In particular,  $(\mathbf{p}_t)$  satisfy (see Lemma 9 and its proof)

$$\mathbf{p}_{t+1} \in \operatorname{argmin}_{\mathbf{p}} \langle \tilde{\mathbf{g}}_t, \mathbf{p} \rangle + D_{\Phi_t}(\mathbf{p}, \mathbf{p}_t), \quad \text{where } \tilde{\mathbf{g}}_t := \mathbf{g}_t - \nabla \phi_t(\mathbf{p}_t) + \nabla \phi_t(\mathbf{w}_t), \quad (7)$$

and  $\phi_t := \Phi_{t-1} - \Phi_t$ . We now give a sketch of how this observation allows us to leverage the MD analysis and extract negative terms as in (4).

**MD analysis for FTRL (a sketch).** To illustrate how (7) helps in our analysis, consider the regularizer  $\Phi_t$  in (2), which we write as  $\Phi_t = \Psi_t + \Theta_t$ , where  $\Psi_t$  is the barrier part  $\Psi_t(\mathbf{p}) := \sum_{i \in [d]} -\eta_{t,i}^{-1} \ln p_i$ . We will further define  $\psi_t := \Psi_{t-1} - \Psi_t$  and note that the fact that  $(\eta_{t,i})$  are non-decreasing, implies that  $(\psi_t)$  are convex, self-concordant functions (the latter fact is all that is needed to generalize the current analysis). A key step in the analysis of the regret involves bounding the sum  $\Sigma_T := \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u} \rangle$  of linearized losses. For simplicity of the exposition, suppose that  $\Theta_t \equiv 0$ , for all  $t$ ; the quadratic terms in  $\Theta_t$  present no difficulty when it comes to bounding  $\Sigma_T$  (see proof of Lem. 10 for a derivation with non-zero  $(\Theta_t)$ ). In this case, the sum  $\Sigma_T$  can be written as

$$\begin{aligned} \Sigma_T &= \sum_{t=1}^T (\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u} \rangle + \langle \nabla \psi_t(\mathbf{p}_t) - \nabla \psi_t(\mathbf{w}_t), \mathbf{p}_t - \mathbf{u} \rangle), \\ &= \sum_{t=1}^T (\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle) \\ &\quad + \sum_{t=1}^T (\langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u} \rangle + D_{\psi_t}(\mathbf{u}, \mathbf{p}_t) - D_{\psi_t}(\mathbf{u}, \mathbf{w}_t) + D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t)), \end{aligned}$$

where we used the definitions of  $\tilde{\mathbf{g}}_t$  and the Bregman divergence. Using this, and the facts that  $\langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u} \rangle \leq D_{\Phi_t}(\mathbf{u}, \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}, \mathbf{p}_{t+1}) - D_{\Phi_t}(\mathbf{p}_{t+1}, \mathbf{p}_t)$  (by optimality of  $\mathbf{p}_{t+1}$ —see proof of (Luo et al., 2018, Lem. 5)) and  $D_{\psi_t}(\mathbf{u}, \mathbf{p}_t) = D_{\Phi_{t-1}}(\mathbf{u}, \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}, \mathbf{p}_t)$ , we can bound  $\Sigma_T$  as

$$\begin{aligned} \Sigma_T &\leq \sum_{t=1}^T (\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t)) \\ &\quad + D_{\Phi_0}(\mathbf{u}, \mathbf{p}_1) - D_{\Phi_T}(\mathbf{u}, \mathbf{p}_{T+1}) + \sum_{t=1}^T (D_{\Phi_t}(\mathbf{u}, \mathbf{w}_t) - D_{\Phi_{t-1}}(\mathbf{u}, \mathbf{w}_t)). \quad (8) \end{aligned}$$

Thus, one can extract negative terms (as in (4)) from the last sum in (8) to cancel the  $O(\beta^{-1} d \ln T)$  term in the regret bound. The remaining terms in (8) can be shown to be small thanks to I)  $(\mathbf{w}_t)$  approximate  $(\mathbf{p}_t)$  well, II) stability of the mirror descent iterates, and III) the fact that  $\psi_t$  is self-concordant together with Lem. 5. Next, we present the full details and guarantee of our algorithm.



## 4. An Efficient Algorithm for Online Portfolio Selection

Our final algorithm (Alg. 3) will consist of a set of base algorithms (instances of Alg. 1) and a meta algorithm (instance of Alg. 2). We analyze these algorithms separately in the next two subsections before combining the results in Subsection 4.3.

### 4.1. Base Algorithm: Damped Online Newton Step (DONS)

To analyze our base algorithm (Alg. 1), we will consider the following sequence of regularizers that are defined in terms of the iterates  $(\mathbf{u}_t)$  and  $(\mathbf{w}_t)$  in Algorithm 1; the sequence of observed return vectors  $(\mathbf{r}_t)$ ; and the gradients  $(\mathbf{g}_t) \equiv (J\mathbf{r}_t / \langle \mathbf{r}_t, \bar{\mathbf{u}}_t \rangle)$ :

$$\Phi_t(\mathbf{u}) := \Psi_t(\mathbf{u}) + \frac{\beta d \|\mathbf{u}\|^2}{8} + \frac{\beta}{8} \sum_{s=1}^t \langle \mathbf{g}_s, \mathbf{u} - \mathbf{w}_s \rangle^2, \quad \forall \mathbf{u} \in \mathcal{C}_{d-1}. \quad (9)$$

$$\text{where } \Psi_t(\mathbf{x}) := - \sum_{i=1}^d \frac{-\ln \bar{x}_i}{\eta_{t,i}}, \quad \eta_{t,i} := \eta \cdot e^{\log T (\rho_{t,i}/d)}, \quad (10)$$

and  $\rho_{t,i}$  is such that  $\rho_{t,i} \in [\max_{s \in [t]} (2\bar{u}_{s,i})^{-1}, \max_{s \in [t]} (\bar{u}_{s,i})^{-1}]$ , for all  $i \in [d]$ , and  $\rho_0 := d\mathbf{1}$ . In particular, for every  $i \in [d]$ ,  $(\rho_{t,i})_t$  satisfies the recursion

$$\rho_{t,i} = \mathbb{I}\{2\rho_{t-1,i} < \frac{1}{\bar{u}_{t,i}}\} \cdot \frac{1}{\bar{u}_{t,i}} + \mathbb{I}\{2\rho_{t-1,i} \geq \frac{1}{\bar{u}_{t,i}}\} \cdot \rho_{t-1,i}. \quad (11)$$

Luo et al. (2018) chose the sequence  $(\rho_{t,i})$  such that  $\rho_{t,i} = \max_{s \in [t]} 1/\bar{u}_{s,i}$ ,  $\forall i \in [d]$  and  $t \in [T]$ . Using our new analysis of the damped Newton steps for MD, this choice leads to a regret bound of order  $\tilde{O}(d^m)$  with  $m > 2$ , which is worse than what we are aiming for. Our choice in (11) ensures that the barrier  $\Psi_t$  changes at most  $O(d \ln T)$  times, which is crucial to proving the desired bound.

For any  $t \in [T]$ , the output  $\mathbf{u}_{t+1}$  of Algorithm 1 at round  $t+1$  is given by  $\mathbf{u}_{t+1} = \mathbf{w}'_{t+1} = (1 - 1/T)\mathbf{w}_{t+1} + \mathbf{1}/(dT)$ , where  $\mathbf{w}_{t+1}$  is the damped Newton step:

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \frac{\nabla^{-2} \Phi_t(\mathbf{w}_t) \nabla_t}{1 + 4\sqrt{e\eta} \|\nabla_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}}, \quad (12)$$

with  $\nabla_t := \nabla \Phi_t(\mathbf{w}_t) + \sum_{s=1}^t (\mathbf{g}_s - \nabla \Psi_s(\mathbf{w}_s) + \nabla \Psi_{s-1}(\mathbf{w}_s))$  and  $\mathbf{w}_1 = \mathbf{1}/d$ . In part due to the fact that  $\text{dom } \Phi_t = \mathcal{C}_{d-1}$ , for all  $t \geq 1$ , the iterates  $(\mathbf{w}_t)$  in Algorithm 1 are only well defined when the update rule in (12) ensures that  $\mathbf{w}_{t+1} \in \mathcal{C}_{d-1}$  for any  $\mathbf{w}_t \in \mathcal{C}_{d-1}$ . This is in fact the case as we show next by leveraging the self-concordant property of  $\Phi_t$ . To simplify notation in the proof of the next lemma (which is in App. B) and in the rest of the paper, we let  $\vartheta_i: \mathcal{C}_{d-1} \rightarrow \mathbb{R}$  be defined by

$$\vartheta_i(\mathbf{x}) := -\ln \bar{x}_i, \quad \forall i \in [d]. \quad (13)$$

Note that the self-concordant barrier in 10 satisfies  $\Psi_t(\cdot) = \sum_{i \in [d]} \vartheta_i(\cdot) / \eta_{t,i}$ .

**Lemma 7** *For all  $t \geq 1$ ,  $\Phi_t$  in (9) is a self-concordant function with constant  $M_{\Phi_t} \leq \sqrt{\eta e}$ .*

---

**Algorithm 1** DONS (Base Algorithm): Damped Online Newton Step for Portfolio Selection.

---

**Require:** Parameters  $\eta, \beta > 0$ .

- 1: Set  $\mathbf{w}_1 = \mathbf{1}/d \in \mathbb{R}^{d-1}$ ,  $\boldsymbol{\rho}_0 = d\mathbf{1} \in \mathbb{R}^d$ ,  $\mathbf{G}_0 = \mathbf{0}$ , and  $V_0 = \beta dI/4 \in \mathbb{R}^{d-1 \times d-1}$ .
  - 2: **for**  $t = 1, 2, \dots$  **do**
  - 3:   Play  $\mathbf{u}_t = (1 - \frac{1}{T})\mathbf{w}_t + \frac{1}{dT}\mathbf{1}$  and observe gradient  $\mathbf{g}_t = \nabla \ell_t(\mathbf{u}_t) = J\mathbf{r}_t / \langle \mathbf{r}_t, \bar{\mathbf{u}}_t \rangle$ .
  - 4:   Set  $\rho_{t,i} = \mathbb{I}\{2\rho_{t-1,i} < \frac{1}{\bar{u}_{t,i}}\} \cdot \frac{1}{\bar{u}_{t,i}} + \mathbb{I}\{2\rho_{t-1,i} \geq \frac{1}{\bar{u}_{t,i}}\} \cdot \rho_{t-1,i}$ , for all  $i \in [d]$ .
  - 5:   Define  $\Psi_t(\mathbf{x}) = -\sum_{i=1}^d \frac{\ln \bar{x}_i}{\eta_{t,i}}$ , where  $\eta_{t,i} := \eta \cdot \exp(\log_T(\rho_{t,i}/d))$ ,  $\forall i \in [d]$ .
  - 6:   Set  $\mathbf{G}_t = \mathbf{G}_{t-1} + \mathbf{g}_t \cdot (1 - \beta \langle \mathbf{g}_t, \mathbf{w}_t \rangle / 4) - \nabla \Psi_t(\mathbf{w}_t) + \nabla \Psi_{t-1}(\mathbf{w}_t)$ .
  - 7:   Set  $V_t = V_{t-1} + \beta \mathbf{g}_t \mathbf{g}_t^\top / 4$  and  $\nabla_t = \mathbf{G}_t + V_t \mathbf{w}_t + \nabla \Psi_t(\mathbf{w}_t) + \beta d \mathbf{w}_t / 4$ .
  - 8:   Set  $\mathbf{w}_{t+1} = \mathbf{w}_t - \frac{1}{1+4\sqrt{\eta e} \|\nabla_t\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1}}} (\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1} \nabla_t$ .
  - 9: **end for**
- 

**Damped Online Newton Steps as Approximate MD Iterates.** A key part of our analysis consists of showing that the intermediate iterates  $(\mathbf{w}_t)$  on Line 8 of Algorithm 1 are close to the mirror descent iterates  $(\mathbf{p}_t)$  with respect to the sequence of regularizers  $(\Phi_t)$  in (9):

$$\mathbf{p}_{t+1} \in \operatorname{argmin}_{\mathbf{p} \in \mathcal{C}_{d-1}} F_{t+1}(\mathbf{p}) := \langle \mathbf{p}, \tilde{\mathbf{g}}_t \rangle + D_{\Phi_t}(\mathbf{p}, \mathbf{p}_t), \quad \text{where} \quad (14)$$

$$\tilde{\mathbf{g}}_t := (1 + \beta \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle / 4) \mathbf{g}_t + \sum_{i=1}^d \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) (\nabla \vartheta_i(\mathbf{p}_t) - \nabla \vartheta_i(\mathbf{w}_t)), \quad (15)$$

and  $\mathbf{p}_1 := \mathbf{1}/d$  (recall  $(\vartheta_i)$  from (13)). Next, we formally state this result (recall  $\lambda(\cdot, \cdot)$  from §2):

**Lemma 8** For any  $\beta \in (0, 1/8)$  and  $\eta \leq 1/2^{14}$ , the iterates  $(\mathbf{w}_t)$  in Algorithm 1 satisfy,

$$\forall t \geq 1, \quad \frac{\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 F_t(\mathbf{w}_t)}}{2^4 \sqrt{e\eta}} \leq \frac{\lambda(\mathbf{w}_t, F_t)}{2^3 \sqrt{e\eta}} \leq \lambda(\mathbf{w}_{t-1}, F_t)^2 \leq C\eta, \quad (16)$$

where  $C := \frac{4e}{4^{-1}\sqrt{(1-1/T)^2}}$ . Further, we have  $\sum_{t=1}^T \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2 \leq 1 + 15\beta^{-1}d \log T$ .

The next lemma, which will be useful in the proof of Lemma 8, essentially shows that the mirror descent iterates in (14) match the FTRL iterates with respect to  $(f_t)$ , where  $f_{t+1}(\mathbf{p}) := \sum_{s=1}^t \mathbf{p}^\top (\mathbf{g}_s - \nabla \Phi_s(\mathbf{w}_s) + \nabla \Phi_{s-1}(\mathbf{w}_s)) + \Phi_t(\mathbf{p})$  (c.f. discussion in §3.3).

**Lemma 9** For all  $t \in [T]$ , we have  $\nabla F_{t+1}(\mathbf{w}_t) = \mathbf{g}_t + \nabla F_t(\mathbf{w}_t)$  ( $F_t$  as in (14)) and  $\forall \mathbf{w} \in \mathcal{C}_{d-1}$ ,

$$\nabla F_{t+1}(\mathbf{w}) = \nabla \Phi_t(\mathbf{w}) + \sum_{s=1}^t (\mathbf{g}_s - \nabla \Psi_s(\mathbf{w}_s) + \nabla \Psi_{s-1}(\mathbf{w}_s)).$$

**Regret Decomposition.** We now present the main regret decomposition. In the proof, which is Appendix C, we follow similar steps as the ones outlined in §3.3.

**Lemma 10** Let  $T > 1$ ,  $c_T := 1 - 1/T$ , and  $\psi_t := \Psi_{t-1} - \Psi_t$ . Further let  $(\mathbf{u}_t)$  and  $(\mathbf{w}_t)$  be as in Algorithm 1 with parameters  $\eta, \beta > 0$ , and  $(\mathbf{p}_t)$  as in (14). For any sequence of returns  $(\mathbf{r}_t)$  and

$\mathbf{u} \in \mathcal{C}_{d-1}$  such that  $\beta \leq 8^{-1} \wedge |8\langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle|^{-1}$  (recall  $\mathbf{u}'$ ,  $\mathbf{u}''$  from §2), we have, for all  $t \in [T]$ ,

$$\sum_{t=1}^T \frac{\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u})}{c_T} \leq \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + \sum_{t=1}^T (D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) + D_{\Psi_t}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{w}_t)) + O\left(\frac{d \ln T}{\eta}\right) + \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \frac{3\beta}{8} \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2. \quad (17)$$

Next, we bound each term in this decomposition starting with  $\sum_{t=1}^T D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t)$ . To show that this term is small, we rely on the fact that  $\psi_t$  is a self-concordant functions, which holds true by our choice of ‘‘doubling’’  $(\rho_{t,i})$ 's in (11). This enables us to relate  $D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t)$  to  $\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}$  via Lem. 5, which we can then bound using Lem. 8.

