Optimistic and Adaptive Lagrangian Hedging

Ryan D'Orazio,[1]
Ruitong Huang [2]
[1] Université de Montréal and MILA *
[2] Borealis AI
ryan.dorazio@mila.quebec, rtonghuang@gmail.com

Abstract

In online learning an algorithm plays against an environment with losses possibly picked by an adversary at each round. The generality of this framework includes problems that are not adversarial, for example offline optimization, or saddle point problems (i.e. min max optimization). However, online algorithms are typically not designed to leverage additional structure present in non-adversarial problems. Recently, slight modifications to well-known online algorithms such as optimism and adaptive step sizes have been used in several domains to accelerate online learning – recovering optimal rates in offline smooth optimization, and accelerating convergence to saddle points or social welfare in smooth games. In this work we introduce optimism and adaptive stepsizes to Lagrangian hedging, a class of online algorithms that includes regret-matching, and hedge (i.e. multiplicative weights). Our results include: a general general regret bound; a path length regret bound for a fixed smooth loss, applicable to an optimistic variant of regret-matching and regret-matching+; optimistic regret bounds for Φ regret, a framework that includes external, internal, and swap regret; and optimistic bounds for a family of algorithms that includes regret-matching+ as a special case.

Introduction

Online optimization is a general framework applicable to various problems such as offline optimization, and finding equilibria in games. Typical algorithms only use first-order information (i.e. a subgradient or gradient), such as online mirror descent (MD) (Nemirovsky and Yudin 1983; Warmuth and Jagota 1997; Beck and Teboulle 2003) which generalizes projected gradient descent (see for example (Orabona 2019)), and follow the regularized leader (FTRL) (Shalev-Shwartz and Singer 2006; Abernethy, Hazan, and Rakhlin 2009; Nesterov 2009).[1]

In general, online learning is adversarial, losses may change almost arbitrarily from one time step to the next. However, most problems of interest including offline optimization, and saddle point optimization can be "predictable." That is, the sequence of losses induced by running an online algorithm in these settings has specific structure and can be predictable under the right conditions, like smoothness (i.e. Lipschitz continuous gradient). When losses are predictable a powerful framework is optimistic online learning (Rakhlin and Sridharan 2013a,b; Chiang et al. 2012). Where algorithms are modified to incorporate a *guess* of the next loss, m_t, into their update.

Combining optimism with MD and FTRL yields their optimistic counterparts, optimistic mirror descent (OMD), and optimistic follow the regularized leader (OFTRL), respectively. OMD and OFTRL both provide tangible benefits when problems are not quite adversarial. For example, faster convergence to a saddle point on the average (Rakhlin and Sridharan 2013b; Syrgkanis et al. 2015; Farina et al. 2019; Farina, Kroer, and Sandholm 2019); faster convergence to optimal social wellfare in n-player games (Syrgkanis et al. 2015); last iterate convergence in games (Daskalakis and Panageas 2018); acceleration in offline or online optimization (Cutkosky 2019; Mohri and Yang 2016; Joulani et al. 2020; Joulani, György, and Szepesvári 2017).

Interestingly, much of the analysis of optimistic algorithms is black-box. For example, most of the results rely on regret bounds being of a particular form, which is satisfied by both OMD and OFTRL. Naturally, one may ask what other classes of algorithms can be combined with optimism to achieve faster rates in predictable problems?

In this paper we extend the idea of optimism to the class of algorithms known as Lagrangian hedging (Gordon 2007). Unfortunately, the regret bounds attained are not consistent with OMD and OFTRL, therefore, immediate theoretical acceleration via the previously mentioned works is not attained. However, our analysis provides interesting regret bounds that should be small given a "good" guess. And in the case for a smooth fixed loss we show a path length bound for the regret. This result, for example, is applicable to an optimistic varaint of the well-known regret-matching algorithm when used to train a linear regressor with L_1 regularization and the least-squares loss (Schuurmans and Zinkevich 2016).

Additionally, our analysis extends beyond the typical regret objectives of MD and FTRL, and includes regret bounds for internal and swap regret (Cesa-Bianchi and Lugosi 2006). To the best of our knowledge, our results provide

*This work is partially done when Ryan D'Orazio was working at Borealis AI.

[1]See Orabona for an excellent historical overview of MD and FTRL.

the first optimistic and adaptive algorithms for minimizing internal regret, with possible applications including finding correlated equilibria in n-player general sum games (Cesa-Bianchi and Lugosi 2006).

Background

Online Linear Optimization

In online convex optimization an algorithm $\mathcal{A}$ interacts with an environment for T rounds (Zinkevich 2003). In each round t, $\mathcal{A}$ selects an iterate x_t within some convex compact set $\mathcal{X}$, afterwhich a convex loss function $\ell_t : \mathcal{X} \to \mathbb{R}$ chosen by the environment is revealed. Furthermore, $\mathcal{A}$ is only allowed to use information from previous rounds. The performance of $\mathcal{A}$ after T rounds is measured by its regret

$$R^T_{\mathcal{X}} = \sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x). \quad (1)$$

The objective is to ensure sublinear regret, $R^T_{\mathcal{X}} \in o(T)$, e.g $R^T_{\mathcal{X}} \in O(\sqrt{T})$. In the most general of settings, no assumptions are made on the sequence of losses $\{\ell_t\}_{t \le T}$, they may be chosen by an adversary with knowledge of $\mathcal{A}$.

If each loss ℓ_t is subdifferentiable at x_t, then there exists a vector $\partial \ell_t(x_t)$ (a subgradient) such that

$$\ell_t(x) \ge \ell_t(x_t) + \langle \partial \ell_t(x_t), x - x_t \rangle \quad \forall x \in \mathcal{X}.$$

Provided $\mathcal{A}$ has access to a subgradient, it is enough to design algorithms for linear losses. The original regret $R^T_{\mathcal{X}}$ is upper bounded by the regret with respect to the linear losses $\{\tilde{\ell}_t\}_{t \le T}$, where $\tilde{\ell}_t(x) = \langle \partial \ell_t(x_t), x \rangle$. For the remainder of the paper we assume linear losses unless specified otherwise.

Lagrangian Hedging

Lagrangian hedging defines a class of algorithms for online linear optimization (Gordon 2007). The class generalizes potential based methods introduced by Cesa-Bianchi and Lugosi for learning with expert advice (Cesa-Bianchi and Lugosi 2003),[2] and includes the well-known Hedge aglorithm (also known as multiplicative weights) (Freund and Schapire 1997), and regret-matching (Hart and Mas-Colell 2000) .

