
Best-case lower bounds in online learning

Cristóbal Guzmán^{1,2} Nishant A. Mehta³ Ali Mortazavi³

¹Department of Applied Mathematics, University of Twente

²Institute for Mathematical and Computational Eng., Pontificia Universidad Católica de Chile

³Department of Computer Science, University of Victoria

c.guzman@utwente.nl, nmehta@uvic.ca, alithemorty@gmail.com

Abstract

Much of the work in online learning focuses on the study of sublinear upper bounds on the regret. In this work, we initiate the study of best-case lower bounds in online convex optimization, wherein we bound the largest *improvement* an algorithm can obtain relative to the single best action in hindsight. This problem is motivated by the goal of better understanding the adaptivity of a learning algorithm. Another motivation comes from fairness: it is known that best-case lower bounds are instrumental in obtaining algorithms for decision-theoretic online learning (DTOL) that satisfy a notion of group fairness. Our contributions are a general method to provide best-case lower bounds in Follow the Regularized Leader (FTRL) algorithms with time-varying regularizers, which we use to show that best-case lower bounds are of the same order as existing upper regret bounds: this includes situations with a fixed learning rate, decreasing learning rates, timeless methods, and adaptive gradient methods. In stark contrast, we show that the linearized version of FTRL can attain negative linear regret. Finally, in DTOL with two experts and binary losses, we fully characterize the best-case sequences, which provides a finer understanding of the best-case lower bounds.

1 Introduction

A typical work in online learning would develop algorithms that provably achieve low regret for some family of problems, where low regret means that a learning algorithm’s cumulative loss is not much larger than that of the best expert (or action) in hindsight. Such a work often focuses on algorithms that exhibit various forms of adaptivity, including *anytime* algorithms, which adapt to an unknown time horizon T ; *timeless* algorithms, which obtain “first-order” regret bounds that replace dependence on the time horizon by the cumulative loss of the best expert; and algorithms like AdaGrad [8], which adapt to the geometry of the data. These examples of adaptivity all involve competing with the best expert in hindsight, but adaptivity comes in many guises. Another form of adaptivity involves upgrading the comparator itself: in the *shifting regret* (also known as the *tracking regret*) [11], the learning algorithm competes with the best sequence of experts that shifts, or switches, k times for some small k . Naturally, an algorithm with low shifting regret can potentially perform much better than the single best expert in hindsight, thereby obtaining classical regret that is substantially negative.

Our work serves as a counterpoint to previous works: we show that a broad class of learning strategies provably fails, against *any* sequence of data, to substantially outperform the best expert in hindsight. Thus, these strategies are unable to obtain low shifting regret and, more generally, low regret against any comparator sequence that can be substantially better than the best expert in hindsight. More concretely, this paper initiates the study of *best-case lower bounds* in online convex optimization (OCO) for the general family of learning strategies known as Follow the Regularized Leader (FTRL) [1]. That is, we study the *minimum* possible regret of a learning algorithm over all possible sequences.

As we will show, many instances of FTRL — including adaptive instances that are anytime, timeless, or adapt to gradients like AdaGrad — never have regret that is much less than the negation of the corresponding regret upper bounds. Thus, while these instances can be adaptive in some ways, they are in a sense prohibited from uniformly obtaining low regret for adaptive notions of regret like the shifting regret. For example, in the setting of decision-theoretic online learning (DTOL) with d experts [9], the well-known anytime version of Hedge (which uses the time-varying learning rate $\eta_t \asymp \sqrt{\log(d)/t}$) enjoys $O(\sqrt{T \log d})$ worst-case regret and, as we show, has $-O(\sqrt{T \log d})$ best-case regret. Moreover, in the same setting under the restriction of two experts and binary losses, we *exactly identify the best-case sequence*, thereby showing that our best-case lower bound for this setting is tight. The structure of this sequence is surprisingly simple, but the arguments we use to pinpoint this sequence are playfully complex, bearing some similarity to the techniques of [20] and [15]. The latter work [15] considers the regret of Thompson Sampling in adversarial bit prediction; they use swapping rules and identify best-case sequences, as do we. However, the algorithms and problem settings have important differences.

A key motivation for our work is a recent result [4] which shows, in the setting of DTOL, that Hedge with constant learning rate has best-case regret lower bounded by $-O(\sqrt{T})$. This result, taken together with worst-case upper bounds of order $O(\sqrt{T})$, is then used to show that if each of finitely many experts approximately satisfies a certain notion of group fairness, then a clever use of the Hedge algorithm (running it separately on each group) also approximately satisfies the same notion of group fairness while still enjoying $O(\sqrt{T})$ regret. However, we stress that their result is very limited in that it applies only to Hedge when run with a known time-horizon. The fixed time horizon assumption also implies that their notion of group fairness also is inherently tied to a fixed time horizon (see Section 4.2 for a detailed discussion), and this latter implication can lead to experts that seem very unfair but which, based on a fixed horizon view of group fairness, are technically considered to be fair. Our best-case lower bounds enable the results of [4] to hold in much greater generality; in particular, our results enable the use of an anytime version of group fairness, which we feel is truly needed.

To our knowledge, our work is the first to study best-case lower bounds for Adaptive FTRL [16], i.e., FTRL with time-varying regularizers that can adapt to the learning algorithms’ past observations. The most closely related work is a paper by Gofer and Mansour (GM) [10] which, in the setting of online linear optimization (OLO) and when using FTRL with a *fixed* regularizer,¹ provides various lower bounds on the regret. For instance, they show that for any sequence of data, the regret is nonnegative; we recover this result as a special case of our analysis, and our analysis extends to OCO as well. GM also lower bound what they call the *anytime regret*, which superficially may seem similar to our providing best-case lower bounds for anytime algorithms. Yet, as we explain in Section 3, these two notions greatly differ. In short, their analysis lower bounds the maximum regret (over all prefixes of a sequence) for fixed horizon algorithms, whereas our analysis lower bounds the regret for all prefixes (including the minimum) for adaptively regularized algorithms, which includes anytime algorithms.

A natural question is whether results similar to our results for FTRL also hold for online mirror descent (OMD). In some situations, such as in OLO when the action space is the probability simplex, the regularizer is the negative Shannon entropy, and the learning rate is constant, our results automatically apply to OMD because the methods are then the same. More generally, it is known that OMD with a time-varying learning rate can fail spectacularly by obtaining linear regret (see Theorem 4 of [18]). Since so much of our work is tied to obtaining anytime guarantees (which would require a time-varying learning rate), we forego providing best-case lower bounds for OMD.

Our main contributions are as follows:

1. We give a general best-case lower bound on the regret for Adaptive FTRL (Section 3). Our analysis crucially centers on the notion of adaptively regularized regret, which serves as a potential function to keep track of the regret.
2. We show that this general bound can easily be applied to yield concrete best-case lower bounds for FTRL with time-varying negative regularizers, one special case being the negative Shannon entropy. We also show that an adaptive gradient FTRL algorithm (which can be viewed as a “non-linearized” version of the dual averaging version of AdaGrad [8]; see Section 4.3 for details) admits a best-case lower bound that is essentially the negation of its upper bound (Section 4).

¹[10] uses learners that follow the gradient of a concave potential, which is essentially equivalent to FTRL.

3. A widely used variant of FTRL for OCO is to first linearize the losses, leading to linearized FTRL. This method works well with respect to upper bounds, as a basic argument involving convexity goes in the right direction. However, with regards to best-case lower bounds, we show a simple construction (Section 5) for which linearized FTRL obtains $-\Omega(T)$ regret.²
4. In the setting of DTOL with 2 experts and binary losses, we explicitly identify the best-case sequence, proving that our best-case lower bounds are tight in this setting (Section 6).

The next section formalizes the problem setting and FTRL. We then develop the main results.

2 Problem Setting and General Prediction Strategies

Before giving the problem setting, we first set some notation. We denote the norm of a vector $w \in \mathcal{W}$ as $\|w\|$, the corresponding dual norm is denoted as $\|\cdot\|_*$, and \log is always the natural logarithm.

Problem setting. We consider the OCO setting. This is a game between Learner and Nature. In each round $t = 1, 2, \dots, T$, Learner selects an action w_t belonging to a closed, bounded, convex set $\mathcal{W} \subseteq \mathbb{R}^d$. Then, with knowledge of w_1, \dots, w_t , Nature responds with a convex loss function $f_t: \mathcal{W} \mapsto \mathbb{R}$. Learner then observes f_t and suffers loss $f_t(w_t)$. Learner's goal is to minimize its regret, defined as

$$\mathcal{R}_T := \sup_{w \in \mathcal{W}} \sum_{t=1}^T [f_t(w_t) - f_t(w)],$$

which is the gap between Learner's cumulative loss and that of the best action in hindsight.

This paper will cover several examples of OCO. The first example is the subclass of OLO problems. In OLO, the loss functions f_t are linear, with $f_t(w) = \langle \ell_t, w \rangle$ for some loss vector $\ell_t \in \mathbb{R}^d$. A noteworthy special case of OLO is DTOL, also known as the Hedge setting. In DTOL, we take \mathcal{W} to be equal to the simplex Δ_d over d outcomes and restrict the loss vectors as $\ell_t \in [0, 1]^d$. We introduce some notation that will be useful in the DTOL setting. For any expert $j \in [d]$, let $L_{t,j} := \sum_{s=1}^t \ell_{s,j}$ denote the cumulative loss of expert j until the end of round t . We denote the loss of Learner in round t as $\hat{\ell}_t := \langle \ell_t, w_t \rangle$ and Learner's cumulative loss at the end of round t as $\hat{L}_t := \sum_{s=1}^t \hat{\ell}_s$.

In this work, we consider the general prediction strategy of FTRL.

FTRL. Let Φ_1, Φ_2, \dots be a possibly data-dependent sequence of regularizers where, for each t , the regularizer Φ_t is a mapping $\Phi_t: \mathcal{W} \rightarrow \mathbb{R}$ which is allowed to depend on $(f_s)_{s \leq t}$. Then FTRL chooses actions according to the past regularized cumulative loss:

$$w_{t+1} = \arg \min_{w \in \mathcal{W}} \left\{ \sum_{s=1}^t f_s(w) + \Phi_t(w) \right\}. \quad (1)$$

We would like to emphasize that this is a very general template. It includes a fixed learning rate regularization, $\Phi_t \equiv \frac{1}{\eta} \Phi$, as well as its variable learning rate counterpart, $\Phi_t = \frac{1}{\eta_t} \Phi$, and arbitrary forms of adaptive choices of Φ_t based on the past. Despite this adaptivity, in the next section we will show that proving best-case lower bounds for this strategy is quite straightforward.

The regret attained by FTRL is summarized in the following known result.

Theorem 1 (Theorem 1 of [16]). *Let $(\Phi_t)_{t \geq 1}$ be a sequence of nonnegative regularizers such that, for each $t \geq 1$, Φ_t is 1-strongly convex with respect to a norm $\|\cdot\|_{(t)}$. The regret of the FTRL algorithm (1) with this sequence of regularizers is upper bounded as*

$$\mathcal{R}_T \leq \Phi_T(w^*) + \frac{1}{2} \sum_{t=1}^T \|\nabla f_t(w_t)\|_{(t-1),*}^2, \quad (2)$$

where $w^* \in \mathcal{W}$ is the best action in hindsight. If we do not require the regularizers to be nonnegative but instead assume that, for each $w \in \mathcal{W}$, the sequence $(\Phi_t(w))_{t \geq 1}$ is non-increasing, then (2) still holds if we replace $\Phi_T(w^*)$ by $\Phi_T(w^*) - \inf_{w \in \mathcal{W}} \Phi_T(w)$.

²This negative construction, combined with the fact that the dual averaging version of AdaGrad is a linearized version of FTRL, is why we only prove best-case lower bounds for the adaptive gradient FTRL algorithm.

3 A general best-case lower bound

We now present our general best-case lower bound for FTRL. Key to our analysis is the concept of *adaptively regularized regret* (hereafter abbreviated as “regularized regret”), defined as

$$\mathcal{R}_t^{\Phi_t} = \sup_{w \in \mathcal{W}} \left\{ \sum_{s=1}^t [f_s(w_s) - f_s(w)] - \Phi_t(w) \right\}. \quad (3)$$

The regularized regret can easily be related to the regret as

$$\mathcal{R}_t^{\Phi_t} \leq \mathcal{R}_t - \inf_{w \in \mathcal{W}} \Phi_t(w). \quad (4)$$

Also, applying (1), the following re-expression of the regularized regret is immediate:

$$\mathcal{R}_t^{\Phi_t} = \sum_{s=1}^t [f_s(w_s) - f_s(w_{t+1})] - \Phi_t(w_{t+1}). \quad (5)$$

Theorem 2 (Best-case lower bound on regret for adaptive FTRL). *Consider the setting of online convex optimization and the adaptive FTRL strategy (1). Suppose that there exists a sequence $(\alpha_t)_{t \in [T]}$ such that $\Phi_t(w_t) \leq \Phi_{t-1}(w_t) + \alpha_t$ for all $t \in [T]$. Then*

$$\mathcal{R}_T \geq \inf_{w \in \mathcal{W}} \Phi_T(w) - \inf_{w \in \mathcal{W}} \Phi_0(w) - \sum_{t=1}^T \alpha_t.$$

Proof. We start by inductively bounding the adaptively regularized regret:

$$\begin{aligned} \mathcal{R}_{t+1}^{\Phi_{t+1}} - \mathcal{R}_t^{\Phi_t} &= \sup_{w \in \mathcal{W}} \left\{ \sum_{s=1}^{t+1} [f_s(w_s) - f_s(w)] - \Phi_{t+1}(w) \right\} - \sum_{s=1}^{t+1} [f_s(w_s) - f_s(w_{t+1})] + \Phi_t(w_{t+1}) \\ &\geq -\Phi_{t+1}(w_{t+1}) + \Phi_t(w_{t+1}) \\ &\geq -\alpha_{t+1}, \end{aligned}$$

where the first equality is from (5). We conclude that $\mathcal{R}_T^{\Phi_T} \geq \mathcal{R}_0^{\Phi_0} - \sum_{t=1}^T \alpha_t$. Next, from (4),

$$\begin{aligned} \mathcal{R}_T &\geq \mathcal{R}_T^{\Phi_T} + \inf_{w \in \mathcal{W}} \Phi_T(w) \geq \inf_{w \in \mathcal{W}} \Phi_T(w) + \mathcal{R}_0^{\Phi_0} - \sum_{t=1}^T \alpha_t \\ &= \inf_{w \in \mathcal{W}} \Phi_T(w) - \inf_{w \in \mathcal{W}} \Phi_0(w) - \sum_{t=1}^T \alpha_t, \end{aligned}$$

where in the last equality we used that $\mathcal{R}_0^{\Phi_0} = \sup_{w \in \mathcal{W}} \{-\Phi_0(w)\}$. \square

The closest results to Theorem 2 of which we are aware are in the intriguing work of Gofer and Mansour (GM) [10], who provided lower bounds for FTRL with a fixed regularizer for OLO. Their Theorem 1 shows that the best-case regret is nonnegative. One wonders if the doubling trick can extend their results to anytime best-case lower bounds; regrettably, the doubling trick’s restarts can allow the algorithm to achieve $-\Omega(T)$ regret (see Appendix A). GM also lower bound a notion they call the anytime regret (see Theorem 5 of [10]). The anytime regret for a sequence as they define it is actually the maximum regret over all prefixes of the sequence (where the sequence has a fixed time horizon); ultimately, the lower bound they obtain depends on the quadratic variation of the sequence as computed on a *fixed* time horizon. Related to this, GM’s analysis is for algorithms that use the same regularizer in all rounds. They lament that it is unclear how to extend their analysis to handle time-varying learning rates³. Our goal is rather different, as taking the maximum regret over all prefixes says nothing about how large (in the negative direction) the regret could be for some particular prefix. Thus, our style of analysis, which provides a lower bound on the regret for all time horizons (which, in particular, provides a lower bound on the *minimum* over all prefixes), differs greatly from theirs and is what is needed. We think the different styles of our work and theirs stems from the regret being nonnegative in their paper, whereas it need not be for time-varying regularizers.

