

Faster Rates for Convex-Concave Games

Jacob Abernethy

Georgia Tech

PROF@GATECH.EDU

Kevin A. Lai

Georgia Tech

KEVINLAI@GATECH.EDU

Kfir Y. Levy

ETH Zurich

YEHUDA.LEVY@INF.ETHZ.CH

Jun-Kun Wang

Georgia Tech

JIMWANG@GATECH.EDU

Editors: Sebastien Bubeck, Vianney Perchet and Philippe Rigollet

Abstract

We consider the use of no-regret algorithms to compute equilibria for particular classes of convex-concave games. While standard regret bounds would lead to convergence rates on the order of $O(T^{-1/2})$, recent work (Rakhlin and Sridharan, 2013b; Syrgkanis et al., 2015) has established $O(1/T)$ rates by taking advantage of a particular class of optimistic prediction algorithms. In this work we go further, showing that for a particular class of games one achieves a $O(1/T^2)$ rate, and we show how this applies to the Frank-Wolfe method and recovers a similar bound (Garber and Hazan, 2015). We also show that such no-regret techniques can even achieve a linear rate, $O(\exp(-T))$, for equilibrium computation under additional curvature assumptions.

Keywords: Online learning, zero-sum games, Frank-Wolfe, fast rates

1. Introduction

A large number of core problems in statistics, optimization, and machine learning, can be framed as the solution of a two-player zero-sum game. Linear programs, for example, can be viewed as a competition between a feasibility player, who selects a point in \mathbb{R}^n , and a constraint player that aims to check for feasibility violations (Adler, 2013). Boosting (Freund and Schapire, 1999) can be viewed as the competition between an agent that selects hard distributions and a weak learning oracle that aims to overcome such challenges (Freund and Schapire, 1996). The hugely popular technique of Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), which produce implicit generative models from unlabelled data, has been framed in terms of a repeated game, with a distribution player aiming to produce realistic samples and a discriminative player that seeks to distinguish real from fake.

While many vanilla supervised learning problems reduce to finding the minimum of an objective function $f(\cdot)$ over some constraint set, tasks that require the search for a *saddle point*—that is, min-max solution of some convex-concave payoff function $g(\cdot, \cdot)$ —don’t easily lend themselves to standard optimization protocols such as gradient descent, Newton’s method, etc. It is not clear, for example, whether successive iterates should even increase or decrease the payoff g . This issue has been noticed in the training of GANs, for example, where the standard update method is a simultaneous gradient descent procedure, and many practitioners have raised concerns about cycling.

On the other hand, what has emerged as a very popular and widely-used trick is the following: simulate a pair of online learning algorithms, each competing in the game with the objective of *minimizing regret*, and return the time-averaged sequence of actions taken by the players as an approximate solution. The method of applying no-regret learning strategies to find equilibria in zero-sum games was explored in [Freund and Schapire \(1999\)](#), yet the idea goes back at least as far as work by [Blackwell \(1956\)](#) and [Hannan \(1957\)](#). This methodology has several major benefits, which include the following. First, this framework “decouples” the optimization into two parallel routines that have very little communication overhead. Second, the use of no-regret learning is ideal in this scenario, as most of the guarantees for such algorithms are robust to even adversarial environments. Third, one is able to bound the approximation error of the returned saddle point simply in terms of the total regret of the two players. Finally, several surprising recent results have suggested that this parallel online learning methodology leads to *even stronger guarantees* than what the naïve theory would tell you. In short, whereas the typical no-regret analysis would lead to an approximation error of $O(T^{-1/2})$ after T iterations, the use of *optimistic learning* strategies ([Chiang et al., 2012](#)) can be shown to guarantee $O(T^{-1})$ convergence; this technique was developed by [Rakhlin and Sridharan \(2013a\)](#) and further expanded by [Syrkanis et al. \(2015\)](#).

In this work we go further, showing that *even faster rates* are achievable for some specific cases of saddle-point problems. In particular:

1. [Abernethy and Wang \(2017\)](#) observed that the optimization method known as *Frank-Wolfe* is simply an instance of the above no-regret framework for solving a particular convex-concave game, leading to a rate of $O(T^{-1})$. In this work we further analyze the Frank-Wolfe game, and show that when the objective function and constraint set have additional structure, and both algorithms use optimistic learning procedures, then we can achieve a rate of $O(T^{-2})$. This generalizes a result of [Garber and Hazan \(2015\)](#) who proved a similar convergence rate for Frank-Wolfe.
2. Additionally, we show that when the game payoff function is suitably curved in both inputs—i.e. it is strongly-convex-concave and smooth—then we can use no-regret dynamics to achieve a *linear rate*, with the error decaying as $O(\exp(-T))$. Applying our technique to the Frank-Wolfe game we are able to recover the linear rate results of [Levitin and Polyak \(1966\)](#); [Dem'yanov and Rubinov \(1970\)](#) and [Dunn \(1979\)](#).

A notable aspect of our work is the combination of several key algorithmic techniques. First, our Frank-Wolfe result relies on regularization using the *squared gauge function*, allowing the learner to need only a single linear optimization call on each round. Second, we introduce a notion of weighted regret minimization, and our rates depend on the careful selection of the weight schedule as well as a careful analysis of what has been called *Optimistic FollowTheRegularizedLeader*. Third, our linear convergence rate leans on a trick developed recently by [Levy \(2017\)](#) that generates an adaptive weighting scheme based on the norm of the observed gradients.

1.1. Preliminaries

We first provide some definitions that are used in this paper. Let $f : \mathcal{X} \mapsto \mathbb{R}$ be some function.

Definition 1 A vector w is a subgradient of f at v for any $u \in \text{dom} f$, $f(u) \geq f(v) + \langle w, u - v \rangle$.

Definition 2 f is L -smooth w.r.t. a norm $\|\cdot\|$ if f is everywhere differentiable and for any $u, v \in \mathcal{X}$ $f(u) \leq f(v) + \langle \nabla f(v), u - v \rangle + \frac{L}{2} \|u - v\|^2$. An equivalent definition of smoothness is that f has Lipschitz continuous gradient, i.e., $\|\nabla f(u) - \nabla f(v)\|_* \leq L \|u - v\|$.

Definition 3 f is σ -strongly convex w.r.t. a norm $\|\cdot\|$ if for any $u, v \in \mathcal{X}$, $f(u) \geq f(v) + \langle \nabla f(v), u - v \rangle + \frac{\sigma}{2} \|u - v\|^2$ for some constant $\sigma > 0$.

Definition 4 For a convex function f , its Fenchel conjugate is $f^*(x) := \sup_{y \in \text{dom} f} \langle x, y \rangle - f(y)$.

Note that if f is convex then so is its conjugate f^* , since it is defined as the maximum over linear functions of x (Boyd (2004)). Moreover, when the function $f(\cdot)$ is strictly convex and the above supremum is attained, we have that $\nabla f^*(x) = \arg \max_y \langle x, y \rangle - f(y)$. Furthermore, the biconjugate f^{**} equals f if and only if f is closed and convex. It is known that f is σ -strongly convex w.r.t. $\|\cdot\|$ if and only if f^* is $1/\sigma$ strongly smooth with respect to the dual norm $\|\cdot\|_*$ (Kakade et al. (2009)), assuming that f is a closed and convex function.

Definition 5 A convex set $\mathcal{K} \subseteq \mathbb{R}^d$ is a λ -strongly convex set w.r.t. a norm $\|\cdot\|$ if for any $u, v \in \mathcal{K}$, any $\theta \in [0, 1]$, the $\|\cdot\|$ ball centered at $\theta u + (1 - \theta)v$ with radius $\theta(1 - \theta)^{\frac{\lambda}{2}} \|u - v\|^2$ is included in \mathcal{K} . For examples of strongly-convex sets, we refer the readers to (Garber and Hazan, 2015).