At each round t a Lagrangian hedging algorithm maintains a regret vector

$$s_{1:t-1} = s_{1:t-2} + \langle \ell_{t-1}, x_{t-1} \rangle u - \ell_{t-1},$$

with the initial vector initialized as $s_{1:0} = s_0 = 0$. The change in the regret vector is denoted as $s_t = \langle \ell_t, x_t \rangle u - \ell_t$, $s_{1:t} = \sum_{k=1}^{t} s_k$. u is a vector such that $\langle u, x \rangle = 1$ for any $x \in \mathcal{X}$. As mentioned by Gordon (Gordon 2007), if no such u can be found then we may append an extra 1 for each $x \in \mathcal{X}$. Then we can take u to be the vector of zeros except for a 1 coinciding with the new dimension added to x, as well as append a 0 to each loss. $s_{1:t}$ is referred to as the regret vector because it tracks how well an algorithm has done so far,

$$\sum_{t=1}^{T} \langle \ell_t, x_t \rangle - \sum_{t=1}^{T} \langle \ell_t, x \rangle = \langle s_{1:T}, x \rangle \quad \forall x \in \mathcal{X}.$$

The regret is then simply $R^T_{\mathcal{X}} = \max_{x \in \mathcal{X}} \langle s_{1:T}, x \rangle$.

Instead of explicitly ensuring the regret to be small, Lagrangian hedging ensures $s_{1:T}$ is not too far from a safe set $\mathcal{S}$. The safe set is defined to be the polar cone to $\mathcal{X}$,

$$\mathcal{S} = \{ s : \forall x \in \mathcal{X} \ \langle s, x \rangle \le 0 \}.$$

Forcing $s_{1:T}$ to be in $\mathcal{S}$ may not be possible, as it would guarantee $R^T_{\mathcal{X}} \le 0$ when it is possible to encounter an adversary that guarantees $\Omega(\sqrt{T})$ regret (Orabona 2019; Hazan et al. 2016). However,

$$\begin{aligned} R^T_{\mathcal{X}} &= \max_{x \in \mathcal{X}} \langle s_T, x \rangle \le \max_{x \in \mathcal{X}} \langle s_T - s, x \rangle \quad \forall s \in \mathcal{S} \\ &\le \inf_{s \in \mathcal{S}} \| s_T - s \| \max_{x \in \mathcal{X}} \| x \|_* . \end{aligned} \quad (2)$$

Therefore, if the distance of s_T to the set $\mathcal{S}$ grows at a sublinear rate then the regret will be sublinear, since by assumption the set $\mathcal{X}$ is bounded, $\|x\|_* \le D$. The norm $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$, defined as $\|x\|_* = \sup\{\langle x, y \rangle | \ \|y\| \le 1\}$.

Additionally, we assume the change in the regret vector is bounded in norm, $\|s_t\|^2 \le C$. This assumption is similar to assuming bounded linear functions $\|\ell_t\| \le C$, or in the convex (possibly non-linear) case $\|\partial \ell_t\| \le C$ (i.e convex Lipschitz continuous functions).

The distance of $s_{1:t}$ to $\mathcal{S}$ is then controlled via a smooth potential function F, with the following conditions:

$$F(s) \le 0 \quad \forall s \in \mathcal{S} \quad (3)$$

$$F(x+y) \le F(x) + \langle \partial F(x), y \rangle + \frac{L}{2} \|y\|^2 \quad (4)$$

$$(F(s) + A)^+ \ge \inf_{s' \in S} B \|s - s'\|^p , \quad (5)$$

for constants $L, B > 0$, $A \ge 0$, and $1 \le p \le 2$. $\partial F(x)$ is a subgradient of F at x. $(x)^+$ refers to the Relu operation which sets all negative values in the vector to 0. In addition to the above conditions we will also assume that F is convex, and therefore differentiable with $\partial F(x) = \nabla F(x)$, the gradient of F at x. When F is differentiable condition (4) is equivalent to Lipschitz continuity of the gradient, $\|\nabla F(x) - \nabla F(y)\|_* \le L \|x - y\|$ (Nesterov 2018)[Theorem 2.1.5].

Once an appropriate potential function is chosen, a Lagrangian hedging algorithm ensures $F(s_{1:T}) \in O(T)$ by picking an iterate at each round t such that

$$\langle \nabla F(s_{1:t-1}), s_t \rangle \le 0, \quad (6)$$

for any possible s_t (s_t can change depending on the loss ℓ_t picked by the environment). The above inequality is also known as the Blackwell condition, often used in potential based expert algorithms (Cesa-Bianchi and Lugosi 2006),

[2] Learning with expert advice resembles online linear optimization where the decision set $\mathcal{X}$ is an n-dimensional simplex Δ^n, which is interpreted as the set of distributions over n experts.

and, as shown by Gordon, is guaranteed if the iterate at time t is chosen by the following rule

$$x_t = \begin{cases} \frac{\nabla F(s_{1:t-1})}{\langle \nabla F(s_{1:t-1}), u\rangle} & \text{if } \langle \nabla F(s_{1:t-1}), u\rangle > 0 \\ \text{arbitrary } x \in \mathcal{X} & o.w. \end{cases} \tag{7}$$

Gordon also showed that procedure (7) always yields a feasible iterate $x_t \in \mathcal{X}$.

Equipped with the Blackwell condition and the smoothness of F (condition 4), the growth of $F(s_{1:t})$ is easily bounded by

$$\begin{aligned} F(s_{1:t}) = F(s_{1:t-1} + s_t) &\leq F(s_{1:t-1}) + \frac{L}{2} \|s_t\|^2 \\ &\leq F(s_{1:t-1}) + \frac{LC}{2}. \end{aligned}$$

Summing across time and with $s_{1:0} = s_0 = 0$,

$$F(s_{1:t}) \leq F(0) + \frac{LCt}{2} \leq \frac{LCt}{2},$$

since $0 \in \mathcal{S}$. With a linear bound on $F(s_{1:t})$ a regret bound follows immediately by condition (5) and (2),

$$R^T_{\mathcal{X}} \leq D \left(\frac{LCT + 2A}{2B} \right)^{1/p}. \tag{8}$$

If $p = 1$ then the regret is linear, however, as mentioned by Gordon, a stepsize can be used to achieve sublinear regret. We can define a new potential function with stepsize η, $F_\eta(s) = F(\eta s)$. The smoothness condition becomes

$$\begin{aligned} F_\eta(x + y) = F(\eta(x + y)) \leq & \\ & F(\eta x) + \eta \langle \nabla F(\eta x), y\rangle + \frac{\eta^2 L}{2} \|y\|^2, \\ & F_\eta(x) + \langle \nabla F_\eta(x), y\rangle + \frac{\eta^2 L}{2} \|y\|^2. \end{aligned}$$

F_η is therefore a valid potential function, with condition (4) now being

$$(F_\eta(s) + A)^+ \geq \inf_{s' \in S} \eta^p B \|s - s'\|^p.^3$$

Following the same arguments as Gordon ((Gordon 2007)[Theorem 3]), the regret becomes

$$R^T_{\mathcal{X}} \leq D \left(\frac{\eta^2 LCT + 2A}{2B\eta^p} \right)^{1/p}. \tag{9}$$

When $p = 1$ a stepsize $\eta \in O(\frac{1}{\sqrt{T}})$ achieves a regret bound of $O(\sqrt{T})$, similar to the standard results in MD and FTRL analysis (see (Orabona 2019) for example). Despite this guarantee on regret, this bound does not hold uniformly over time, one must have knowledge of the horizon T to select a stepsize. However, the standard doubling trick can be applied to achieve a regret guarantee for all time steps, requiring algorithm resets after exponentially growing time intervals (Cesa-Bianchi and Lugosi 2006).