Theorem 2 possesses considerable generality, owing to its applying to FTRL with adaptive regularizers. We highlight just a few applications in the next section.

³A time-varying learning rate gives rise to perhaps the most basic form of an adaptive regularizer.

4 Best-case lower bounds in particular settings

Theorem 2 presented in last section, despite its simplicity, is an extremely powerful method, capable of addressing several of the settings where FTRL attains sublinear regret. Next, we proceed to enumerate some important examples of instances of FTRL, comparing existing worst-case upper bounds on the regret with our best-case lower bounds.

4.1 Non-increasing learning rates and timeless algorithms

We first present several examples in which the elements of the sequence $(\Phi_t)_{t \geq 1}$ take the form $\Phi_t = \frac{1}{\eta_t} \Phi$ for a fixed regularizer Φ and a time-varying learning rate η_t .

Constant learning rate. The simplest example is that of a fixed learning rate $\eta_t \equiv \eta$, which means $\Phi_t(w) = \frac{1}{\eta} \Phi(w)$. Taking $\alpha_t = 0$ for all t , Theorem 2 immediately implies that the regret is always nonnegative; this implication was previously shown by Gofer and Mansour (see Thm. 1 of [10]). A simple consequence is that the Follow the Leader (FTL) strategy (i.e., $\Phi_t \equiv 0$) has nonnegative regret, which also can be inferred from the results of [10]. Although we cannot find a precise reference, we believe that it was already known that FTL always obtains nonnegative regret (even prior to [10]).

Time-varying learning rate. More generally, taking a time-varying learning rate, Theorem 2 gives

$$\mathcal{R}_T \geq \left(\frac{1}{\eta_T} - \frac{1}{\eta_0} \right) \inf_{w \in \mathcal{W}} \Phi(w) - \sum_{t=1}^T \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \Phi(w_{t+1}). \quad (6)$$

A typical strategy to obtain sublinear anytime worst-case regret is to set the learning rate as $\eta_t = \eta / \sqrt{t+1}$ for some constant η that depends on various known problem-dependent constants. Continuing from (6) with this setting further implies that

$$\mathcal{R}_T \geq \frac{1}{\eta} (\sqrt{T+1} - 1) \inf_{w \in \mathcal{W}} \Phi(w) - \frac{1}{\eta} \sup_{w \in \mathcal{W}} \Phi(w) \sum_{t=1}^T \frac{1}{2\sqrt{t+1}}.$$

If Φ is nonpositive — as holds when \mathcal{W} is the d -dimensional simplex and Φ is the negative Shannon entropy — the above is further lower bounded by

$$\frac{1}{\eta} (\sqrt{T+1} - 1) \inf_{w \in \mathcal{W}} \Phi(w) - \frac{1}{\eta} \sup_{w \in \mathcal{W}} \Phi(w) \sqrt{T+1}. \quad (7)$$

A particularly interesting example is the DTOL setting. In this setting, when we take Φ to be the negative Shannon entropy $\Phi(w) = \sum_{j=1}^d w_j \log w_j$ and set $\eta = 2\sqrt{\log d}$ so that $\eta_t = 2\sqrt{(\log d)/(t+1)}$ recovers the standard anytime version of Hedge which has also been called Decreasing Hedge [17]. In round t , this algorithm plays w_t such that $w_{t,j} \propto \exp(-\eta_{t-1} L_{t-1,j})$ for $j \in [d]$, and we have the following anytime best-case lower bound and worst-case upper bound on the regret:

$$-\frac{1}{2} \sqrt{T \log d} \leq \mathcal{R}_T \leq \sqrt{T \log d}.$$

The lower bound holds from (7) combined with $\sqrt{T+1} - 1 \leq \sqrt{T}$ and $-\log d \leq \Phi \leq 0$, while the upper bound is from Theorem 2 of [6]. The upper bound is minimax optimal in terms of the rate and, asymptotically (letting both d and T go to infinity) has a constant that is optimal up to a factor of $\sqrt{2}$. As we show in Section 6, in the case of $d = 2$ the lower bound also has the optimal rate.

Timeless algorithms. In the context of DTOL, it is straightforward to adapt our analysis for Decreasing Hedge to Hedge with any non-increasing learning rate. One example of interest is

$$\eta_t = -\log \left(1 - \min \left\{ \frac{1}{4}, \sqrt{\frac{2 \log d}{L_t^*}} \right\} \right),$$

where $L_T^* = \min_{j \in [d]} L_{T,j}$ is the cumulative loss of the best expert. This choice of adaptive learning rate yields an anytime upper bound on the regret of $O(\sqrt{L_T^* \log d} + \log d)$ in the DTOL setting (see Theorem 2.1 of [2], who actually prove this result in the more general setting of prediction with

expert advice). Such a bound is called timeless because rounds in which all experts suffer the same loss have no effect on the bound [7]. This is a natural property to have in this setting, and with the above choice of learning rate, we have the following timeless⁴ best-case lower bound:

$$\mathcal{R}_T \geq \min \left\{ 0, -\sqrt{\frac{L_T^* \log d}{2}} + 4 \log d \right\}; \quad (8)$$

a brief derivation is in Appendix B. Again, notice the similarity between the upper and lower bounds.

4.2 Group fairness in online learning

In a pioneering work, Blum, Gunasekar, Lykouris, and Srebro (BGLS) [4] considered a notion of group fairness in DTOL. In their setup, the DTOL protocol is augmented so that, at the start of each round t , Nature selects and reveals to Learner a group g_t belonging to a set of groups \mathcal{G} prior to Learner's playing its action w_t . They assume that for a known, fixed time horizon T , each expert $j \in [d]$ has balanced mistakes across groups in the following sense: For any $g \in \mathcal{G}$, let $\mathcal{T}(g) := \{t \in [T] : g_t = g\}$ denote the rounds belonging to group g , and let $L_{\mathcal{T}(g),j} := \sum_{t \in \mathcal{T}(g)} \ell_{t,j}$ denote the cumulative loss of expert $j \in [d]$ when considering only the rounds in $\mathcal{T}(g)$; then we say that expert j is fair in isolation if

$$\frac{L_{\mathcal{T}(g),j}}{|\mathcal{T}(g)|} = \frac{L_{\mathcal{T}(g'),j}}{|\mathcal{T}(g')|} \quad \text{for all } g, g' \in \mathcal{G}.$$

BGLS devised a strategy based on the multiplicative weights algorithm which simultaneously satisfies an approximate notion of group fairness while still enjoying $O(\sqrt{T \log d})$ worst-case regret. The multiplicative weights algorithm (which we hereafter refer to as Hedge as the algorithms are equivalent for constant learning rate) sets w_t as $w_{t,j} \propto (1 - \tilde{\eta})^{L_{t-1,j}}$ for $j \in [d]$ for a learning rate parameter $\tilde{\eta}$. Their strategy is to run a separate copy of Hedge for each group, so that for any group g , the copy corresponding to group g is run on the subsequence corresponding to the rounds in $\mathcal{T}(g)$. For brevity, we call this ‘‘Interleaved Hedge’’. In BGLS's analysis (see the proof of their Theorem 3), they first give worst-case regret upper and lower bounds for Hedge. Specifically, they show that if Hedge is run with a constant learning rate, then

$$(1 - 4\tilde{\eta}) \cdot L_T^* \leq \hat{L}_T \leq (1 + \tilde{\eta})L_T^* + (\log d)/\tilde{\eta}. \quad (9)$$

An optimal, non-anytime worst-case tuning of $\tilde{\eta}$ then yields matching-magnitude regret lower and upper bounds of $-O(\sqrt{T \log d})$ and $O(\sqrt{T \log d})$ respectively. Then on the one hand, the regret of Interleaved Hedge satisfies $\mathcal{R}_T = O(\sqrt{|\mathcal{G}|T \log d})$. In addition, by virtue of BGLS's $-O(\sqrt{T \log d})$ lower bound for Hedge when fed T rounds (along with the standard upper bound), their Interleaved Hedge enjoys the following group fairness guarantee:

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g')}}{|\mathcal{T}(g')|} = O\left(\sqrt{\frac{\log d}{T_0}}\right) \quad \text{for all } g, g' \in \mathcal{G} \text{ and } T_0 := \min_g |\mathcal{T}(g)|,$$

where we adopt the notation $\hat{L}_{\mathcal{T}(g)} := \sum_{t \in \mathcal{T}(g)} \hat{\ell}_t$.

By using our improved nonnegative best-case lower bound for Hedge with constant learning rate (which again, is not a new result) together with an optimal, non-anytime worst-case tuning of η , we can obtain the following improvement over (9):

$$0 \leq \hat{L}_T \leq \sqrt{T \log d}.$$

As explained in the appendix (the remaining steps are essentially due to BGLS), we can then obtain the following improved group fairness guarantee:

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g')}}{|\mathcal{T}(g')|} \leq \sqrt{\frac{\log d}{T_0}} \quad \text{for all } g, g' \in \mathcal{G} \text{ and } T_0 := \min_g |\mathcal{T}(g)|.$$

The above result holds when the learning rate for each group g is set as $\eta^{(g)} = \sqrt{\frac{\log d}{|\mathcal{T}(g)|}}$. Of course, running Hedge instances for each group with the correct constant learning rates means that the

⁴Technically, the upper and lower bounds as stated are not timeless, but it is easy to see that any round for which all experts have the same loss can be replaced by a round where all experts have zero loss, with no change in the algorithm's behavior nor its regret. Our regret bounds on the modified loss sequence then become timeless.

algorithm must know $|\mathcal{T}(g)|$ for each group $g \in \mathcal{G}$, at least within a reasonable constant factor. This is a far stronger assumption than the already strong assumption of a known time horizon.

Using our anytime best-case lower bound for Decreasing Hedge, combined with the analysis of BGLS, it is straightforward to vastly extend their results in each of the following ways:

1. Using BGLS's fixed horizon notion of fairness and a copy of Decreasing Hedge for each group's instance, we can drop the assumption that each group's cardinality $|\mathcal{T}(g)|$ is known.
2. The most interesting extension, whose possibility was the original basis of our entire work, is that we can now upgrade BGLS's notion of group fairness to its anytime sibling. This involves measuring, for every prefix of the length- T game, the discrepancy between the error rates of any pair of groups. This is an arguably more natural notion of fairness, as it avoids situations where an expert purports to be fair while having all of its mistakes for one group occur in the first half of the game. Since we now have an anytime best-case lower bound for Decreasing Hedge, we have the requisite piece needed to show that Interleaved (Decreasing) Hedge satisfies the same notion of anytime group fairness. Our timeless best-case lower bounds also apply here, giving that extension as well.

We now briefly sketch each of these extensions, leaving more detailed derivations to the appendix.

Decreasing Hedge (Anytime results). Suppose now that Interleaving Hedge uses copies of Decreasing Hedge with time-varying learning rate $\eta_t = 2\sqrt{(\log d)/(t+1)}$. Note that in this case, the copy for group g increments its internal round only each time a new round for group g appears. We can then automatically apply our anytime lower bound on the regret of Decreasing Hedge to obtain

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} \leq \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} + \frac{1}{2} \sqrt{\frac{\log d}{|\mathcal{T}(g^*)|}} \quad \text{for } g^* \text{ such that } \frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} = \min_{g \in \mathcal{G}} \frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|}.$$

Therefore, if in hindsight we have $T_0 = \min_{g \in \mathcal{G}} |\mathcal{T}(g)|$, then the following fairness guarantee holds:

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g')}}{|\mathcal{T}(g')|} \leq \frac{3}{2} \sqrt{\frac{\log d}{T_0}} \quad \text{for all } g, g' \in \mathcal{G}.$$

Note that even though Learner need not know the time horizon T nor $|\mathcal{T}(g)|$ for each g , the above guarantee still can only hold for a fixed time horizon T . This is because the definition of fairness in isolation only holds for a fixed time horizon.

Anytime fairness. As mentioned earlier, the fixed horizon view of fairness in isolation can be very limiting. Rather than opting for fairness in isolation to hold for a fixed time horizon, we argue that it is more natural for this definition to hold in the following anytime sense:

$$\frac{L_{\mathcal{T}(g),j}}{|\mathcal{T}(g)|} = \frac{L_{\mathcal{T}(g'),j}}{|\mathcal{T}(g')|} \quad \text{for all } g, g' \in \mathcal{G} \text{ and for all } T;$$

note that the time horizon T is implicit in each $\mathcal{T}(g)$.