**Lemma 11** *Let  $(\Psi_t)$  be as in (10). Then, the function  $\psi_t := \Psi_{t-1} - \Psi_t$  is a self-concordant function with constant  $\sqrt{\epsilon \eta \log_2 T}$ . Furthermore, the mirror descent iterates  $(\mathbf{p}_t)$  in (14) and the iterates  $(\mathbf{w}_t)$  in Algorithm 1 satisfy  $\sum_{t=1}^T D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) \leq O(d \ln T)$ .*

We move to the stability term  $\sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle$ , which is the most technical one to bound due to the modified gradients  $(\tilde{\mathbf{g}}_t)$ . We will use Hölder's inequality and the triangle inequality to bound  $\langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle$  in terms of  $\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 \Phi_t(\mathbf{w}_t)}$ ,  $\|\mathbf{g}_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}$ , and  $\|\nabla \vartheta_i(\mathbf{p}_t) - \nabla \vartheta_i(\mathbf{w}_t)\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}$ . Then, we use the self-concordance property in Lem. 5 to relate the latter term to  $\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_t(\mathbf{w}_t)}$ , which (thanks to Lem. 8) will allow us to show that the stability term is small.

**Lemma 12** *Let  $T > 1$  and  $(\tilde{\mathbf{g}}_t)$  be as in (15). If  $\eta \leq 1/2^{14}$  and  $\beta \in (0, 1/8)$ , then the iterates  $(\mathbf{p}_t)$  in (14) satisfy  $\sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle \leq \frac{18d \ln T}{\beta} + O(d \ln T)$ .*

We now bound the sum divergences which will allow us to cancel the undesirable  $O(\beta^{-1} d \ln T)$  term in the regret bound as discussed in §3.1.

**Lemma 13** *Let  $T > 1$  and  $(\mathbf{u}_t)$  be the iterates of Alg. 1 with parameters  $\beta \in (0, 1/8)$  and  $\eta \leq 1/2^{14}$ . For any sequence  $(\mathbf{r}_t)$ , the iterates  $(\mathbf{p}_t)$  in (14) satisfy (recall  $\mathbf{u}'$  and  $\mathbf{u}''$  from §2)  $\sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{w}_t)) \leq \frac{-1}{16\eta \ln T} \sum_{i=1}^d \max_{t \leq T} \frac{\bar{u}_i''}{\bar{u}_{t,i}} + O(d/\eta)$ , for all  $\mathbf{u} \in \mathcal{C}_{d-1}$ .*

It remains to upper bound the sums  $\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^i$ , for  $i \in \{1, 2\}$  which are expected to be small since  $(\mathbf{w}_t)$  are close to  $(\mathbf{p}_t)$  by Lemma 8:

**Lemma 14** *Let  $T > 1$ ,  $c_T := 1 - 1/T$ , and  $S_T := 1 + 4\sqrt{15}\beta^{-1} d \ln T$ . Further, let  $(\mathbf{w}_t)$  be the iterates in Alg. 1 with parameters  $\beta \in (0, 1/8)$  and  $\eta \leq 1/2^{14}$ . For any sequence of returns  $(\mathbf{r}_t)$ , the mirror descent iterates in (14) satisfy  $\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle \leq S_T$  and  $\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2 \leq \frac{64e^2 \eta^2}{c_T^3} S_T$ .*

Combining these results, we obtain the following regret bound for our base algorithm:

**Theorem 15 (Base Algorithm Regret)** *Let  $T > 1$ , and  $(\mathbf{u}_t)$  be the iterates of Algorithm 1 with parameters  $\beta \in (0, 1/8)$  and  $\eta \leq 1/2^{14}$ . For any sequence of returns  $(\mathbf{r}_t)$  and  $\mathbf{u} \in \mathcal{C}_{d-1}$  such that, for all  $t \in [T]$ ,  $\beta \leq 8^{-1} \wedge |8\langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle|^{-1}$  (recall  $\mathbf{u}'$  and  $\mathbf{u}''$  from §2), we have*

$$\sum_{t=1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u})) \leq O\left(\frac{d \ln T}{\eta}\right) + \frac{34d \ln T}{\beta} - \frac{1}{32\eta \ln T} \sum_{i=1}^d \max_{t \leq T} \frac{\bar{u}_i''}{\bar{u}_{t,i}}.$$

The regret bound in Theorem 15 is the same as that of BARRONS up to constant factors. We are now going to describe the adaptive meta-algorithm which allows us to emulate the effect of restarts in Ada-BARRONS.

---

**Algorithm 2** PAE: Adaptive Meta-Algorithm for Exp-Concave Losses (Zhang et al., 2019).

---

**Require: I** Grid  $\mathcal{G}$  of  $\beta$  values, horizon  $T$ , and  $\eta > 0$ ; and **II** Instances  $(A^{\beta,I})_{\beta \in \mathcal{G}, I \in \mathcal{I}|_T}$ , where  $A^{\beta,I}$  is active during  $I$

Set  $\mathcal{J}_0 = \emptyset$ .

**for**  $t = 1, \dots, T$  **do**

Identify the set of newly active intervals  $\tilde{\mathcal{J}}_t = \{I \in \mathcal{I}|_T : \min I = t\}$ .

Set  $\mathcal{F}_{t-1}^{\beta,I} = 0$  for all  $\beta \in \mathcal{G}$  and  $I \in \tilde{\mathcal{J}}_t$ .

Update the set of active intervals  $\mathcal{J}_t = \mathcal{J}_{t-1} \cup \tilde{\mathcal{J}}_t$ .

Receive  $\mathbf{u}_t^{\beta,I} \in \mathcal{U}$  from each  $A^{\beta,I}$  such that  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .

Set  $q_{t-1}^{\beta,I} = e^{-\eta \mathcal{F}_{t-1}^{\beta,I}} / Z_t$  for  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ , where  $Z_t = \sum_{\beta \in \mathcal{G}, I \in \mathcal{J}_t} e^{-\eta \mathcal{F}_{t-1}^{\beta,I}}$ .

Play  $\mathbf{u}_t = \sum_{\beta, I: t \in I} q_{t-1}^{\beta,I} \mathbf{u}_t^{\beta,I}$ .

Observe  $f_t$  and set  $\mathcal{F}_t^{\beta,I} = \mathcal{F}_{t-1}^{\beta,I} + f_t(\mathbf{u}_t^{\beta,I}) - f_t(\mathbf{u}_t)$ , for all  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .

Send  $f_t$  to each  $A^{\beta,I}$  such that  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .

Update the set of active intervals  $\mathcal{J}_t = \mathcal{J}_t \setminus \{I \in \mathcal{I}|_T : \max I = t\}$ .

**end for**

---

## 4.2. Adaptive Meta-Algorithm for Exp-Concave Losses (PAE)

As discussed in §3.2, we will use the adaptive meta-algorithm PAE due to Zhang et al. (2019) for exp-concave losses to circumvent the restarts required by Ada-BARRONS. In App. E, we present a version of this meta-algorithm that yields an adaptive regret for the larger class of mixable losses (see Def. 6). We note that using the Follow-the-Leading-History algorithm by Hazan and Seshadhri (2007) (which is also an adaptive algorithm for exp-concave losses) instead of PAE would lead to a similar final result for our application. However, we use PAE due to a slightly simpler analysis.

We now introduce some notation to describe the PAE algorithm (Alg. 2). Let  $\mathcal{I}$  be the set of geometric covering intervals:

$$\mathcal{I} := \bigcup_{i,k \in \mathbb{N}} \{[2^k i, 2^k(i+1) - 1]\}.$$

Such a set was first used by Daniely et al. (2015) in the context of strongly adaptive Online Learning. We further define the “restriction” of  $\mathcal{I}$  to  $[T]$  as  $\mathcal{I}|_T := \{I \cap [T] : I \in \mathcal{I}\} \cup \{[T]\}$ . The PAE algorithm maintains a set of base algorithms/experts  $(A^{\beta,I})$ , where for expert  $A^{\beta,I}$ ,  $\beta$  represents a parameter in some predefined grid  $\mathcal{G} \subset \mathbb{R}$  and  $I \in \mathcal{I}|_T$  represents the interval on which the expert is active; in this case, expert  $A^{\beta,I}$  is initialized at round  $t = \min I$  and terminates after round  $t = \max I$ . If  $(\mathbf{u}_t^{\beta,I})_{t \in I}$  are the outputs of expert  $A^{\beta,I}$  within Alg. 2, then the outputs  $(\mathbf{u}_t)$  of Alg. 2 are given by

$$\mathbf{u}_t := \frac{\sum_{\beta, I: t \in I} \exp(-\mathcal{L}_{t-1}^{\beta,I}) \mathbf{u}_t^{\beta,I}}{\sum_{\beta, I: t \in I} \exp(-\mathcal{L}_{t-1}^{\beta,I})}, \quad \text{where } \mathcal{L}_t^{\beta,I} := \sum_{s \in I \cap [t]} (\ell_s(\mathbf{u}_s^{\beta,I}) - \ell_s(\mathbf{u}_s)). \quad (18)$$

We now state the guarantee of PAE (Alg. 2) relative to the performance of the experts  $(A^{\beta,I})$ :

**Proposition 16** *Let  $\eta > 0$  and  $\mathcal{G}$  be a set s.t.  $|\mathcal{G}| \leq M$ . Further, for  $\beta \in \mathcal{G}$  and  $I \in \mathcal{I}$ , let  $(\mathbf{u}_t^{\beta,I})_{t \in I}$  be the outputs of the subroutine  $A^{\beta,I}$  within Alg. 2 in response to a sequence of  $\eta$ -exp-concave losses*

( $f_t$ ). Then, the outputs ( $\mathbf{u}_t$ ) of Algorithm 2 guarantee

$$\sum_{s \in I \cap [t]} (f_s(\mathbf{u}_s) - f_s(\mathbf{u}_s^{\beta, I})) \leq (2 \ln t + \ln M) / \eta, \quad \text{for all } I \in \mathcal{I}, \beta \in \mathcal{G}, \text{ and } t \in I.$$

This regret guarantee can be derived from results by Zhang et al. (2019), and follows directly from Proposition 25 in Appendix E, which generalizes the result of Proposition 16 to mixable losses. Next, we combine PAE and DONS to get an algorithm with our desired guarantee for the Portfolio Selection problem.

### 4.3. Final Algorithm and Guarantee

For our Portfolio Selection application, we instantiate Alg. 2 with Cover’s loss and the base algorithms ( $A^{\beta, I}$ ) set as instances of DONS (Alg. 1) with  $\eta = 1/(286^2 d \ln^3 T)$  and  $\beta \in \mathcal{G}$ , where

$$\mathcal{G} := \left\{ \frac{1}{d^{2^{i+3}}} : i \in [\lceil \log_2 T \rceil] \right\}. \quad (19)$$

This specific instance of Alg. 2 is displayed in Alg. 3.

Since the regret of DONS is the same as that of BARRONS (see discussion after Thm. 15), and the adaptive regret enabled by Alg. 2 allows us to emulate the restarts of Ada-BARRONS (see §3.2), the regret of our final Alg. 3 will be the same as that of Ada-BARRONS up to log factors:

**Theorem 17** *The regret of Algorithm 3 is bounded by  $O(d^2 \ln^5 T)$ . Furthermore, the algorithm runs in  $O(d^3 \ln^2 T)$  per round and requires  $O(d^2 \ln^2 T)$  total space.*

---

**Algorithm 3** PAE + DONS: Adaptive Meta-Algorithm for Online Portfolio Selection.

---

Instantiate of Alg. 2 with **I** ( $f_t \equiv (\ell_t)$  ( $\ell_t$  as in (1)) with  $\mathcal{U} = \mathcal{C}_{d-1}$ ; and **II**) ( $A^{\beta, I}$ ) set as instances of Alg. 1 with  $\eta = \frac{1}{286^2 d \ln^3 T}$ ,  $\beta \in \mathcal{G}$  with  $\mathcal{G}$  as in (19), and  $I \in \mathcal{I}_T$ .

---

## Acknowledgments

We thank anonymous reviewers for comments that helped improve the presentation of the paper. We acknowledge support from the ONR through awards N00014-20-1-2336 and N00014-20-1-2394.

## References

- Jacob D Abernethy, Elad Hazan, and Alexander Rakhlin. Interior-point methods for full-information and bandit online learning. *IEEE Transactions on Information Theory*, 58(7):4164–4175, 2012.
- Dmitry Adamskiy, Wouter M Koolen, Alexey Chernov, and Vladimir Vovk. A closer look at adaptive regret. In *International Conference on Algorithmic Learning Theory*, pages 290–304. Springer, 2012.
- Alekh Agarwal, Haipeng Luo, Behnam Neyshabur, and Robert E Schapire. Corraling a band of bandit algorithms. In *Conference on Learning Theory*, pages 12–38. PMLR, 2017.
- Amit Agarwal and Elad Hazan. Efficient algorithms for online game playing and universal portfolio management. *ECCC, TR06-033*, 2005.
- Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- Sébastien Bubeck, Ronen Eldan, and Joseph Lehec. Sampling from a log-concave distribution with projected langevin monte carlo. *Discrete & Computational Geometry*, 59(4):757–783, 2018.
- Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- Thomas M Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- Amit Daniely, Alon Gonen, and Shai Shalev-Shwartz. Strongly adaptive online learning. In *International Conference on Machine Learning*, pages 1405–1411. PMLR, 2015.
- Dylan J Foster, Claudio Gentile, Mehryar Mohri, and Julian Zimmert. Adapting to misspecification in contextual bandits. *Advances in Neural Information Processing Systems*, 33:11478–11489, 2020.
- Pierre Gaillard, Gilles Stoltz, and Tim Van Erven. A second-order bound with excess losses. In *Conference on Learning Theory*, pages 176–196. PMLR, 2014.
- Elad Hazan and Comandur Seshadhri. Adaptive algorithms for online decision problems. In *Electronic colloquium on computational complexity (ECCC)*, volume 14, 2007.
- Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- David P Helmbold, Robert E Schapire, Yoram Singer, and Manfred K Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8(4):325–347, 1998.

- Adam Tauman Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, pages 423–440, 2002.
- Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. *arXiv preprint arXiv:1805.07430*, 2018.
- Hariharan Narayanan and Alexander Rakhlin. Efficient sampling from time-varying log-concave distributions. *The Journal of Machine Learning Research*, 18(1):4017–4045, 2017.
- Arkadi S Nemirovski and Michael J Todd. Interior-point methods for optimization. *Acta Numerica*, 17:191–234, 2008.
- Yurii Nesterov et al. *Lectures on convex optimization*, volume 137. Springer, 2018.
- Laurent Orseau, Tor Lattimore, and Shane Legg. Soft-bayes: Prod for mixtures of experts with log-loss. In *International Conference on Algorithmic Learning Theory*, pages 372–399. PMLR, 2017.
- Alexander Rakhlin and Karthik Sridharan. Sequential probability assignment with binary alphabets and large classes of experts. *arXiv preprint arXiv:1501.07340*, 2015.
- Tim Van Erven, Dirk Van der Hoeven, Wojciech Kotłowski, and Wouter M Koolen. Open problem: Fast and optimal online portfolio selection. In *Conference on Learning Theory*, pages 3864–3869. PMLR, 2020.
- Chen-Yu Wei and Haipeng Luo. More adaptive algorithms for adversarial bandits. In *Conference On Learning Theory*, pages 1263–1291. PMLR, 2018.
- Lijun Zhang, Guanghui Wang, Wei-Wei Tu, and Zhi-Hua Zhou. Dual adaptivity: A universal algorithm for minimizing the adaptive regret of convex functions. *arXiv preprint arXiv:1906.10851*, 2019.
- Julian Zimmert, Naman Agarwal, and Satyen Kale. Pushing the efficiency-regret pareto frontier for online learning of portfolios and quantum states. *arXiv preprint arXiv:2202.02765*, 2022.
- Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003.