[3]See Gordon for details.

In MD, FTRL, and the potential based approaches from expert problems, however, a stepsize schedule of $\eta_t \in O(\frac{1}{\sqrt{t}}))$ is enough to achieve a $O(\sqrt{T})$ regret bound that holds uniformly over time (applies to any time horizon). Given that Lagrangian hedging generalizes potential based methods, a similar result likely should hold. Indeed we show with the help of the following simple yet important lemma, that the same learning rate schedule would suffice for Lagrangian Hedging algorithms with potential functions that need a learning rate (i.e $p = 1$). This result is interesting as it makes no additional assumptions on the potential function; whereas, for example when viewing multiplicative weights as a potential based method and a specific instance of Lagrangian hedging, inequalities particular to the algorithm (a specific potential function) are used to derive the regret bounds that hold uniformly over time (Cesa-Bianchi and Lugosi 2006).

First we extend the Lagrangian hedging framework with an arbitrary sequence of stepsizes $\{\eta_t\}_{t \leq T}$, where the potential function $F(\eta_t s)$ is used at round t to construct the iterate x_t,

$$x_t = \begin{cases} \frac{\nabla F(\eta_t s_{1:t-1})}{\langle \nabla F(\eta_t s_{1:t-1}), u\rangle} & \text{if } \langle \nabla F(\eta_t s_{1:t-1}), u\rangle > 0 \\ \text{arbitrary } x \in \mathcal{X} & o.w. \end{cases} \tag{10}$$

Lemma 1. *Assume F is a convex function satisfying condition (3), consider step sizes $0 < \eta_t \leq \eta_{t-1}$, then*

$$F(\eta_t s) \leq \frac{\eta_t}{\eta_{t-1}} F(\eta_{t-1} s).$$

Proof.

$$\begin{aligned} F(\eta_t s) &= F\left(\frac{\eta_t}{\eta_{t-1}} \eta_{t-1} s + 0\right) = F\left(\frac{\eta_t}{\eta_{t-1}} \eta_{t-1} s + \left(1 - \frac{\eta_t}{\eta_{t-1}}\right) 0\right) \\ &\leq \frac{\eta_t}{\eta_{t-1}} F(\eta_{t-1} s) + (1 - \frac{\eta_t}{\eta_{t-1}}) F(0) \\ &\leq \frac{\eta_t}{\eta_{t-1}} F(\eta_{t-1} s), \quad \text{since } 0 \in \mathcal{S}. \end{aligned}$$

□

Coupling the above lemma with the algorithm (10) and the Blackwell condition yields a bound on the growth of $F(\eta_t s_t)$ and therefore a regret bound. Such a bound is a special case of optimistic Lagrangian hedging when the prediction is 0, and so we defer the presentation to the next section.

Adaptivity and Optimism in Lagrangian Hedging

In this section we present the optimistic Lagrangian hedging algorithm along with adaptive stepsizes and the regret guarantees. Optimistic Lagangian hedging leverages a prediction m_t at round t to construct the iterate x_t. In the optimistic and adaptive variants of MD and FTRL, one hopes to have $m_t \approx \ell_t$ since the regret bounds attained are usually of the form $O\left(\sqrt{\sum_{t=1}^T \|m_t - \ell_t\|^2}\right)$, with adaptive stepsizes (in the case of MD) similar to

$$\eta_t = \frac{1}{\sqrt{\sum_{s=1}^{t-1} \|m_s - \ell_s\|^2}}. \tag{11}$$

In optimistic Lagrangian hedging we hope the prediction m_t to be a good predictor of the change in the regret vector $m_t \approx s_t$, with the provable regret bound of $O\left(\sqrt{\sum_{t=1}^{T} \|m_t - s_t\|^2}\right)$. Interestingly for the case of $p = 2$ no adaptive step size is needed!

General Optimistic Bound

Given a prediction m_t we define optimistic Lagrangian hedging with stepsizes η_t as the following rule

$$x_t = \begin{cases} \frac{\nabla F(\eta_t(s_{1:t-1}+m_t))}{\langle \nabla F(\eta_t(s_{1:t-1}+m_t)), u\rangle} & \text{if } \langle \nabla F(\eta_t(s_{1:t-1}+m_t)), u\rangle > 0 \\ \text{arbitrary } x \in \mathcal{X} & o.w. \end{cases} \tag{12}$$

Optimistic Lagrangian hedging then guarantees the general upper bound on the growth of the potential function.

Theorem 1. *An optimistic Lagrangian hedging algorithm with a convex potential function F satisfying conditions (3-4) and positive decreasing stepsizes $0 < \eta_t \le \eta_{t-1}$, ensures*

$$F(\eta_T s_{1:T}) \le \frac{L}{2} \sum_{t=1}^{T} \eta_T \eta_t \|s_t - m_t\|^2.$$

Proof. From the same arguments as Gordon, we have the following Blackwell condition

$$\langle \nabla F(\eta_t(s_{1:t-1} + m_t)), s_t)\rangle \le 0.$$

By the smoothness of F we have

$$\begin{aligned} &F(\eta_t s_{1:t}) = F(\eta_t(s_{t-1} + m_t + s_t - m_t)) \le \\ &F(\eta_t(s_{1:t-1} + m_t)) \\ &+ \langle \nabla F(\eta_t(s_{1:t-1} + m_t)), \eta_t(s_t - m_t)\rangle \\ &+ \eta_t^2 \frac{L}{2} \|s_t - m_t\|^2 \\ &= F(\eta_t(s_{1:t-1} + m_t)) - F(\eta_t(s_{1:t-1})) + F(\eta_t(s_{1:t-1})) \\ &+ \langle \nabla F(\eta_t(s_{1:t-1} + m_t)), -\eta_t m_t\rangle + \eta_t^2 \frac{L}{2} \|s_t - m_t\|^2 \\ &\le F(\eta_t(s_{1:t-1})) + \eta_t^2 \frac{L}{2} \|s_t - m_t\|^2 \text{ (by convexity)}^4 \\ &\le \frac{\eta_t}{\eta_{t-1}} F(\eta_{t-1} s_{1:t-1}) + \eta_t^2 \frac{L}{2} \|s_t - m_t\|^2 \text{ (by Lemma 1).} \end{aligned}$$