The attentive reader may notice that this definition can be overly restrictive on Nature (perhaps impossibly so), and therefore it is natural to allow the equality to hold only approximately (which BGLS did explore). Just as in BGLS's work, it is possible to extend our results to cases where fairness holds only approximately, and such an extension is certainly warranted in the case of anytime fairness. This extension requires only straightforward modifications to the analysis and so we do not give further details here. In the case of anytime fairness, it further makes sense for approximate fairness to be defined according to a rate. We leave this extension to a future, longer paper.

Using the above anytime version of fairness in isolation, it is straightforward to extend the previous result to the following new result. Just like above, suppose now that Interleaved Hedge uses copies of Decreasing Hedge with time-varying learning rate. Then, for all time horizons T ,⁵

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g')}}{|\mathcal{T}(g')|} \leq \frac{3}{2} \sqrt{\frac{\log d}{T_0}} \quad \text{for all } g, g' \in \mathcal{G}.$$

Mini-conclusion. Whereas a key message of [4] is adaptive algorithms (in the sense of low shifting regret) cannot satisfy group fairness, at least when using the interleaved approach, our best-case lower bounds do cover many other types of adaptivity. This shows that with regards to group fairness and the interleaved strategy, the group-fair adversarial online learning tent is actually quite large.

⁵Note that in the below, each $\mathcal{T}(g)$ (and hence T_0) implicitly depends on the time horizon T .

4.3 Adaptive gradient FTRL

Inspired by the *adaptive gradient* (AdaGrad) algorithm for OCO [8], we consider an adaptive gradient FTRL algorithm with quadratic regularizers, $\Phi_t(w) = \frac{1}{2\eta} \langle w, H_t w \rangle$: this corresponds to the adaptive FTRL (1) with regularizers Φ_t as above. We include its two variants. In the below, we use the notation $g_t := \nabla f_t(w_t)$ and $\delta > 0$ is a fixed number:

- (a) Diagonal: $H_t = \delta I + \text{Diag}(s_t)$, where $s_t = \left(\left(\sqrt{\sum_{\tau=1}^t g_{j,\tau}^2} \right)_{j \in [d]} \right)$,
- (b) Full-matrix: $H_t = \delta I + G_t^{1/2}$ where $G_t = \sum_{\tau=1}^t g_\tau g_\tau^\top$.

We emphasize that the proposed algorithm does not exactly match AdaGrad from [8]. To resolve this inconsistency, we use the regret upper bounds from Theorem 1, combined with upper bounds on the gradient norms that appear in such upper bounds, proved in [8]. This leads to the following result.

Theorem 3 (From [8]). *Consider the setting of OCO. Given $1 \leq p \leq \infty$, we denote by D_p the diameter of \mathcal{W} in the $\|\cdot\|_p$ norm, and M_p is a bound on the $\|\cdot\|_p$ norm of the gradients for any possible loss. Then the adaptive gradient FTRL algorithm satisfies the following bounds:*

- (a) Diagonal: If $\delta = M_\infty$, then $\sum_{t=1}^T \|\nabla f_t(w_t)\|_{(t-1),*}^2 \leq 2\eta \sum_{j=1}^d \sqrt{\sum_{t=1}^T g_{t,j}^2}$. In particular, setting $\eta = D_\infty$, we obtain an upper bound on the regret $\frac{D_\infty^2}{D_\infty} M_\infty + D_\infty \sum_{j=1}^d \sqrt{\sum_{t=1}^T g_{t,j}^2}$.
- (b) Full matrix: If $\delta = M_2$, then $\sum_{t=1}^T \|\nabla f_t(w_t)\|_{(t-1),*}^2 \leq 2\eta \text{Tr}(G_T^{1/2})$. In particular, setting $\eta = D_2$, we obtain an upper bound on the regret $D_2[M_2/2 + \text{Tr}(G_T^{1/2})]$.

We proceed to best-case lower bounds, deferring the proof to Appendix D. The proof strategy follows similar telescopic recursions to those used in [8, 16]. Notice that these bounds closely match the respective worst-case upper bounds.

Proposition 4. *In the OCO setting, the adaptive gradient FTRL strategy with parameter tuning as in Theorem 3 attains best-case lower bounds on the regret of $-(D_\infty/2) \sum_{j=1}^d \sqrt{\sum_{t \in [T]} g_{t,j}^2}$ in the diagonal case and $-(D_2/2)\text{Tr}(G_T^{1/2})$ in the full-matrix case.*

5 On best-case lower bounds for other algorithms

So far, our study of best-case lower bounds has focused on adaptive FTRL algorithms. A major drawback of FTRL is that the iteration cost of each subproblem grows linearly with the number of rounds. It is then tempting to consider more efficient algorithms attaining comparable regret upper bounds. We discuss such possibilities for two natural algorithms: linearized FTRL and OMD.

5.1 Linearized FTRL

In this section, we give a simple construction in which linearized FTRL obtains negative linear regret, i.e., regret which is $-\Omega(T)$. The problem is a forecasting game with binary outcomes, the action space \mathcal{D} equal to the 2-simplex $[0, 1]$, and the squared loss $\ell(p, y) = \frac{1}{2}(p - y)^2$. This can be cast as an OCO problem by taking $\mathcal{W} = [0, 1]$ and, for each $t \in [T]$, setting $f_t(p) = (p - y_t)^2$ for some outcome $y_t \in \{0, 1\}$ selected by Nature.

We consider linearized FTRL using the negative Shannon entropy regularizer; this is equivalent to exponentiated gradient descent [12] and uses the update $p_t = \frac{1}{1 + \exp(\eta_t G_{t-1})}$, where $G_{t-1} = \sum_{s=1}^{t-1} g_s$ and each $g_s = p_s - y_s$ is the gradient of the loss with respect to p_s under outcome y_s . In this situation, an $O(\sqrt{T})$ anytime regret upper bound can be obtained by employing the time-varying learning rate $\eta_t = 1/\sqrt{t}$ (see Theorem 2.3 of [5] or Theorem 2 of [6]).

Let Nature's outcome sequence be the piecewise constant sequence consisting of $T/2$ zeros followed by $T/2$ ones. Therefore, the best action in hindsight is $p^* = 0.5$. We now state our negative result and briefly sketch a proof; the full proof can be found in Appendix E.

Theorem 5. *In the construction above, linearized FTRL obtains regret $\mathcal{R}_T = -\Omega(T)$.*

Proof (sketch). The core idea behind the proof is that linearized FTRL exhibits a type of switching behavior. Let q_0 and q_1 be constants satisfying $0 < q_0 < 1/2 < q_1 < 1$. Observe that in the first half of the game (considered in isolation), the optimal action is 0. Thus, if p_t drops below q_0 in the first half of the game, then for all rounds thereafter in the first half, Learner picks up negative constant regret in each round. We show that this drop happens after only a constant number of rounds for a simple reason: the gradient must be at least q_0 whenever $p_t > q_0$ in the first half, and p_t changes according to the exponentiated cumulative gradient. Thus, Learner picks up negative linear regret in the first half of the game. Next, consider the second half of the game. For this half (considered in isolation), the optimal action is 1. Thus, once p_t surpasses q_1 , for all rounds thereafter Learner will again pick up negative constant regret. Via an arguably coarse analysis, we show that the number of rounds in the second half before Learner's action p_t surpasses q_1 is at most a small constant fraction of T . To show this, we rely on the fact that whenever $p_t < q_1$ in the second half, we have that the gradient must be at least $-q_1$ (combined with the exponential effect of the gradients on p_t). That is, Learner quickly switches from having p_t below q_0 to having p_t above q_1 . Thus, for most of the second half, Learner again picks up negative linear regret. \square

5.2 Online mirror descent

For OLO over the simplex, the online mirror descent method with entropic regularization and *constant learning rate* attains best-case lower bounds $-O(\sqrt{T} \log d)$ [4]. This algorithm is known to attain upper bounds on the regret of the same order. The extension of this result for *non-increasing learning rates*,⁶ albeit possible, seems of limited interest, as this algorithm is known to achieve linear regret in the worst case [18], which is attributed to the unboundedness of the Bregman divergence.

6 Best-case loss sequence for binary DTOL with two experts

In this section, we characterize the binary loss sequence for DTOL with two experts in which the regret for Decreasing Hedge is minimized. Recall that Decreasing Hedge chooses expert i at round t with probability $p_{i,t} = \frac{\exp(-\eta_t L_{i,t-1})}{\sum_{j \in \{1,2\}} \exp(-\eta_t L_{j,t-1})}$, where $\eta_t = \sqrt{1/t}$. Losses are the form $\ell_t = \begin{pmatrix} \ell_{1,t} \\ \ell_{2,t} \end{pmatrix}$ where $\ell_{1,t}, \ell_{2,t} \in \{0, 1\}$. Note that $\eta_{t+1} < \eta_t$ for all t .

Our approach is to introduce a set of operations on a sequence of binary loss vectors $\ell_1, \ell_2, \dots, \ell_T$ that do not increase the regret of Decreasing Hedge. Using this set of operations, we will show that any loss sequence with $T = 2K$ rounds $\ell_1, \ell_2, \dots, \ell_{2K}$, can be converted to a specific loss sequence which we call the canonical best-case sequence.

Definition 6. For any length $T = 2K \geq 16$, the canonical best-case loss sequence is defined as:

$$\text{Canonical}(2K) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{K-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{K-1}. \quad (10)$$

We now sketch our proof that (10) is indeed the best-case sequence. The full proof is in Appendix F.

First, consider an arbitrary binary loss sequence for two experts ℓ_1, \dots, ℓ_T . We can substitute all $(1, 1)$ loss vectors with $(0, 0)$ loss vectors as it can be easily shown that this operation does not change the regret. Thus, there is no need for $(1, 1)$ loss vectors in the best-case sequence.

Next, we introduce the notion of a leader change and show that there is no leader change in the best-case sequence. Therefore, we only need to consider loss sequences without leader changes. Formally, we say expert j is a leader at round t if $L_{j,t} = \min_{i \in \{1,2\}} L_{i,t}$. Moreover, if j is the *only* leader at round t , then we call $i_t^* = j$ the strict leader in round t . We say that a sequence has a leader change if there exists times t_1 and t_2 where in both times the strict leader is defined⁷ and $i_{t_1}^* \neq i_{t_2}^*$. Defining $\Delta_t := L_{2,t} - L_{1,t}$, observe that if a sequence has a leader change, then the sign of Δ should change at least once. We show (Lemma 15) that any loss sequence $\ell_1, \ell_2, \dots, \ell_T$ can be converted, without increasing the regret, to a sequence $\ell'_1, \ell'_2, \dots, \ell'_T$ where Δ is always nonnegative (hence, the resulting sequence has no leader change). Removing all leader changes facilitates the next operation to successfully convert loss sequences.

Next, consider the operation of swapping two consecutive loss vectors. When we swap a loss sequence $\ell_1, \dots, \ell_{t-1}, \ell_t, \ell_{t+1}, \ell_{t+2}, \dots, \ell_T$ at round $(t, t+1)$, the resulting sequence becomes

⁶In fact, we have found a simple argument showing the regret is nonnegative, strengthening the result in [4].

⁷If in round t , $L_{1,t} = L_{2,t}$, then the strict leader is not defined in that round.

$\ell_1, \dots, \ell_{t-1}, \ell_{t+1}, \ell_t, \ell_{t+2}, \dots, \ell_T$. It can be easily shown that this swap only changes the loss of Decreasing Hedge at rounds t and $t + 1$. Since we only consider loss sequences with no leader change, we can assume without loss of generality that expert 1 is the only leader. Now based on the swapping rule lemma (Lemma 16), we can always move $\binom{0}{1}$ to the earlier rounds and $\binom{1}{0}$ to the later rounds. By repeatedly applying this swapping rule, the resulting loss sequence will be of the form $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$. Note that as we have shown that Δ is always nonnegative, $a \geq b$.

So far, we have shown that any loss sequence can be converted to a loss sequence of the form $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$. We then show in Lemma 17 that if c is not even, then the sequence $\binom{0}{1}^{a-1} \binom{0}{0}^{c+1} \binom{1}{0}^b$ has smaller regret and $a - 1 \geq b$. Therefore, we only need to consider loss sequences of the form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where $a \geq b$ and c is even. As we assume T is even, $a + b$ is also even. Now it is shown (in Lemma 18) that among all sequence of losses of form $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$, where $a + b = 2K$, the loss sequence $\binom{0}{1}^K \binom{0}{0}^c \binom{1}{0}^K$ has the least regret. Therefore, we only need to consider loss sequences of this latter form where $2K + c = T$.

Finally, among all possible (K, c) for loss sequences $\binom{0}{1}^K \binom{0}{0}^c \binom{1}{0}^K$ such that $2K + c = T$, we show that the regret is minimized when $c = 2$. This form coincides with the canonical best-case sequence.

Bounding the regret. As shown in Appendix F.6, the regret on the canonical best-case loss sequence can be lower and upper bounded as follows:

$$-\frac{e^2}{(1-\frac{1}{e})} \cdot \sqrt{T} - \frac{1}{2} \leq \mathcal{R}(T) \leq -\frac{1}{1+e\sqrt{e}} \sqrt{T} + \frac{12}{\sqrt{e}} + \frac{1}{2}.$$

As both the lower and upper bounds are $-\Theta(\sqrt{T})$, the analysis in Section 3 is tight in this case.