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Preliminaries</b>	<b>3</b>
<b>3</b>	<b>Background, Challenges, and Solution Sketch</b>	<b>5</b>
3.1	The Ada-BARRONS Algorithm . . . . .	5
3.2	Avoiding the FTRL Computation . . . . .	6
3.3	Damped Newton Step for Mirror Descent . . . . .	7
<b>4</b>	<b>An Efficient Algorithm for Online Portfolio Selection</b>	<b>9</b>
4.1	Base Algorithm: Damped Online Newton Step (DONS) . . . . .	9
4.2	Adaptive Meta-Algorithm for Exp-Concave Losses (PAE) . . . . .	12
4.3	Final Algorithm and Guarantee . . . . .	13
<b>A</b>	<b>Technical Lemmas</b>	<b>17</b>
<b>B</b>	<b>Proof of Lemmas 5, 7, 8, and 9</b>	<b>19</b>
<b>C</b>	<b>Proof of Theorem 15 (Regret of Base Algorithm)</b>	<b>22</b>
<b>D</b>	<b>Proof of Theorem 17 (Meta-Algorithm Regret)</b>	<b>28</b>
D.1	Proof of Theorem 17 . . . . .	28
D.2	Proofs of Lemma 23 and Corollary 24 . . . . .	30
<b>E</b>	<b>Adaptive Meta-Algorithm for Mixable Losses (AdaMix)</b>	<b>31</b>
<b>F</b>	<b>Why the Ada-BARRONS Restarts Work</b>	<b>34</b>



## Appendix A. Technical Lemmas

**Lemma 18** *Let  $\Psi(\mathbf{x}) := -\sum_{i=1}^d \ln \bar{x}_i$  and  $f(\mathbf{w}) := -\log\langle \mathbf{r}, \bar{\mathbf{w}} \rangle$ , for  $\mathbf{r} \in [0, 1]^d$ , and  $\mathbf{x}, \mathbf{w} \in \mathcal{C}_{d-1}$ . Then, for all  $\mu \in (0, 1)$  and  $\tilde{\mathbf{w}} \in \mathcal{C}_{d-1}$ , we have*

$$\nabla f(\mathbf{w})\nabla f(\mathbf{w})^\top \preceq (1 - \mu)^{-2}\nabla^2\Psi(\tilde{\mathbf{w}}), \quad \text{where } \mathbf{w} := (1 - \mu)\tilde{\mathbf{w}} + \mu\mathbf{1}/d. \quad (20)$$

Furthermore,  $\|\nabla f(\mathbf{w})\|_{\nabla^{-2}\Psi(\tilde{\mathbf{w}})}^2 \leq (1 - \mu)^{-2}$ .

**Proof .** Since  $\|\nabla f(\mathbf{w})\|_{\nabla^{-2}\Psi(\tilde{\mathbf{w}})}^2 = \nabla f(\mathbf{w})^\top \nabla^{-2}\Psi(\tilde{\mathbf{w}})\nabla f(\mathbf{w})$ , the second claim follows from (20). Thus, it suffices to show that  $\nabla f(\mathbf{w})\nabla f(\mathbf{w})^\top \preceq (1 - \mu)^{-2}\nabla^2\Psi(\tilde{\mathbf{w}})$ . Letting  $\tilde{\mathbf{r}} := (r_1, \dots, r_{d-1})$ , we have

$$\begin{aligned} \nabla f(\mathbf{w})\nabla f(\mathbf{w})^\top &= (\tilde{\mathbf{r}} - r_d\mathbf{1})(\tilde{\mathbf{r}} - r_d\mathbf{1})^\top / \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^2, \\ &= (\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top + r_d^2\mathbf{1}\mathbf{1}^\top - r_d\tilde{\mathbf{r}}\mathbf{1}^\top - r_d\mathbf{1}\tilde{\mathbf{r}}^\top) / \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^2, \\ &\preceq (\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top + r_d^2\mathbf{1}\mathbf{1}^\top) / \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^2, \\ &\preceq (\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top + r_d^2\mathbf{1}\mathbf{1}^\top) / \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^2, \\ &\preceq \tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top / \langle \tilde{\mathbf{r}}, \mathbf{w} \rangle^2 + \mathbf{1}\mathbf{1}^\top / w_d^2, \\ &\preceq \tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top / \langle \tilde{\mathbf{r}}, \mathbf{w} \rangle^2 + (1 - \mu)^{-2}\mathbf{1}\mathbf{1}^\top / \tilde{w}_d^2. \end{aligned} \quad (21)$$

We will now show that  $\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top / \langle \tilde{\mathbf{r}}, \mathbf{w} \rangle^2 \preceq \text{diag}(1/w_1^2, \dots, 1/w_{d-1}^2)$ . For this, it suffices to show that for any vector  $\mathbf{u} \in \mathbb{R}^d$ , we have  $(\sum_{i=1}^{d-1} r_i u_i)^2 / (\sum_{i=1}^{d-1} r_i w_i)^2 \leq \sum_{i=1}^{d-1} u_i^2 / w_i^2$ . This is indeed the case; by Cauchy Schwarz, we have

$$\begin{aligned} \left( \sum_{i=1}^{d-1} r_i u_i \right)^2 &\leq \left( \sum_{i=1}^{d-1} r_i^2 w_i^2 \right) \left( \sum_{i=1}^{d-1} u_i^2 / w_i^2 \right), \\ &\leq \left( \sum_{i=1}^{d-1} r_i w_i \right)^2 \left( \sum_{i=1}^{d-1} u_i^2 / w_i^2 \right), \end{aligned}$$

where the last inequality follows by the fact that  $r_i, w_i \geq 0$  for all  $i$ . Therefore, we have  $\tilde{\mathbf{r}}\tilde{\mathbf{r}}^\top / \langle \tilde{\mathbf{r}}, \mathbf{w} \rangle^2 \preceq \text{diag}(1/w_1^2, \dots, 1/w_{d-1}^2)$ . Plugging this into (21) implies that

$$\nabla f(\mathbf{w})\nabla f(\mathbf{w})^\top \preceq \text{diag}(1/w_1^2, \dots, 1/w_{d-1}^2) + (1 - \mu)^{-2}\mathbf{1}\mathbf{1}^\top / \tilde{w}_d^2 \preceq (1 - \mu)^{-2}\nabla^2\Psi(\tilde{\mathbf{w}}). \quad \blacksquare$$

**Lemma 19** *Let  $\beta \leq 1/8$  and  $\eta \in (0, 1)$ . If  $T > 1$ , then the iterates  $(\mathbf{w}_t)$  and  $(\mathbf{u}_t)$  in Alg. 1, satisfy*

$$\sum_{t=1}^T \|\mathbf{g}_t\|_{\nabla^{-2}F_t(\mathbf{w}_t)}^2 \leq \frac{16d \ln T}{\beta},$$

where  $(F_t)$  are as in (14) and  $\mathbf{g}_t = \nabla \ell_t(\mathbf{u}_t)$ .

**Proof .** Let  $c_T := 1 - 1/T$ . First note that by Lemma 18, we have  $c_T^2 \mathbf{g}_t \mathbf{g}_t^\top \preceq \nabla^2 \Psi(\mathbf{w}_t)$ , where  $\Psi(\mathbf{x}) := -\sum_{i=1}^d \ln \bar{x}_i$ . Combining this with the fact that  $\eta_t \in [\eta, \eta e]$  and the range assumptions on  $\beta, \eta$ , and  $T$ , we get

$$\|\mathbf{g}_t\|_{\nabla^{-2}F_t(\mathbf{w}_t)}^2 = \mathbf{g}_t^\top (\nabla^2 \Psi_t(\mathbf{w}_t) + \beta Q_{t-1}/4)^{-1} \mathbf{g}_t \leq 4\beta^{-1} \mathbf{g}_t^\top Q_t^{-1} \mathbf{g}_t. \quad (22)$$

where  $Q_t := dI + \sum_{s=1}^t \mathbf{g}_s \mathbf{g}_s^\top$ . Thus, by (22) and (Hazan et al., 2007, Lemma 11), we have

$$\|\mathbf{g}_t\|_{\nabla^{-2}F_t(\mathbf{w}_t)}^2 \leq \frac{4}{\beta} \sum_{t=1}^T \mathbf{g}_t^\top Q_t^{-1} \mathbf{g}_t \leq \frac{4}{\beta} \ln \frac{|Q_T|}{|Q_0|} \leq \frac{4d \ln(1+T^3)}{\beta} \leq \frac{16d \ln T}{\beta},$$

where the second inequality uses the fact  $\ln |Q_0| = d \ln d$  and by AM-GM inequality  $\ln |Q_T| \leq d \ln \frac{\text{Tr}(Q_T)}{d} \leq d \ln \left( d + \sum_{t=1}^T \|\mathbf{g}_t\|_2^2/d \right) \leq d \ln(d + dT^3)$  since

$$\|\mathbf{g}_t\|_2^2 \leq d^2 T^2 \frac{\sum_{i=1}^d r_{t,i}^2}{\left( \sum_{i=1}^d r_{t,i} \right)^2} \leq d^2 T^2.$$

This completes the proof.  $\blacksquare$

**Lemma 20** Suppose that  $T > 1$  and define  $c_T := 1 - 1/T$  and  $\alpha_T := 1 + \beta \eta c_T^{-2}/4$ . If  $\eta, \beta \in (0, 1)$ , then the iterates  $(\mathbf{w}_t)$  and gradients  $(\mathbf{g}_t)$  in Algorithm 1, and the regularizers  $(\Phi_t)$  in (9) are such that, for all  $t \in [T]$ ,

$$\nabla^2 \Phi_t(\mathbf{w}_t) \preceq \alpha_T \nabla^2 \Phi_{t-1}(\mathbf{w}_t), \quad \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_t(\mathbf{w}_t)}^2 \leq \frac{e\eta}{c_T^2}, \quad \text{and} \quad \sum_{s=1}^t \|\mathbf{g}_s\|_{\nabla^{-2}\Phi_s(\mathbf{w}_s)}^2 \leq \frac{16d \ln T}{\beta}.$$

**Proof .** By Lemma 18, we have  $\nabla^2 \Psi(\mathbf{w}_t) \succeq c_T^2 \mathbf{g}_t \mathbf{g}_t^\top$ , where  $\Psi(\mathbf{x}) := -\sum_{i=1}^d \ln \bar{x}_i$ . Using this and the fact that  $(\eta_{t,i}) \subset [\eta, \eta e]$ , we get that  $\beta \mathbf{g}_t \mathbf{g}_t^\top / 4 \leq \beta \eta c_T^{-2} \nabla^2 \Psi_{t-1}(\mathbf{w}_t) / 4 \preceq \beta \eta c_T^{-2} \nabla^2 \Phi_{t-1}(\mathbf{w}_t) / 4$ . Thus, adding  $\nabla^2 \Psi_{t-1}(\mathbf{w}_t)$  on both sides and using that  $\alpha_t = 1 + \beta \eta c_T^{-2}/4$ , we get that

$$\alpha_T \nabla^2 \Phi_{t-1}(\mathbf{w}_t) \succeq \beta \mathbf{g}_t \mathbf{g}_t^\top / 4 + \nabla^2 \Phi_{t-1}(\mathbf{w}_t) \succeq \nabla^2 \Phi_t(\mathbf{w}_t),$$

where the last inequality follows by the fact that  $\eta_{t,i} \geq \eta_{t-1,i}$ , for all  $i \in [d]$ . The remaining inequalities follow from Lemmas 18 and 19.  $\blacksquare$

**Lemma 21** Let  $\mathbf{u}, \mathbf{w} \in \mathcal{C}_{d-1}$ . For any  $\mathbf{r} \in [0, 1]^d$  and  $f(\mathbf{x}) := -\log \langle \mathbf{r}, \bar{\mathbf{x}} \rangle$ , we have

$$|\langle \nabla f(\mathbf{w}), \mathbf{u} - \mathbf{w} \rangle| \leq 1 \vee \max_{i \in [d]} \frac{\bar{u}_i}{\bar{w}_i}.$$

**Proof .** We have

$$\langle \nabla f(\mathbf{w}), \mathbf{u} - \mathbf{w} \rangle = \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^{-1} \langle J^\top \mathbf{r}, \mathbf{u} - \mathbf{w} \rangle = \langle \mathbf{r}, \bar{\mathbf{w}} \rangle^{-1} \langle \mathbf{r}, \bar{\mathbf{u}} - \bar{\mathbf{w}} \rangle = \frac{\langle \mathbf{r}, \bar{\mathbf{u}} \rangle}{\langle \mathbf{r}, \bar{\mathbf{w}} \rangle} - 1.$$

Now since the function  $\mathbf{r} \mapsto \frac{\langle \mathbf{r}, \bar{\mathbf{u}} \rangle}{\langle \mathbf{r}, \bar{\mathbf{w}} \rangle}$  is quasi-convex (Boyd et al., 2004, Example 3.32), its maximum is reached at the boundary of  $[0, 1]^d$ . Thus, the previous display implies that

$$|\langle \nabla f(\mathbf{w}), \mathbf{u} - \mathbf{w} \rangle| \leq 1 \vee \sup_{\tilde{\mathbf{r}} \in [0, 1]^d} \frac{\langle \tilde{\mathbf{r}}, \bar{\mathbf{u}} \rangle}{\langle \tilde{\mathbf{r}}, \bar{\mathbf{w}} \rangle} \leq 1 \vee \max_{i \in [d]} \frac{\bar{u}_i}{\bar{w}_i}. \quad \blacksquare$$

## Appendix B. Proof of Lemmas 5, 7, 8, and 9

**Proof of Lemma 5.** Since  $\nabla f(\mathbf{w}) - \nabla f(\mathbf{p}) = \int_0^1 \nabla^2 f(\mathbf{w}_\mu)(\mathbf{w} - \mathbf{p}) d\mu$ , where  $\mathbf{w}_\mu := \mu\mathbf{w} + (1-\mu)\mathbf{p}$ , we have,

$$\begin{aligned} & \|\nabla f(\mathbf{w}) - \nabla f(\mathbf{p})\|_{\nabla^{-2}f(\mathbf{w})}^2 \\ &= (\nabla f(\mathbf{w}) - \nabla f(\mathbf{p}))^\top \nabla^{-2}f(\mathbf{w})(\nabla f(\mathbf{w}) - \nabla f(\mathbf{p})), \\ &= (\mathbf{w} - \mathbf{p})^\top \left( \int_0^1 \nabla^2 f(\mathbf{w}_\mu) d\mu \right)^\top \nabla^{-2}f(\mathbf{w}) \left( \int_0^1 \nabla^2 f(\mathbf{w}_\nu) d\nu \right) (\mathbf{w} - \mathbf{p}). \end{aligned} \quad (23)$$

On the other hand, since  $f$  is self-concordant with constant  $M_f$  and  $r := \|\mathbf{p} - \mathbf{w}\|_{\nabla^2 f(\mathbf{w})} \leq 1/M_f$  by assumption, we have (see e.g. (Nesterov et al., 2018, Corollary 5.1.5))

$$(1 - M_f r + 1/3 M_f^2 r^2) \nabla^2 f(\mathbf{w}) \preceq H \preceq \frac{1}{1 - M_f r} \nabla^2 f(\mathbf{w}), \quad (24)$$

where  $H := \int_0^1 \nabla^2 f(\mathbf{w}_\mu) d\mu$ . Since  $\nabla^2 f(\mathbf{w})$  is definite positive for all  $\mathbf{w}$ , (24) further implies that  $H \succeq (1 - M_f r) \nabla^2 f(\mathbf{w})$ . Combining these facts with (23), we get that

$$\|\nabla f(\mathbf{w}) - \nabla f(\mathbf{p})\|_{\nabla^{-2}f(\mathbf{w})}^2 \preceq \frac{1}{(1 - M_f r)^2} \|\mathbf{p} - \mathbf{w}\|_{\nabla^2 f(\mathbf{w})}^2. \quad \blacksquare$$

**Proof of Lemma 7.** We prove the claim by induction. We start with the base case  $t = 1$ . For  $t = 1$ , we have  $\mathbf{w}_1 = 1/d \in \mathcal{C}_{d-1}$ . We now check that  $\Phi_1$  is self-concordant with constant  $\sqrt{e\eta}$ . For any  $i \in [d]$ , the function  $\vartheta_i: \mathbf{x} \rightarrow -\ln \bar{x}_i$  defined on  $\mathcal{C}_{d-1}$  is self-concordant with constant 1. Furthermore, the function  $\Theta_1: \mathbf{u} \mapsto \Phi_1(\mathbf{u}) - \Psi_1(\mathbf{u})$  is self-concordant with constant 0 (since it is a quadratic). Thus, by (Nesterov et al., 2018, Theorem 5.1.1) and the fact that  $\Phi_1(\cdot) = \Theta_1(\cdot) + \Psi_1(\cdot) = \Theta_1(\cdot) + \sum_{i=1}^d \vartheta_i(\cdot)/\eta_{1,i}$ , we have that  $\Phi_1$  is self-concordant with constant less than  $0 \vee \max_{i \in [d]} \sqrt{\eta_{1,i}} = \sqrt{\eta} \leq \sqrt{\eta e}$ .