Definition 6 Let \mathcal{K} be any closed convex set which contains the origin. Then the gauge function of \mathcal{K} is $\gamma_{\mathcal{K}}(x) := \inf\{c \geq 0 : \frac{x}{c} \in \mathcal{K}\}$

One can show that the gauge function is a convex function (e.g. Rockafellar (1996)). It is known that several closed convex sets can lead to the same gauge function (Bach (2013)). But if a closed convex set \mathcal{K} contains the origin, then the gauge function is unique and one has $\mathcal{K} = \{x \in \mathbb{R}^d : \gamma_{\mathcal{K}}(x) \leq 1\}$. Furthermore, $\text{int } \mathcal{K} = \{x \in \mathbb{R}^d : \gamma_{\mathcal{K}}(x) < 1\}$.

Next we provide a characterization of sets based on their gauge function.

Definition 7 (β -Gauge set) Let \mathcal{K} be a closed convex set which contains the origin. We say that \mathcal{K} is β -Gauge if its squared gauge function, $\gamma_{\mathcal{K}}^2(\cdot)$, is β -strongly-convex.

This property captures a wide class of constraints. Among these are l_p balls, Schatten p balls, and the Group (s, p) ball. We refer the reader to Appendix B for more details. Curiously, all of these Gauge sets are also known to be strongly-convex. We conjecture that strong-convexity and the Gauge property are equivalent.

2. Minimizing Regret to Solve Games

Let us now turn our attention to a now-classical trick: using sequential no-regret learning strategies to find equilibria in zero-sum games.

2.1. Weighted Regret Minimization

We begin by briefly defining the standard online learning setup. We imagine a learner who must make a sequence of decisions, selecting at each round t a point x_t that lies within a convex and compact *decision space* \mathcal{K} . After selecting x_t she is charged $\ell_t(x_t)$ for her action, where $\ell_t(\cdot)$ is the *loss function* in round t , and she proceeds to the next round. Typically it is assumed that when the learner selects x_t in round t , she has observed all loss functions $\ell_1(\cdot), \dots, \ell_{t-1}(\cdot)$ up to, but not including, time t . However, we will also consider learners that are *prescient*, i.e. that can choose x_t with knowledge of the loss functions up to *and including* time t .

The standard objective for adversarial online learning is the *regret*, defined as the difference between the learner's loss over the sequence, discounted by the loss of the best fixed action in hindsight. However, for the purposes of this paper we consider a generalized notion which we call the *weighted regret*, where every time period has an importance weight that can differ from round to round. More precisely, we assume that the learning process is characterized by a sequence of weights $\alpha := \alpha_1, \alpha_2, \dots, \alpha_T$, where $\alpha_t > 0$ for every t . Now we define the *weighted regret* according to

$$\alpha\text{-REG} := \sum_{t=1}^T \alpha_t \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^T \alpha_t \ell_t(x).$$

(Note that when we drop the α —, this implies that $\alpha_t = 1$ for all t). The sequence of α_t 's can be arbitrary, and indeed we will consider scenarios under which these weights can be selected in an online fashion, according to the observed loss sequence. The learners also observe α_t at the end of each round. Throughout the paper we will use A_t to denote the cumulative sum $\sum_{s=1}^t \alpha_s$, and of particular importance will be the weighted average regret $\overline{\alpha\text{-REG}} := \frac{\alpha\text{-REG}}{A_T}$.

2.2. Algorithms

In this section we present several of the classical, and a few more recent, algorithms with well-established regret guarantees. For the most part, we present these algorithms in *unweighted* form, without reference to the weight sequence α . In later sections we specify more precisely their weighted counterparts.

One of the most well-known online learning strategies is known as FOLLOWTHEREGULARIZEDLEADER (FTRL), in which the decision point x_t is chosen as the “best” point over the previous loss functions, with some additional regularization penalty according to some convex $R(\cdot)$. Precisely, given a parameter $\eta > 0$, the learner chooses on round t the point

$$x_t = \arg \min_{x \in \mathcal{K}} \{ \eta \sum_{s=1}^{t-1} \ell_s(x) + R(x) \}. \quad (1)$$

For convenience, let ∇_t be the gradient $\nabla \ell_t(x_t)$. If we assume that $R(\cdot)$ is a strongly convex function with respect to some norm $\|\cdot\|$, then a well-known regret analysis grants the following bound:

$$\text{REG}_T \leq \frac{D}{\eta} + \frac{\eta}{2} \sum_{t=1}^T \|\nabla_t\|_*^2, \quad (2)$$

where $D := \sup_{x \in \mathcal{K}} R(x)$. With an appropriately-tuned η , one achieves $\text{REG}_T \leq \sqrt{D \sum_{t=1}^T \|\nabla_t\|_*^2}$, which is $O(\sqrt{T})$ as long the gradients have bounded norm. See, e.g., [Shalev-Shwartz et al. \(2012\)](#); [Hazan \(2014\)](#); [Rakhlin and Sridharan \(2016\)](#) for further details on this analysis.

The FOLLOWTHELEADER (FTL) strategy minimizes the objective (1), but without the regularization penalty; i.e. $x_t = \arg \min_{x \in \mathcal{K}} \sum_{s=1}^{t-1} \ell_s(x)$. Another way to formalize this is to consider

$\eta \rightarrow \infty$. Given that the above bound has a $\frac{1}{\eta}$ term, it is clear we can not simply apply the same analysis of FOLLOWTHEREGULARIZEDLEADER, and indeed one can find examples where linear regret is unavoidable [Cesa-Bianchi and Lugosi \(2006\)](#); [Shalev-Shwartz et al. \(2012\)](#). On the other hand, it has been shown that a strong regret guarantee is achievable even without regularization, as long as the sequence of loss functions are *strongly convex*. In particular, [Kakade and Shalev-Shwartz \(2009\)](#) show the following result:

Lemma 8 (Corollary 1 from [Kakade and Shalev-Shwartz \(2009\)](#)) *Let ℓ_1, \dots, ℓ_T be a sequence of functions such that for all $t \in [T]$, ℓ_t is σ_t -strongly convex. Assume that the FTL algorithm runs on this sequence and for each $t \in [T]$, let v_t be in $\partial \ell_t(x_t)$. Then,*

$$\sum_{t=1}^T \ell_t(x_t) - \min_x \sum_{t=1}^T \ell_t(x) \leq \frac{1}{2} \sum_{t=1}^T \frac{\|v_t\|^2}{\sum_{\tau=1}^t \sigma_\tau} \quad (3)$$

Furthermore, let $G = \max_t \|v_t\|$ and assume that for all $t \in [T]$, $\sigma_t \geq \sigma$. Then, the regret is bounded by $\frac{G^2}{2\sigma} (\log(T) + 1)$.

In the context of solving zero-sum games, the online learning framework allows for one of the two players to be prescient, so she has access to one additional loss function $\ell_t(\cdot)$ before selecting her x_t . In such a case it is much easier to achieve low regret, and we present three standard prescient algorithms:

$$\text{BESTRESPONSE} \quad x_t = \arg \min_{x \in \mathcal{K}} \ell_t(x), \quad (4)$$

$$\text{BETHELEADER} \quad x_t = \arg \min_{x \in \mathcal{K}} \sum_{s=1}^t \ell_s(x), \quad (5)$$

$$\text{BETHEREGULARIZEDLEADER} \quad x_t = \arg \min_{x \in \mathcal{K}} \sum_{s=1}^t \ell_s(x) + \frac{1}{\eta} R(x). \quad (6)$$

Indeed it is easy to show that, for the first two of these prescient strategies, one easily obtains $\text{REG}_T \leq 0$ ([Kalai and Vempala, 2005b](#)). The regret of BETHEREGULARIZEDLEADER is no more than $\frac{1}{\eta} \sup_{x, x' \in \mathcal{K}} R(x) - R(x')$. We also consider optimistic algorithms, which we discuss in [Appendix A](#).