7 Discussion

In this work, we provided a systematic treatment of best-case lower bounds in online learning for adaptive FTRL algorithms, discussed the impossibility of certain natural extensions, and also provided a tighter analysis of such lower bounds in the binary prediction experts setting. As one application, we have shown that our results for adaptive FTRL enable the use a broad class of adaptive online learning algorithms that satisfy a balanced mistakes notion of group fairness. Naturally, many questions still remain open, and we hope this work motivates further research in this intriguing topic. A first question relates to algorithms that can achieve negative regret: can we characterize the conditions under which this happens? The goal would be to reveal, beyond the particular example we gave in Section 5.1, structural properties of an instance that lead to substantial outperformance of the best fixed action. Returning to that example, another question arises. We have shown an example of OCO with a strongly convex loss (the squared loss) for which linearized FTRL obtains negative linear regret, whereas our best-case lower bounds for (non-linearized) FTRL with constant learning rate imply $-O(\sqrt{T})$ regret and known upper bounds imply $O(\log T)$ regret. Thus, while in this situation, linearized FTRL pays a price with respect to regret upper bounds, it can exhibit a switching behavior that might allow it to compete with a shifting sequence of comparators. Investigating this “blessing” of linearization would be a fascinating direction for future work. Finally, for DTOL with two experts, we showed that our best-case lower bounds are tight by constructing the best-case sequence. It would be interesting to show tightness without an explicit construction, as we could then say that the best-case lower bounds are achievable more generally. For instance, while our swapping-based technique is most related to that of [20] for worst-case sequences and [15] (who give both worst- and best-case sequences), it could also be worthwhile to draw from previous works [19, 13, 3, 14] which explicitly work out the minimax regret and minimax strategies. Even so, our results on identifying the structure of the best-case loss sequence show that Decreasing Hedge has the lowest regret for loss sequences for which the best shifting sequence of experts has precisely one shift. One research direction would be to see if this pattern holds more generally, for other FTRL-type algorithms.

We conclude our discussion pointing out that the theoretical nature of our work does not entail direct negative societal consequences. As far as our discussions on group fairness are concerned, one should be careful when applying these methods as the notion of balanced mistakes may be susceptible to strong forms of discrimination (see, e.g., the Discussion section in [4]). In particular, balanced mistakes is a sound notion of non-discrimination when all examples that contribute towards the performance also contribute towards fairness.

Acknowledgments and Disclosure of Funding

CG's work was partially supported by FONDECYT Regular project 1210362, INRIA through the INRIA Associate Teams project, and ANID – Millennium Science Initiative Program – NCN17 059. NM and AM were supported by the NSERC Discovery Grant RGPIN-2018-03942. We also thank Sajjad Azami for his involvement in the earlier stages of this work.

References

- [1] Jacob Abernethy, Elad E Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *21st Annual Conference on Learning Theory, COLT 2008*, pages 263–273, 2008.
- [2] Peter Auer, Nicolo Cesa-Bianchi, and Claudio Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002.
- [3] Peter L Bartlett, Wouter M Koolen, Alan Malek, Eiji Takimoto, and Manfred K Warmuth. Minimax fixed-design linear regression. In *Conference on Learning Theory*, pages 226–239. PMLR, 2015.
- [4] Avrim Blum, Suriya Gunasekar, Thodoris Lykouris, and Nati Srebro. On preserving non-discrimination when combining expert advice. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL <https://proceedings.neurips.cc/paper/2018/file/2e855f9489df0712b4bd8ea9e2848c5a-Paper.pdf>.
- [5] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- [6] Alexey Chernov and Fedor Zhdanov. Prediction with expert advice under discounted loss. In *International Conference on Algorithmic Learning Theory*, pages 255–269. Springer, 2010.
- [7] Steven De Rooij, Tim Van Erven, Peter D Grünwald, and Wouter M Koolen. Follow the leader if you can, hedge if you must. *The Journal of Machine Learning Research*, 15(1):1281–1316, 2014.
- [8] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.
- [9] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.
- [10] Eyal Gofer and Yishay Mansour. Lower bounds on individual sequence regret. *Machine Learning*, 103(1):1–26, 2016.
- [11] Mark Herbster and Manfred K Warmuth. Tracking the best expert. *Machine learning*, 32(2): 151–178, 1998.
- [12] Jyrki Kivinen and Manfred K Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *information and computation*, 132(1):1–63, 1997.
- [13] Wouter M Koolen, Alan Malek, and Peter L Bartlett. Efficient minimax strategies for square loss games. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL <https://proceedings.neurips.cc/paper/2014/file/8d420fa35754d1f1c19969c88780314d-Paper.pdf>.
- [14] Wouter M Koolen, Alan Malek, Peter L Bartlett, and Yasin Abbasi Yadkori. Minimax time series prediction. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015. URL <https://proceedings.neurips.cc/paper/2015/file/4dcf435435894a4d0972046fc566af76-Paper.pdf>.

- [15] Yuval Lewi, Haim Kaplan, and Yishay Mansour. Thompson sampling for adversarial bit prediction. In *Algorithmic Learning Theory*, pages 518–553. PMLR, 2020.
- [16] H. Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *Journal of Machine Learning Research*, 18(90):1–50, 2017. URL <http://jmlr.org/papers/v18/14-428.html>.
- [17] Jaouad Mourtada and Stéphane Gaïffas. On the optimality of the hedge algorithm in the stochastic regime. *Journal of Machine Learning Research*, 20:1–28, 2019.
- [18] Francesco Orabona and Dávid Pál. Scale-free online learning. *Theoretical Computer Science*, 716:50–69, 2018.
- [19] Eiji Takimoto and Manfred K Warmuth. The minimax strategy for gaussian density estimation. pp. In *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory*, pages 100–106, 2000.
- [20] Tim Van Erven, Wojciech Kotłowski, and Manfred K Warmuth. Follow the leader with dropout perturbations. In *Conference on Learning Theory*, pages 949–974. PMLR, 2014.

A Inadequacy of the doubling trick for best-case lower bounds

In this appendix, we briefly sketch in the setting of DTOL why a fixed horizon best-case lower bound (based on a constant regularizer) cannot be combined with the doubling trick in order to obtain an anytime best-case lower bound of the same order. Recall that in the standard analysis used for the doubling trick, the algorithm actually competes with the best expert over each epoch (where epoch sizes double). Consequently, it is possible that, over each epoch, an algorithm has (say) $-O(\sqrt{T})$ regret, but as compared to the single best expert over the entire sequence of rounds, the regret is $-\Omega(T)$. This can happen as follows. Let the best expert in all but the last epoch be the same expert j_1 , while the best expert over the last epoch is another expert j_2 . Finally, assume that (considering the cumulative loss over the full game) expert j_1 is linearly better than expert j_2 , whereas in the last epoch expert j_2 is linearly better than expert j_1 . In this situation, the doubling trick allows for sublinear regret with respect to the 1-shifting regret, and the best 1-shifting sequence of experts (a sequence of experts that can change experts only once) is linearly better than the best 0-shifting sequence of experts (i.e., the best constant sequence, also known as the best expert).

B Derivations for (8)

Recall that we set η_t as

$$\eta_t = -\log \left(1 - \min \left\{ \frac{1}{4}, \sqrt{\frac{2 \log d}{L_t^*}} \right\} \right).$$

Starting from (6) and using the fact that $(\eta_t)_{t \geq 0}$ is non-increasing and Φ (the negative Shannon entropy) is non-positive, we have

$$\mathcal{R}_T \geq \left(\frac{1}{\eta_T} - \frac{1}{\eta_0} \right) \inf_{w \in \mathcal{W}} \Phi(w).$$

Next, since $-\log(1-x) \geq x$ for $x < 1$, we have that $\frac{1}{\eta_T} \leq \max \left\{ 4, \sqrt{\frac{L_T^*}{2 \log d}} \right\}$. We also have $\eta_0 = 4$ since $L_0^* = 0$. Using these two facts, together with $\inf_{w \in \Delta_d} \Phi(w) = -\log d$ yields

$$\mathcal{R}_T \geq \left(4 - \max \left\{ 4, \sqrt{\frac{L_T^*}{2 \log d}} \right\} \right) \log d.$$

The result (8) follows from some basics manipulations.

C Fairness

In this appendix, for the convenience of the reader we repeat the results from Section 4.2 while giving derivations for them along the way.

In their work, Blum, Gunasekar, Lykouris, and Srebro (BGLS) [4] express regret bounds in a somewhat different language than we do. Rather than dealing directly with the regret, they instead analyze the *approximate regret*. The ε -approximate regret relative to expert $j \in [d]$ is defined as

$$\hat{L}_T - (1 + \varepsilon)L_{T,j}. \tag{11}$$

In order to more easily compare to their results, and to interpret their results in the large body of literature that focuses on the actual (non-approximate) regret, we first show how to modify some of their analysis to give results based on the actual regret (hereafter simply referred to as the “regret”, just like in the main text).

First, recall that BGLS use the multiplicative weights algorithm, which sets w_t as $w_{t,j} \propto (1 - \tilde{\eta})^{L_{t-1,j}}$ for $j \in [d]$ for a learning rate parameter $\tilde{\eta}$. In BGLS’s proof of their Theorem 3, they first give worst-case regret upper and lower bounds for Hedge (“multiplicative weights” in their work). Specifically, they show that

$$(1 - 4\tilde{\eta}) \cdot L_T^* \leq \hat{L}_T \leq (1 + \tilde{\eta})L_T^* + \frac{\log d}{\tilde{\eta}}.$$

An optimal, non-anytime worst-case tuning of $\hat{\eta}$ then yields matching-magnitude regret lower and upper bounds of $-O(\sqrt{T \log d})$ and $O(\sqrt{T \log d})$ respectively. We note in passing that the anytime version of multiplicative weights (which uses incremental updates) is equivalent to OMD with a time-varying learning rate and is known to achieve linear regret ([18, Theorem 4]).

We can obtain a similar bound for constant learning rate Hedge with an optimal, non-anytime worst case tuning of η (defined as in the main text of our paper), with the improvement⁸

$$0 \leq \hat{L}_T \leq \sqrt{T \log d}.$$

Next, we briefly explain how to modify BGLS's proof of their Theorem 3. We first need some notation (largely derived from BGLS, but with small modifications to integrate more nicely into our notation). For any group g , let $j^*(g)$ be the best expert when considering the rounds involving group g (i.e., $\mathcal{T}(g)$). Therefore, $j^*(g) \in \arg \min_{j \in [d]} L_{\mathcal{T}(g), j}$. Next, recall that Interleaving Hedge runs a separate copy of Hedge for each group. Let g^* be the group whose copy of Hedge obtains the lowest average loss. That is, g^* is such that

$$\frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} = \min_{g \in \mathcal{G}} \frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|}.$$

We now pick up at the last math display in Appendix C of BGLS [4]; this is the step where the lower and upper regret bounds are used. Adjusting their analysis using our bounds, we have for a fixed time horizon T and when the copy of Hedge running on group g uses learning rate⁹ $\eta^{(g)} = \sqrt{\frac{\log d}{|\mathcal{T}(g)|}}$:

$$\begin{aligned} \frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} &\leq \frac{L_{\mathcal{T}(g), j^*(g)}}{|\mathcal{T}(g)|} + \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} - \frac{L_{\mathcal{T}(g^*), j^*(g^*)}}{|\mathcal{T}(g^*)|} \\ &\leq \frac{L_{\mathcal{T}(g), j^*(g^*)}}{|\mathcal{T}(g)|} - \frac{L_{\mathcal{T}(g^*), j^*(g^*)}}{|\mathcal{T}(g^*)|} + \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} \\ &= \sqrt{\frac{\log d}{|\mathcal{T}(g)|}}, \end{aligned}$$

where the first inequality uses the regret lower and upper bounds, the second inequality is based on the optimality of $j^*(g)$ for group g , and the equality uses fairness in isolation (for any pair of groups $g, g' \in \mathcal{G}$ and expert j (including $j^*(g^*)$), we have that $\frac{L_{\mathcal{T}(g), j}}{|\mathcal{T}(g)|} = \frac{L_{\mathcal{T}(g'), j}}{|\mathcal{T}(g')|}$).

Decreasing Hedge (Anytime analysis). Suppose now that Interleaving Hedge uses copies of Decreasing Hedge with time-varying learning rate $\eta_t = 2\sqrt{(\log d)/(t+1)}$. Note that in this case, the copy for group g increments its internal round only each time a new round for group g appears. We can then automatically apply our anytime lower bound on the regret of Decreasing Hedge (together with the already well-known regret upper bound) to obtain

$$\begin{aligned} \frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} &\leq \frac{L_{\mathcal{T}(g), j^*(g)}}{|\mathcal{T}(g)|} + \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} - \frac{L_{\mathcal{T}(g^*), j^*(g^*)}}{|\mathcal{T}(g^*)|} + \frac{1}{2} \sqrt{\frac{\log d}{|\mathcal{T}(g^*)|}} \\ &\leq \frac{L_{\mathcal{T}(g), j^*(g^*)}}{|\mathcal{T}(g)|} - \frac{L_{\mathcal{T}(g^*), j^*(g^*)}}{|\mathcal{T}(g^*)|} + \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} + \frac{1}{2} \sqrt{\frac{\log d}{|\mathcal{T}(g^*)|}} \\ &= \sqrt{\frac{\log d}{|\mathcal{T}(g)|}} + \frac{1}{2} \sqrt{\frac{\log d}{|\mathcal{T}(g^*)|}}. \end{aligned}$$

Therefore, if in hindsight we have $T_0 := \min_{g \in \mathcal{G}} |\mathcal{T}(g)|$, then the following fairness guarantee holds:

$$\frac{\hat{L}_{\mathcal{T}(g)}}{|\mathcal{T}(g)|} - \frac{\hat{L}_{\mathcal{T}(g^*)}}{|\mathcal{T}(g^*)|} \leq \frac{3}{2} \sqrt{\frac{\log d}{T_0}}.$$

⁸Since this is a fixed horizon setting for now, the constant in the upper bound can be improved, but for simplicity we will just use the constant of 1.

⁹Yes, it is a very strong assumption to assume that the $\mathcal{T}(g)$'s are known ahead of time, but this is precisely our point. We will relax this soon.

Note that even though Learner need not know the time horizon T nor $|\mathcal{T}(g)|$ for each g , the above guarantee still can only hold for a fixed time horizon T . This is because the definition of fairness in isolation only holds for a fixed time horizon. By adjusting this definition as we did in the main text (which we believe is sensible), we can extend this result to hold for anytime fairness.

D Proof of Proposition 4

Before proceeding to the proof of this result, notice that without loss of generality, we may assume that $0 \in \mathcal{W}$ (if not, shift the regularizers by centering them at any fixed point $\bar{w} \in \mathcal{W}$). This implies that $\inf_{w \in \mathcal{W}} \Phi_t(w) = 0$, so we only need to focus on the coefficients α_t .