Now, suppose the claim of the lemma holds for all  $t \leq s$ . We will show that it holds for  $t = s + 1$ . Since  $\Phi_s$  is self-concordant with constant  $M_{\Phi_s} \leq \sqrt{\eta e}$  and  $\mathbf{w}_s \in \mathcal{C}_{d-1}$  (by the induction hypothesis), we have that the *Dikin ellipsoid*

$$\mathcal{W}_s := \{\mathbf{x} \in \mathbb{R}^{d-1} : \|\mathbf{x} - \mathbf{w}_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)} < 1/\sqrt{\eta e}\}$$

is contained within  $\text{dom } \Phi_s = \mathcal{C}_{d-1}$  (see e.g. (Nesterov et al., 2018, Thm. 5.1.5)). Thus, to show that  $\mathbf{w}_{s+1} \in \mathcal{C}_{d-1}$ , it suffices to show that  $\mathbf{w}_{s+1} \in \mathcal{W}_s$ . By definition of  $\mathbf{w}_{s+1}$ , we have

$$\|\mathbf{w}_{s+1} - \mathbf{w}_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)} = \frac{\|\nabla^{-2} \Phi_s(\mathbf{w}_s) \nabla_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)}}{1 + 4\sqrt{\eta e} \|\nabla_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)}} = \frac{\|\nabla_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)}}{1 + 4\sqrt{\eta e} \|\nabla_s\|_{\nabla^2 \Phi_s(\mathbf{w}_s)}} < \frac{1}{\sqrt{\eta e}},$$

where  $\nabla_s$  is as in (12). This shows that  $\mathbf{w}_{s+1} \in \mathcal{W}_s$  and so  $\mathbf{w}_{s+1} \in \mathcal{C}_{d-1}$ . As a consequence, the output  $\mathbf{u}_{s+1}$  of Algorithm 1 satisfies  $\bar{u}_{s+1,i} \geq 1/(dT)$ , for all  $i \in [d]$ , which in turn implies that  $\eta_{s+1,i} \in [\eta, e\eta]$ , for all  $i \in [d]$ . Using this and (Nesterov et al., 2018, Theorem 5.1.1), we have that  $\Phi_{s+1}(\cdot) = \Phi_{s+1}(\cdot) - \Psi_{s+1}(\cdot) + \sum_{i=1}^d \vartheta_i(\cdot)/\eta_{s+1,i}$  is self-concordant with constant less than  $0 \vee \max_{i \in [d]} \sqrt{\eta_{s+1,i}} \leq \sqrt{\eta e}$ .  $\blacksquare$

**Proof of Lemma 8.** For any twice differentiable function  $F: \mathcal{W} \rightarrow \mathbb{R}$  and  $\mathbf{w} \in \mathcal{W}$ , we recall the definition of the Newton decrement  $\lambda(\mathbf{w}, F) := \|\nabla F(\mathbf{w})\|_{\nabla^{-2} F(\mathbf{w})}$  which will be useful in this proof. First, note that if  $T = 1$ , then the result holds trivially since  $\mathbf{w}_1 = \mathbf{p}_1 = \mathbf{1}/d$ . Assume that  $T > 1$  and let  $c_T := 1 - 1/T$ . Next, we will show by induction that

$$\frac{1}{16\sqrt{e\eta}} \|\mathbf{w}_s - \mathbf{p}_s\|_{\nabla^2 F_s(\mathbf{w}_s)} \leq \frac{1}{8\sqrt{e\eta}} \lambda(\mathbf{w}_s, F_s) \leq \lambda(\mathbf{w}_{s-1}, F_s)^2 \leq C\eta, \quad (25)$$

for all  $s \geq 1$ , where  $C := 4e/c_T^2$  with the convention that  $\mathbf{w}_0 = \mathbf{1}/d$ . The base case follows trivially since  $\nabla F_1(\mathbf{w}_0) = \nabla F_1(\mathbf{w}_1) = \mathbf{0}$  and  $\mathbf{w}_1 = \mathbf{p}_1$ . Suppose that (25) holds for  $s = t$ . We will show that it holds for  $s = t + 1$ . By Lemma 9, we have  $\nabla F_{t+1}(\mathbf{w}_t) = \mathbf{g}_t + \nabla F_t(\mathbf{w}_t)$ , and so by the fact that  $(a + b)^2 \leq 2a^2 + 2b^2$ , we get

$$\begin{aligned} \lambda(\mathbf{w}_t, F_{t+1})^2 &= \|\nabla F_{t+1}(\mathbf{w}_t)\|_{\nabla^{-2} F_t(\mathbf{w}_t)}^2, \\ &\leq 2\|\nabla F_t(\mathbf{w}_t)\|_{\nabla^{-2} F_t(\mathbf{w}_t)}^2 + 2\|\mathbf{g}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{w}_t)}^2, \\ &= 2\lambda(\mathbf{w}_t, F_t)^2 + 2\|\mathbf{g}_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}^2, \end{aligned} \quad (26)$$

$$\leq 2^T e C^2 \eta^3 + 2e\eta/c_T^2 \leq C\eta, \quad (27)$$

where in the penultimate inequality we used the induction hypothesis in (25) for  $s = t$  and the bound on  $\|\mathbf{g}_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}^2$  from Lemma 20. The last inequality in (27) uses the range assumptions on  $\eta$ . Now, by the expression of  $\nabla F_{t+1}(\mathbf{w}_t)$  in Lemma 9, one can verify that the iterate  $\mathbf{w}_{t+1}$  in Algorithm 1 satisfies

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \frac{1}{1 + 4\sqrt{e\eta} \lambda(\mathbf{w}_t, F_{t+1})} \nabla^{-2} F_{t+1}(\mathbf{w}_t) \nabla F_{t+1}(\mathbf{w}_t),$$

which is the damped Newton step with respect to the function  $F_{t+1}$ . Therefore, by Lemma 3 and the fact that  $\lambda(\mathbf{w}_t, F_{t+1}) \leq 1/(8\sqrt{e\eta})$  (which follows from (27) and the range assumption on  $\eta$ ), we have  $\lambda(\mathbf{w}_{t+1}, F_{t+1}) \leq 8\sqrt{e\eta} \lambda(\mathbf{w}_t, F_{t+1})^2$ . Furthermore, since  $\mathbf{p}_{t+1}$  is the minimizer of  $F_{t+1}$  and  $\lambda(\mathbf{w}_{t+1}, F_{t+1}) \leq 1/(2\sqrt{e\eta})$ , we have  $\|\mathbf{w}_{t+1} - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{w}_t)} \leq 2\lambda(\mathbf{w}_{t+1}, F_{t+1})$  (by Lemma 3 again). Combining these facts with (27), implies (25) for  $s = t + 1$ , which concludes the induction. This shows (16).

We now use (25) together with (26) to bound the sum  $\sum_{t=1}^T \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2$ . Using that  $\lambda(\mathbf{w}_{t+1}, F_{t+1}) \leq 8\sqrt{e\eta} \lambda(\mathbf{w}_t, F_{t+1})^2$  (as argued above) and (26), we get

$$\lambda(\mathbf{w}_{t+1}, F_{t+1}) \leq 16\sqrt{e\eta} \lambda(\mathbf{w}_t, F_t)^2 + 16\sqrt{e\eta} \|\mathbf{g}_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}^2. \quad (28)$$

Summing (28), for  $t = 1, \dots, T$ , rearranging, and using that  $\lambda(\mathbf{w}_{T+1}, F_{T+1}) \geq 0$ , we get

$$\sum_{t=2}^T (\lambda(\mathbf{w}_t, F_t) - 16\sqrt{e\eta}\lambda(\mathbf{w}_t, F_t)^2) \leq 16\sqrt{e\eta}\lambda(\mathbf{w}_1, F_1)^2 + 16\sqrt{e\eta} \sum_{t=2}^T \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_t(\mathbf{w}_t)}^2.$$

Using (26) and the range assumption on  $\eta$ , we have  $0 \leq 16\sqrt{e\eta}\lambda(\mathbf{w}_t, F_t) \leq 2^7 eC\eta^2 \leq 1/4$ . Therefore, we have

$$\begin{aligned} \frac{3}{4} \sum_{t=1}^T \lambda(\mathbf{w}_t, F_t) &\leq \lambda(\mathbf{w}_1, F_1) + 16\sqrt{e\eta} \sum_{t=2}^T \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_t(\mathbf{w}_t)}^2, \\ &\leq \frac{1}{16} + 16\sqrt{e\eta} \sum_{t=1}^T \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_t(\mathbf{w}_t)}^2 \leq \frac{3}{8} + \frac{45d \ln T}{8\beta}, \end{aligned}$$

where the last inequality follows by Lemma 20 and the range assumption on  $\eta$ . Now, using the fact that  $\mathbf{p}_t$  is the minimizer of  $F_t$ , we have  $\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 F_t(\mathbf{w}_t)} \leq 2\lambda(\mathbf{w}_t, F_t)$ . Combining this with the previous display, we get the desired result.  $\blacksquare$

**Proof of Lemma 9.** Let  $t \in [T]$ ,  $\vartheta_i: \mathbf{x} \mapsto -\ln \bar{x}_i$ , and  $\delta_{t,i} := (1/\eta_{t,i} - 1/\eta_{t-1,i})$  with  $\eta_0 = \eta_1$ , for all  $i \in [d]$ . We will first show that  $\nabla F_{t+1}(\mathbf{w}) = \tilde{\mathbf{g}}_t(\mathbf{w}) + \nabla F_t(\mathbf{w})$ , where

$$\tilde{\mathbf{g}}_t(\mathbf{w}) := \mathbf{g}_t \cdot (1 + \beta \langle \mathbf{g}_t, \mathbf{w} - \mathbf{w}_t \rangle / 4) + \sum_{i=1}^d \delta_{t,i} (\nabla \vartheta_i(\mathbf{w}) - \nabla \vartheta_i(\mathbf{w}_t)).$$

By definition of  $(F_t)$  and  $(\mathbf{p}_t)$ , we have

$$\begin{aligned} \nabla F_{t+1}(\mathbf{w}) &= \tilde{\mathbf{g}}_t + \nabla \Phi_t(\mathbf{w}) - \nabla \Phi_t(\mathbf{p}_t), \\ &= \tilde{\mathbf{g}}_t + \beta \mathbf{g}_t \cdot (\langle \mathbf{g}_t, \mathbf{w} - \mathbf{w}_t \rangle - \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle) / 4 + \sum_{i=1}^d \delta_{t,i} (\nabla \vartheta_i(\mathbf{w}) - \nabla \vartheta_i(\mathbf{p}_t)) \\ &\quad + \nabla \Phi_{t-1}(\mathbf{w}) - \nabla \Phi_{t-1}(\mathbf{p}_t), \\ &= \mathbf{g}_t \cdot (1 + \beta \langle \mathbf{g}_t, \mathbf{w} - \mathbf{w}_t \rangle / 4) + \sum_{i=1}^d \delta_{t,i} (\nabla \vartheta_i(\mathbf{w}) - \nabla \vartheta_i(\mathbf{w}_t)) \\ &\quad + \mathbf{g}_{t-1} + \nabla \Phi_{t-1}(\mathbf{w}) - \nabla \Phi_{t-1}(\mathbf{p}_{t-1}), \end{aligned} \tag{29}$$

$$= \mathbf{g}_t \cdot (1 + \beta \langle \mathbf{g}_t, \mathbf{w} - \mathbf{w}_t \rangle / 4) + \sum_{i=1}^d \delta_{t,i} (\nabla \vartheta_i(\mathbf{w}) - \nabla \vartheta_i(\mathbf{w}_t)) + \nabla F_t(\mathbf{w}), \tag{30}$$

where (29) follows by definition of  $\tilde{\mathbf{g}}_t$  in (15) and the fact that  $\mathbf{0} = \nabla F_t(\mathbf{p}_t) = \mathbf{g}_{t-1} + \nabla \Phi_{t-1}(\mathbf{p}_t) - \nabla \Phi_t(\mathbf{p}_{t-1})$ . Setting  $\mathbf{w} = \mathbf{w}_t$  in (30) shows the first equality of the lemma. Now, by induction, we get

$$\begin{aligned} \nabla F_{t+1}(\mathbf{w}) &= \sum_{s=1}^t \mathbf{g}_s \cdot (1 + \beta \langle \mathbf{g}_s, \mathbf{w} - \mathbf{w}_s \rangle / 4) + \sum_{s=1}^t \sum_{i=1}^d \delta_{s,i} (\nabla \vartheta_i(\mathbf{w}) - \nabla \vartheta_i(\mathbf{w}_s)) + \nabla F_1(\mathbf{w}), \\ &= \sum_{s=1}^t \mathbf{g}_s \cdot (1 + \beta \langle \mathbf{g}_s, \mathbf{w} - \mathbf{w}_s \rangle / 4) + \nabla \Psi_t(\mathbf{w}) - \sum_{s=1}^t \sum_{i=1}^d \delta_{s,i} \vartheta_i(\mathbf{w}_s) + \beta d \mathbf{w} / 4. \end{aligned}$$

Using that  $\nabla\Phi_t(\mathbf{w}) = \beta d\mathbf{w}/4 + \sum_{s=1}^t \beta \mathbf{g}_s \cdot \langle \mathbf{g}_s, \mathbf{w} - \mathbf{w}_s \rangle / 4 + \nabla\Psi_t(\mathbf{w})$  completes the proof. ■

### Appendix C. Proof of Theorem 15 (Regret of Base Algorithm)

We present the proof of Theorem 15 before proving Lemmas 10-14.