Gauge Function FTRL. While the analysis of FOLLOWTHEREGULARIZEDLEADER is natural and leads to a simple intuitive bound (2), it requires solving a non-linear optimization problem on each round even when the loss functions $\ell_t(\cdot)$ are themselves linear – a very common scenario. From a computational perspective, it is often impractical to solve the FOLLOWTHEREGULARIZEDLEADER objective. Nevertheless, in many scenarios a (computationally feasible) *linear optimization oracle* is at hand. In such instances, much attention has been focused on a *perturbed* version of FOLLOWTHELEADER, where one solves the unregularized optimization problem but with a linear noise term added to the objective; there is much work analyzing these algorithms and we refer the reader to [Kalai and Vempala \(2005b\)](#); [Cesa-Bianchi and Lugosi \(2006\)](#); [Abernethy et al. \(2014\)](#) among many others. The main downside of such randomized approaches is that they have good expected regret but suffer in variance, which makes them less suitable in various reductions.

In this work, we introduce a family of FOLLOWTHEREGULARIZEDLEADER algorithms that rely solely on a linear oracle, and we believe this is a novel approach to online linear optimization problems. The restriction we require is that the regularizer $R(\cdot)$ is chosen as the *squared gauge function* $\gamma_{\mathcal{K}}^2(\cdot)$ for the decision set \mathcal{K} of the learner. Here we will assume¹ for every t that $\ell_t(\cdot) =$

1. One can reduce any arbitrary convex loss to the linear loss case by convexity $\ell_t(x) - \ell_t(x^*) \leq \langle \partial \ell_t(x), x - x^* \rangle$. ([Shalev-Shwartz et al. \(2012\)](#); [Hazan \(2014\)](#); [Rakhlin and Sridharan \(2016\)](#)).

$\langle l_t, \cdot \rangle$ for some vector l_t , hence the objective (1) reduces to

$$x_t = \arg \min_{x \in \mathcal{K}} \eta \langle L_{t-1}, x \rangle + \gamma_{\mathcal{K}}^2(x), \quad (7)$$

where $L_{t-1} = l_1 + \dots + l_{t-1}$. Denote $\text{bndry}(\mathcal{K})$ as the boundary of the constraint set \mathcal{K} . We can reparameterize the above optimization, by observing that any point $x \in \mathcal{K}$ can be written as ρz where $z \in \text{bndry}(\mathcal{K})$, and $\rho \in [0, 1]$. Hence we have

$$\min_{\rho \in [0, 1]} \min_{z \in \text{bndry}(\mathcal{K})} \eta \langle L_{t-1}, \rho z \rangle + \gamma_{\mathcal{K}}^2(\rho z) = \min_{\rho \in [0, 1]} \left(\min_{z \in \text{bndry}(\mathcal{K})} \eta \langle L_{t-1}, z \rangle \right) \rho + \rho^2. \quad (8)$$

We are able to remove the dependence on the gauge function since it is homogeneous, $\gamma_{\mathcal{K}}(\rho x) = |\rho| \gamma_{\mathcal{K}}(x)$, and is identically 1 on the boundary of \mathcal{K} . The inner minimization reduces to the linear optimization $z^* := \arg \min_{z \in \mathcal{K}} \langle L_{t-1}, z \rangle$, and the optimal ρ is $\max(0, \min(1, -(\eta/2) \langle L_{t-1}, z^* \rangle))$.

2.3. Solving zero-sum convex-concave games

Let us now apply the tools described above to the problem of solving a particular class of zero-sum games; these are often referred to as convex-concave saddle point problems. Assume we have convex and compact sets $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{Y} \subset \mathbb{R}^m$, known as the *action spaces* for the two players. We are given a convex-concave *payoff function* $g : \mathcal{X} \times \mathcal{Y}$; that is, $g(\cdot, y)$ is convex in its first argument for every fixed $y \in \mathcal{Y}$, and $g(x, \cdot)$ is concave in its second argument for every fixed $x \in \mathcal{X}$. We say that a pair $(\hat{x}, \hat{y}) \in \mathcal{X} \times \mathcal{Y}$ is an ϵ -equilibrium for $g(\cdot, \cdot)$ if $\sup_{y \in \mathcal{Y}} g(\hat{x}, y) - \inf_{x \in \mathcal{X}} g(x, \hat{y}) \leq \epsilon$.

The celebrated *minimax theorem*, first proven by von Neumann for a simple class of bilinear payoff functions (v. Neumann, 1928; Neumann et al., 1944) and generalized by Sion (1958) and others, states that there exist 0-equilibria for convex-concave games under reasonably weak conditions. Another way to state this is $\inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} g(x, y) = \sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} g(x, y)$, and we tend to call this quantity V^* , the *value of the game* $g(\cdot, \cdot)$.

The method of computing an ϵ -equilibrium using a pair of no-regret algorithms is reasonably straightforward, although here we will emphasize the use of weighted regret, which has been much less common in the literature. Algorithm 1 describes a basic template used throughout the paper.

Algorithm 1 Computing equilibria using no-regret algorithms

- 1: Input: a T -length sequence α
 - 2: **for** $t = 1, 2, \dots, T$ **do**
 - 3: x -player selects $x_t \in \mathcal{X}$ using no-regret algorithm OAlg^x
 - 4: y -player selects $y_t \in \mathcal{Y}$ using (possibly-different) no-regret algorithm OAlg^y
 - 5: x -player suffers loss $\ell_t(x_t)$ with weight α_t , where $\ell_t(\cdot) = g(\cdot, y_t)$
 - 6: y -player suffers loss $h_t(y_t)$ with weight α_t , where $h_t(\cdot) = -g(x_t, \cdot)$
 - 7: **end for**
 - 8: Output $(\bar{x}_\alpha, \bar{y}_\alpha) := \left(\frac{\sum_{s=1}^T \alpha_s x_s}{A_T}, \frac{\sum_{s=1}^T \alpha_s y_s}{A_T} \right)$
-

Theorem 9 Assume that a convex-concave game payoff $g(\cdot, \cdot)$ and a T -length sequence α are given. Assume that we run Algorithm 1 using no-regret procedures OAlg^x and OAlg^y , and the α -weighted average regret of each is $\overline{\alpha\text{-REG}}^x$ and $\overline{\alpha\text{-REG}}^y$, respectively. Then the output $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an ϵ -equilibrium for $g(\cdot, \cdot)$, with $\epsilon = \overline{\alpha\text{-REG}}^x + \overline{\alpha\text{-REG}}^y$.

The theorem can be restated in terms of V^* , where we get the following “ ϵ -sandwich”:

$$V^* - \epsilon \leq \inf_{x \in X} g(x, \bar{y}_\alpha) \leq V^* \leq \sup_{y \in Y} g(\bar{x}_\alpha, y) \leq V^* + \epsilon \quad (9)$$

But the key insight is that the regret analysis leads immediately to a convergence rate for the algorithmic template presented above. We provide the proof of Theorem 9 in Appendix C.

2.4. Application: the Frank-Wolfe Algorithm

We can tie the above set of tools together with an illustrative application, describing a natural connection to the Frank-Wolfe (FW) method (Frank and Wolfe, 1956) for constrained optimization. The ideas presented here summarize the work of Abernethy and Wang (2017), but in Section 3 we significantly strengthen the result for a special case.