Proof. 1. **Diagonal case.** First, for the diagonal case, let s_t be the d -dimensional vector with coefficients $s_{t,j} = \sqrt{\sum_{s=1}^t g_{t,j}^2}$. Then, if we let $\mathbf{1} \in \mathbb{R}^d$ be the all-ones vector,

$$\begin{aligned} \alpha_{t+1} &= \langle w_{t+1}, (H_{t+1} - H_t)w_{t+1} \rangle = \langle w_{t+1}, \text{Diag}(s_{t+1} - s_t)w_{t+1} \rangle \\ &\leq \max_{j \in [d]} (w_j^{t+1})^2 \|s_{t+1} - s_t\|_1 = \max_{j \in [d]} (w_j^{t+1})^2 \langle s_{t+1} - s_t, \mathbf{1} \rangle, \end{aligned}$$

where we used that s_t is coordinate-wise nondecreasing. Next,

$$\sum_{t=1}^T \alpha_t = \frac{1}{2\eta} \sum_{t=0}^{T-1} \langle w_{t+1}, (H_{t+1} - H_t)w_{t+1} \rangle \leq \frac{1}{2\eta} \max_{t \in [T]} \|w_t\|_\infty^2 \langle s_T - s_0, \mathbf{1} \rangle.$$

Hence, choosing $\eta = D_\infty$, we obtain a lower bound on regret $-\frac{1}{2} D_\infty \sum_{j=1}^d \sqrt{\sum_{t=1}^T g_{t,j}^2}$.

2. **Full-matrix case.** Now for the full-matrix update, we can proceed similarly. First,

$$\begin{aligned} 2\eta\alpha_{t+1} &= \langle w_{t+1}, (H_{t+1} - H_t)w_{t+1} \rangle = \langle w_{t+1}, (G_{t+1}^{1/2} - G_t^{1/2})w_{t+1} \rangle \\ &\leq \|w_{t+1}\|_2^2 \lambda_{\max}(G_{t+1}^{1/2} - G_t^{1/2}) = \|w_{t+1}\|_2^2 \text{Tr}(G_{t+1}^{1/2} - G_t^{1/2}), \end{aligned}$$

where we used that the matrix $G_{t+1}^{1/2} - G_t^{1/2}$ is positive semidefinite. Now, summing over t ,

$$\sum_{t=1}^T \alpha_t = \frac{1}{2\eta} \sum_{t=1}^T \langle w_{t+1}, (H_{t+1} - H_t)w_{t+1} \rangle \leq \frac{1}{2\eta} \max_{t \in [T]} \|w_t\|_2^2 \left(\text{Tr}(G_T^{1/2}) - \text{Tr}(G_0^{1/2}) \right).$$

To conclude, for $\eta = D_2$, the regret is lower bounded by

$$\mathcal{R}^T \geq -\frac{D_2}{2} \text{Tr}(G_T^{1/2}),$$

which proves the result. \square

E Negative result for Linearized FTRL

In this section, we give a full proof of Theorem 5.

Proof (of Theorem 5). Let q_0 and q_1 be constants satisfying $0 < q_0 < 1/2 < q_1 < 1$; we will tune these constants later. For the analysis, we divide the T rounds into 4 segments: (i) the rounds in the first half for which $p_t \geq q_0$; (ii) the remaining rounds in the first half; (iii) the rounds in the second half for which $p_t \leq q_1$; (iv) the remaining rounds in the second half.

The basic idea of the proof is to show that $p_t < q_0$ after a constant number of rounds (where the constant depends on q_0) and hence the first segment is of constant length. In the second segment, the algorithm picks up negative linear regret since $p_t < q_0 < 1/2 = p^*$ and $y_t = 0$ in the second segment. In the third segment, where the outcomes now satisfy $y_t = 1$, we show that the algorithm can take at most a linear number of rounds (but with a suitably small coefficient) before satisfying

$p_t > q_1$ and so picks up at most linear positive regret (but again, with a suitably small coefficient). Finally, in the last segment, the algorithm once again picks up negative linear regret since in this segment $y_t = 1$ and $p_t > q_1 > 1/2 = p^*$ and also, as we show, this segment is linear in length.

In more detail, we will upper bound the regret via the following claims, whose proofs appear after this proof.

Claim 1: In the first half of the game, we always have $p_t \leq \frac{1}{2}$.

A simple consequence of Claim 1 is that the first segment's contribution to the regret is nonpositive (recall that $y_t = 0$ in the first half).

Claim 2: The number of rounds in the first segment is at most $t_1 = O(1)$.

Claim 3: The number of rounds in the third segment is at most $t_3 = \frac{q_0}{2(1-q_1)}T + O(\sqrt{T})$.

Putting the 3 claims together and recalling that the contribution to the regret is nonpositive in the first segment, we have that the regret of the algorithm is at most:

$$-\underbrace{\left(\frac{T}{2} - t_1\right)}_{\text{second segment}} \cdot \underbrace{\left(\frac{1}{8} - \frac{q_0^2}{2}\right)}_{\text{third segment}} + t_3 \cdot \underbrace{\left(\frac{1}{2} - \frac{1}{8}\right)}_{\text{third segment}} - \underbrace{\left(\frac{T}{2} - t_3\right)}_{\text{fourth segment}} \cdot \underbrace{\left(\frac{1}{8} - \frac{(1-q_1)^2}{2}\right)}_{\text{fourth segment}}. \quad (12)$$

Substituting the values of t_1 and t_3 , grouping terms, and rearranging, the above is equal to

$$\frac{T}{4} \left(-\frac{1}{2} + q_0^2 + \frac{q_0}{1-q_1} + (1-q_1)^2 - q_0(1-q_1) \right) + O(\sqrt{T}).$$

A suitable choice of q_0 is $q_0 = (1-q_1)^2$, which yields

$$\begin{aligned} & \frac{T}{4} \left(-\frac{1}{2} + (1-q_1)^4 + (1-q_1) + (1-q_1)^2 - (1-q_1)^3 \right) + O(\sqrt{T}) \\ & \leq \frac{T}{4} \left(-\frac{1}{2} + (1-q_1) + (1-q_1)^2 \right) + O(\sqrt{T}). \end{aligned}$$

We can take the moderate choice $q_1 = \frac{3}{4}$ for example (and hence $q_0 = \frac{1}{16}$), yielding the upper bound $-\frac{3T}{64} + O(\sqrt{T})$. \square

We now prove the claims.

Proof (of Claim 1). Consider an arbitrary round t in the first half. Observe that $g_s = p_s \geq 0$ for all $s < t$; consequently, $G_{t-1} \geq 0$. The claim follows from the definition of p_t . \square

Proof (of Claim 2). We will show that the number of rounds in the first segment is at most

$$t_1 := \left\lceil \left(\frac{1}{2q_0} \left(\log \frac{1-q_0}{q_0} + \sqrt{\log^2 \frac{1-q_0}{q_0} + 4q_0^2} \right) \right)^2 \right\rceil.$$

The idea behind the proof is that whenever $p_t \geq q_0$, the gradient g_t will be sufficiently positive and, consequently, when p_t is updated it will make sufficient progress towards dropping below q_0 . We now formalize this intuition.

Let $\mathcal{T}_1 \subseteq [T/2]$ be the set of round indices in the first segment. By definition of the first segment, we have $p_t \geq q_0$ for all $t \in \mathcal{T}_1$. Therefore, for t such that $t-1 \in \mathcal{T}_1$,

$$\begin{aligned} p_t &= \frac{1}{1 + \exp\left(\eta_t \sum_{s=1}^{t-1} g_s\right)} \\ &\leq \frac{1}{1 + \exp\left(\eta_t \sum_{s=1}^{t-1} q_0\right)} = \frac{1}{1 + \exp\left(q_0 \frac{t-1}{\sqrt{t}}\right)}. \end{aligned} \quad (13)$$

It suffices to find the smallest t such that (13) is at most q_0 . This is equivalent to finding the smallest t such that

$$q_0 \frac{t-1}{\sqrt{t}} \geq \log \frac{1-q_0}{q_0}.$$

Rearranging and making the replacement $u := \sqrt{t}$, it suffices to solve the quadratic equation

$$q_0 u^2 - \left(\log \frac{1-q_0}{q_0} \right) u - q_0 = 0.$$

Taking the positive solution and rounding up yields that the smallest such t is at most

$$\left\lceil \left(\frac{\log \frac{1-q_0}{q_0} + \sqrt{\log^2 \frac{1-q_0}{q_0} + 4q_0^2}}{2q_0} \right)^2 \right\rceil = t_1.$$

□

Proof (of Claim 3). Let \mathcal{T}_3 be the set of round indices in the third segment. Let t be such that $t-1 \in \mathcal{T}_3$ (note: this implies that $t-1 \geq T/2$). Then Claims 1 and 2, together with the definition of segment 3, imply that

$$\begin{aligned} p_t &= \frac{1}{1 + \exp\left(\eta_t \sum_{s=1}^{t-1} g_s\right)} \\ &\geq \frac{1}{1 + \exp\left(\eta_t \left(t_1 \cdot \frac{1}{2} + \left(\frac{T}{2} - t_1\right) \cdot q_0 - \left(t-1 - \frac{T}{2}\right) \cdot (1-q_1)\right)\right)}. \end{aligned} \quad (14)$$

It suffices to find the smallest t such that (14) is at least q_1 . Proceeding similarly to the proof of Claim 2 (and once again introducing $u := \sqrt{t}$ yields the quadratic equation

$$-(1-q_1)u^2 + \left(\log \frac{q_1}{1-q_1}\right)u + \left(\frac{1-q_1+q_0}{2} \cdot T + t_1 \left(\frac{1}{2} - q_0\right) + 1 - q_1\right) = 0$$

Solving for u yields

$$u = \frac{\log \frac{q_1}{1-q_1} + \sqrt{\log^2 \frac{q_1}{1-q_1} + 4(1-q_1) \left(\frac{1-q_1+q_0}{2} \cdot T + t_1 \left(\frac{1}{2} - q_0\right) + 1 - q_1\right)}}{2(1-q_1)}.$$

From the above, we see that the smallest t satisfying (14) is equal to

$$\frac{1-q_1+q_0}{2(1-q_1)} \cdot T + O\left(\sqrt{T}\right),$$

which proves the claim. □

F Best-case loss sequence for binary DTOL with 2 experts

In this section, we characterize the binary loss sequence for DTOL with two experts in which the regret for Decreasing Hedge is minimized. Recall that Decreasing Hedge chooses expert i at round t with probability $p_{i,t} = \frac{\exp(-\eta_t L_{i,t-1})}{\sum_{j \in \{1,2\}} \exp(-\eta_t L_{j,t-1})}$, where $\eta_t = \sqrt{1/t}$. We denote the loss vector for round t as $\ell_t = \begin{pmatrix} \ell_{1,t} \\ \ell_{2,t} \end{pmatrix}$, where $\ell_{1,t}, \ell_{2,t} \in \{0, 1\}$. Note that $\eta_{t+1} < \eta_t$ for all t . We denote $\begin{pmatrix} L_{t,1} \\ L_{t,2} \end{pmatrix}$ by \mathbf{L}_t .

Our approach is to introduce a set of operations on a sequence of binary losses $\ell_1, \ell_2, \dots, \ell_T$ that do not increase the regret of Decreasing Hedge. Using this set of operations, we will show that any (suitably long) loss sequence with $T = 2K$ rounds $\ell_1, \ell_2, \dots, \ell_{2K}$, can be converted to a specific loss sequence which we call the canonical best-case sequence.

Definition 7. For any length $2K \geq 16$, the canonical best case loss sequence will be defined as follows.

$$\text{Canonical}(2K) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{K-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{K-1}. \quad (15)$$

In this appendix, we will give a complete presentation of Section 6 along with complete proofs.

F.1 Converting all (1,1) to (0,0)

In this part, we will show that adding a constant term to a loss sequence at round t would not have any effect on the regret. This is because in Decreasing Hedge, the weights in each round t depend only on the difference between the cumulative loss of two experts up until the end of round $t - 1$. Adding a constant loss to both experts at any round would not change the difference between cumulative losses at any round; therefore, the weight vector remains the same for all rounds. Moreover, in round t , the best expert as well as Decreasing Hedge incur the same additional loss; therefore, the regret would be the same in that round.

Lemma 8. *Adding constant value c to the loss of all experts at round t would not have any effect on the regret.*

Using Lemma 8, we can substitute all (1,1) loss vectors with (0,0) loss vectors. This means that there is no need for (1,1) loss vectors in the best-case sequence.

F.2 Dealing with Leader Change

We define $\bar{i} := \begin{cases} 2 & i = 1 \\ 1 & i = 2 \end{cases}$ and let $L_{i,t}$ denote the cumulative loss of expert i up until the end of round t . We say expert j is a leader at round t if $L_{j,t} = \min_{i \in \{1,2\}} L_{i,t}$. Moreover, if j is the *only* leader at round t , then we define the *strict leader* at round t to be $i_t^* = j$.¹⁰

We say that a sequence has a leader change if there exists times t_1 and t_2 such that in both times the strict leader is defined and $i_{t_1}^* \neq i_{t_2}^*$. Defining $\Delta_t := L_{2,t} - L_{1,t}$, observe that if a sequence has a leader change, then the sign of Δ should change at least once. In the following, we will show that any loss sequence $\ell_1, \ell_2, \dots, \ell_T$ can be converted, without increasing the regret, to a sequence $\ell'_1, \ell'_2, \dots, \ell'_T$ where Δ is always nonnegative; therefore the resulting sequence does not have any leader change.

We first need to define an operation called switching the loss at round t .

Definition 9. The operation of *switching loss sequence* $\ell_1, \ell_2, \dots, \ell_T$ at round t yields a new loss sequence $\ell'_1, \ell'_2, \dots, \ell'_T$ where for all $s < t$, we have $\ell'_{i,s} = \ell_{i,s}$ and for all $s \geq t$, we have $\ell'_{i,s} = \ell_{\bar{i},s}$.

Obviously switching loss sequence $\ell_1, \ell_2, \dots, \ell_T$ at round 1 only swaps the indices of the experts; therefore, the regret remains the same for the switched loss sequence. Similarly, we will show that for any t , where $\Delta_{t-1} = 0$, switching loss sequence $\ell_1, \ell_2, \dots, \ell_T$ at round t does not change the regret.