**Proof of Theorem 15.** Our starting point is the regret decomposition in Lemma 10. Using Lemmas 11, 12, and 13 to bound the first two sums on the RHS of the regret decomposition (17), we get

$$\begin{aligned} \sum_{t=1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u}'')) &\leq \frac{18d \ln T}{\beta} + O\left(\frac{d \ln T}{\eta}\right) - \frac{1}{32\eta \ln T} \sum_{i=1}^d \max_{t \leq T} \frac{\bar{u}_i''}{\bar{u}_{t,i}} \\ &\quad + c_T \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \frac{3\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2. \end{aligned} \quad (31)$$

Now using Lemma 14 to bound the last two sums in (31), we get the desired result. ■

In addition to Lemmas 10-14 in the main body, we also need the following result (which follows from the proof of (Luo et al., 2018, Lemma 5)) to prove Theorem 15:

**Lemma 22** *Let  $(\mathbf{p}_t)$  and  $(\tilde{\mathbf{g}}_t)$  be as in (14) and (15), respectively. Then,  $\forall t \in [T], \forall \mathbf{u} \in \mathcal{C}_{d-1}$ ,*

$$\sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle \leq O\left(\frac{d \ln T}{\eta}\right) + \sum_{t=1}^T \left( \frac{\beta}{8} \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{u}' \rangle^2 + D_{\Psi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{p}_t) \right).$$

**Proof.** In this proof, we let  $\Theta_t := \Phi_t - \Psi_t$ , for all  $t \geq 1$ . Since  $\mathbf{p}_{t+1}$  is the minimizer of  $F_{t+1}$ , which is a self-concordant barrier for the set  $\mathcal{C}_{d-1}$ , we have  $\mathbf{0} = \nabla F_{t+1}(\mathbf{p}_{t+1}) = \tilde{\mathbf{g}}_t + \nabla\Phi_t(\mathbf{p}_{t+1}) - \nabla\Phi_t(\mathbf{p}_t)$ . Therefore, we have

$$\begin{aligned} \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle &\leq \langle \nabla\Phi_t(\mathbf{p}_{t+1}) - \nabla\Phi_t(\mathbf{p}_t), \mathbf{u}' - \mathbf{p}_{t+1} \rangle, \\ &= D_{\Phi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}', \mathbf{p}_{t+1}) - D_{\Phi_t}(\mathbf{p}_{t+1}, \mathbf{p}_t), \end{aligned} \quad (32)$$

$$\leq D_{\Phi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}', \mathbf{p}_{t+1}), \quad (33)$$

$$= D_{\Psi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Psi_t}(\mathbf{u}', \mathbf{p}_{t+1}) + D_{\Theta_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Theta_t}(\mathbf{u}', \mathbf{p}_{t+1}), \quad (34)$$

where (32) follows by the definition of the Bregman divergence, and (33) follows by the positivity of the Bregman divergence. Summing (34) for  $t = 1$  to  $T$ , we get

$$\begin{aligned} \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle &\leq D_{\Psi_0}(\mathbf{u}', \mathbf{w}_1) + D_{\Theta_0}(\mathbf{u}', \mathbf{w}_1) + \sum_{t=1}^T (D_{\Theta_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Theta_{t-1}}(\mathbf{u}', \mathbf{p}_t)) \\ &\quad + \sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{p}_t)), \\ &= D_{\Psi_0}(\mathbf{u}', \mathbf{w}_1) + D_{\Theta_0}(\mathbf{u}', \mathbf{w}_1) + \sum_{t=1}^T \frac{\beta}{8} \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{u}' \rangle^2 \\ &\quad + \sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{p}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{p}_t)). \end{aligned} \quad (35)$$

By definition of  $\mathbf{u}'$  and  $\mathbf{w}_1$ , we have  $D_{\Psi_0}(\mathbf{u}', \mathbf{w}_1) + D_{\Theta_0}(\mathbf{u}', \mathbf{w}_1) \leq O(\eta^{-1}d \ln T)$ , which combined with (35) implies the desired result.  $\blacksquare$

**Proof of Lemma 11.** Denote by  $\mathcal{T} \in [T]$  the subset of rounds  $t$  where any of  $(\rho_{t,i})_{i \in [d]}$  change. For  $t \notin \mathcal{T}$ , we have  $\psi_t \equiv 0$ , which is self-concordant with any constant. Now, let  $t \in \mathcal{T}$ . Since  $\psi_t$  is the sum of self-concordant functions, we have by (Nesterov et al., 2018, Theorem 5.1.1) that  $\psi_t$  is a self-concordant function with constant less than

$$\begin{aligned} \max_{i \in [d]} \sqrt{\left(\frac{1}{\eta_{t-1,i}} - \frac{1}{\eta_{t,i}}\right)^{-1}} &= \max_{i \in [d]} \sqrt{\frac{\eta_{t,i}}{e^{\log_T(\rho_{t,i}/\rho_{t-1,i})} - 1}} \\ &\stackrel{(*)}{\leq} \max_{i \in [d]} \sqrt{\frac{e\eta}{\log_T(\rho_{t,i}/\rho_{t-1,i})}} \stackrel{(**)}{\leq} \sqrt{e\eta \log_2 T}, \end{aligned}$$

where  $(*)$  follows by the fact that  $e^x - 1 \geq x$ , for all  $x \in \mathbb{R}$ , and that  $\rho_{t,i} > 2\rho_{t-1,i}$  since  $t \in \mathcal{T}$ . This shows the first claim of the lemma. We now show the second claim. Let  $t \in \mathcal{T}$  and define  $\mathbf{w}_{t,\mu} := \mu \mathbf{p}_t + (1 - \mu)\mathbf{w}_t$ , for  $\mu \in [0, 1]$ . By Lemmas 4, 7, and 8, we have

$$(1 - \sqrt{e\eta}r_t)^2 \nabla^{-2} \Phi_{t-1}(\mathbf{w}_{t,\mu}) \preceq \nabla^{-2} \Phi_{t-1}(\mathbf{w}_t) \preceq (1 - \sqrt{e\eta}r_t)^{-2} \nabla^{-2} \Phi_{t-1}(\mathbf{w}_{t,\mu}), \quad (36)$$

where  $r_t := \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}$ . By Taylor's theorem, there exists  $\mu_* \in [0, 1]$  such that

$$D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) = \frac{1}{2} \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2 \psi_t(\mathbf{w}_{t,\mu_*})}^2 \leq \frac{1}{2} \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_{t,\mu_*})}^2, \quad (37)$$

$$\leq \frac{1}{2(1 - \sqrt{e\eta}r_t)^2} \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2. \quad (38)$$

where (37) follows by the fact that  $\nabla^2 \psi_t \preceq \nabla^2 \Psi_{t-1} \preceq \nabla^2 \Phi_{t-1}$  and the last inequality follows by (36). Plugging the bound on  $r_t = \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2 \leq 64(e\eta)^{3/2}/c_T^2$  (from Lemma 8) into (38) and using the facts that  $\mathbf{I} |\mathcal{T}| \leq O(d \ln T)$  (by definition of  $(\rho_{t,i})$  in (11) and the fact that  $(\bar{u}_{t,i}) \subset [1/(dT), 1]$ ); and  $\mathbf{II} D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) = 0$  if  $t \notin \mathcal{T}$ , we get that  $\sum_{t=1}^T D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) \leq O(\eta^{3/2}d \ln T) \leq O(d \ln T)$ .  $\blacksquare$

**Proof of Lemma 10.** Let  $c_T := 1 - 1/T$  and  $\mathbf{u}$  be as in the lemma's statement and recall the definition of  $(\hat{\mathbf{g}}_t)$  from (15). First, we note that by Lemma 1, we have  $\sum_{t=1}^T \ell_t(\mathbf{u}) \leq \sum_{t=1}^T \ell_t(\mathbf{u}'') + 4$ , and so it suffices to bound the regret against  $\mathbf{u}''$ . Let  $\hat{\mathbf{g}}_t := \mathbf{g}_t(1 + \beta \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle / 4)$ . Using that Cover's loss is 1-exp-concave and  $\beta \leq 8^{-1} \wedge |8 \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle|^{-1}$ , for all  $t \geq 1$ , we have (see e.g. proof of (Luo et al., 2018, Lemma 5) for the first inequality)

$$\begin{aligned} &\sum_{t=1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u}'')) \\ &\leq \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle - \frac{\beta}{2} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle^2, \\ &\leq \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u}_t - \mathbf{u}'' \rangle + \frac{\beta}{4} \sum_{t=1}^T |\langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle| - \frac{\beta}{2} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle^2, \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{u}_t - \mathbf{u}'' \rangle + \frac{\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2 - \frac{3\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{u}_t - \mathbf{u}'' \rangle^2, \\
 &= c_T \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{u}' \rangle + \frac{\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2 - \frac{3c_T^2\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u}' \rangle^2, \\
 &= c_T \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{p}_t \rangle + \frac{\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2 - \frac{3c_T^2\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u}' \rangle^2 \\
 &\quad + c_T \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u}' \rangle, \\
 &\leq c_T \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle - \frac{3c_T^2\beta}{8} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{u}' \rangle^2 + c_T \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u}' \rangle, \tag{39}
 \end{aligned}$$

where in the last inequality we used the fact that  $\langle \hat{\mathbf{g}}_t, \mathbf{w}_t - \mathbf{p}_t \rangle = \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle - \beta \langle \mathbf{g}_t, \mathbf{w}_t - \mathbf{p}_t \rangle^2 / 4$  and  $c_T \geq 1/2$ . We now focus on the last sum in (39). By definition of  $(\tilde{\mathbf{g}}_t)$  and  $(\hat{\mathbf{g}}_t)$ , we have

$$\begin{aligned}
 \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u}' \rangle &= \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u}' \rangle \\
 &\quad + \sum_{t=1}^T \sum_{i=1}^d \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) \langle \nabla \vartheta_i(\mathbf{w}_t) - \nabla \vartheta_i(\mathbf{p}_t), \mathbf{p}_t - \mathbf{u}' \rangle, \\
 &= \sum_{t=1}^T \sum_{i=1}^d \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) (D_{\vartheta_i}(\mathbf{u}', \mathbf{w}_t) - D_{\vartheta_i}(\mathbf{u}', \mathbf{p}_t) - D_{\vartheta_i}(\mathbf{p}_t, \mathbf{w}_t)) \\
 &\quad + \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle, \tag{40} \\
 &= \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle \\
 &\quad + \sum_{t=1}^T \sum_{i=1}^d \left( \frac{1}{\eta_{t-1,i}} - \frac{1}{\eta_{t,i}} \right) D_{\vartheta_i}(\mathbf{p}_t, \mathbf{w}_t) \\
 &\quad + \sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_t}(\mathbf{u}', \mathbf{p}_t) + D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{p}_t)).
 \end{aligned}$$

where (40) uses the definition of the Bregman divergence. Plugging in the bound on  $\sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_{t+1} - \mathbf{u}' \rangle$  from Lemma 22 and letting  $\psi_t: \mathbf{p} \mapsto \sum_{i=1}^d \left( \frac{1}{\eta_{t-1,i}} - \frac{1}{\eta_{t,i}} \right) \vartheta_i(\mathbf{p})$ , we get

$$\begin{aligned}
 \sum_{t=1}^T \langle \hat{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{u}' \rangle &\leq \sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle + \sum_{t=1}^T D_{\psi_t}(\mathbf{p}_t, \mathbf{w}_t) + \sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{w}_t)) \\
 &\quad + O\left(\frac{d \ln T}{\eta}\right) + \sum_{t=1}^T \frac{\beta}{8} \langle \mathbf{g}_t, \mathbf{u}' - \mathbf{p}_t \rangle^2. \tag{41}
 \end{aligned}$$



Now, by the fact  $c_T \geq 1/2$  and that  $(a+b)^2 \leq 3/2a^2 + 3b^2$ , for all  $T \geq 1$ , we get

$$\sum_{t=1}^T \frac{\beta}{8} \langle \mathbf{g}_t, \mathbf{u}' - \mathbf{p}_t \rangle^2 \leq \frac{3c_T\beta}{8} \langle \mathbf{g}_t, \mathbf{u}' - \mathbf{w}_t \rangle^2 + \frac{3\beta}{8} \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle^2.$$

By combining this with (41), (39), and the fact that  $\sum_{t=1}^T \ell_t(\mathbf{u}) \leq \sum_{t=1}^T \ell_t(\mathbf{u}'') + 4$  (see beginning of the proof), we obtain the desired result.  $\blacksquare$

**Proof of Lemma 12.** Let  $(F_t)$  be as in (14) and  $\alpha_T := 1 + \beta\eta c_T^{-2}/4$ , where  $c_T := 1 - 1/T$ . We start by bounding the Newton decrement  $\lambda(\mathbf{p}_t, F_{t+1}) = \|\nabla F_{t+1}(\mathbf{p}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{p}_t)}$ . Since  $F_{t+1}$  is equal to a linear function plus  $\Phi_t$ , Lemma 7 implies that  $F_{t+1}$  is  $\sqrt{e\eta}$ -self-concordant. On the other hand, by Lemmas 20 and Lemma 8, we have

$$\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_t(\mathbf{w}_t)} \leq \sqrt{\alpha_T} \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_{t-1}(\mathbf{w}_t)} \leq \frac{2^6\sqrt{\alpha_T}(e\eta)^{3/2}}{c_T^2} < \frac{1}{2\sqrt{e\eta}}, \quad (42)$$

where the last inequality follows by the range assumptions on  $\eta$  and  $\beta$ . This, together with the fact that  $F_{t+1}$  is  $\sqrt{e\eta}$ -self-concordant and Lemma 5, we have

$$\|\nabla F_{t+1}(\mathbf{p}_t) - \nabla F_{t+1}(\mathbf{w}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)} \leq \frac{\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_t(\mathbf{w}_t)}}{1 - \sqrt{e\eta}\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_t(\mathbf{w}_t)}} \leq \frac{2^7\sqrt{\alpha_T}}{c_T^2(e\eta)^{-3/2}}. \quad (43)$$

Thus, by the triangle inequality, we get that

$$\begin{aligned} \|\nabla F_{t+1}(\mathbf{p}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)} &\leq \|\nabla F_{t+1}(\mathbf{w}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)} + \|\nabla F_{t+1}(\mathbf{p}_t) - \nabla F_{t+1}(\mathbf{w}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)}, \\ &= \lambda(\mathbf{w}_t, F_{t+1}) + \|\nabla F_{t+1}(\mathbf{p}_t) - \nabla F_{t+1}(\mathbf{w}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)}, \\ &\leq 2\sqrt{e\eta}/c_T + 2^7\sqrt{\alpha_T}(e\eta)^{3/2}/c_T^2, \end{aligned}$$

where the last inequality follows by (43) and Lemma 8. Now, by (42) and Lemma 4, we have

$$(1 - \sqrt{e\eta}r_t)^2 \nabla^{-2}F_{t+1}(\mathbf{p}_t) \preceq \nabla^{-2}F_{t+1}(\mathbf{w}_t) \preceq (1 - \sqrt{e\eta}r_t)^{-2} \nabla^{-2}F_{t+1}(\mathbf{p}_t), \quad (44)$$

where  $r_t := \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_t(\mathbf{w}_t)}$ , and so

$$\begin{aligned} \|\nabla F_{t+1}(\mathbf{p}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{p}_t)} &\leq (1 - \sqrt{e\eta}\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2\Phi_t(\mathbf{w}_t)})^{-1} \|\nabla F_{t+1}(\mathbf{p}_t)\|_{\nabla^{-2}F_{t+1}(\mathbf{w}_t)}, \\ &\leq 4\sqrt{e\eta}/c_T + 2^8\sqrt{\alpha_T}(e\eta)^{3/2}/c_T^2 \leq \frac{1}{2\sqrt{e\eta}}, \end{aligned} \quad (45)$$

where the last inequality follows by the range assumption on  $\eta$  and  $\beta$ . Combining (45) with the fact that  $\mathbf{p}_{t+1}$  is the minimizer of  $F_{t+1}$  (which is  $\sqrt{e\eta}$ -self-concordant as we argued above) and Lem. 3, we get

$$\tilde{r}_t := \|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2F_{t+1}(\mathbf{p}_t)} \leq 2\lambda(\mathbf{p}_t, F_{t+1}) \leq \frac{8\sqrt{e\eta}}{c_T} + \frac{2^9\sqrt{\alpha_T}(e\eta)^{3/2}}{c_T^2} < \frac{1}{2\sqrt{e\eta}}, \quad (46)$$

where the last inequality follows by the range assumption on  $\eta$  and  $\beta$ . Thus, using Lemma 4 again, we get that

$$(1 - \sqrt{e\eta}\tilde{r}_t)^2 \nabla^2F_{t+1}(\mathbf{p}_t) \preceq \nabla^2F_{t+1}(\mathbf{p}_{t+1}) \preceq (1 - \sqrt{e\eta}\tilde{r}_t)^{-2} \nabla^2F_{t+1}(\mathbf{p}_t). \quad (47)$$

Using this and (46), we have

$$\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})} \leq \frac{\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_t)}}{1 - \sqrt{e\eta}\tilde{r}_t} \leq \frac{16\sqrt{e\eta}}{c_T} + \frac{2^{14}\sqrt{\alpha_T}(e\eta)^{3/2}}{c_T^2}. \quad (48)$$