We have a convex set \mathcal{K} , an L -smooth convex function $f : \mathcal{K} \rightarrow \mathbb{R}$, and some initial point $w_0 \in \mathcal{K}$. The FW algorithm makes repeated calls to a linear optimization oracle over \mathcal{K} , followed by a convex averaging step:

$$(\text{linear opt}) \quad v_t = \arg \min_{v \in \mathcal{K}} \langle v, \nabla f(w_{t-1}) \rangle; \quad (\text{update}) \quad w_t = (1 - \eta_t)w_{t-1} + \eta_t v_t,$$

where the parameter η_t is a learning rate, and following the standard analysis one sets $\eta_t = \frac{2}{t+2}$. A well-known result is that $f(w_T) - \inf_{w \in \mathcal{K}} f(w) \leq 2LD^2/T$.

Let us leverage Theorem 9 to obtain a convergence rate from a no-regret perspective. With a brief inspection, one can verify that FW is indeed a special case of Algorithm 1, assuming that (a) the game payoff is $g(x, y) := f^*(x) - \langle x, y \rangle$, where f^* is the Fenchel conjugate of f ; (b) the sequence α is $1, 2, \dots, T$; (c) the x -player and y -player employ FOLLOWTHELEADER and BESTRESPONSE, respectively; we output the final iterate as $w_T := \bar{y}_\alpha$. We refer to Abernethy and Wang (2017) for a thorough exposition, but it is striking that this use of Algorithm 1 leads to Frank-Wolfe even up to $\eta_t = \frac{2}{t+2}$.

As we have reframed FW in terms of our repeated game, we can now appeal to our main theorem to obtain a rate. We must first observe, using the duality of Fenchel conjugation, that

$$V^* = \sup_y \inf_x g(x, y) = -(\inf_y \sup_x \langle x, y \rangle - f^*(x)) = -(\inf_y f(y)). \quad (10)$$

Using (9) and the above equality, we can obtain $f(\bar{y}_\alpha) \leq \inf_{y \in \mathcal{K}} f(y) + \overline{\alpha\text{-REG}}^x + \overline{\alpha\text{-REG}}^y$.

The convergence rate of FW thus boils down to bounding the regret of the two players. We note first that the y -player is prescient and employs BESTRESPONSE, hence we conclude that $\overline{\alpha\text{-REG}}^y \leq 0$. The x -player on the other hand will suffer the α -weighted regret of FOLLOWTHELEADER. But notice, critically, that the choice of payoff $g(x, y) = f^*(x) - \langle x, y \rangle$ happens to be *strongly convex* in x , as L -smoothness of f implies L^{-1} -strong convexity in f^* . We may thus use Lemma 8 to obtain:

$$\overline{\alpha\text{-REG}}^x \stackrel{\text{Lemma 8}}{\leq} \frac{1}{2A_T} \sum_{t=1}^T \frac{\alpha_t^2 D}{\sum_{\tau=1}^t \alpha_\tau (1/L)} \leq O(\sum_{\tau=1}^T LD^2/A_T) = O(\frac{LD^2}{T}),$$

where we use the fact that the x -player observes an α_t/L strongly convex function, $\alpha_t \ell_t(\cdot)$, and that $\|v_t\|^2$ in Lemma 8 is $\|\alpha_t \nabla \ell_t(\cdot)\|^2 \leq \alpha_t^2 D^2$, where D is the diameter of \mathcal{Y} . We conclude by noting that the absence of the $\log T$ term, which tends to arise from the regret of online strongly convex optimization, was removed by carefully selecting the sequence of weights α .

3. Fast convergence in the FW game

In this section, we introduce a new FW-like algorithm that achieves a $O(1/T^2)$ convergence rate on β -Gauge sets \mathcal{Y} accessed using a linear optimization oracle. The design and analysis are based on a reweighting scheme and OPTIMISTIC-FTRL, taking advantage of recent tools developed for fast rates in solving games (Chiang et al., 2012; Rakhlin and Sridharan, 2013b; Syrgkanis et al., 2015).

In Theorem 10 we give an instantiation of Algorithm 1 that finds an approximate saddle point for the FW game $g(x, y) = f^*(x) - \langle x, y \rangle$. In this instantiation the x -player plays OPTIMISTIC-FTRL and the y -player plays BETHEREGULARIZEDLEADER. With an appropriate weighting, the weighted regret guarantees of these two algorithms imply that we can find an $O(\frac{1}{T^2})$ -approximate saddle point solution of the FW game in T rounds. Recalling that $\min_{x \in \mathcal{X}} \{f^*(x) - \langle x, y \rangle\} = -f(y)$, this immediately translates to a convergence rate of $O(\frac{1}{T^2})$ for the problem $\min_{y \in \mathcal{Y}} f(y)$.

The algorithm that we describe in Theorem 10 does not immediately yield a FW-like algorithm—in general, we may not be able to compute the y -player’s BETHEREGULARIZEDLEADER iterates using only a linear optimization oracle. However, if the y -player uses the squared gauge function of \mathcal{Y} as a regularizer, then the y iterates are computable using a linear optimization oracle, as shown in Section 2.2. This fact immediately implies that for β -Gauge sets and upon choosing the gauge function as regularizer, Algorithm 2 instantiates a *projection-free* procedure which provides a convergence rate of $O(1/T^2)$ for the problem $\min_{y \in \mathcal{Y}} f(y)$ (see Corollary 11). In Appendix G, we discuss how to get a faster rate than $O(1/T)$ for arbitrary convex sets if BETHEPERTURBEDLEADER rather than BETHEREGULARIZEDLEADER is used by the y -player in the FW game.

3.1. Solving the FW game with OPTIMISTIC-FTRL and BETHEREGULARIZEDLEADER

In this section, we present our algorithm for finding $O(1/T^2)$ -saddle point solutions to the FW game. We instantiate Algorithm 1 using the FW objective $g(x, y) = f^*(x) - \langle x, y \rangle$, where we assume f is L -smooth and σ -strongly convex. The x -player plays OPTIMISTIC-FTRL and the y -player plays BETHEREGULARIZEDLEADER.

Theorem 10 *Assume that we instantiate Algorithm 1 with the FW game $g(x, y) = f^*(x) - \langle x, y \rangle$, weight sequence $\alpha_t = t$, and the following strategies for the players. The x -player plays OPTIMISTIC-FTRL:*

$$x_t = \arg \min_{x \in \mathcal{X}} \sum_{s=1}^{t-1} \alpha_s \ell_s(x) + m_t(x) \text{ with } m_t(x) = \alpha_t \ell_{t-1}(x) \quad (11)$$

where $\ell_t(x) = g(x, y_t)$, and the y -player plays BETHEREGULARIZEDLEADER:

$$y_t = \arg \min_{y \in \mathcal{Y}} \sum_{s=1}^t \alpha_s h_s(y) + \frac{1}{\eta} R(y) \quad (12)$$

with a β -strongly-convex regularizer $R(\cdot)$ and $\eta = \beta/16L(1 + \frac{L}{\sigma})$, where $h_t(x) = -g(x_t, y)$. Then the output $(\bar{x}_\alpha, \bar{y}_\alpha)$ of Algorithm 1 is an $O\left(\frac{L(R(y^*) - R(z))(1 + \frac{L}{\sigma})}{\beta T^2}\right)$ -approximate saddle point solution to the FW game, where $z = \arg \min_{y \in \mathcal{Y}} R(y)$.