We now proceed by induction. Observe that $F(\eta_0 s_0) = F(0) \le 0$ by assumption. So, for any $\eta_0 \le \eta_1$,[5]

$$F(\eta_1 s_{1:1}) \le \frac{\eta_1}{\eta_0} F(\eta_0 s_0) + \eta_1^2 \frac{L}{2} \|s_1 - m_1\|^2 \le \eta_1^2 \frac{L}{2} \|s_1 - m_1\|^2.$$

Assume that

$$F(\eta_{t-1} s_{1:t-1}) \le \frac{L}{2} \sum_{k=1}^{t-1} \eta_{t-1}\eta_k \|s_k - m_k\|^2.$$

Then we have

$$\begin{aligned} F(\eta_t s_t) &\le \frac{\eta_t}{\eta_{t-1}} F(\eta_{t-1} s_{t-1}) + \eta_t^2 \frac{L}{2} \|r_t - m_t\|^2 \\ &\le \frac{L}{2} \sum_{s=1}^{t-1} \eta_t \eta_s \|r_s - m_s\|^2 + \eta_t^2 \frac{L}{2} \|r_t - m_t\|^2 \end{aligned}$$

□

Taking no stepsize or constant stepsize and setting $m_t = 0$ recovers the original results by Gordon. When $p = 1$ and $m_t = 0$, and applying the assumed upper bound on s_t, Theorem 1 gives

$$R_{\mathcal{X}}^T \le D \left(\frac{LC \sum_{t=1}^{T} \eta_t}{2B} + \frac{A}{B\eta_T} \right). \tag{13}$$

Therefore, taking $\eta_t \in O(\frac{1}{\sqrt{t}})$ gives a regret bound holding uniformly over time that is of the order $O(\sqrt{T})$.

For the case of when $p > 1$ where no stepsize is needed the following regret bound is immediate

$$R_{\mathcal{X}}^T \le D \left(\frac{L(\sum_{t=1}^{T} \|s_t - m_t\|^2) + 2A}{2B} \right)^{1/p}. \tag{14}$$

In the case of regret-matching, where $B = 1$, $A = 0$, $L = 2$, and $p = 2$ (Gordon 2007), we get

$$R_{\mathcal{X}}^T \le D \sqrt{\sum_{t=1}^{T} \|s_t - m_t\|^2}. \tag{15}$$

Adaptive Stepsizes

For the case of $p = 1$ we can still achieve a regret bound similar to (15) by taking adaptive stepsizes. Intuitively, the stepsizes account for how well previous predictions have done, or in the case of no predictions, how large in norm s_t have been.

Unlike the typical adaptive stepsize scheme for mirror descent (11), the stepsizes for Lagrangian hedging will be similar to adaptive FTRL methods (Mohri and Yang 2016), including the initial stepsize η_1,

$$\eta_t = \frac{1}{\sqrt{\frac{1}{\eta_1^2} + \sum_{k=1}^{t-1} \|s_k - m_k\|^2}} \quad t > 1. \tag{16}$$

Our result is a direct application of the following Lemma, which is a slight modification of a similar result by Orabona (Orabona 2019)[Lemma 4.13], we provide the proof in the appendix.

Lemma 2. *Let $a_0 \ge 0$ and $0 \le a_i \le C$ for $i > 0$. If f is a non-negative decreasing function then*

$$\sum_{t=1}^{T} a_t f(a_0 + \sum_{i=1}^{t-1} a_i) \le (C - a_0) f(a_0) + \int_{a_0}^{s_{T-1}} f(x)dx.$$

[4] Using the subgradient inequality $F(x) - F(y) \le \langle \partial F(x), x - y\rangle$.

[5] η_0 is not used to construct x_1 and is only used for the analysis.

Following adaptive stepsize scheme (16) yields the following regret bound.

Theorem 2. *An optimistic Lagrangian hedging algorithm with a convex potential function F satisfying conditions (3-5), with $p = 1$ and stepsizes*

$$\eta_t = \frac{1}{\sqrt{\frac{1}{\eta_1^2} + \sum_{k=1}^{t-1} \|s_k - m_k\|^2}} \quad t > 1,$$

and $\eta_1 \leq \sqrt{\frac{3}{C}}$, attains the following regret bound

$$R_{\mathcal{X}}^T \leq \frac{D}{B}\left((L + A)\sqrt{\frac{1}{\eta_1^2} + \sum_{t=1}^{T-1} \|s_t - m_t\|^2}\right).$$

See appendix for proof.

Path Length Bound with Smooth Losses

Optimism and adaptivity have found useful applications in improving rates for several smooth problems. For example, faster rates in smooth games (Rakhlin and Sridharan 2013b; Syrgkanis et al. 2015; Farina et al. 2019; Farina, Kroer, and Sandholm 2019), and faster rates for offline optimization (Cutkosky 2019; Joulani et al. 2020).

Unfortunately, these results strongly depend on the regret bound having the same form as OMD and OFTRL. However, in Lagrangian hedging we can attain a path length regret bound when the loss is fixed and smooth; the regret is upper bounded by the change in iterates.

The new path length bound is a direct application of the general optimistic results of the previous section combined with the assumption of a fixed Lipschitz continuous smooth convex loss (possibly non-linear), that is $\ell_t = \ell, K \geq \|\nabla \ell(x)\|$, and $\|\nabla \ell(x) - \nabla \ell(y)\| \leq L\|x - y\|_*$. If we take the typical martingale prediction $m_t = s_{t-1}$ then we have that $\|s_t - m_t\|^2 \leq \tilde{C}\|x_t - x_{t-1}\|_*^2$.