The RHS of (48) is at most 1/2 due to the range assumptions on  $\eta$  and  $\beta$ . This, combined with (Nesterov et al., 2018, Theorem 5.1.8) and the optimality of  $\mathbf{p}_{t+1}$  implies that for  $\omega(x) := x - \ln(1+x)$ , we have

$$\begin{aligned} F_{t+1}(\mathbf{p}_t) - F_{t+1}(\mathbf{p}_{t+1}) &\geq \nabla F_{t+1}(\mathbf{p}_{t+1})^\top (\mathbf{p}_t - \mathbf{p}_{t+1}) + \omega(\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})}), \\ &\geq \frac{1}{3}\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})}^2, \end{aligned}$$

where the last inequality follows by the fact that  $\omega(x) \geq x^2/3$ , for  $x \in [0, 1/2]$ . On the other hand, by the definition of  $F_{t+1}$ , non-negativity of Bregman divergence, and Hölder inequality,

$$\begin{aligned} F_{t+1}(\mathbf{p}_t) - F_{t+1}(\mathbf{p}_{t+1}) &= \langle \mathbf{p}_t - \mathbf{p}_{t+1}, \tilde{\mathbf{g}}_t \rangle - D_{\Phi_t}(\mathbf{p}_{t+1}, \mathbf{p}_t), \\ &\leq \|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})} \cdot \|\tilde{\mathbf{g}}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{p}_{t+1})}. \end{aligned}$$

Combining the above two inequalities we get

$$\|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})} \leq 3\|\tilde{\mathbf{g}}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{p}_{t+1})}.$$

Using this and Hölder's inequality leads to

$$\begin{aligned} \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle &\leq \|\tilde{\mathbf{g}}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{p}_{t+1})} \cdot \|\mathbf{p}_t - \mathbf{p}_{t+1}\|_{\nabla^2 F_{t+1}(\mathbf{p}_{t+1})} \leq 3\|\tilde{\mathbf{g}}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{p}_{t+1})}^2, \\ &\leq 4\|\tilde{\mathbf{g}}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{w}_t)}^2, \quad (49) \end{aligned}$$

where the last inequality follows by (47), (44), and the range assumptions on  $\eta$  and  $\beta$ . Let  $\psi_t := \Psi_{t-1} - \Psi_t$  and recall the definition of  $(\vartheta_i)$  in (13). Let  $\psi_t := \Psi_{t-1} - \Psi_t$ . By the definition of  $\tilde{\mathbf{g}}_t$  in (15), the fact that  $(a+b)^2 \leq (1+\gamma)a^2 + (1+1/\gamma)b^2$ , for all  $\gamma > 0$ , and (49), we have

$$\begin{aligned} \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle &\leq 9(1 + \beta \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle / 4)^2 \cdot \|\mathbf{g}_t\|_{\nabla^{-2} F_{t+1}(\mathbf{w}_t)}^2 / 8 \\ &\quad + 9 \left\| \sum_{i=1}^d \left( \frac{1}{\eta_{t-1,i}} - \frac{1}{\eta_{t,i}} \right) (\nabla \vartheta_i(\mathbf{p}_t) - \nabla \vartheta_i(\mathbf{w}_t)) \right\|_{\nabla^{-2} F_{t+1}(\mathbf{w}_t)}^2, \\ &\leq 9(1 + \beta \|\mathbf{g}_t\|_{\nabla^{-2} \Phi_{t-1}(\mathbf{w}_t)} \cdot \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)} / 4)^2 \cdot \|\mathbf{g}_t\|_{\nabla^{-2} \Phi_t(\mathbf{w}_t)}^2 / 8 \\ &\quad + 9e \|\nabla \psi_t(\mathbf{p}_t) - \nabla \psi_t(\mathbf{w}_t)\|_{\nabla^{-2} \psi_t(\mathbf{w}_t)}^2, \quad (50) \end{aligned}$$

where in the last inequality we used Hölder's inequality and the fact that  $e\nabla^2 F_{t+1} \succeq \nabla^2 \Psi_{t-1} \succeq \nabla^2 \psi_t$  (since  $\eta_{t,i} \in [\eta, \eta e]$ , for all  $i \in [d]$  and  $t \in [T]$ ). On the other hand, since  $\nabla^2 \psi_t \preceq \nabla^2 \Psi_{t-1} \preceq \nabla^2 \Phi_{t-1}$ , we have  $\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \psi_t(\mathbf{w}_t)}^2 \leq \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2$ . Using this, (42), and Lemmas 11 and 5, we get

$$\begin{aligned} \|\nabla \psi_t(\mathbf{p}_t) - \nabla \psi_t(\mathbf{w}_t)\|_{\nabla^{-2} \psi_t(\mathbf{w}_t)}^2 &\leq \mathbb{I}\{\rho_{t,i} \neq \rho_{t-1,i}\} \frac{\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \psi_t(\mathbf{w}_t)}^2}{(1 - \sqrt{e\eta} \log_2 T \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \psi_t(\mathbf{w}_t)})^2}, \\ &\leq \mathbb{I}\{\rho_{t,i} \neq \rho_{t-1,i}\} \frac{\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)}^2}{(1 - \sqrt{e\eta} \log_2 T \|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 \Phi_{t-1}(\mathbf{w}_t)})^2}, \\ &\leq 2^{14} \mathbb{I}\{\rho_{t,i} \neq \rho_{t-1,i}\} (e\eta)^3 / c_T^4, \quad (51) \end{aligned}$$

where the last two inequalities follow by (42) and the range assumptions on  $\eta$  and  $\beta$ . Since  $\sum_{t=1}^T \mathbb{I}\{\rho_{t,i} \neq \rho_{t-1,i}\} \leq O(d \ln T)$  (by definition of  $(\rho_{t,i})$  in (11) and the fact that  $(\bar{u}_{t,i}) \subset [1/(dT), 1]$ ), (51) implies that  $\sum_{t=1}^T \|\nabla \psi_t(\mathbf{p}_t) - \nabla \psi_t(\mathbf{w}_t)\|_{\nabla^{-2}\psi_t(\mathbf{w}_t)}^2 \leq O(\eta^3 d \ln T) \leq O(d \ln T)$ . Using this together with (42), (50), and Lemma 20 (and the range assumptions on  $\eta$  and  $\beta$ ), we get

$$\sum_{t=1}^T \langle \tilde{\mathbf{g}}_t, \mathbf{p}_t - \mathbf{p}_{t+1} \rangle \leq \frac{18d \ln T}{\beta} + O(d \ln T).$$

This completes the proof.  $\blacksquare$

**Proof of Lemma 14.** Let  $c_T := 1 - 1/T$ . By Cauchy Schwarz inequality and Lemmas 8 and 19,

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle &\leq \sum_{t=1}^T |\langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle| \leq \sqrt{\sum_{t=1}^T \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_{t-1}(\mathbf{w}_t)}^2} \sqrt{\sum_{t=1}^T \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2\Phi_{t-1}(\mathbf{w}_t)}^2}, \\ &\leq \sqrt{4^2 d \beta^{-1} \ln T} \cdot \sqrt{1 + 15d \beta^{-1} \ln T}, \\ &\leq 1 + 4\sqrt{15} d \beta^{-1} \ln T. \end{aligned} \quad (52)$$

On the other hand, by Hölder's inequality, we have

$$|\langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle| \leq \|\mathbf{g}_t\|_{\nabla^{-2}\Phi_{t-1}(\mathbf{w}_t)} \cdot \|\mathbf{p}_t - \mathbf{w}_t\|_{\nabla^2\Phi_{t-1}(\mathbf{w}_t)} \leq 64e^2 \eta^2 / c_T^3,$$

where the last equality follows by Lemmas 8 and 20. Combining this with the range assumptions on  $\eta$  and  $\beta$ , we get

$$\sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle^2 \leq \frac{64e^2 \eta^2}{c_T^3} \sum_{t=1}^T |\langle \mathbf{g}_t, \mathbf{p}_t - \mathbf{w}_t \rangle| \leq \frac{64e^2 \eta^2}{c_T^3} \left(1 + 4\sqrt{15} d \beta^{-1} \ln T\right),$$

where the last inequality follows by (52).  $\blacksquare$

**Proof of Lemma 13.** Let  $c_T := 1 - 1/T$ . We have

$$\sum_{t=1}^T (D_{\Psi_t}(\mathbf{u}', \mathbf{w}_t) - D_{\Psi_{t-1}}(\mathbf{u}', \mathbf{w}_t)) = \sum_{t=2}^T \sum_{i=1}^d \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) h \left( \frac{\bar{u}'_i}{\bar{w}_{t,i}} \right),$$

where  $h(x) := x - 1 - \ln x$  and  $(\eta_{t,i})$  are defined in (10). Fix  $i \in [d]$  and let  $\mathcal{T} := \{t : \rho_{t,i} \neq \rho_{t-1,i}\}$ . First, suppose that  $\bar{u}_{t,i} \geq \frac{1}{2d} \wedge (c_T \bar{u}'_i)$ , for all  $t \in [T]$ . In this case, we have  $\max_{t \in [T]} \frac{\bar{u}'_i}{\bar{u}_{t,i}} \leq 2d \bar{u}'_i + c_T^{-1}$ . Therefore, by positivity of  $h$  and the fact that  $\eta_{t+1,i} \geq \eta_{t,i}$ , we have

$$\sum_{t=2}^T \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) h \left( \frac{\bar{u}'_i}{\bar{w}_{t,i}} \right) \leq 0 \leq c_T^{-1} + 2d \bar{u}'_i - \max_{t \in [T]} \frac{\bar{u}'_i}{\bar{u}_{t,i}}. \quad (53)$$

Now, assume that  $\bar{u}_{t,i} \leq \frac{1}{2d} \wedge (c_T \bar{u}'_i)$ . Note that this implies that  $\mathcal{T} \neq \emptyset$  (due to  $\bar{u}_{t,i} \leq 1/(2d)$ ) and the definition of  $(\rho_{t,i})$ . Let  $\tau := \max \mathcal{T}$ . For notational convenience, we let

$$\mathbf{u} := \bar{\mathbf{u}}'_i, \quad u_\tau := \bar{u}_{\tau,i}, \quad \mathbf{w}_\tau := \bar{\mathbf{w}}_{\tau,i}, \quad \text{and} \quad \eta_\tau := \eta_{\tau,i}.$$

By definition of  $(\eta_{t,i})$ , we have  $\eta_\tau = \eta \exp(-\log_T du_\tau) \leq \eta \exp(\log_T T) = \eta e$ . Thus, by the positivity of  $h$ , we have

$$\begin{aligned} \sum_{t=2}^T \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) h \left( \frac{\bar{u}'_i}{\bar{w}_{t,i}} \right) &\leq \left( \frac{1}{\eta_\tau} - \frac{1}{\eta_{\tau-1}} \right) h \left( \frac{u}{w_\tau} \right), \\ &= \frac{1 - e^{\log_T(u_{\tau-1}/u_\tau)}}{\eta_\tau} \cdot h \left( \frac{u}{w_\tau} \right), \\ &\leq \frac{-\log_T \frac{u_{\tau-1}}{u_\tau}}{\eta_\tau} \cdot h \left( \frac{u}{w_\tau} \right), \\ &\leq -\frac{\log_2 \frac{u_{\tau-1}}{u_\tau}}{e\eta \ln_2 T} \cdot h \left( \frac{u}{w_\tau} \right), \end{aligned}$$

By definition of  $(u_{t,i})$ , we have

$$\frac{u_\tau}{w_\tau} = \frac{1/(Td) + (1 - 1/T)w_\tau}{w_\tau} \geq 1 - \frac{1}{T} = c_T. \quad (54)$$

Hence,  $h(u/w_\tau) \geq h(c_T u/u_\tau)$ , since  $c_T u/u_\tau \geq 1$  by assumption and  $h$  is positive and decreasing on  $[1, +\infty[$ . On the other hand, by construction of  $(\rho_{t,i})$  and the definition of  $\tau$ , we have  $u_\tau \leq u_{\tau-1}/2$ , which implies that  $-\log_2(u_{\tau-1}/u_\tau) \leq -1$ . Plugging these bounds into (54), we get that

$$\begin{aligned} \sum_{t=2}^T \left( \frac{1}{\eta_{t,i}} - \frac{1}{\eta_{t-1,i}} \right) h \left( \frac{\bar{u}'_i}{\bar{w}_{t,i}} \right) &\leq -\frac{1}{e\eta \log_2 T} h \left( \frac{c_T u}{u_\tau} \right), \\ &= -\frac{1}{e\eta \log_2 T} \left( \frac{c_T u}{u_\tau} - 1 - \ln \left( \frac{c_T u}{u_\tau} \right) \right), \\ &= -\frac{c_T}{e\eta \log_2 T} \max_{t \in [T]} \frac{\bar{u}'_i}{\bar{u}_{t,i}} + \mathcal{O} \left( \frac{\ln(dT \bar{u}'_i)}{\eta \ln T} \right), \end{aligned} \quad (55)$$

Thus, by (53), (55), and the fact that  $\bar{u}''_i \leq 2\bar{u}'_i$ , for all  $i \in [d]$ , we get the desired result after summing over  $i \in [d]$ .  $\blacksquare$

## Appendix D. Proof of Theorem 17 (Meta-Algorithm Regret)

### D.1. Proof of Theorem 17

To prove Theorem 17, we define

$$\mathbf{u}_* \in \operatorname{argmin}_{\mathbf{u} \in \mathcal{C}_{d-1}} \sum_{t=1}^T \ell_t(\mathbf{u}), \quad \text{and} \quad \alpha_t^{\beta, I} := 8^{-1} \wedge \min_{s \in I \cap [t], i \in [d]} (\bar{u}_{t,i}^{\beta, I} / (8\bar{u}_{*,i}'')). \quad (56)$$

With this, we start by stating a regret guarantee against  $\mathbf{u}_*$  for the subroutines of Algorithm 3, which follows from Lemma 8 and the base algorithm's regret bound in Theorem 15.

**Lemma 23** Let  $I \in \mathcal{I}$ ,  $\beta \in \mathcal{G}$ , and  $(\mathbf{u}_t^{\beta, I})_{t \in I}$  be the outputs of the subroutine  $\mathbf{A}^{\beta, I}$  within Algorithm 3 in response to Cover's losses  $(\ell_t)$ . Further, let  $\mathbf{u}_*$  be as in (56). If there exists  $t \in I \cap [T-1]$  such that  $\alpha_t^{\beta, I} \geq \beta > \alpha_{t+1}^{\beta, I}$ , and  $t > \min I$  then,

$$\sum_{s \in I \cap [t]} (\ell_s(\mathbf{u}_s^{\beta, I}) - \ell_s(\mathbf{u}_*)) \leq O(d^2 \ln^4 T) - \frac{204d \ln^2 T}{\beta}. \quad (57)$$

Otherwise, if  $\alpha_t^{\beta, I} \geq \beta$  for all  $t \in I$ , then  $\sum_{s \in I \cap [T]} (\ell_s(\mathbf{u}_s^{\beta, I}) - \ell_s(\mathbf{u}_*)) \leq O(d^2 \ln^4 T) + \frac{34d \ln T}{\beta}$ .

For the proofs of Lemma 23 and Theorem 15, we need the following Corollary to Lemma 8:

**Corollary 24** In the setting of Lemma 8, the iterate of  $(\mathbf{w}_t)$  in Alg. 1 and  $(\mathbf{p}_t)$  in (14) satisfy:

$$\forall i \in [d], \forall t \in [T], \quad 1 - 64(\epsilon\eta)^2 \leq \frac{\bar{p}_{t,i}}{\bar{w}_{t,i}} \leq 1 + 64(\epsilon\eta)^2, \quad \frac{3}{4} \leq \frac{\bar{w}_{t+1,i}}{\bar{w}_{t,i}} \leq \frac{5}{4}, \quad \text{and} \quad \frac{\bar{u}_{t,i}}{\bar{u}_{t+1,i}} \leq \frac{4}{3}.$$

The proofs of Lemma 23 and Corollary 24 are in Appendix D.2.