Now recall that for the FW setting, we are interested in y -players that may only employ a linear optimization oracle. In general it is impossible to solve Equation (12) within $O(1)$ calls to such oracles in each round. Nevertheless, recall that for β -Gauge sets, choosing $R(y) = \gamma_{\mathcal{Y}}^2(y)$ induces a β -strongly-convex regularizer, while enabling us to solve Equation (12) with a single call to the

linear oracle, as shown in Equation 8. The proof of Theorem 10 shows that the x -player's strategy is the gradient of the primal objective f at the point $\bar{y}_{\alpha'_{1:t-1}}$, where α' is a weight vector such that $\alpha'_s = \alpha_s$ for $s = 1, \dots, t-1$ and $\alpha'_{t-1} = \alpha_{t-1} + \alpha_t$ and $\bar{y}_{\alpha'_{1:t-1}}$ is the α' -weighted average of y_1, \dots, y_{t-1} (See Equation 18). This leads to Algorithm 2 and Corollary 11.

Algorithm 2 A new FW algorithm

- 1: Let α be a T -length weight sequence
 - 2: **for** $t = 1, 2, \dots, T$ **do**
 - 3: Define the T -length sequence α' as $\alpha'_s = \alpha_s$ for $s = 1, \dots, t-1$, and $\alpha'_t = \alpha_{t-1} + \alpha_t$
 - 4: Set $x_t = \nabla f(\frac{\sum_{s=1}^{t-1} \alpha'_s y_s}{A_t})$
 - 5: Set $(\hat{y}_t, \rho_t) = \arg \min_{y \in \mathcal{Y}, \rho \in [0,1]} \sum_{s=1}^t \rho \langle y, \alpha_s x_s \rangle + \frac{1}{\eta} \rho^2$ for $\eta = \frac{\beta}{16L(1+\frac{L}{\sigma})}$
 - 6: Set $y_t = \rho_t \hat{y}_t$
 - 7: **end for**
 - 8: Output $\bar{y}_T := \frac{\sum_{s=1}^T \alpha_s y_s}{A_T}$, where $A_T = \sum_{s=1}^T \alpha_t$
-

We get the following corollary of the above theorem. The full proof is in Appendix F.

Corollary 11 *Let $f : \mathcal{Y} \mapsto \mathbb{R}$ be L -smooth and σ -strongly-convex. Also assume that \mathcal{Y} is a β -Gauge set. Let $\alpha_t = t$. Then the output \bar{y}_T in Algorithm 2 is an $O\left(\frac{L(R(y^*)-R(z))(1+\frac{L}{\sigma})}{\beta T^2}\right)$ -approximate optimal solution to the optimization problem $\min_{y \in \mathcal{Y}} f(y)$. Moreover, Algorithm 2 only requires a single linear optimization oracle call in each round.*

3.2. Proof of Theorem 10

Proof [Proof of Theorem 10] In the FW game, we observe that the loss functions $\alpha_t \ell_t(\cdot)$ seen by the x -player are $\frac{\alpha_t}{L}$ -strongly convex, since the function $f(\cdot)$ is L smooth, which implies that $f^*(\cdot)$ is $\frac{1}{L}$ -strongly convex.

The x -player chooses x_t based on OPTIMISTIC-FTL: $x_t = \arg \min_{x \in \mathcal{X}} \sum_{s=1}^{t-1} \alpha_s \ell_s(x) + m_t(x)$, where $m_t(x) = \alpha_t \ell_{t-1}(x)$. To analyze the regret of the x -player, let us first denote the update of the standard FOLLOWTHELEADER as

$$z_t = \arg \min_x \sum_{s=1}^{t-1} \alpha_s \ell_s(x). \quad (13)$$

Denote $x^* := \arg \min_{x \in \mathcal{X}} \sum_{t=1}^T \alpha_t \ell_t(x)$.² Now we are going to analyze the α -weighted regret of the x -player, which is

$$\begin{aligned} \alpha\text{-REG}^x &:= \sum_{t=1}^T \alpha_t \ell_t(x_t) - \alpha_t \ell_t(x^*) \\ &= \sum_{t=1}^T \alpha_t \ell_t(x_t) - \alpha_t \ell_t(z_{t+1}) - m_t(x_t) + m_t(z_{t+1}) + \sum_{t=1}^T m_t(x_t) - m_t(z_{t+1}) \\ &\quad + \sum_{t=1}^T \alpha_t \ell_t(z_{t+1}) - \alpha_t \ell_t(x^*). \\ &\leq \sum_{t=1}^T \langle \alpha_t \nabla \ell_t(x_t) - \nabla m_t(z_{t+1}), x_t - z_{t+1} \rangle + \sum_{t=1}^T m_t(x_t) - m_t(z_{t+1}) \\ &\quad + \sum_{t=1}^T \alpha_t \ell_t(z_{t+1}) - \alpha_t \ell_t(x^*) - \frac{2\alpha_t}{L} \|x_t - z_{t+1}\|^2, \end{aligned} \quad (14)$$

2. The following analysis actually holds for any $x^* \in \mathcal{X}$.

where the last inequality uses strong convexity of $\ell_t(\cdot)$ so that

$$\alpha_t \ell_t(x_t) - \alpha_t \ell_t(z_{t+1}) \leq \langle \alpha_t \nabla \ell_t(x_t), x_t - z_{t+1} \rangle - \frac{\alpha_t}{L} \|x_t - z_{t+1}\|^2, \quad (15)$$

and that

$$-m_t(x_t) + m_t(z_{t+1}) \leq \langle \nabla m_t(z_{t+1}), z_{t+1} - x_t \rangle - \frac{\alpha_t}{L} \|x_t - z_{t+1}\|^2. \quad (16)$$

There are three sums in (14). Note that the second sum should be small because the expression for x_t “exploits” $m_t(\cdot) = \alpha_t \ell_{t-1}(\cdot)$ more than the expression for z_{t-1} does. The third sum is the regret of BETHELEADER, which is non-positive. In Lemma 14, we show that the second and third sums in (14) are in total non-positive. For the proof, please see Appendix D.

Since $m_t(\cdot) := \alpha_t \ell_{t-1}(\cdot)$, each term in the first sum in (14) can be bounded by

$$\begin{aligned} \langle \alpha_t \nabla \ell_t(x_t) - \nabla m_t(z_{t+1}), x_t - z_{t+1} \rangle &= \alpha_t \langle \nabla \ell_t(x_t) - \nabla \ell_{t-1}(z_{t+1}), x_t - z_{t+1} \rangle \\ &= \alpha_t \langle -y_t + \nabla f^*(x_t) + y_{t-1} - \nabla f^*(z_{t+1}), x_t - z_{t+1} \rangle \\ &\leq \alpha_t \left(\|y_t - y_{t-1}\|_* \|x_t - z_{t+1}\| + \frac{1}{\sigma} \|x_t - z_{t+1}\|^2 \right), \end{aligned} \quad (17)$$

where the last inequality uses Hölder’s inequality and the fact that f is σ -strongly convex so that f^* is $\frac{1}{\sigma}$ smooth. Let us analyze $\|x_t - z_{t+1}\|^2$. Note that, by Fenchel conjugacy, $z_{t+1} := \arg \min_x \sum_{s=1}^t \alpha_s (-y_s^\top x + f^*(x)) = \nabla f(\bar{y}_{\alpha_{1:t}})$, where $\bar{y}_{\alpha_{1:t}}$ is the α -weighted average of y_1, \dots, y_t . For notational simplicity, let us define a new weight vector α' , where $\alpha'_s = \alpha_s$ for $s = 1, \dots, t-1$ and $\alpha'_{t-1} = \alpha_{t-1} + \alpha_t$. Similarly, for x_t , we have

$$x_t := \arg \min \{ \alpha_t (-y_{t-1}^\top x + f^*(x)) + \sum_{s=1}^{t-1} \alpha_s (-y_s^\top x + f^*(x)) \} = \nabla f(\bar{y}_{\alpha'_{1:t-1}}), \quad (18)$$

where $\bar{y}_{\alpha'_{1:t-1}}$ is the α' -weighted average of y_1, \dots, y_{t-1} . According to (18),

$$\begin{aligned} \|x_t - z_{t+1}\|^2 &= \|\nabla f(\bar{y}_{\alpha_{1:t}}) - \nabla f(\bar{y}_{\alpha'_{1:t-1}})\|^2 \leq L^2 \|\bar{y}_{\alpha_{1:t}} - \bar{y}_{\alpha'_{1:t-1}}\|^2 \\ &= \frac{L^2}{A_t^2} \left\| \sum_{s=1}^t \alpha_s y_s - \sum_{s=1}^{t-1} \alpha'_s y_s \right\|^2 = \frac{L^2}{A_t^2} \|\alpha_{t-1} y_{t-1} + \alpha_t y_t - \alpha'_{t-1} y_{t-1}\|^2 \\ &= \frac{L^2}{A_t^2} \|\alpha_t (y_{t-1} - y_t)\|^2 = \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_{t-1} - y_t\|^2. \end{aligned} \quad (19)$$