Proof. In the fixed loss case we have $s_t = \langle \nabla \ell(x_t), x_t \rangle u - \nabla \ell(x_t)$. Therefore,

$$\begin{aligned}
&\|s_t - m_t\| = \|s_t - s_{t-1}\| = \\
&\|\nabla \ell(x_{t-1}) - \nabla \ell(x_t) + \langle \nabla \ell(x_t), x_t \rangle u - \langle \nabla \ell(x_{t-1}), x_{t-1} \rangle u\| \\
&\leq L\|x_{t-1} - x_t\|_* + \|u\| \, |\langle \nabla \ell(x_t), x_t \rangle - \langle \nabla \ell(x_{t-1}), x_{t-1}| \\
&= L\|x_{t-1} - x_t\|_* \\
&+ \|u\| \, |\langle \nabla \ell(x_t) - \nabla \ell(x_{t-1}), x_t \rangle + \langle \nabla \ell(x_{t-1}), x_t - x_{t-1})| \\
&\leq (L + \|u\| DL + K)\|x_t - x_{t-1}\|_*.
\end{aligned}$$

Taking $\tilde{C} = (L + \|u\| DL + K)^2$ gives the result. □

Generalization to Φ-Regret in Experts

When $\mathcal{X} = \Delta^n$, the n-dimensional simplex, online linear optimization becomes a problem of learning with expert advice. At each round t an iterate $x_t \in \Delta^n$ is a distribution over n actions, interpreted as weightings over recommendations by n experts. Similar to before, regret will compare the total loss with the best $x^* \in \mathcal{X}$. However, this is equal to comparing with the best action (best expert recommendation) and is referred to as external regret,

$$R_{\mathcal{X}}^T = \max_{a \in A} \sum_{t=1}^{T} \langle \ell_t, x_t \rangle - \langle \ell_t, \delta_a \rangle. \tag{17}$$

Where δ_a is a distribution over A with full weight on action a. The regret can be interpreted as considering an alternative sequence of iterates $\{\tilde{x}_t\}_{t \leq T}$, where each $\tilde{x}_t = \phi(x_t)$, for some transformation of the form $\phi(x) = \delta_a$. More generally, we can measure regret with respect to a set of linear transformations Φ, referred to as Φ regret

$$R_{\Phi}^T = \max_{\phi \in \Phi} \sum_{t=1}^{T} \langle \ell_t, x_t \rangle - \langle \ell_t, \phi(x_t)) \rangle. \tag{18}$$

Similar to Lagrangian hedging, we seek to force a vector to some safe set. More precisely, we consider the Φ regret vector $s_{1:t}^{\Phi}$ that keeps track of how an algorithm is doing with respect to the set Φ,

$$s_{1:t}^{\Phi} = s_{1:t-1}^{\Phi} + s_t^{\Phi}. \tag{19}$$

Where $s_t^{\Phi} = \{\langle \ell_t, x_t \rangle - \langle \ell_t, \phi(x_t) \rangle\}_{\phi \in \Phi} \in \mathbb{R}^{|\Phi|}$. If $s_{1:t}^{\Phi}$ has all non-positive entries then $R_{\Phi}^T \leq 0$, therefore the safe set is chosen to be $\mathbb{R}_{\leq 0}^{|\Phi|}$, the negative orthant. This Φ regret framework, though abstract, includes other interesting forms of regret such as internal and swap regret (Greenwald, Li, and Marks 2006). Internal regret is interesting as it allows for efficient computation of a correlated equilibrium in game theory (Cesa-Bianchi and Lugosi 2006).

Similar to Lagrangian hedging, and as proposed by Greenwald, Li, and Marks, the algorithms will use a potential function F to measure how far $s_{1:t}^{\Phi}$ is from the safe set and slow down its growth with the Blackwell condition

$$\langle \nabla F(s_{1:t-1}^{\Phi}), s_t^{\Phi} \rangle \leq 0. \tag{20}$$

As shown by Greenwald, Li, and Marks, the generalized Blackwell condition with respect to Φ is achieved if an algorithm plays a fixed point of a linear operator M_t^{Φ},

$$M_t^{\Phi}(x) = \frac{\sum_{\phi \in \Phi} (\nabla F(s_{1:t-1}^{\Phi}))_{\phi} \phi(x)}{\langle \nabla F(s_{1:t-1}^{\Phi}), \mathbf{1} \rangle} \tag{21}$$

where $(\nabla F(s_{t-1}^{\Phi}))_{\phi}$ denotes the component of the vector $\nabla F(s_{t-1}^{\Phi}) \in \mathbb{R}^{|\Phi|}$ associated with the transformation $\phi \in \Phi$, and $\mathbf{1} = (1, \cdots, 1) \in \mathbb{R}^{|\Phi|}$. This fixed point exactly coincides with the Lagrangian hedging method when $\mathcal{X}$ is a simplex, and $\Phi = \{\phi : \exists a \in A \, \forall x \, \phi(x) = \delta_a\}$. In other words, the rule (7) is a fixed point of $M_t^{\Phi}(x)$ for external regret.

If an upperbound on F provides a regret bound, as in the previous sections, then optimistic Lagrangian hedging can be generalized to the Φ-regret setting, by defining a new operator,

$$\tilde{M}_t^{\Phi}(x) = \frac{\sum_{\phi \in \Phi} (\nabla F(\eta_t(s_{t-1}^{\Phi} + m_t)))_{\phi} \phi(x)}{\langle \nabla F(\eta_t(s_{t-1}^{\Phi} + m_t)), \mathbf{1} \rangle}. \tag{22}$$

The main result is a theorem analogous to Theorem 1, except with the new regret vector $s_{1:t}^{\Phi}$.

Theorem 3. *An optimistic Lagrangian hedging algorithm playing a fixed point of $\tilde{M}_t^{\Phi}$, with a convex potential function F satisfying conditions (3-4) and positive decreasing stepsizes $0 < \eta_t \le \eta_{t-1}$, ensures*

$$F(\eta_T s^{\Phi}_{1:T}) \le \frac{L}{2} \sum_{t=1}^{T} \eta_T \eta_t \left\| s_t^{\Phi} - m_t \right\|^2 .$$

The proof is identical to Theorem 1 except we use the Blackwell condition

$$\langle \nabla F(\eta_t(s^{\Phi}_{1:t-1} + m_t)), s_t^{\Phi} \rangle \le 0,$$

see appendix for more details.

To the best of our knowledge, this results in the first set of optimistic and adaptive algorithms for minimizing internal and swap regret.

Lagrangian Hedging+

In this section we extend optimistic Lagrangian hedging in the Φ-regret setting to use a modified regret vector

$$s^{\Phi+}_{1:t} = (s^{\Phi+}_{1:t-1} + s^{\phi}_t)^+.$$

This modification is inspired by the regret-matching+ algorithm, which has been successfully used to solve large two-player zero-sum games and play poker at an expert-level (Tammelin 2014; Tammelin et al. 2015; Burch 2017). Indeed this framework generalizes regret-matching+, beyond external regret and beyond regret-matching.