**Proof of Theorem 17.** Recall that  $\mathcal{G} := \{\frac{1}{d2^{i+3}} : i \in [\lceil \log_2 T \rceil]\}$ . We note that the outputs of  $(\mathbf{u}_t^{\beta, I})$  of the base algorithms are all in the set

$$\bar{\mathcal{C}}_{d-1} := \{\mathbf{u} \in \mathcal{C}_{d-1} : \bar{u}_i \geq 1/(dT), \forall i \in [d]\}.$$

This is because of the mixing with the uniform distribution on Line 3 of Algorithm 1. By (18), the outputs  $(\mathbf{u}_t)$  of Algorithm 3 are convex combinations of  $(\mathbf{u}_t^{\beta, I})$ , and so  $(\mathbf{u}_t) \subset \bar{\mathcal{C}}_{d-1}$ . This fact will be useful below.

For any  $\beta \in \mathcal{G}$ , we define the set

$$\mathcal{J}^\beta := \{t \in [T] : \alpha_t^{\beta, I_0} < \beta, I_0 := [T]\}.$$

Note that  $I_0 = [T] \in \mathcal{I}|_T$ , by definition, where  $\mathcal{I}|_T$  is the set of intervals indexing the base algorithms of Alg. 3. Furthermore, for  $\beta' = \min \mathcal{G}$ , we have  $\mathcal{J}^{\beta'} = \emptyset$  since  $\beta' \leq \frac{1}{8dT} \stackrel{(*)}{\leq} \alpha_t^{\beta', I}$  for any  $I \in \mathcal{I}|_T$  and  $t \geq 1$ , where  $(*)$  follows by the fact that all the outputs  $(\mathbf{u}_t^{\beta', I})$  are in the set  $\bar{\mathcal{C}}_{d-1}$ . If  $\mathcal{J}_\beta = \emptyset$ , for all  $\beta \in \mathcal{G}$ , then by Proposition 16 and the second inequality in Lemma 23, instantiated with  $\beta_0 := \max \mathcal{G} = \frac{1}{2^4 d}$  and  $I_0 = [T]$ , we get

$$\sum_{t=1}^T (\ell_t(\mathbf{u}) - \ell_t(\mathbf{u}_*)) \leq \sum_{t=1}^T (\ell_t(\mathbf{u}_t^{\beta_0, I_0}) - \ell_t(\mathbf{u}_*)) + O(\ln T) \leq O(\beta_0^{-1} d \ln T) = O(d^2 \ln T),$$

which implies the desired regret bound. Now, suppose that there exists  $\beta \in \mathcal{G}$  such that  $\mathcal{J}^\beta \neq \emptyset$  and let  $\beta_* := \min\{\beta \in \mathcal{G} : \mathcal{J}^\beta \neq \emptyset\}$  and  $\tau := \min \mathcal{J}^{\beta_*}$ . By the definition of the base Algorithm 1, we have  $\mathbf{u}_1^{\beta_*, I_0} = \mathbf{1}/d$ , and thus,

$$2\beta_0 = \frac{1}{2^3 d} \leq \alpha_1^{\beta_*, I_0} \quad \text{and} \quad \beta_* \leq \beta_0 \leq \alpha_2^{\beta_*, I_0},$$

where the last inequality follows by the fact that  $2\beta_0 \leq \alpha_1^{\beta_*, I_0}$  and the bound on  $\bar{u}_{1,i}/\bar{u}_{2,i}$  for  $i \in [d]$  from Corollary 24. Therefore, by definition of  $\tau$  we must have that  $\tau > 2$  and  $\alpha_{\tau-1}^{\beta_*, I_0} \geq \beta_* > \alpha_\tau^{\beta_*, I_0}$ . Thus, by invoking Proposition 16 and Lemma 23 (in particular (57)), we get

$$\begin{aligned} \sum_{t=1}^{\tau-1} (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u}_*)) &\leq \sum_{t=1}^{\tau-1} (\ell_t(\mathbf{u}_t^{\beta_*, I_0}) - \ell_t(\mathbf{u}_*)) + O(\ln T), \\ &\leq O(d^2 \ln^4 T) - \frac{204d \ln^2 T}{\beta_*}. \end{aligned} \quad (58)$$

Now, let  $\beta' := \beta_*/2$  and  $\mathcal{S} = \mathcal{S}([\tau + 1, T]) \subset \mathcal{I}$  denote any disjoint partition of  $[\tau + 1, T]$  that contains at most  $2 \log_2 T$  elements in  $\mathcal{I}$ ; this is guaranteed to exist by (Daniely et al., 2015, Lemma 1.2). The fact that  $\mathcal{J}^{\beta'} = \emptyset$  together with Proposition 16 and the second inequality in Lemma 23 implies

$$\begin{aligned} \sum_{t=\tau+1}^T (\ell_t(\mathbf{u}_t) - \ell_t(\mathbf{u}_*)) &\leq \sum_{I \in \mathcal{S}} \sum_{t \in I} (\ell_t(\mathbf{u}_t^{\beta', I}) - \ell_t(\mathbf{u}_*)) + O(|\mathcal{S}| \ln T), \\ &\leq O(|\mathcal{S}| d^2 \ln^4 T) + \frac{34|\mathcal{S}| d \ln T}{\beta'}, \\ &\leq O(d^2 \ln^5 T) + \frac{204d \ln^2 T}{\beta_*}, \end{aligned}$$

where the last inequality, we used the fact that  $\beta' = \beta_*/2$  and  $|\mathcal{S}| \leq 2 \ln_2 T \leq 3 \ln T$ . Combining this with (58) and using the fact that  $\ell_\tau(\mathbf{u}_\tau) - \ell_\tau(\mathbf{u}_*) \leq \ln(dT)$  (since  $\mathbf{u}_\tau \in \bar{\mathcal{C}}_{d-1}$ ) completes the proof.  $\blacksquare$

## D.2. Proofs of Lemma 23 and Corollary 24

**Proof of Corollary 24.** We start by the first inequality. By Lemma 8, we have

$$\|\mathbf{w}_t - \mathbf{p}_t\|_{\nabla^2 F_t(\mathbf{w}_t)} \leq 64(e\eta)^{3/2}. \quad (59)$$

Note also that for any  $i \in [d-1]$ , we have  $\nabla^2 F_t(\mathbf{w}_t) \succeq \nabla^2 \Psi_t(\mathbf{w}_t) \succeq \mathbf{e}_i \mathbf{e}_i^\top / (\eta e w_{t,i}^2)$ , since  $\eta_{t,i} \in [\eta, \eta e]$ . Therefore, (59) implies that  $(w_{t,i} - p_{t,i})^2 / (\eta e w_{t,i}^2) \leq 64^2 (e\eta)^3$ . Rearranging this implies the bounds on  $\bar{p}_{t,i}/\bar{w}_{t,i}$  for the case where  $i \in [d-1]$ . Also, we have that  $\nabla^2 F_t(\mathbf{w}_t) \succeq \nabla^2 \Psi_t(\mathbf{w}_t) \succeq \mathbf{1}\mathbf{1}^\top / (\eta e \bar{w}_{t,d}^2)$ . Therefore, (59) implies that

$$64^2 (e\eta)^3 \geq (\langle \mathbf{w}_t, \mathbf{1} \rangle - \langle \mathbf{p}_t, \mathbf{1} \rangle)^2 / (\eta e \bar{w}_{t,d}^2) = (\bar{w}_{t,d} - \bar{p}_{t,d})^2 / (\eta e \bar{w}_{t,d}^2).$$

Rearranging this inequality yields the desired bounds on  $\bar{p}_{t,i}/\bar{w}_{t,i}$  for  $i = d$ . We now bound  $\bar{w}_{t+1,i}/\bar{w}_{t,i}$  for  $i \in [d]$ . By definition of  $\mathbf{w}_{t+1}$ , we have

$$\begin{aligned} \|\mathbf{w}_{t+1} - \mathbf{w}_t\|_{\nabla^2 \Psi_t(\mathbf{w}_t)} &\leq \|\mathbf{w}_{t+1} - \mathbf{w}_t\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)}, \\ &= \left\| \frac{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1} \nabla_t}{1 + 4\sqrt{e\eta} \|\nabla_t\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1}}} \right\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)} \\ &= \frac{\|\nabla_t\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1}}}{1 + 4\sqrt{e\eta} \|\nabla_t\|_{(\nabla^2 \Psi_t(\mathbf{w}_t) + V_t)^{-1}}} \leq \frac{1}{4\sqrt{e\eta}}. \end{aligned}$$

Thus, since  $\nabla^2 \Psi_t(\mathbf{w}_t) \succeq \mathbf{e}_i \mathbf{e}_i^\top / (\eta e \bar{w}_{t,i}^2)$ , for all  $i \in [d]$ , we get that (as we did above)

$$\frac{(\bar{w}_{t+1,i} - \bar{w}_{t,i})^2}{\eta e \bar{w}_{t,i}^2} \leq \frac{1}{16\eta e}, \quad \forall i \in [d].$$

Thus, after rearranging, we obtain the desired bounds on  $\bar{w}_{t+1,i}/\bar{w}_{t,i}$ . Finally, we have

$$\frac{\bar{u}_{t,i}}{\bar{u}_{t+1,i}} = \frac{(1 - 1/T)\bar{w}_{t,i} + 1/(dT)}{(1 - 1/T)\bar{w}_{t+1,i} + 1/(dT)} \leq 1 \vee \frac{\bar{w}_{t,i}}{\bar{w}_{t+1,i}} \leq \frac{4}{3},$$

where the last inequality follows by the fact that  $3/4 \leq \bar{w}_{t+1,i}/\bar{w}_{t,i} \leq 5/4$ .  $\blacksquare$

**Proof of Lemma 23.** Let  $\mathbf{u}_*$  be as in (56). We first show (57). Since  $\beta \leq \alpha_t^{\beta,I}$ , Lemma 21 implies that

$$\beta \leq 8^{-1} \wedge |8\langle \mathbf{g}_s^{\beta,I}, \mathbf{u}_*'' - \mathbf{u}_s^{\beta,I} \rangle|^{-1},$$

for all  $s \in I \cap [t]$ , where  $\mathbf{g}_s^{\beta,I} := \nabla \ell_s(\mathbf{u}_s^{\beta,I})$ . Therefore, by the assumption that  $t > \min I$  and Theorem 15, we have

$$\sum_{s \in I \cap [t]} (\ell_s(\mathbf{u}_s^{\beta,I}) - \ell_s(\mathbf{u}_*)) \leq O\left(\frac{d \ln T}{\eta}\right) + \frac{34d \ln T}{\beta} - \frac{a_t}{32\eta \ln T}, \quad (60)$$

where  $a_t := \max_{s \in I \cap [t]} \sum_{i=1}^d \bar{u}_{*,i}'' / \bar{u}_{s,i}$ . Now the fact that  $\beta > \alpha_{t+1}^{\beta,I}$  implies

$$\frac{1}{\beta} \leq 8 \vee \max_{s \in I \cap [t+1], i \in [d]} \frac{8\bar{u}_{*,i}''}{\bar{u}_{s,i}} \leq 8a_{t+1} + 8 \leq 32a_t/3 + 8,$$

where the last step follows by the bound on  $\bar{u}_{t,i}/\bar{u}_{t+1,i}$  from Corollary 24. Further, combining this with (60), we get that

$$\begin{aligned} \sum_{s \in I \cap [t]} (\ell_s(\mathbf{u}_s^{\beta,I}) - \ell_s(\mathbf{u}_*)) &\leq \mathcal{O}\left(\frac{d \ln T}{\eta}\right) + \frac{34d \ln T}{\beta} - \frac{1}{32\eta \ln T} \left(\frac{3}{32\beta} - \frac{3}{4}\right), \\ &\leq \mathcal{O}(d^2 \ln^4 T) - \frac{204d \ln^2 T}{\beta}. \end{aligned}$$

where we used that  $\eta = 1/(286^2 d \ln^3 T)$ . When,  $\beta \leq \alpha_t^{\beta,I}$ , for all  $t \in I$ , we can simply invoke (60) and discard the negative term to obtain the second claimed bound of the lemma.  $\blacksquare$

## Appendix E. Adaptive Meta-Algorithm for Mixable Losses (AdaMix)

In this section, we present a version of PAE (Zhang et al., 2019) that yields an adaptive, logarithmic regret in the expert setting with mixable losses (see Definition 6). That is, we present an algorithm that enjoys a logarithmic regret on any interval  $I \subseteq [T]$  when the losses are mixable.

Recall the sets  $\mathcal{I}$  and  $\mathcal{I}|_T$  from §4.2 which are defined as

$$\mathcal{I} := \bigcup_{i,k \in \mathbb{N}} \{[2^k i, 2^k(i+1) - 1]\} \quad \text{and} \quad \mathcal{I}|_T := \{I \cap [T] : I \in \mathcal{I}\} \cup \{[T]\}.$$

---

**Algorithm 4** AdaMix: Adaptive Meta-Algorithm for Mixable Losses.

---

**Require:** **I**) Grid  $\mathcal{G}$  of  $\beta$  values, horizon  $T$ , and  $\eta > 0$ ; **II**) Instances  $(A^{\beta,I})_{\beta \in \mathcal{G}, I \in \mathcal{I}|_T}$ , where  $A^{\beta,I}$  is active during  $I$ ; **III**) a substituting function  $\Upsilon: \bigcup_{\mathcal{J} \subseteq \mathcal{I}|_T} \Delta(\mathcal{G} \times \mathcal{J}) \times \mathcal{U}^{|\mathcal{G}| \times |\mathcal{J}|} \rightarrow \mathcal{U}$  for the losses  $(f_t: \mathcal{U} \rightarrow \mathbb{R})$ . // The substitution function  $\Upsilon$  will be chosen to satisfy (62).  
 Set  $\mathcal{J}_0 = \emptyset$ .  
**for**  $t = 1, \dots, T$  **do**  
     Identify the set of newly active intervals  $\tilde{\mathcal{J}}_t = \{I \in \mathcal{I}|_T: \min I = t\}$ .  
     Set  $\mathcal{F}_{t-1}^{\beta,I} = 0$  for all  $\beta \in \mathcal{G}$  and  $I \in \tilde{\mathcal{J}}_t$ .  
     Update the set of active intervals  $\mathcal{J}_t = \mathcal{J}_{t-1} \cup \tilde{\mathcal{J}}_t$ .  
     Receive  $\mathbf{u}_t^{\beta,I} \in \mathcal{U}$  from each  $A^{\beta,I}$  such that  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .  
     Set  $q_{t-1}^{\beta,I} = e^{-\eta \mathcal{F}_{t-1}^{\beta,I}} / Z_t$  for  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ , where  $Z_t = \sum_{\beta \in \mathcal{G}, I \in \mathcal{J}_t} e^{-\eta \mathcal{F}_{t-1}^{\beta,I}}$ .  
     Play  $\mathbf{u}_t = \Upsilon(Q_{t-1}, U_t)$ , where  $Q_{t-1} := (q_{t-1}^{\beta,I})_{\beta \in \mathcal{G}, I \in \mathcal{J}_t}$  and  $U_t := (\mathbf{u}_t^{\beta,I})_{\beta \in \mathcal{G}, I \in \mathcal{J}_t}$ .  
     Observe  $f_t$  and set  $\mathcal{F}_t^{\beta,I} = \mathcal{F}_{t-1}^{\beta,I} + f_t(\mathbf{u}_t^{\beta,I}) - f_t(\mathbf{u}_t)$ , for all  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .  
     Send  $f_t$  to each  $A^{\beta,I}$  such that  $\beta \in \mathcal{G}$  and  $I \in \mathcal{J}_t$ .  
     Update the set of active intervals  $\mathcal{J}_t = \mathcal{J}_t \setminus \{I \in \mathcal{I}|_T: \max I = t\}$ .  
**end for**