Combining (17) and (19), we get

$$\begin{aligned} (\alpha_t \nabla \ell_t(x_t) - \nabla m_t(z_{t+1}))^\top (x_t - z_{t+1}) &\leq \alpha_t (\|y_t - y_{t-1}\|_* \|x_t - z_{t+1}\| + \frac{1}{\sigma} \|x_t - z_{t+1}\|^2) \\ &\leq \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right). \end{aligned} \quad (20)$$

Therefore, we have shown that the first sum in (14) is bounded by

$$\sum_{t=1}^T \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right). \quad (21)$$

By (14), (21), and Lemma 14, we get the upper bound of the regret of the x -player,

$$\alpha\text{-REG}^x \leq \sum_{t=1}^T \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right). \quad (22)$$

Now let us switch to analyze the regret of the y -player, which is defined as $\alpha\text{-REG}^y := \sum_{t=1}^T -\alpha_t h(y_t) + \alpha_t h(y^*) = \sum_{t=1}^T -\alpha_t (-\langle x_t, y_t \rangle + f^*(x_t)) + \alpha_t (-\langle x_t, y^* \rangle + f^*(x_t))$, which equals $\sum_{t=1}^T \langle y_t - y^*, \alpha_t x_t \rangle$. This means that the y -player actually observes the linear loss $\alpha_t x_t$ in each round t , due to the fact that the y -player plays after the x -player plays. We can reinterpret BETHEREGULARIZEDLEADER as OPTIMISTIC-FTRL (Syrkanis et al. (2015)) when the learner is fully informed as to the loss function for the current round. That is, we may write the update as $y_t = \arg \min_{y \in \mathcal{Y}} \langle y, \sum_{s=1}^{t-1} \alpha_s x_s + m_t \rangle + \frac{1}{\eta} R(y)$, where $m_t := \alpha_t x_t$ and $R(\cdot)$ is β -strongly convex with respect to a norm $\|\cdot\|$ on \mathcal{Y} .

For loss vectors θ_t , Theorem 16 gives the regret of OPTIMISTIC-FTRL as

$$\sum_{t=1}^T \langle y_t - y^*, \theta_t \rangle \leq \frac{R(y^*) - R(z) - \frac{\beta}{2} (\sum_{t=1}^T \|y_t - z_t\|^2 + \sum_{t=1}^T \|y_t - z_{t+1}\|^2)}{\eta} + \sum_{t=1}^T \frac{\eta}{\beta} \|\theta_t - m_t\|_*^2, \quad (23)$$

where z_t is FTRL update, defined as $z_t = \arg \min_{z \in \mathcal{Y}} \langle z, \sum_{s=1}^{t-1} \theta_s \rangle + \frac{1}{\eta} R(y)$, while y_t is OPTIMISTIC-FTRL update, defined as $y_t = \arg \min_{z \in \mathcal{Y}} \langle z, (\sum_{s=1}^{t-1} \theta_s) + m_t \rangle + \frac{1}{\eta} R(y)$. We prove Theorem 16 in Appendix E.

Since in our case $\theta_t = m_t = \alpha_t x_t$, $y_t = z_{t+1}$ we get the bound of the regret of the y -player in the FW game,

$$\alpha\text{-REG}^y \leq \frac{R(y^*) - R(z) - \frac{\beta}{2} \sum_{t=1}^T \|y_t - y_{t-1}\|^2}{\eta}. \quad (24)$$

Combining (22) and (24), we get

$$\begin{aligned} \alpha\text{-REG}^x + \alpha\text{-REG}^y &\leq \sum_{t=1}^T \left\{ \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right) \right. \\ &\quad \left. + \frac{R(y^*) - R(z) - \frac{\beta}{2} \sum_{t=1}^T \|y_{t+1} - y_t\|^2}{\eta} \right\} \end{aligned} \quad (25)$$

The coefficient of $\|y_t - y_{t-1}\|^2$ is $\frac{\alpha_t^2 L}{A_t} + \frac{\alpha_t}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 - \frac{\beta}{2\eta}$ and, as $\alpha_t = t$ and if we set $\eta = \frac{\beta}{16L(1+\frac{L}{\sigma})}$, the quantity becomes negative. So,

$$\overline{\alpha\text{-REG}}^x + \overline{\alpha\text{-REG}}^y \leq \frac{\alpha\text{-REG}^x + \alpha\text{-REG}^y}{A_t} = O\left(\frac{L(R(y^*) - R(z))(1+\frac{L}{\sigma})}{\beta T^2}\right). \quad (26)$$

Combining this with Theorem 9 completes the proof. \blacksquare

4. Linear convergence in some strongly-convex and strongly-concave game

The $O(1/T^2)$ rate in the previous section is nice, as it shows that one can achieve ‘‘acceleration’’ under certain conditions using no-regret learning dynamics. But when can we get even faster rates using the template Algorithm 1? In the present section we describe scenarios in which the two players can compute an approximate equilibrium in *linear time*, i.e. with a rate exponentially decaying in T . This requires stronger assumptions on the payoff $g(\cdot, \cdot)$, and we will explore three such scenarios. But at a minimum we assume, throughout this section, that $g(x, y)$ is (a) σ_x -strongly convex in x , (b) L -smooth in x for every $y \in \mathcal{Y}$.

The key to obtaining a fast rate is to consider the function $s(x) = \sup_{y \in \mathcal{Y}} g(x, y)$. This function reports the payoff/loss value given to x when the y -player plays the best response $y_x =$

$\arg \max_{y \in \mathcal{Y}} g(x, y)$; i.e. $s(x) = g(x, y_x)$. We make two important insights about s . First, the supremum of strongly convex functions remains strongly convex, so s is σ_x strongly convex in x . Second, we make an important insight, established below, that the gradient of $g(\cdot, y_x)$ at the point x must be a subgradient of $s(\cdot)$ also at x . This is formally stated and proven in Appendix I.

The key property needed for $s(\cdot)$ however, is *smoothness*. It is easy to construct games for which $g(\cdot, \cdot)$ is smooth in a very strong sense, but the resulting $s(\cdot)$ is non-smooth. We need additional structure, and we give three scenarios where the condition holds. $s(\cdot)$ is smooth when any of the following are true:

1. We can write $g(x, y) = a(x) + x^\top M y - b(y)$, where $a(x)$ is any σ_x -strongly convex function of x , M is a matrix, and $b(y)$ is any σ_y -strongly convex function of y ; (See Appendix I)
2. The function $g(x, y)$ is σ_y -strongly-concave in y , the best response function $y_x := \arg \max_{y \in \mathcal{Y}} g(x, y)$ is always achieved in the interior of \mathcal{Y} , and we have that for every $w, z \in \mathcal{X}$, $\|\nabla_{y_z} g(w, \cdot) - \nabla_{y_z} g(z, \cdot)\| \leq L\|w - z\|$; (See Appendix J)
3. Assume we have the Frank-Wolfe game, where $g(x, y) := f^*(x) - \langle x, y \rangle$, the objective $f(\cdot)$ is smooth, and we are guaranteed that $\|\nabla f(y)\|_2$ is lower bounded by some positive constant B for every $y \in \mathcal{Y}$.