With the modified regret vector $s^{\Phi+}_{1:t}$, the safe set remains as $\mathbb{R}^{|\Phi|}_{\le 0}$, because of the component wise inequality

$$s^{\Phi}_{1:t} \le s^{\Phi+}_{1:t}.$$

Therefore, $s^{\Phi+}_{1:t} \in \mathbb{R}^{|\Phi|}_{\le 0}$ implies $s^{\Phi}_{1:t} \in \mathbb{R}^{|\Phi|}_{\le 0}$, and we have

$$R^T_{\phi} = \max_{\phi \in \Phi} (s^{\Phi}_{1:t})_{\phi} \le \max_{\phi \in \Phi} (s^{\Phi+}_{1:t})_{\phi}.$$

As one would expect, we define optimistic Lagrangian hedging+ with the operator M_t but modified to use the regret vector $s^{\Phi+}_{1:t}$ and a prediction m_t.

$$\tilde{M}_t^{\Phi+}(x) = \frac{\sum_{\phi \in \Phi} (\nabla F(\eta_t(s^{\Phi+}_{t-1} + m_t)))_{\phi} \phi(x)}{\langle \nabla F(\eta_t(s^{\Phi+}_{t-1} + m_t)), \mathbf{1} \rangle}. \quad (23)$$

Like the previous section, if we can control the growth of $F(s^{\Phi,+}_{1:t})$ and F indeed provides an upper bound to the safe set, then a regret bound is attainable. However, we must make an additional assumption on F, for which we call *positive invariant and smooth* (D'Orazio 2020). That is

$$F((x+y)^+) \le F(x) + \langle \partial F(x), y \rangle + \frac{L}{2} \|y\|^2 .$$

Once again, equipped with the new smoothness condition and the Blackwell condition

$$\langle \nabla F(\eta_t(s^{\Phi+}_{1:t-1} + m_t)), s_t^{\Phi} \rangle \le 0,$$

which is guaranteed by playing the fixed point (23) (see appendix for details), we have the following bound on F.

Theorem 4. *An optimistic Lagrangian hedging+ algorithm playing a fixed point of $\tilde{M}_t^{\Phi+}$, with a convex potential function F that is positive invariant and smooth and satisfying condition (3), with positive decreasing stepsizes $0 < \eta_t \le \eta_{t-1}$, ensures*

$$F(\eta_T s^{\Phi+}_{1:T}) \le \frac{L}{2} \sum_{t=1}^{T} \eta_T \eta_t \left\| s_t^{\Phi} - m_t \right\|^2 .$$

Φ Regret Examples

Inspired by Lagrangian hedging, Greenwald, Li, and Marks present different potential functions that are appropriate to minimizing Φ regret. The functions include a polynomial family of algorithms, with regret-matching as special case, and an exponential variant which amounts to the hedge algorithm when external regret is minimized.

Polynomial

$$F(x) = \left\| x^+ \right\|_p^2 \quad p \ge 2, \quad (24)$$

F is smooth with $L = 2(p-1)$, and with respect to the p-norm $\|\cdot\|_p$. Greenwald, Li, and Marks showed that an upper bound on $F(s_{1:T}) \le K$ amounts to the regret bound $R^T_{\Phi} \le \sqrt{K}$. In the case of when an algorithm is using the modified regret vector $s^{\Phi+}_{1:T}$ it is easy to show $\max_{\phi \in \Phi} (s^{\Phi+}_{1:t})_{\phi} \le \sqrt{F(s^{\Phi+}_{1:t})} \le \sqrt{K}$, since F is positive invariant and smooth because $F(x^+) = F(x)$.

When $p = 2$ and Φ is taken to be equivalent to external regret, we have the gradient of $F(x)$ is x^+, which gives the regret-matching algorithm when if $m_t = 0$, and the regret-matching+ algorithm if $s^{\Phi+}_{1:t}$ is used with $m_t = 0$. Notice that we exactly recover the regret-matching bound (15) in this case by applying the upper bound from Theorem 3 and with no stepsize ($\eta_t = 1$).

Greenwald, Li, and Marks also showed that the polynomial case can be extended to $1 < p < 2$ with the potential function

$$F(x) = \left\| x^+ \right\|_p^p \quad 1 < p < 2.$$

However, the smoothness condition (4) must be modified by replacing $\|\cdot\|^2$ with $\|\cdot\|^p$. This does not change the analysis but the bounds need to be changed accordingly. More importantly the regret bound degrades as p approaches 1, $R^T_{\Phi} \le K^{1/p}$.

Similar to the case of $p \ge 2$, if $1 < p < 2$ bounds on $\max_{\phi \in \Phi} (s^{\Phi+}_{1:t})_{\phi}$ are attainable since $F(x^+) = F(x)$ and hence is positive invariant and smooth.[6]

Exponential In addition to the polynomial family, we can pick the exponential variant with potential function

$$F(\eta x) = \ln \sum_i e^{\eta x_i} - \ln(d).$$

Where $x \in \mathbb{R}^d$, $L = 1$, and $\|\cdot\|^2 = \|\cdot\|_{\infty}^2$ for the smoothness condition. It can also be shown that $\max_i x_i$ for some vector

[6] See (D'Orazio 2020) for more details.

$x \in \mathbb{R}^d$ is upper bounded by $\frac{1}{\eta}(F(\eta x) + \ln(d))$, therefore the bound on F from Theorem 3 gives an upper bound on $\max_\phi(s^\Phi_{1:t})_\phi$ with $d = |\Phi|$.

The gradient of $F(x)$ is the softmax function and gives the hedge algorithm if Φ is chosen to correspond with external regret.

Related Work

An important instance of Lagrangian hedging is regret-matching, an algorithm typically used within the game theory community (Hart and Mas-Colell 2000; Zinkevich et al. 2008), and is a special case of Blackwell's algorithm (Blackwell et al. 1956). At the same time of writing Farina, Kroer, and Sandholm have also analyzed an optimistic variant of regret-matching and its popular variant regret-matching+, named predictive regret-matching and predictive regret-matching+, respectively (Farina, Kroer, and Sandholm 2020). On the surface, our analysis provides more generality as it includes both of their variants of regret-matching and more. However, it is conceivable that the main tool used in their paper, the equivalence of Blackwell approachability and online linear optimization (Abernethy, Bartlett, and Hazan 2011), provides generality to analyze other optimistic Blackwell style algorithms. More importantly, we do not believe that the tools from Abernethy, Bartlett, and Hazan to be equivalent to Lagrangian hedging. Further investigation is left to future work.

Conclusion

In this paper we extend Lagrangian hedging to include optimism, a guess m_t of how the regret vector will change, and adaptive stepsizes. The regret bounds attained for optimistic and adaptive Lagrangian hedging lead to a path length bound in constrained smooth convex optimization. Furthermore, we devise optimistic and adaptive algorithms to minimize Φ regret, a generalization of external regret that includes internal regret, and include a new class of algorithms that generalizes regret-matching+.

The analysis in this paper provides new algorithms, experimental evaluation is left to future work. For example, do the new optimistic and adaptive algorithms for internal regret provide better convergence to correlated equilibria then their non-optimistic counterparts? Additionally, in this work the step size scheme (16) is prescribed for potential functions with parameter $p = 1$, which amounts to a new step size scheme for the well-known hedge algorithm, with many preexisting adaptive variants, does this scheme provide any benefits over other adaptive schemes in practice?