Lemma 10. *For a loss sequence $\ell_1, \ell_2, \dots, \ell_T$, if $\Delta_{t-1} = 0$ then switching loss sequence at round t does not change the regret.*

Proof. We show that after switching at round t where $\Delta_{t-1} = L_{2,t-1} - L_{1,t-1} = 0$, the cumulative loss of Decreasing Hedge and the best expert remains the same. Therefore, the regret does not change.

For any round $k \geq t$,

$$\begin{aligned} L'_{i,k} &:= \sum_{s=1}^{t-1} \ell'_{i,s} + \sum_{s=t}^k \ell'_{i,s} = \sum_{s=1}^{t-1} \ell_{i,s} + \sum_{s=t}^k \ell_{\bar{i},s} && \text{(switched loss definition)} \\ &= L_{i,t-1} + \sum_{s=t}^k \ell_{\bar{i},s} = L_{\bar{i},t-1} + \sum_{s=t}^k \ell_{\bar{i},s} && (\Delta_{t-1} = 0) \\ &= L_{\bar{i},k}. \end{aligned}$$

Hence, for $k = T$ we have $L'_{1,T} = L_{2,T}$, and $L'_{2,T} = L_{1,T}$. Therefore, the cumulative loss of the best expert for both loss sequences is the same, but the index of the best expert is changed. Moreover, observe that for any $k \geq t$,

$$\Delta'_k := L'_{2,k} - L'_{1,k} = L_{1,k} - L_{2,k} = -\Delta_k.$$

¹⁰If in round t , $L_{1,t} = L_{2,t}$, then the strict leader is not defined in that round.

Also for round t we know that

$$\Delta'_{t-1} := L'_{2,t-1} - L'_{1,t-1} = L_{2,t-1} - L_{1,t-1} = \Delta_{t-1} = 0 = -\Delta_{t-1}.$$

Therefore, for s such that $s \geq t$, we have $\Delta'_{s-1} = -\Delta_{s-1}$.

Next, denote the loss of Decreasing Hedge for loss sequence ℓ_1, \dots, ℓ_T (and ℓ'_1, \dots, ℓ'_T) at round s by $\hat{\ell}_s$ (and $\hat{\ell}'_s$ respectively). Obviously, as there is no difference up until round t , for all rounds $s < t$, we have $\hat{\ell}_s = \hat{\ell}'_s$. Now, for any round $s \geq t$:

$$\begin{aligned} \hat{\ell}'_s &= \sum_{i \in \{1,2\}} \frac{\exp(-\eta_s(L'_{i,s-1}))}{\exp(-\eta_s(L'_{i,s-1})) + \exp(-\eta_s(L'_{\bar{i},s-1}))} \ell'_{i,s} \\ &= \sum_{i \in \{1,2\}} \frac{1}{1 + \exp(\eta_s(L'_{i,s-1} - L'_{\bar{i},s-1}))} \ell'_{i,s} \\ &= \frac{1}{1 + \exp(\eta_s(L'_{1,s-1} - L'_{2,s-1}))} \ell'_{1,s} + \frac{1}{1 + \exp(\eta_s(L'_{2,s-1} - L'_{1,s-1}))} \ell'_{2,s} \\ &= \frac{1}{1 + \exp(\eta_s(-\Delta'_{s-1}))} \ell'_{1,s} + \frac{1}{1 + \exp(\eta_s(\Delta'_{s-1}))} \ell'_{2,s} \\ &= \frac{1}{1 + \exp(\eta_s(\Delta_{s-1}))} \ell_{2,s} + \frac{1}{1 + \exp(\eta_s(-\Delta_{s-1}))} \ell_{1,s} \\ &= \sum_{i \in \{1,2\}} \frac{1}{1 + \exp(\eta_s(L_{i,s-1} - L_{\bar{i},s-1}))} \ell_{i,s} = \hat{\ell}_s. \end{aligned}$$

Therefore, the cumulative loss of Decreasing Hedge for both loss sequences is the same. \square

We also need to formally define the notion of a leader change at round t .

Definition 11. We say a loss sequence has a leader change at round t if there exists $k < t$ such that for all s satisfying $t - k < s < t$, it holds that $\Delta_s = 0$ and $\Delta_t \cdot \Delta_{t-k} < 0$.

Definition 12. Consider a sequence $\ell_1, \ell_2, \dots, \ell_T$ where t is the smallest value such that $\Delta_t \neq 0$. We define the first strict leader in this sequence to be expert 1 if $\Delta_t > 0$ and expert 2 if $\Delta_t < 0$.

We are now ready to show that any sequence with leader changes can be converted, without increasing the regret, to a sequence with no leader change. We only consider loss sequences in which the first strict leader is expert 1 because if the first strict leader for a sequence $\ell_1, \ell_2, \dots, \ell_T$ is expert 2, then we can simply swap the experts' indices.

Now, if the first strict leader is expert 1, then Δ first becomes positive before it gets any chance to become negative. Our goal now is to modify the loss sequence without increasing the regret so that Δ stays nonnegative for all rounds. The number of times the sign of Δ changes corresponds to the number of leader changes.

Observation 13. The number of leader changes in a sequence is the number of times the sign of Δ_t changes from negative to positive or vice versa.

Lemma 14. *If a loss sequence ℓ_1, \dots, ℓ_T has at least one leader change, and the first leader change happens at round t , then switching the loss sequence at round t would remove the first leader change without changing the regret. Moreover, this switch operation ensures that the first leader change (if any) in resulting loss sequence ℓ'_1, \dots, ℓ'_T happens at time $s > t$.*

Proof. We assume that the first strict leader is expert 1, therefore $\Delta_s \geq 0$ for $s < t$. Now, in round t where the first leader change happens, we know that $\Delta_t < 0$. Moreover, $\Delta_{t-1} = 0$. Now by Lemma 10 we can switch the loss sequence at round t without increasing the regret. In the resulting loss sequence, for all $s < t$, $\Delta'_s = \Delta_s$ and for all $s \geq t$, $\Delta'_s = -\Delta_s$. This ensures that for all $s \leq t$, $\Delta'_s \geq 0$. Therefore, in the resulting loss sequence, no leader change could happen at any time $s \leq t$. Therefore, the first leader change (if any) in the resulting loss sequence happens at some time $s > t$. \square

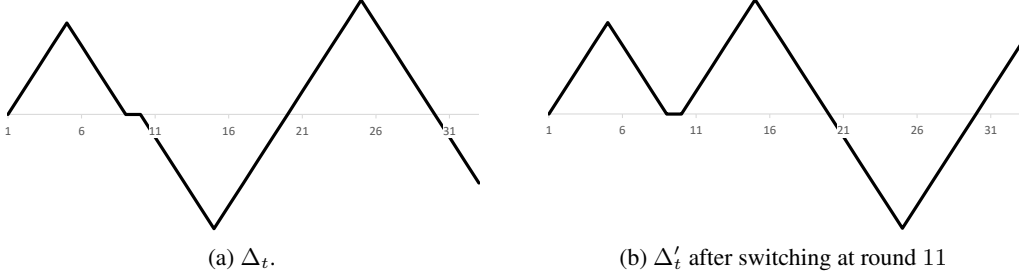


Figure 1: The plots show Δ_t and Δ'_t for a loss sequence ℓ_1, \dots, ℓ_T and ℓ'_1, \dots, ℓ'_T . The first leader change in ℓ_1, \dots, ℓ_T happens at round 11. Switching at round 11 decreases the number of leader changes. The first leader change in the resulting sequence ℓ'_1, \dots, ℓ'_T happens at round 21 $>$ 11.

Lemma 15. *Any loss sequence with a leader change can be converted to a loss sequence with no leader change without changing the regret.*

Proof. If a sequence has at least one leader change, we can remove the first leader change at round t using Lemma 14. By applying this lemma, the first leader change (if any) in the resulting loss sequence happens at round $s > t$. Because the number of rounds T is bounded, we can use this operation finitely many times to reach a sequence without any leader change. \square

F.3 Swapping two consecutive losses

In this section, we will consider the operation of swapping two consecutive loss vectors. When we swap a loss sequence $\ell_1, \dots, \ell_{t-1}, \ell_t, \ell_{t+1}, \ell_{t+2}, \dots, \ell_T$ at rounds $(t, t+1)$, the resulting sequence becomes $\ell_1, \dots, \ell_{t-1}, \ell_{t+1}, \ell_t, \ell_{t+2}, \dots, \ell_T$. Because the behavior of Decreasing Hedge at round s only depends on L_{s-1} , swapping two consecutive loss vectors at rounds $(t, t+1)$ does not change L_{s-1} for $s \in [T] \setminus \{t, t+1\}$ and therefore does not affect the behavior of the algorithm at any rounds other than rounds t and $t+1$. Therefore, in order to see the effect of swapping two consecutive loss vectors at rounds $(t, t+1)$, it is enough to compare $\hat{\ell}_t + \hat{\ell}_{t+1}$ with $\hat{\ell}'_t + \hat{\ell}'_{t+1}$ where $\hat{\ell}_t$ is defined to be the loss incurred by Decreasing Hedge at round t in the modified sequence.

In Section F.1, we had shown that we can replace all $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ loss vectors with $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ without changing the regret. Thus, we only consider all possible sequences with loss vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and characterize when swapping two consecutive loss vectors of the above form does not increase the regret.

Moreover, by Section F.2, we only need to consider loss sequences ℓ_1, \dots, ℓ_T that do not have a leader change prior to the swap. Moreover, without loss of generality, we can assume the expert 1 is always the leader in the loss sequence.

We will now show that in a sequence where expert 1 is always the leader, any swap that pushes the $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ vector to the later rounds or any swap that pushes the $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to the earlier rounds does not increase the regret.

Lemma 16 (Swapping rules). *In a loss sequence ℓ_1, \dots, ℓ_T , where $\Delta_{t-1} \geq 0$, swapping at round $(t, t+1)$ does not increase the regret in any of the following cases:*

- (a) $\ell_t, \ell_{t+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$;
- (b) $\ell_t, \ell_{t+1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$;
- (c) $\ell_t, \ell_{t+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Proof. Let p_t (and p'_t) be the probability of choosing arm 1 in round t in loss sequence ℓ_1, \dots, ℓ_T (and ℓ'_1, \dots, ℓ'_T respectively). Define an increasing function $f(x) := \frac{1}{1 + \exp(-x)}$. Now, we can rewrite p_t as follows:

$$p_t = \frac{1}{1 + \exp(-\eta_t \Delta_{t-1})} = f(\eta_t \Delta_{t-1}).$$

- (a) In this case, the loss incurred by Decreasing Hedge for loss sequence ℓ_1, \dots, ℓ_T at rounds t and $t + 1$ can be expressed as

$$\hat{\ell}_t + \hat{\ell}_{t+1} = \underbrace{\binom{p_t}{1-p_t} \cdot \binom{1}{0}}_{\text{round } t} + \underbrace{\binom{p_{t+1}}{1-p_{t+1}} \cdot \binom{0}{0}}_{\text{round } t+1} = p_t.$$

Similarly the loss incurred by Decreasing Hedge for loss sequence ℓ'_1, \dots, ℓ'_T at rounds t and $t + 1$ can be written as

$$\hat{\ell}'_t + \hat{\ell}'_{t+1} = \underbrace{\binom{p'_t}{1-p'_t} \cdot \binom{0}{0}}_{\text{round } t} + \underbrace{\binom{p'_{t+1}}{1-p'_{t+1}} \cdot \binom{1}{0}}_{\text{round } t+1} = p'_{t+1}.$$

We know that Decreasing Hedge incurs the same loss at any round $s \in [T] \setminus \{t, t + 1\}$. Therefore,

$$\begin{aligned} \hat{L}'_T - \hat{L}_T &= \sum_{s=1}^T \hat{\ell}'_s - \sum_{s=1}^T \hat{\ell}_s \\ &= (\hat{\ell}'_t + \hat{\ell}'_{t+1}) - (\hat{\ell}_t + \hat{\ell}_{t+1}) \\ &= p'_{t+1} - p_t \\ &= f(\eta_{t+1} \Delta'_t) - f(\eta_t \Delta_{t-1}) \\ &= f(\eta_{t+1} (\Delta'_{t-1} + 0)) - f(\eta_t \Delta_{t-1}) \\ &= f(\eta_{t+1} \Delta'_{t-1}) - f(\eta_t \Delta_{t-1}) \\ &= f(\eta_{t+1} \Delta_{t-1}) - f(\eta_t \Delta_{t-1}) \\ &\leq 0, \end{aligned}$$

where the last inequality comes from the fact in Decreasing Hedge, $\eta_{t+1} < \eta_t$ and f is an increasing function.

- (b) In this case, the loss incurred by Decreasing Hedge for loss sequence ℓ_1, \dots, ℓ_T at rounds t and $t + 1$ can be expressed as

$$\hat{\ell}_t + \hat{\ell}_{t+1} = \underbrace{\binom{p_t}{1-p_t} \cdot \binom{0}{0}}_{\text{round } t} + \underbrace{\binom{p_{t+1}}{1-p_{t+1}} \cdot \binom{0}{1}}_{\text{round } t+1} = 1 - p_{t+1}.$$

Similarly the loss incurred by Decreasing Hedge for loss sequence ℓ'_1, \dots, ℓ'_T at rounds t and $t + 1$ can be written as

$$\hat{\ell}'_t + \hat{\ell}'_{t+1} = \underbrace{\binom{p'_t}{1-p'_t} \cdot \binom{0}{1}}_{\text{round } t} + \underbrace{\binom{p'_{t+1}}{1-p'_{t+1}} \cdot \binom{0}{0}}_{\text{round } t+1} = 1 - p'_t.$$

We know that Decreasing Hedge incurs the same loss at any round $s \in [T] \setminus \{t, t + 1\}$; therefore,

$$\begin{aligned} \hat{L}'_T - \hat{L}_T &= \sum_{s=1}^T \hat{\ell}'_s - \sum_{s=1}^T \hat{\ell}_s \\ &= (\hat{\ell}'_t + \hat{\ell}'_{t+1}) - (\hat{\ell}_t + \hat{\ell}_{t+1}) \\ &= (1 - p'_t) - (1 - p_{t+1}) \\ &= p_{t+1} - p'_t \\ &= f(\eta_{t+1} \Delta_t) - f(\eta_t \Delta'_{t-1}) \\ &= f(\eta_{t+1} (\Delta_{t-1} + 0)) - f(\eta_t \Delta'_{t-1}) \\ &= f(\eta_{t+1} \Delta_{t-1}) - f(\eta_t \Delta_{t-1}) \\ &\leq 0, \end{aligned}$$

where the last inequality comes from the fact in Decreasing Hedge, $\eta_{t+1} < \eta_t$ and f is an increasing function.