Any of the above conditions is suitable to obtain our main result of this section, which is that the framework of no-regret equilibrium computation allows one to obtain a *linear convergence rate* to solve certain games. This combines a careful analysis of the game, with a recent online-to-batch conversion trick given by Levy (2017). The key tool is to set the sequence of weights according to the inverse squared norm of the loss gradients. The result relies on the x -player using an algorithm known as SC-AdaGrad, a simple adaptive gradient descent procedure (Algorithm 3). We postpone the description of this algorithm, and the proof of the following theorem, to Appendix K.

Theorem 12 *Assume g satisfies any conditions such that the resulting $s(\cdot)$ is L -smooth and σ -strongly convex function for constants L and σ and moreover that $(\arg \min_{x \in \mathbb{R}^d} s(x)) \in \mathcal{X}$. Suppose we instantiate Algorithm 1, where x -player uses SC-AdaGrad (Algorithm 3) for loss functions $\alpha_t \ell_t(\cdot) = \alpha_t g(\cdot, y_t)$, y -player uses BESTRESPONSE, and the sequence α is defined as $\alpha_t := \|\nabla \ell_t(x_t)\|^{-2}$. Let $T \geq \frac{L}{\sigma} \log 3$, and constant G upper bounded $\|\nabla \ell_t\|$. Then the output $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an ϵ -equilibrium of g where $\epsilon = O\left(\frac{G^2 T}{L} e^{-\frac{\sigma}{L} T}\right)$.*

Notice that one of our scenarios stated above is the Frank-Wolfe game with an additional strong convexity assumption on the objective f , as well as a necessary lower bound on the norm of ∇f within the feasible region. That we achieve a linear rate for this scenario is indeed not surprising, as this was previously established (Levitin and Polyak, 1966; Demyanov and Rubinov, 1970; Dunn, 1979). Unfortunately, Theorem 12 does *not* provide an FW-like algorithm, as the combination of the SC-AdaGrad and BESTRESPONSE subroutines do not reduce to a simple linear optimization. An alternative algorithm we propose called SC-AFTL (Algorithm 4), which is akin to an adaptive version of FOLLOWTHELEADER, does reduce to a linear oracle. We state the main theorem below, yet its proof and Algorithm 4 are described in full detail in Appendix L.

Theorem 13 *Consider the FW game in which $g(x, y) = -\langle x, y \rangle + f^*(x)$. Suppose that $f(\cdot)$ is a L -smooth convex function and that \mathcal{Y} is a λ -strongly convex set. Also assume that the gradients of the f in \mathcal{Y} are bounded away from 0, i.e., $\max_{y \in \mathcal{Y}} \|\nabla f(y)\| \geq B$. Then, there exists a FW-like algorithm that has $O(\exp(-\frac{\lambda B}{L} T))$ rate.*

Acknowledgments

J.A., K.A.L., and J.W. gratefully acknowledge financial support from the National Science Foundation, award IIS 1453304. K.Y.L. is supported by the ETH Zurich Postdoctoral Fellowship and Marie Curie Actions for People COFUND program. J.W. also thanks the support from ARC-TRIAD Fellowship.

References

- Jacob Abernethy and Jun-Kun Wang. Frank-wolfe and equilibrium computation. *NIPS*, 2017.
- Jacob Abernethy, Chansoo Lee, Abhinav Sinha, and Ambuj Tewari. Online linear optimization via smoothing. In *COLT*, pages 807–823, 2014.
- Ilan Adler. The equivalence of linear programs and zero-sum games. *International Journal of Game Theory*, 42(1):165–177, Feb 2013. ISSN 1432-1270. doi: 10.1007/s00182-012-0328-8. URL <https://doi.org/10.1007/s00182-012-0328-8>.
- Francis Bach. Convex relaxations of structured matrix factorizations. <https://arxiv.org/pdf/1309.3117.pdf>, 2013.
- Dimitri P. Bertsekas. Stochastic optimization problems with nondifferentiable cost functionals. *Journal of Optimization Theory and Applications*, 1973.
- David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.
- Jonathan Borwein and Adrian S Lewis. Convex analysis and nonlinear optimization theory and examples. *Springer*, 2006.
- Stephen Boyd. Convex optimization. *Cambridge University Press*, 2004.
- Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, , and Shenghuo Zhu. Online optimization with gradual variations. 2012.
- Vladimir F. Demyanov and Aleksandr M. Rubinov. Approximate methods in optimization problems. *Elsevier Publishing Company*, 1970.
- John Duchi, Peter L. Bartlett, and Martin Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization (SIOPT)*, 2012.
- Joseph C. Dunn. Rates of convergence for conditional gradient algorithms near singular and non-singular extremals. *SIAM Journal on Control and Optimization*, 1979.
- Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

- Yoav Freund and Robert E Schapire. Game theory, on-line prediction and boosting. In *Proceedings of the ninth annual conference on Computational learning theory*, pages 325–332. ACM, 1996.
- Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.
- Dan Garber and Elad Hazan. Faster rates for the frank-wolfe method over strongly-convex sets. *ICML*, 2015.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.
- Elad Hazan. Introduction to online convex optimization. 2014.
- Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- Sham Kakade and Shai Shalev-Shwartz. Mind the duality gap: Logarithmic regret algorithms for online optimization. 2009.
- Sham M. Kakade, Shai Shalev-shwartz, and Ambuj Tewari. On the duality of strong convexity and strong smoothness: Learning applications and matrix regularization. 2009.
- Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 2005a.
- Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005b.
- Evgeny S Levitin and Boris T Polyak. Constrained minimization methods. *USSR Computational mathematics and mathematical physics*, 1966.
- Kfir Levy. Online to offline conversions, universality and adaptive minibatch sizes. 2017.
- Haipeng Luo. Introduction to online learning: Lecture 6. 2017.
- J von Neumann, Oskar Morgenstern, et al. Theory of games and economic behavior, 1944.
- E. S. Polovinkin. Strongly convex analysis. *Sbornik: Mathematics* 187(2):259, 1996.
- Alexander Rakhlin and Karthik Sridharan. Online learning with predictable sequences. *COLT*, 2013a.
- Alexander Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. *NIPS*, 2013b.
- Alexander Rakhlin and Karthik Sridharan. Statistical learning theory and sequential prediction. 2016.

- Tyrrell Rockafellar. Convex analysis. *Princeton University Press*, 1996.
- Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.
- Maurice Sion. On general minimax theorems. *Pacific J. Math*, 8(1):171–176, 1958.
- Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E. Schapire. Fast convergence of regularized learning in games. *NIPS*, 2015.
- J v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.
- Linli Xu, James Neufeld, Bryce Larson, and Dale Schuurmans. Maximum margin clustering. 2005.
- Yuchen Zhang and Lin Xiao. Stochastic primal-dual coordinate method for regularized empirical risk minimization. 2015.