References

Abernethy, J.; Bartlett, P. L.; and Hazan, E. 2011. Blackwell approachability and no-regret learning are equivalent. In *Proceedings of the 24th Annual Conference on Learning Theory*, 27–46.

Abernethy, J. D.; Hazan, E.; and Rakhlin, A. 2009. Competing in the dark: An efficient algorithm for bandit linear optimization .

Beck, A.; and Teboulle, M. 2003. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters* 31(3): 167–175.

Blackwell, D.; et al. 1956. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics* 6(1): 1–8.

Burch, N. 2017. *Time and Space: Why Imperfect Information Games are Hard*. Ph.D. thesis, University of Alberta.

Cesa-Bianchi, N.; and Lugosi, G. 2003. Potential-based algorithms in on-line prediction and game theory. *Machine Learning* 51(3): 239–261.

Cesa-Bianchi, N.; and Lugosi, G. 2006. *Prediction, learning, and games*. Cambridge university press.

Chiang, C.-K.; Yang, T.; Lee, C.-J.; Mahdavi, M.; Lu, C.-J.; Jin, R.; and Zhu, S. 2012. Online optimization with gradual variations. In *Conference on Learning Theory*, 6–1.

Cutkosky, A. 2019. Anytime online-to-batch, optimism and acceleration. In *International Conference on Machine Learning*, 1446–1454.

Daskalakis, C.; and Panageas, I. 2018. Last-iterate convergence: Zero-sum games and constrained min-max optimization. *arXiv preprint arXiv:1807.04252* .

D'Orazio, R. 2020. *Regret Minimization with Function Approximation in Extensive-Form Games*. Master's thesis, University of Alberta.

Farina, G.; Kroer, C.; Brown, N.; and Sandholm, T. 2019. Stable-Predictive Optimistic Counterfactual Regret Minimization. In *International Conference on Machine Learning*, 1853–1862.

Farina, G.; Kroer, C.; and Sandholm, T. 2019. Optimistic regret minimization for extensive-form games via dilated distance-generating functions. In *Advances in Neural Information Processing Systems*, 5221–5231.

Farina, G.; Kroer, C.; and Sandholm, T. 2020. Faster Game Solving via Predictive Blackwell Approachability: Connecting Regret Matching and Mirror Descent. *arXiv preprint arXiv:2007.14358* .

Freund, Y.; and Schapire, R. E. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences* 55(1): 119–139.

Gordon, G. J. 2007. No-regret algorithms for online convex programs. In *Advances in Neural Information Processing Systems*, 489–496.

Greenwald, A.; Li, Z.; and Marks, C. 2006. Bounds for Regret-Matching Algorithms. In *ISAIM*.

Hart, S.; and Mas-Colell, A. 2000. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica* 68(5): 1127–1150.

Hazan, E.; et al. 2016. Introduction to Online Convex Optimization. *Foundations and Trends® in Optimization* 2(3-4): 157–325.

Joulani, P.; György, A.; and Szepesvári, C. 2017. A Modular Analysis of Adaptive (Non-) Convex Optimization: Optimism, Composite Objectives, and Variational Bounds. *Journal of Machine Learning Research* 1: 40.

Joulani, P.; Raj, A.; Gyorgy, A.; and Szepesvari, C. 2020. A simpler approach to accelerated optimization: iterative averaging meets optimism. In III, H. D.; and Singh, A., eds., *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, 4984–4993. PMLR. URL http://proceedings.mlr.press/v119/joulani20a.html.

Mohri, M.; and Yang, S. 2016. Accelerating online convex optimization via adaptive prediction. In *Artificial Intelligence and Statistics*, 848–856.

Nemirovsky, A. S.; and Yudin, D. B. 1983. Problem complexity and method efficiency in optimization. .

Nesterov, Y. 2009. Primal-dual subgradient methods for convex problems. *Mathematical programming* 120(1): 221–259.

Nesterov, Y. 2018. *Lectures on convex optimization*, volume 137. Springer.

Orabona, F. 2019. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213* .

Rakhlin, A.; and Sridharan, K. 2013a. Online Learning with Predictable Sequences. In *Conference on Learning Theory*, 993–1019.

Rakhlin, S.; and Sridharan, K. 2013b. Optimization, Learning, and Games with Predictable Sequences. In Burges, C. J. C.; Bottou, L.; Welling, M.; Ghahramani, Z.; and Weinberger, K. Q., eds., *Advances in Neural Information Processing Systems 26*, 3066–3074. Curran Associates, Inc. URL http://papers.nips.cc/paper/5147-optimization-learning-and-games-with-predictable-sequences.pdf.

Schuurmans, D.; and Zinkevich, M. A. 2016. Deep learning games. In *Advances in Neural Information Processing Systems*, 1678–1686.

Shalev-Shwartz, S.; and Singer, Y. 2006. Online learning meets optimization in the dual. In *International Conference on Computational Learning Theory*, 423–437. Springer.

Syrgkanis, V.; Agarwal, A.; Luo, H.; and Schapire, R. E. 2015. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems*, 2989–2997.

Tammelin, O. 2014. Solving large imperfect information games using CFR+. *arXiv preprint arXiv:1407.5042* .

Tammelin, O.; Burch, N.; Johanson, M.; and Bowling, M. 2015. Solving heads-up limit Texas Hold'em. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Warmuth, M. K.; and Jagota, A. K. 1997. Continuous and discrete-time nonlinear gradient descent: Relative loss bounds and convergence. In *Electronic proceedings of the 5th International Symposium on Artificial Intelligence and Mathematics*, volume 326. Citeseer.

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

Zinkevich, M.; Johanson, M.; Bowling, M.; and Piccione, C. 2008. Regret minimization in games with incomplete information. In *Advances in neural information processing systems*, 1729–1736.