- (c) In this case, as we assume that expert 1 is always the leader, $\Delta_s \geq 0$ for all s . Therefore, $\Delta_t = \Delta_{t-1} - 1 \geq 0$. The loss incurred by Decreasing Hedge for loss sequence ℓ_1, \dots, ℓ_T at rounds t and $t+1$ can be expressed as

$$\hat{\ell}_t + \hat{\ell}_{t+1} = \underbrace{\binom{p_t}{1-p_t} \cdot \binom{1}{0}}_{\text{round } t} + \underbrace{\binom{p_{t+1}}{1-p_{t+1}} \cdot \binom{0}{1}}_{\text{round } t+1} = (p_t) + (1 - p_{t+1}).$$

Similarly, the loss incurred by Decreasing Hedge for loss sequence ℓ'_1, \dots, ℓ'_T at rounds t and $t+1$ can be written as

$$\hat{\ell}'_t + \hat{\ell}'_{t+1} = \underbrace{\binom{p'_t}{1-p'_t} \cdot \binom{0}{1}}_{\text{round } t} + \underbrace{\binom{p'_{t+1}}{1-p'_{t+1}} \cdot \binom{1}{0}}_{\text{round } t+1} = (1 - p'_t) + (p'_{t+1}).$$

We know that Decreasing Hedge incurs the same loss at any round $s \in [T] \setminus \{t, t+1\}$; therefore,

$$\begin{aligned} \hat{L}'_T - \hat{L}_T &= \sum_{s=1}^T \hat{\ell}'_s - \sum_{s=1}^T \hat{\ell}_s \\ &= (\hat{\ell}'_t + \hat{\ell}'_{t+1}) - (\hat{\ell}_t + \hat{\ell}_{t+1}) \\ &= (p'_{t+1} + p_{t+1}) - (p'_t + p_t) \\ &= (f(\eta_{t+1}\Delta'_t) + f(\eta_{t+1}\Delta_t)) - (f(\eta_t\Delta'_{t-1}) + f(\eta_t\Delta_{t-1})) \\ &= (f(\eta_{t+1}\Delta'_t) + f(\eta_{t+1}\Delta_t)) - (2 \cdot f(\eta_t\Delta_{t-1})) \\ &= (f(\eta_{t+1}(\Delta_{t-1} + 1)) + f(\eta_{t+1}(\Delta_{t-1} - 1))) - 2 \cdot f(\eta_t\Delta_{t-1}) \\ &= 2 \cdot \left(\frac{1}{2} f(\eta_{t+1}(\Delta_{t-1} + 1)) + \frac{1}{2} f(\eta_{t+1}(\Delta_{t-1} - 1)) \right) - 2 \cdot f(\eta_t\Delta_{t-1}) \\ &\leq 2 \cdot f\left(\frac{1}{2}\eta_{t+1}(\Delta_{t-1} + 1) + \frac{1}{2}\eta_{t+1}(\Delta_{t-1} - 1)\right) - 2 \cdot f(\eta_t\Delta_{t-1}) \\ &= 2 \cdot f(\eta_{t+1}\Delta_{t-1}) - 2 \cdot f(\eta_t\Delta_{t-1}) \\ &= 2 \cdot [f(\eta_{t+1}\Delta_{t-1}) - f(\eta_t\Delta_{t-1})] \\ &\leq 0, \end{aligned}$$

where the first inequality is Jensen's inequality applied for function f which is concave for nonnegative domain and $\Delta_{t-1} - 1$ and $\Delta_{t-1} + 1$ are nonnegative. The second inequality comes from the fact that in Decreasing Hedge, $\eta_{t+1} < \eta_t$ and f is an increasing function. □

Observe that by Lemma 16 we can always move $\binom{0}{1}$ to the earlier rounds and $\binom{1}{0}$ to the later rounds. If we apply this swapping rule until for no $(t, t+1)$ the conditions in the lemma hold, the resulting loss sequence will be of the form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$. Note that $a \geq b$ as there is no leader change in the sequence.

F.4 Regret is minimized when both experts have the same cumulative loss

So far we have shown that any loss sequence can be converted to a loss sequence with form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where $a \geq b$ while having the same or even less regret for Decreasing Hedge. First we show that when T is even, the number of zeros in the middle should be even.

Lemma 17. *If a sequence has the form $\ell_1, \dots, \ell_T = \binom{0}{1}^{a-1} \binom{0}{0} \binom{0}{0}^{2k-1} \binom{1}{0}^b$ where $a \geq b$ and T is even, then converting it to a loss sequence $\ell'_1, \dots, \ell'_T = \binom{0}{1}^{a-1} \binom{0}{0} \binom{0}{0}^{2k-1} \binom{1}{0}^b$ decreases the regret.*

Proof. As $T = a + b + 2k - 1$, and T is even, $a + b$ should be odd. Therefore, $a \neq b$. Which implies that $a > b$. Thus $a - 1 \geq b$. Therefore, the best expert in both loss sequences has the same loss:

$$\min_i L'_{i,T} = \min_i L_{i,T} = b.$$

Also, $\hat{\ell}'_t = \hat{\ell}_t$ for all $t < a$ as the loss sequences are the same for all rounds $t < a$.

In round $t = a$, Decreasing Hedge incurs some positive loss in ℓ_1, \dots, ℓ_T while it incurs zero loss in sequence ℓ'_1, \dots, ℓ'_T . In other words, $\hat{\ell}'_a = 0 < \hat{\ell}_a$.

For rounds $t \in \{a + 1, \dots, a + 2k - 1\}$, Decreasing Hedge incurs zero loss for both loss sequences ℓ_1, \dots, ℓ_T and ℓ'_1, \dots, ℓ'_T . In other words, $\hat{\ell}'_t = \hat{\ell}_t = 0$.

Since $\Delta'_{t-1} = L'_{2,t-1} - L'_{1,t-1} = (L_{2,t-1} - 1) - L_{1,t-1} = \Delta_{t-1} - 1$ for any round $t \in \{a + (2k - 1) + 1, \dots, a + (2k - 1) + b\}$,

$$\hat{\ell}'_t = \frac{1}{1 + \exp(-\eta_t \Delta'_{t-1})} = \frac{1}{1 + \exp(-\eta_t (\Delta_{t-1} - 1))} \leq \frac{1}{1 + \exp(-\eta_t \Delta_{t-1})} = \hat{\ell}_t.$$

As a result, $\hat{\ell}'_t \leq \hat{\ell}_t$ for all $1 \leq t \leq T$. Therefore,

$$\begin{aligned} \mathcal{R}'(T) &= \sum_{t=1}^T \hat{\ell}'_t - \min_i L'_{i,T} = \sum_{t=1}^T \hat{\ell}'_t - \min_i L_{i,T} \\ &\leq \sum_{t=1}^T \hat{\ell}_t - \min_i L_{i,T} = \mathcal{R}(T). \end{aligned}$$

□

Now, we only consider loss sequences of form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where c is even. As $a + b + c = T$ and T is even, $a + b = 2K$ is even.

Next, we will show in the following lemma that any loss sequence $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where $a + b = 2K$ and $a \geq b$ can be converted to a loss sequence $\binom{0}{1}^K \binom{0}{0}^c \binom{1}{0}^K$ while having the same or even less regret.

Lemma 18. *Among all loss sequences of form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where $a + b = 2K$ and $a \geq b$, loss sequence $\binom{0}{1}^K \binom{0}{0}^c \binom{1}{0}^K$ has the least regret for Decreasing Hedge.*

Proof. Consider a sequence $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ where $a \neq b$. Therefore, $a > b$ which means that $a \geq b + 1$. If $a = b + 1$, then $a + b = 2b + 1 \neq 2K$. This means that $a \neq b + 1$. Therefore, $a \geq b + 2$.

Now, consider $\ell'_1, \dots, \ell'_T = \binom{0}{1}^{a-1} \binom{1}{0} \binom{0}{0}^c \binom{1}{0}^b$. We know that the best expert in ℓ_1, \dots, ℓ_T incurs $L_{1,T} = b < a = L_{2,T}$ whereas the best expert in ℓ'_1, \dots, ℓ'_T incurs additional unit loss, since $L'_{1,T} = b + 1 \leq a - 1 = L'_{2,T}$. Therefore, $\min_i L'_{i,T} = b + 1 = \min_i L_{i,T} + 1$.

We now show that the cumulative loss of Decreasing Hedge for loss sequence ℓ'_1, \dots, ℓ'_T is at most 1 unit greater than cumulative loss for loss sequence ℓ_1, \dots, ℓ_T . Therefore, $\mathcal{R}'(T) \leq \mathcal{R}(T)$.

For rounds $t < a$, Decreasing Hedge incurs the same loss for both loss sequences, i.e., $\hat{\ell}'_t = \hat{\ell}_t$ for all $t < a$. In round $t = a$, we have $\hat{\ell}'_a = \frac{1}{1 + \exp(-\eta_a(a-1))} \geq 0$ and $\hat{\ell}_a = \frac{1}{1 + \exp(\eta_a(a-1))} \geq 0$. Observe that $\hat{\ell}'_a + \hat{\ell}_a = \frac{1}{1 + \exp(-\eta_a(a-1))} + \frac{1}{1 + \exp(\eta_a(a-1))} = 1$. Therefore,

$$\hat{\ell}'_a - \hat{\ell}_a = (1 - \hat{\ell}_a) - \hat{\ell}_a = 1 - 2\hat{\ell}_a \leq 1.$$

As a result, in round a Decreasing Hedge incurs at most 1 unit of additional loss in loss sequence ℓ'_1, \dots, ℓ'_T as compared to loss sequence ℓ_1, \dots, ℓ_T .

For rounds $t \in \{a + 1, \dots, a + c\}$, Decreasing Hedge incurs zero loss in both loss sequences.

Note that since there is a change from $\binom{0}{1}$ to $\binom{1}{0}$ in round a , for any round $t \in \{a + 1, \dots, a + c + b\}$, we have $\Delta'_{t-1} = L'_{2,t-1} - L'_{1,t-1} = \Delta_{t-1} - 2$. Therefore, for any round t in this period, Decreasing Hedge incurs less loss in ℓ'_1, \dots, ℓ'_T than ℓ_1, \dots, ℓ_T :

$$\hat{\ell}'_t = \frac{1}{1 + \exp(-\eta_t \Delta'_{t-1})} = \frac{1}{1 + \exp(-\eta_t (\Delta_{t-1} - 2))} \leq \frac{1}{1 + \exp(-\eta_t \Delta_{t-1})} = \hat{\ell}_t.$$

As a result, the cumulative loss of Decreasing Hedge in loss sequence ℓ'_1, \dots, ℓ'_T is at most 1 unit greater than Decreasing Hedge's cumulative loss in loss sequence ℓ_1, \dots, ℓ_T , i.e.,

$$\sum_{t=1}^T \hat{\ell}'_t \leq 1 + \sum_{t=1}^T \hat{\ell}_t.$$

Therefore,

$$\begin{aligned} \mathcal{R}'(T) &= \sum_{t=1}^T \hat{\ell}'_t - \min_i L'_{i,T} \\ &= \sum_{t=1}^T \hat{\ell}'_t - (\min_i L_{i,T} + 1) \\ &\leq (1 + \sum_{t=1}^T \hat{\ell}_t) - (\min_i L_{i,T} + 1) \\ &= \sum_{t=1}^T \hat{\ell}_t - \min_i L_{i,T} = \mathcal{R}(T). \end{aligned}$$

Now by Lemma 16, we can move the loss vector $(1, 0)$ in ℓ'_1, \dots, ℓ'_t in round a to the right without increasing the regret. Therefore, the regret for $\ell''_1, \dots, \ell''_t = \binom{0}{1}^{a-1} \binom{0}{0}^c \binom{1}{0}^{b+1}$ is no greater than the regret for ℓ_1, \dots, ℓ_T . This means that when $a \geq b + 1$, converting loss sequence $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$ to $\binom{0}{1}^{a-1} \binom{0}{0}^c \binom{1}{0}^{b+1}$ does not increase the regret. We can apply this rule many times to equalize the numbers a and b . As $a + b = 2K$, this implies $a = b = K$ and the lemma follows. \square

F.5 The optimum number of $(0, 0)$ loss vectors in the middle

So far, we know that the best-case sequence has the form $\binom{0}{1}^K \binom{0}{0}^c \binom{1}{0}^K$ where c is even. Is the existence of $\binom{0}{0}^c$ in the middle necessary to have the loss sequence with the least possible regret?

In order to examine this, we consider a simple modification to the loss sequence. Consider a loss sequence in which one $(0, 0)$ is replaced with $(0, 1)$ and another $(0, 0)$ is replaced with $(1, 0)$. Then by the swapping lemma (Lemma 16), we can move the $(1, 0)$ to the later rounds and the $(0, 1)$ to the earlier rounds without increasing the regret. If we denote the original sequence as $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$, the modified loss sequence becomes $\ell'_1, \dots, \ell'_T = \binom{0}{1}^{a+1} \binom{0}{0}^{c-2} \binom{1}{0}^{b+1}$.

Lemma 19. *Consider a loss sequence $\ell_1, \dots, \ell_T = \binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$, where c is even and $a \geq 4$. Then modifying loss sequence to $\ell'_1, \dots, \ell'_T = \binom{0}{1}^{a+1} \binom{0}{0}^{c-2} \binom{1}{0}^{b+1}$ does not increase the regret if $c \geq 4$ and does increase the regret if $c \leq 3$.*

Proof. First of all, observe that both experts incur one additional unit loss in loss sequence ℓ'_1, \dots, ℓ'_T in comparison to ℓ_1, \dots, ℓ_T . Therefore,

$$\max_i L'_{i,T} = \max_i L_{i,T} + 1.$$

Note that in any round s except rounds $t = a + 1$ and $t' = a + c$, we have $\hat{\ell}_s = \hat{\ell}'_s$.