---

Much like PAE, AdaMix (Alg. 4) takes in a set of base algorithms/experts  $(A^{\beta,I})$ , where for expert  $A^{\beta,I}$ ,  $\beta$  represents a parameter in some predefined grid  $\mathcal{G} \subset \mathbb{R}$  and  $I \in \mathcal{I}|_T$  represents the interval on which the expert is active; in this case, expert  $A^{\beta,I}$  is initialized at round  $t = \min I$  and terminates after round  $t = \max I$ . Alg. 2 assumes access to a *substitution function*  $\Upsilon: \bigcup_{\mathcal{J} \subseteq \mathcal{I}|_T} \Delta(\mathcal{G} \times \mathcal{J}) \times \mathcal{U}^{|\mathcal{G}| \times |\mathcal{J}|} \rightarrow \mathcal{U}$  for the losses  $f_t: \mathcal{U} \rightarrow \mathbb{R}$ , such that, for all  $\mathcal{J} \subseteq \mathcal{I}|_T$ ,  $Q \in \Delta(\mathcal{G} \times \mathcal{J})$ , and  $U := (\mathbf{u}^{\beta,I}) \in \mathcal{U}^{|\mathcal{G}| \times |\mathcal{J}|}$ ,

$$f_t(\Upsilon(Q, U)) \leq -\eta^{-1} \log \sum_{\beta \in \mathcal{G}, I \in \mathcal{J}} Q^{\beta,I} e^{-\eta f_t(\mathbf{u}^{\beta,I})}, \forall t \geq 1. \quad (62)$$

Such a substitution function is guaranteed to exist when the losses  $(f_t)$  are  $\eta$ -mixable (see Def. 6). For the case of Cover's loss (1), which is 1-mixable (in fact 1-exp-concave), the substitution function can be set to  $\Upsilon(Q, U) = \sum_{\beta, I} Q^{\beta,I} U^{\beta,I}$  (this recovers PAE), which satisfies (62) with  $\mathcal{U} = \mathcal{C}_{d-1}$  and  $(f_t) \equiv (\ell_t)$ . We now state the guarantee of AdaMix (Alg. 4) relative to the performance of the experts  $(A^{\beta,I})$ :

**Proposition 25** *Let  $\eta > 0$  and  $\mathcal{G}$  be a set s.t.  $|\mathcal{G}| \leq M$ . Further, let  $(\mathbf{u}_t^{\beta,I})_{t \in I}$ ,  $\beta \in \mathcal{G}$  and  $I \in \mathcal{I}$ , be the outputs of the subroutine  $A^{\beta,I}$  within Alg. 4 in response to a sequence of  $\eta$ -mixable losses  $(f_t)$ . Then, the outputs  $(\mathbf{u}_t)$  of Algorithm 4 with a substitution function  $\Upsilon$  satisfying (62), guarantee*

$$\sum_{s \in I \cap [t]} (f_s(\mathbf{u}_s) - f_s(\mathbf{u}_s^{\beta,I})) \leq (2 \ln t + \ln M) / \eta, \quad \text{for all } I \in \mathcal{I}, \beta \in \mathcal{G}, \text{ and } t \in I. \quad (63)$$

Now, an adaptive, logarithmic regret for mixable losses follows easily from this proposition. To see this, let  $J$  be any interval in  $[T]$ . Then, by (Daniely et al., 2015, Lemma 1.2) we know that there exist disjoint sets  $I_1, \dots, I_N \in \mathcal{I}$  such that  $\bigcup_{i \in [N]} I_i = J$  and  $|N| \leq O(\ln T)$ . Now, if any subroutines  $A^{\beta,I}$  achieves a logarithmic regret  $\mathcal{R}_I^\beta(\mathbf{u}) \leq O(\eta^{-1} \ln T)$  within the interval  $I$  against any comparator



$\mathbf{u}$ , (63) implies that  $\sum_{s \in J} (f_s(\mathbf{u}_s) - f_s(\mathbf{u})) \leq \frac{1}{\eta} (2|N| \ln T + |N| \ln M) + \sum_{i \in [N]} \mathcal{R}_{I_i}^\beta(\mathbf{u}) \leq O(\frac{\ln^2 T}{\eta})$ .

**Proof of Proposition 25.** The proof of the proposition follows from an extension of the proof of (Zhang et al., 2019, Lemma 6).

In this proof, we denote by  $\mathcal{A}_t$  the set of active experts at round  $t$ ; that is,  $\mathcal{A}_t := \{A^{\beta, I} : I \in \mathcal{J}_t, \beta \in \mathcal{G}\}$ . For notational convenience, we will write  $\sum_{A^{\beta, I} \in \mathcal{A}_t}$  to mean  $\sum_{\beta \in \mathcal{G}, I \in \mathcal{J}_t}$ . By the assumption on the substitution function in (62), we have

$$e^{-\eta f_t(\mathbf{u}_t)} = e^{-\eta f_t(\Upsilon(Q_{t-1}, U_t))} \geq \sum_{A^{\beta, I} \in \mathcal{A}_t} q_{t-1}^{\beta, I} e^{-\eta f_t(\mathbf{w}_t^{\beta, I})}, \quad (64)$$

where  $Q_{t-1} = (q_{t-1}^{\beta, I})_{\beta \in \mathcal{G}, I \in \mathcal{J}_t}$ ,  $U_t = (\mathbf{u}_t^{\beta, I})_{\beta \in \mathcal{G}, I \in \mathcal{J}_t}$ , and

$$q_{t-1}^{\alpha, J} = \frac{e^{-\eta \mathcal{F}_{t-1}^{\alpha, J}}}{\sum_{A^{\beta, I} \in \mathcal{A}_t} e^{-\eta \mathcal{F}_{t-1}^{\beta, I}}}.$$

for any  $J \in \mathcal{I}_T$  and  $\alpha \in \mathcal{G}$ . Rearranging (64) we get

$$\sum_{A^{\beta, I} \in \mathcal{A}_t} e^{-\eta \mathcal{F}_t^{\beta, I}} = \sum_{A^{\beta, I} \in \mathcal{A}_t} e^{-\eta \mathcal{F}_{t-1}^{\beta, I}} \cdot e^{\eta f_t(\mathbf{u}_t) - \eta f_t(\mathbf{w}_t^{\beta, I})} \leq \sum_{A^{\beta, I} \in \mathcal{A}_t} e^{-\eta \mathcal{F}_{t-1}^{\beta, I}}. \quad (65)$$

Summing (65) over  $t = 1, \dots, s$ , we have

$$\sum_{t=1}^s \sum_{A^{\beta, I} \in \mathcal{A}_t} \exp(-\eta \mathcal{F}_t^{\beta, I}) \leq \sum_{t=1}^s \sum_{A^{\beta, I} \in \mathcal{A}_t} \exp(-\eta \mathcal{F}_{t-1}^{\beta, I})$$

which can be rewritten as

$$\begin{aligned} & \sum_{A^{\beta, I} \in \mathcal{A}_s} \exp(-\eta \mathcal{F}_s^{\beta, I}) + \sum_{t=1}^{s-1} \left( \sum_{A^{\beta, I} \in \mathcal{A}_t \setminus \mathcal{A}_{t+1}} \exp(-\eta \mathcal{F}_t^{\beta, I}) + \sum_{A^{\beta, I} \in \mathcal{A}_t \cap \mathcal{A}_{t+1}} \exp(-\eta \mathcal{F}_t^{\beta, I}) \right) \\ & \leq \sum_{A^{\beta, I} \in \mathcal{A}_1} \exp(-\eta \mathcal{F}_0^{\beta, I}) + \sum_{t=2}^s \left( \sum_{A^{\beta, I} \in \mathcal{A}_t \setminus \mathcal{A}_{t-1}} \exp(-\eta \mathcal{F}_{t-1}^{\beta, I}) + \sum_{A^{\beta, I} \in \mathcal{A}_t \cap \mathcal{A}_{t-1}} \exp(-\eta \mathcal{F}_{t-1}^{\beta, I}) \right) \end{aligned}$$

implying

$$\begin{aligned} & \sum_{A^{\beta, I} \in \mathcal{A}_s} \exp(-\eta \mathcal{F}_s^{\beta, I}) + \sum_{t=1}^{s-1} \sum_{A^{\beta, I} \in \mathcal{A}_t \setminus \mathcal{A}_{t+1}} \exp(-\eta \mathcal{F}_t^{\beta, I}) \\ & \leq \sum_{A^{\beta, I} \in \mathcal{A}_1} \exp(-\eta \mathcal{F}_0^{\beta, I}) + \sum_{t=2}^s \sum_{A^{\beta, I} \in \mathcal{A}_t \setminus \mathcal{A}_{t-1}} \exp(-\eta \mathcal{F}_{t-1}^{\beta, I}) \\ & = \sum_{A^{\beta, I} \in \mathcal{A}_1} \exp(0) + \sum_{t=2}^s \sum_{A^{\beta, I} \in \mathcal{A}_t \setminus \mathcal{A}_{t-1}} \exp(0) \\ & = |\mathcal{A}_1| + \sum_{t=2}^s |\mathcal{A}_t \setminus \mathcal{A}_{t-1}|. \end{aligned} \quad (66)$$

Note that  $|\mathcal{A}_1| + \sum_{t=2}^s |\mathcal{A}_t \setminus \mathcal{A}_{t-1}|$  is the total number of experts created until round  $s$ . From the structure of geometric covering intervals and the fact that  $|\mathcal{G}| \leq M$ , we have

$$|\mathcal{A}_1| + \sum_{t=2}^s |\mathcal{A}_t \setminus \mathcal{A}_{t-1}| \leq Ms (\lceil \log_2 s \rceil + 1) \leq Ms^2. \quad (67)$$

From (66) and (67), we have

$$\sum_{A^{\beta,I} \in \mathcal{A}_s} \exp(-\eta \mathcal{F}_s^{\beta,I}) + \sum_{t=1}^{s-1} \sum_{A^{\beta,I} \in \mathcal{A}_t \setminus \mathcal{A}_{t+1}} \exp(-\eta \mathcal{F}_t^{\beta,I}) \leq Ms^2.$$

Thus, for any interval  $I \in \mathcal{I}$  and  $s \in I$ , we have

$$\exp(-\eta \mathcal{F}_s^{\beta,I}) = \exp\left(\eta \sum_{t \in I \cap [s]} (f_t(\mathbf{u}_t) - f_t(\mathbf{u}_t^{\beta,I}))\right) \leq Ms^2,$$

which completes the proof.  $\blacksquare$

## Appendix F. Why the Ada-BARRONS Restarts Work

Before discussing the Ada-BARRONS restarts, we first give some more details on the regret bound of the base algorithm. In particular, we highlight that the RHS of (4) (which is a bound on the sum of divergences in the regret bound of BARRONS) can cancel the bound  $O(\beta^{-1}d \ln T)$  on the stability term as long as  $\beta \geq \alpha_T(\mathbf{u})/2$  and  $\eta \leq O(1/(d \ln T))$ . This is because  $\alpha_T(\mathbf{u}) \geq \frac{1}{8} \wedge \min_{i \in [d], t \in [T]} \frac{p_{t,i}}{8u_{t,i}}$ . In fact, by choosing  $\eta$  small enough and as long as  $\alpha_T(\mathbf{u})/2 \leq \beta \leq \alpha_T(\mathbf{u})$  one can ensure that what remains in the regret bound of BARRONS is

$$R_T(\mathbf{u}) \leq \tilde{O}(d^2) - \frac{Cd \ln^2 T}{\beta},$$

for some  $C > 0$ . Before discussing how one might ensure that  $\alpha_T(\mathbf{u})/2 \leq \beta \leq \alpha_T(\mathbf{u})$ , we reiterate that what enables the cancellation of the term  $O(\beta^{-1}d \ln T)$  is the sum  $\sum_{t=1}^T (D_{\Phi_t}(\mathbf{u}, \mathbf{p}_t) - D_{\Phi_t}(\mathbf{u}, \mathbf{p}_{t+1}))$  in (3) that comes out of the mirror descent analysis. This idea of canceling terms thanks to negative Bregman divergence terms originated from the work by Agarwal et al. (2017) on combining Bandit algorithms. Foster et al. (2020) recently showed that one can use FTRL with slightly modified gradients instead of mirror descent to achieve the same goal in a Bandit setting. The approach of the latter fails to lead to an efficient algorithm in our setting since it is not possible to show that the Newton decrement of the damped Newton step is small on rounds where the iterates  $(u_{t,i})$  reach new minima.

As we already discussed in §3.1, since the sequence of returns  $(r_t)$  is not known up-front, it is not possible for any algorithm to pick  $\beta$  so that the condition  $\alpha_T(\mathbf{u})/2 \leq \beta \leq \alpha_T(\mathbf{u})$  is always satisfied. Aggregating multiple instances of BARRONS with different  $\beta$ 's also fails since  $\alpha_T(\mathbf{u})$  depends on the outputs of the algorithm; and so changing  $\beta$  changes the target  $\alpha_T(\mathbf{u})$  for the base algorithm (see also discussion in (Luo et al., 2018)). Instead of aggregating base algorithms, the

approach taken by [Luo et al. \(2018\)](#) consists of restarting the base algorithm on round  $t$  if the current estimate for  $\beta$  satisfies  $\beta > \alpha_t(\mathbf{u}_t)$ , where  $\mathbf{u}_t$  is the regularized leader:

$$\mathbf{u}_t \in \operatorname{argmin}_{\mathbf{u} \in \bar{\Delta}_d} \sum_{i=1}^d \frac{-\ln u_i}{\eta_{t,i}} + \sum_{s=\tau}^t \ell_s(\mathbf{u}).$$

and  $\tau$  is the round where the current instance of the base algorithm was initialized. By stability of the iterates  $(\mathbf{u}_t)$  and  $(\mathbf{p}_t)$  it is possible to show that  $\alpha_{t-1}(\mathbf{u}_{t-1}) \geq \beta \geq \alpha_{t-1}(\mathbf{u}_{t-1})/2$ , and so invoking the regret bound of the base algorithm for the current epoch yields  $\sum_{s=\tau}^{t-1} (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}_t)) \leq O(d^2 \ln^2 T) - C\beta^{-1}d \ln T$ . When there are no restarts, the regret bound of the algorithm is simply  $\sum_{s=\tau}^{t-1} (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}_t)) \leq O(d^2 \ln T) + C\beta^{-1}d \ln^2 T$ . Starting with  $\beta = \beta_0$  and halving  $\beta$  every time there is a restart and letting  $\tau_i$  [resp.  $\beta_i$ ] be the start round [resp. the  $\beta$ ] of epoch  $i$  and  $M$  be the total number of epochs, the regret of the meta-algorithm is bound by

$$\begin{aligned} R_T(\mathbf{u}) &\leq \sum_{i=1}^M |\ell_{\tau_i}(\mathbf{p}_{\tau_i}) - \ell_{\tau_i}(\mathbf{u})| + \sum_{i=1}^M \sum_{t=\tau_i+1}^{\tau_{i+1}-1} (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u})), \\ &\leq \sum_{i=1}^M |\ell_{\tau_i}(\mathbf{p}_{\tau_i}) - \ell_{\tau_i}(\mathbf{u})| + \sum_{i=1}^M \sum_{t=\tau_i+1}^{\tau_{i+1}-1} (\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}_t)), \\ &\leq \sum_{i=1}^{M-1} \tilde{O}(1) + \sum_{i=1}^{M-1} \left( \tilde{O}(d^2) - \frac{Cd \ln^2 T}{\beta_i} \right) + \tilde{O}(d^2) + \frac{Cd \ln^2 T}{\beta_M}, \end{aligned} \quad (68)$$

$$= \tilde{O}(d^2) - \sum_{i=1}^{M-1} \frac{Cd \ln^2 T}{2^{1-i}\beta_0} + \frac{Cd \ln^2 T}{2^{1-N}\beta_0} = \tilde{O}(d^2) + \frac{Cd \ln^2 T}{\beta_0}, \quad (69)$$

where (68) follows from the regret bound of the base algorithm and the fact that  $\mathbf{p}_t, \mathbf{u} \in \bar{\Delta}_d$ , for all  $t$  (the latter ensures that  $\ell_t(\mathbf{p}_t) - \ell_t(\mathbf{u}) \leq \tilde{O}(1)$ ), and the last inequality follows by the fact that  $M \leq O(\ln T)$  since  $\inf_{\mathbf{u} \in \bar{\Delta}_d} \alpha_T(\mathbf{u}) \geq 1/(dT)$  (also because  $(\mathbf{p}_t) \subset \bar{\Delta}_d$ ). By choosing  $\beta_0 = \Omega(1)$ , (69) leads to the desired regret bound.