Appendix

Proof of Lemma 2

Let $a_0 \geq 0$ and $0 \leq a_i \leq C$ for $i > 0$ f. If f is a non-negative decreasing function then

$$\sum_{t=1}^{T} a_t f(a_0 + \sum_{i=1}^{t-1} a_i) \leq (C - a_0) f(a_0) + \int_{a_0}^{s_{T-1}} f(x)dx$$

Proof. Let $s_t = \sum_{i=0}^{t} a_i$.

$$\begin{aligned} a_t f(a_0 + \sum_{i=1}^{t-1} a_i) = a_t f(s_{t-1}) &= (a_t - a_{t-1}) f(s_{t-1}) + a_{t-1} f(s_{t-1}) \\ &= (a_t - a_{t-1}) f(s_{t-1}) + \int_{s_{t-2}}^{s_{t-1}} f(s_{t-1}) dx \\ &\leq (a_t - a_{t-1}) f(s_{t-1}) + \int_{s_{t-2}}^{s_{t-1}} f(x) dx \end{aligned}$$

So we have that

$$\sum_{t=1}^{T} a_t f(a_0 + \sum_{i=1}^{t-1} a_i) \leq \sum_{t=1}^{T} (a_t - a_{t-1}) f(s_{t-1}) + \int_{a_0}^{s_{T-1}} f(x)dx.$$

Now we will apply the summation by parts formula to analyze the first sum.

$$\begin{aligned} &\sum_{t=1}^{T} (a_t - a_{t-1}) f(s_{t-1}) \\ &= f(s_{T-1}) a_T - f(a_0) a_0 - \sum_{t=2}^{T} a_{t-1}(f(s_{t-1}) - f(s_{t-2})) \\ &= f(s_{T-1}) a_T - f(a_0) a_0 + \sum_{t=1}^{T-1} a_t (f(s_{t-1}) - f(s_t)) \\ &\leq f(s_{T-1}) a_T - f(a_0) a_0 + \sum_{t=1}^{T-1} C (f(s_{t-1}) - f(s_t)) \\ &= f(s_{T-1}) a_T - f(a_0) a_0 + C f(a_0) - C f(s_{T-1}) \\ &\leq (C - a_0) f(a_0) \end{aligned}$$

The first inequality is due to f being a decreasing function, hence $f(s_{t-1}) - f(s_t) \geq 0$, and because $0 \leq a_i \leq C$. The last inequality also follows because $a_T \leq C$. □

Proof of Theorem 2

An optimistic Lagrangian hedging algorithm with a convex potential function F satisfying conditions (1-3), with $p = 1$ and stepsizes

$$\eta_t = \frac{1}{\sqrt{\frac{1}{\eta_1^2} + \sum_{k=1}^{t-1} \|s_k - m_k\|^2}} \quad t > 1,$$

and $\eta_1 \leq \sqrt{\frac{3}{C}}$, attains the following regret bound

$$R_{\mathcal{X}}^T \leq \frac{D}{B}\left((L+A)\sqrt{\frac{1}{\eta_1^2} + \sum_{t=1}^{T-1} \|s_t - m_t\|^2}\right).$$

Proof. Recall assumption (5) with $p = 1$, and inequality (2), then a non-negative upperbound on $F(\eta_T s_{1:T}) \leq K$ translates into the regret bound

$$R_{\mathcal{X}}^T \leq D\left(\frac{K}{B\eta_T} + \frac{A}{B\eta_T}\right).$$

From Theorem 1 we have that

$$0 \leq K = \frac{L}{2}\sum_{t=1}^{T} \eta_T \eta_t \|s_t - m_t\|^2,$$

is a valid upper bound. Therefore

$$R_{\mathcal{X}}^T \leq \frac{D}{B}\left(\frac{L}{2}\sum_{t=1}^{T} \eta_t \|s_t - m_t\|^2 + \frac{A}{\eta_T}\right).$$

We now apply Lemma 2 on the sum across T rounds by noticing that $\eta_t = f(a_0 + \sum_{i=1}^{t-1} a_i)$, where $a_i = \|s_i - m_i\|^2$, $f(x) = \frac{1}{\sqrt{x}}$, and $a_0 = \frac{1}{\eta_1^2}$. By assumption we also have that $0 \leq a_i = \|s_i - m_i\|^2 \leq C$.

Therefore, by Lemma 2

$$\begin{aligned}
&\sum_{t=1}^{T} \eta_t \|s_t - m_t\|^2 \leq \\
&\left(C - \frac{1}{\eta_1^2}\right)\eta_1 + 2\left(\sqrt{\frac{1}{\eta_1^2} + \sum_{t=1}^{T} \|s_t - m_t\|^2} - \frac{1}{\eta_1}\right) \\
&= C\eta_1 - \frac{3}{\eta_1} + 2\sqrt{\frac{1}{\eta_1^2} + \sum_{t=1}^{T} \|s_t - m_t\|^2} \\
&\leq 2\sqrt{\frac{1}{\eta_1^2} + \sum_{t=1}^{T} \|s_t - m_t\|^2} \quad \text{if } \eta_1 \leq \sqrt{\frac{3}{C}}.
\end{aligned}$$

□

Proof of Theorem 4

An optimistic Lagrangian hedging+ algorithm playing a fixed point of $\tilde{M}^{\Phi+}$, with a convex potential function that is positive invariant and smooth F and satisfying conditions (3-4), with positive decreasing stepsizes $0 < \eta_t \leq \eta_{t-1}$, ensures

$$F(\eta_T s_{1:T}^{\Phi+}) \leq \frac{L}{2}\sum_{t=1}^{T} \eta_T \eta_t \left\|s_t^{\Phi} - m_t\right\|^2.$$

Proof. The proof resembles closely to that of Theorem 1.

$$\begin{aligned}
&F(\eta_t s_{1:t}^{\Phi+}) = F(\eta_t(s_{1:t-1}^{\Phi+} + s_t^{\Phi} + m_t - m_t)^+) \\
&\leq F(\eta_t(s_{1:t-1}^{\Phi+} + m_t)) + \langle \nabla F(\eta_t(s_{1:t-1}^{\Phi+} + m_t)), \eta_t(s_t^{\Phi} - m_t)\rangle \\
&+ \frac{L}{2}\left\|s_t^{\Phi} - m_t\right\|^2 \\
&\leq F(\eta_t(s_{1:t-1}^{\Phi+} + m_t)) - F(\eta_t s_{1:t-1}^{\Phi+}) + F(\eta_t s_{1:t-1}^{\Phi+}) \\
&+ \langle \nabla F(\eta_t(s_{1:t-1}^{\Phi+} + m_t)), -\eta_t m_t\rangle + \frac{L}{2}\left\|s_t^{\Phi} - m_t\right\|^2.
\end{aligned}$$

The rest follows from the same arguments as Theorem 1. □

The Φ Regret Fixed Point

Greenwald, Li, and Marks showed that x_t that is a fixed point of M_t is guaranteed to satisfy the generalized Blackwell condition

$$\langle \nabla F(s_{1:t-1}^{\Phi}), s_t^{\Phi}\rangle \leq 0.$$

Our results reuse this observation by modifiying the regret vector $s_{1:t-1}^{\Phi}$ with a prediction and possibly using the modified $s_{1:t-1}^{\Phi+}$ vector. More generally, we use the following result that follows directly from Greenwald, Li, and Marks; the following inequality holds

$$\langle \nabla F(z), s_t^{\Phi}\rangle \leq 0,$$

if the operator used to construct the fixed point is defined to be

$$M_t^{\Phi}(x) = \frac{\sum_{\phi\in\Phi}(\nabla F(z))_\phi \phi(x)}{\langle \nabla F(z), \mathbf{1}\rangle}.$$