On the other hand, for loss ℓ'_1, \dots, ℓ'_T , in round $t = a + 1$ and $t' = a + c$, Decreasing Hedge incurs loss

$$\begin{aligned}\hat{\ell}'_{a+1} &= \frac{1}{1 + \exp(-\eta_t(a))}, \\ \hat{\ell}'_{a+c} &= \frac{1}{1 + \exp(\eta_{t'}(a+1))}.\end{aligned}$$

while for ℓ_1, \dots, ℓ_T , the algorithm does not incur any loss in those rounds (i.e. $\hat{\ell}_{a+1} + \hat{\ell}_{a+c} = 0$). Therefore, the difference in regret would be

$$\mathcal{R}'(T) - \mathcal{R}(T) = \frac{1}{1 + \exp(\eta_t(a))} + \frac{1}{1 + \exp(-\eta_{t'}(a+1))} - 1.$$

Let $f(x) := \frac{1}{1 + \exp(-x)}$. Observe that $f(x) + f(-x) = 1$. Therefore,

$$\begin{aligned}\mathcal{R}'(T) - \mathcal{R}(T) &= \frac{1}{1 + \exp(\eta_t(a))} + \frac{1}{1 + \exp(-\eta_{t'}(a+1))} - 1 \\ &= f(-\eta_t(a)) + f(\eta_{t'}(a+1)) - 1 \\ &= f(-\eta_t(a)) + f(\eta_{t'}(a+1)) - (f(\eta_t(a)) + f(-\eta_t(a))) \\ &= f(\eta_{t'}(a+1)) - f(\eta_t(a)).\end{aligned}$$

As f is an increasing function, $f(\eta_{t'}(a+1)) - f(\eta_t(a)) \leq 0$ if and only if

$$\eta_{t'}(a+1) \leq \eta_t(a). \quad (16)$$

Therefore, if inequality (16) holds, then the replacement does not increase the regret. Substituting the learning rate, inequality (16) becomes

$$\frac{a+1}{\sqrt{a+c}} \leq \frac{a}{\sqrt{a+1}},$$

which is equivalent to

$$c \geq \frac{3a^2 + 3a + 1}{a^2} = 3 + \frac{3}{a} + \frac{1}{a^2}. \quad (17)$$

Note that in case $a \geq 4$ and $c \geq 4$, the inequality $\frac{3}{a} + \frac{1}{a^2} < 1$ holds; therefore, (17) holds which means that the regret does not increase.

In the case $c \leq 3$, then (17) never holds which means that this replacement increases the regret. \square

Now, let us consider a loss sequence ℓ_1, \dots, ℓ_T of length $T = 2K \geq 16$ of the form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^b$. If $a > b$, then we can and will use Lemma 18 to convert this sequence into a sequence for which $a = b$. The resulting sequence has the form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^a$. We will now show how this sequence can be modified, without increasing the regret, to a sequence for which $c = 2$ (while maintaining the invariant $a = b$). There are two cases.

In the first case, the sequence is of the form $\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^a$ where $a \geq 4$ and $c \geq 4$ (recall that c must be even, and note that if $c = 2$, we are done). We may then apply Lemma 19 to decrease the number of $(0, 0)$ vectors from the middle by two while maintaining the invariant $a = b$. As we have shown in Lemma 17 that c is even, the last time where we can apply this lemma is when $c = 4$ and the resulting sequence is of the form $\binom{0}{1}^{a+\frac{c-2}{2}} \binom{0}{0}^2 \binom{1}{0}^{\frac{c-2}{2}+a}$

In the second case, we have $a \in \{1, 2, 3\}$. Since $a + b + c = T \geq 16$ and as we may assume that $a = b$, it follows that $c \geq 10$. In this case, inequality (17) holds, which implies that modifying the loss sequence to $\ell'_1, \dots, \ell'_T = \binom{0}{1}^{a+1} \binom{0}{0}^{c-2} \binom{1}{0}^{a+1}$ does not increase the regret. After applying this modification at most three times, the resulting loss sequence will be of the form $\ell''_1, \dots, \ell''_T =$

$\binom{0}{1}^a \binom{0}{0}^c \binom{1}{0}^a$ where $a \geq 4$. We are now in the first case and so again can arrive at a sequence with two $(0, 0)$ vectors in the middle.

Observe that $\binom{0}{1}^{a+\frac{c-2}{2}} \binom{0}{0}^2 \binom{1}{0}^{\frac{c-2}{2}+a}$ is of the form $\binom{0}{1}^K \binom{0}{0}^2 \binom{1}{0}^K$. This form coincides with the canonical best-case sequence we mentioned earlier.

F.6 Bounding the regret

So far, we have shown that among all loss sequence of size $T = 2K + 2 \geq 16$, the loss sequence $\left(\binom{0}{1}^K \binom{0}{0}^2 \binom{1}{0}^K\right)$ yields the minimum regret for Decreasing Hedge. We will now give upper and lower bounds on the regret for this sequence. The algorithm's loss for this sequence is

$$\begin{aligned} \hat{L}_T &= \sum_{t=1}^T \hat{\ell}_t = \hat{\ell}_1 + \underbrace{\sum_{t=2}^K \hat{\ell}_t}_{\binom{0}{1}^K} + \underbrace{\hat{\ell}_{K+1} + \hat{\ell}_{K+2}}_{\binom{0}{0}^2} + \hat{\ell}_{K+3} + \underbrace{\sum_{t=2}^K \hat{\ell}_{2K+4-t}}_{\binom{1}{0}^K} \\ &= \hat{\ell}_1 + \hat{\ell}_{K+3} + \sum_{t=2}^K \hat{\ell}_t + \hat{\ell}_{2K+4-t} \\ &= \frac{1}{2} + \frac{1}{1 + \exp(-\eta_{K+3}(K))} + \sum_{t=2}^K \left(\frac{1}{1 + \exp(\eta_t(t-1))} + \frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} \right). \end{aligned}$$

Therefore,

$$\begin{aligned} \mathcal{R}(T) &= \hat{L}_T - \min_i L_{i,T} \\ &= \hat{L}_T - K \\ &= \sum_{t=2}^K \left(\frac{1}{1 + \exp(\eta_t(t-1))} + \frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} \right) \\ &\quad + \frac{1}{2} + \frac{1}{1 + \exp(-\eta_{K+3}(K))} - K \\ &= \sum_{t=2}^K \left(\frac{1}{1 + \exp(\eta_t(t-1))} + \frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} - 1 \right) + \sum_{t=2}^K 1 \\ &\quad + \frac{1}{2} + \frac{1}{1 + \exp(-\eta_{K+3}(K))} - K \\ &= \sum_{t=2}^K \left(\frac{1}{1 + \exp(\eta_t(t-1))} + \frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} - 1 \right) \\ &\quad + \frac{1}{2} + \frac{1}{1 + \exp(-\eta_{K+3}(K))} - 1. \end{aligned}$$

As we know that $\frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} + \frac{1}{1 + \exp(+\eta_{2K+4-t}(t-1))} = 1$, it follows that

$$\frac{1}{1 + \exp(\eta_t(t-1))} + \frac{1}{1 + \exp(-\eta_{2K+4-t}(t-1))} - 1 = \frac{1}{1 + \exp(\eta_t(t-1))} - \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))}.$$

Hence, we can rewrite the regret as

$$\mathcal{R}(T) = \underbrace{\sum_{t=2}^K \frac{1}{1 + \exp(\eta_t(t-1))}}_{\text{A = First term}} - \underbrace{\sum_{t=2}^K \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))}}_{\text{B = Second term}} + \underbrace{\frac{1}{2} + \frac{1}{1 + \exp(-\eta_{K+3}(K))} - 1}_{\text{C = Third term}}.$$

We will give upper and lower bounds for each term separately. Note that $K = \frac{T-2}{2}$. In the development of the bounds below, we assume that T is sufficiently large. It suffices to have $T \geq 16$.

Now, for the A (First term), we have

$$\begin{aligned}
0 \leq A &= \sum_{t=2}^K \frac{1}{1 + \exp(\eta_t(t-1))} = \sum_{t=2}^K \frac{1}{1 + \exp(\sqrt{t} - \frac{1}{\sqrt{t}})} \\
&\leq \sum_{t=2}^K \frac{1}{1 + \exp(\frac{1}{2}\sqrt{t})} \leq \sum_{t=2}^K \frac{1}{\exp(\frac{1}{2}\sqrt{t})} \\
&\leq \int_1^K \frac{1}{\exp(\frac{1}{2}\sqrt{x})} dx \\
&= (-4\sqrt{K} - 8) \exp(-0.5\sqrt{K}) - (-4\sqrt{1} - 8) \exp(-0.5\sqrt{1}) \\
&\leq (4\sqrt{1} + 8) \exp(-0.5\sqrt{1}) \\
&= \frac{12}{\sqrt{e}}.
\end{aligned}$$

To upper bound B (Second term), we divide the T rounds into $\lceil \sqrt{T} \rceil$ intervals. We then show that sum of the contributions of the intervals is upper bounded by sum of a geometric series that converge to a constant value.

The second term can be written as the sum of $\lceil \sqrt{T} \rceil$ intervals as

$$\begin{aligned}
\sum_{t=2}^K \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))} &= \sum_{t=2}^{\frac{T}{2}-1} \frac{1}{1 + \exp(\eta_{T+2-t}(t-1))} \\
&\leq \sum_{t=1}^T \frac{1}{1 + \exp(\frac{t-1}{\sqrt{T+2-t}})} \\
&= \sum_{k=0}^{\lceil \sqrt{T} \rceil - 1} \sum_{t=k\lceil \sqrt{T} \rceil + 1}^{k\lceil \sqrt{T} \rceil + \lceil \sqrt{T} \rceil} \frac{1}{1 + \exp(\frac{t-1}{\sqrt{T+2-t}})} \\
&\leq \sum_{k=0}^{\lceil \sqrt{T} \rceil - 1} \sum_{t=k\lceil \sqrt{T} \rceil + 1}^{k\lceil \sqrt{T} \rceil + \lceil \sqrt{T} \rceil} \frac{1}{1 + \exp(\frac{k\sqrt{T}}{\sqrt{T+2}})} \\
&\leq \lceil \sqrt{T} \rceil \sum_{k=0}^{\lceil \sqrt{T} \rceil - 1} \frac{1}{1 + \exp(\frac{k\sqrt{T}}{2+\sqrt{T}})}.
\end{aligned}$$

The term $\sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{1 + \exp(\frac{k\sqrt{T}}{2+\sqrt{T}})}$ can be upper bounded as

$$\begin{aligned}
\sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{1 + \exp(\frac{k\sqrt{T}}{2+\sqrt{T}})} &= \sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{1 + \exp(k - \frac{2k}{2+\sqrt{T}})} \leq \sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{1 + \exp(k-2)} \\
&\leq \sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{\exp(k-2)} \\
&= e^2 \sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \left(\frac{1}{e}\right)^k \\
&\leq e^2 \left(\frac{1 - \frac{1}{e}^{\lceil\sqrt{T}\rceil}}{1 - \frac{1}{e}}\right) \\
&\leq \frac{e^2}{1 - \frac{1}{e}}.
\end{aligned}$$

As a result,

$$\sum_{t=2}^K \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))} \leq \lceil\sqrt{T}\rceil \cdot \sum_{k=0}^{\lceil\sqrt{T}\rceil-1} \frac{1}{1 + \exp(\frac{k\sqrt{T}}{2+\sqrt{T}})} \leq \frac{e^2}{1 - \frac{1}{e}} \cdot \lceil\sqrt{T}\rceil.$$

Moreover, B (Second term) can be lower bounded as follows:

$$\begin{aligned}
\sum_{t=2}^K \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))} &= \sum_{t=2}^K \frac{1}{1 + \exp(\eta_{T+2-t}(t-1))} \\
&= \sum_{t=2}^{\frac{T}{2}-1} \frac{1}{1 + \exp(\frac{t-1}{\sqrt{T+2-t}})} \\
&\geq \sum_{t=2}^{\lceil\sqrt{T}\rceil+1} \frac{1}{1 + \exp(\frac{t-1}{\sqrt{T+2-t}})} \\
&\geq \sum_{t=2}^{\lceil\sqrt{T}\rceil+1} \frac{1}{1 + \exp\left(\frac{\lceil\sqrt{T}\rceil}{\sqrt{T+2-t}}\right)} \\
&\geq \sum_{t=2}^{\lceil\sqrt{T}\rceil+1} \frac{1}{1 + \exp\left(\frac{\sqrt{T}}{\sqrt{T+1-\lceil\sqrt{T}\rceil}}\right)} \\
&\geq \sum_{t=2}^{\lceil\sqrt{T}\rceil+1} \frac{1}{1 + \exp\left(\frac{\lceil\sqrt{T}\rceil}{\sqrt{\frac{1}{2}\lceil\sqrt{T}\rceil}}\right)} \\
&= \sum_{t=2}^{\lceil\sqrt{T}\rceil+1} \frac{1}{1 + \exp(\sqrt{2})} \\
&\geq \frac{\sqrt{T}}{1 + \exp(\sqrt{2})}.
\end{aligned}$$

Therefore,

$$\frac{1}{1 + e\sqrt{2}}\sqrt{T} \leq B = \sum_{t=2}^K \frac{1}{1 + \exp(\eta_{2K+4-t}(t-1))} \leq \frac{e^2}{(1 - \frac{1}{e})} \lceil \sqrt{T} \rceil.$$

For the C (Third term), as we know $0 \leq \hat{\ell}_{K+3} = \frac{1}{1 + \exp(-\eta_{K+3}(K))} \leq 1$, it follows that

$$-\frac{1}{2} \leq C \leq \frac{1}{2}.$$

As a result, we have

$$-\frac{e^2}{(1 - \frac{1}{e})} \cdot \lceil \sqrt{T} \rceil - \frac{1}{2} \leq A - B + C \leq \frac{12}{\sqrt{e}} + \frac{1}{2} - \frac{1}{1 + e\sqrt{2}}\sqrt{T},$$

which for suitably defined constants is

$$-c_1\sqrt{T} + c_2 \leq \mathcal{R}(T) \leq c_3 - c_4\sqrt{T}.$$

Therefore,

$$\mathcal{R}(T) = -\Theta(\sqrt{T}).$$