Appendix A. Discussion of Optimistic Algorithms

Here we discuss optimistic algorithms for online learning. Assume that the loss functions are linear and let $\ell_t(\cdot) = \langle l_t, \cdot \rangle$. When there exists some pattern in the loss sequence $\{l_t\}$, the online learning problem may not be that adversarial. That is, the loss vectors may be predictable. If a “good guess” of l_t is available, then one might hope to have a better bound than $O(\sqrt{T})$. Let m_t be the algorithm’s guess for l_t . If m_t is close to l_t , then the regret should be significantly better than $O(\sqrt{T})$. There have been several papers that obtain $O(\sqrt{\sum_{t=1}^T \|l_t - m_t\|^2})$ regret in recent years (e.g. Chiang et al. (2012); Rakhlin and Sridharan (2013a); Syrgkanis et al. (2015)). Among these, perhaps OPTIMISTIC-FTRL (Syrgkanis et al. (2015)) has the simplest update rule:

$$\text{OPTIMISTIC-FTRL} \quad x_t = \arg \min_{x \in \mathcal{K}} \left\langle m_t + \sum_{s=1}^{t-1} l_s, x \right\rangle + \frac{1}{\eta} R(x) \quad (27)$$

Appendix B. Examples of β -Gauge sets

Garber and Hazan (2015) lists three known classes of strongly convex sets. They are all in the form of norm ball constraints. These are all β -Gauge sets.

1. ℓ_p balls: $\|x\|_p \leq r, \forall p \in (1, 2]$. The strong convexity of the set is $\lambda = \frac{p-1}{r}$ and its squared of gauge function is $\frac{1}{r^2} \|x\|_p^2$, which is a $\frac{2(p-1)}{r^2}$ -strongly convex function with respect to norm $\|\cdot\|_p$ by Lemma 4 in Garber and Hazan (2015).
2. Schatten p balls: $\|\sigma(X)\|_p \leq r$ for $p \in (1, 2]$, where $\sigma(X)$ is the vector consisting of singular values of the matrix X . The strong convexity of the set is $\lambda = \frac{p-1}{r}$ and the squared gauge function is $\frac{1}{r^2} \|\sigma(X)\|_p^2$, which is a $\frac{2(p-1)}{r^2}$ -strongly convex function with respect to norm $\|\sigma(\cdot)\|_p$ by Lemma 6 in Garber and Hazan (2015).
3. Group (s,p) balls: $\|X\|_{s,p} = \|(\|X_1\|_s, \|X_2\|_s, \dots, \|X_m\|_s)\|_p \leq r$ where $X \in \mathbb{R}^{m \times n}$, X_j represents the j -th row of X , and $s, p \in (1, 2]$. The strong convexity of the set is $\lambda = \frac{2(s-1)(p-1)}{r((s-1)+(p-1))}$ and its squared gauge function is $\frac{1}{r^2} \|X\|_{s,p}^2$, which is a $\frac{2(s-1)(p-1)}{r^2((s-1)+(p-1))}$ -strongly convex function with respect to norm $\|\cdot\|_{s,p}$ by Lemma 8 in Garber and Hazan (2015).

Appendix C. Proof of Theorem 9

Theorem 9 Assume that a convex-concave game payoff $g(\cdot, \cdot)$ and a T -length sequence α are given. Assume that we run Algorithm 1 using no-regret procedures OAlg^x and OAlg^y , and the α -weighted average regret of each is $\overline{\alpha\text{-REG}}^x$ and $\overline{\alpha\text{-REG}}^y$, respectively. Then the output $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an ϵ -equilibrium for $g(\cdot, \cdot)$, with

$$\epsilon = \overline{\alpha\text{-REG}}^x + \overline{\alpha\text{-REG}}^y.$$

Proof The proof basically replicates the one in Abernethy and Wang (2017) except that the loss functions of both learners are re-weighted here. It will turn out that the re-weighting trick is one of the keys to getting fast rate for some games.

In each round, the weighted loss function of the x -player is $\alpha_t \ell_t(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$, where $\ell_t(\cdot) := g(\cdot, y_t)$. The y -player, on the other hand, observes her own sequence of loss functions $\alpha_t h_t(\cdot) : \mathcal{Y} \rightarrow \mathbb{R}$, where $h_t(\cdot) := -g(x_t, \cdot)$.

$$\begin{aligned} \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t g(x_t, y_t) &= \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T -\alpha_t h_t(y_t) \\ &= -\frac{1}{\sum_{s=1}^T \alpha_s} \inf_{y \in \mathcal{Y}} \left\{ \sum_{t=1}^T \alpha_t h_t(y) \right\} - \frac{\text{Regret}_T^y}{\sum_{s=1}^T \alpha_s} \\ &= \sup_{y \in \mathcal{Y}} \left\{ \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t g(x_t, y) \right\} - \overline{\alpha\text{-REG}}^y \\ \text{(Jensen)} \quad &\geq \sup_{y \in \mathcal{Y}} g \left(\frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t x_t, y \right) - \overline{\alpha\text{-REG}}^y \\ &\geq \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} g(x, y) - \overline{\alpha\text{-REG}}^y \end{aligned} \tag{28}$$

Let us now apply the same argument on the right hand side, where we use the x -player's regret guarantee.

$$\begin{aligned} \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t g(x_t, y_t) &= \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t \ell_t(x_t) \\ &= \inf_{x \in \mathcal{X}} \left\{ \sum_{t=1}^T \frac{1}{\sum_{s=1}^T \alpha_s} \alpha_t \ell_t(x) \right\} + \frac{\text{Regret}_T^x}{\sum_{s=1}^T \alpha_s} \\ &= \inf_{x \in \mathcal{X}} \left\{ \sum_{t=1}^T \frac{1}{\sum_{s=1}^T \alpha_s} \alpha_t g(x, y_t) \right\} + \overline{\alpha\text{-REG}}^x \\ &\leq \inf_{x \in \mathcal{X}} g \left(x, \sum_{t=1}^T \frac{1}{\sum_{s=1}^T \alpha_s} \alpha_t y_t \right) + \overline{\alpha\text{-REG}}^x \\ &\leq \sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} g(x, y) + \overline{\alpha\text{-REG}}^x \end{aligned} \tag{29}$$

Note that the game value $V^* = \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} g(x, y) = \sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} g(x, y)$. Combining (28) and (29), we see that:

$$\sup_{y \in \mathcal{Y}} g(\bar{x}_\alpha, y) - \overline{\alpha\text{-REG}}^y \leq \inf_{x \in \mathcal{X}} g(x, \bar{y}_\alpha) + \overline{\alpha\text{-REG}}^x$$

which implies that $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an $\overline{\alpha\text{-REG}}^x + \overline{\alpha\text{-REG}}^y$ approximate solution to the saddle point problem. \blacksquare

Appendix D. Proof of Lemma 14 from Proof of Theorem 10

We prove the following stronger statement.

Lemma 14 *Let x_t be the OPTIMISTIC-FTL update $x_t = \arg \min_{x \in \mathcal{X}} \sum_{s=1}^{t-1} \alpha_s \ell_s(x) + m_t(x)$, where $m_t(x) = \alpha_t \ell_{t-1}(x)$. Let z_t be the standard FOLLOWTHELEADER update given by $z_t = \arg \min_x \sum_{s=1}^{t-1} \alpha_s \ell_s(x)$. Let $\alpha_t = t$. Let ℓ_t be $(1/L)$ -strongly convex for all t .*

$$\sum_{t=1}^T m_t(x_t) - m_t(z_{t+1}) + \sum_{t=1}^T \alpha_t \ell_t(z_{t+1}) - \alpha_t \ell_t(w) \leq D_T \quad (30)$$

for any $w \in \mathcal{X}$, where

$$D_T = - \sum_{t=1}^T \left(\frac{A_{t-1}}{2L} \|z_t - x_t\|^2 + \frac{A_t}{2L} \|z_{t+1} - x_t\|^2 \right).$$

We use the following standard lemma for strongly convex functions.

Lemma 15 *For a σ -strongly convex function $F(\cdot)$, suppose p^* is the minimizer. We have*

$$F(p^*) \leq F(q) + \langle \nabla F(p^*), p^* - q \rangle - \frac{\sigma}{2} \|p^* - q\|^2,$$

which leads to

$$F(p^*) \leq F(q) - \frac{\sigma}{2} \|p^* - q\|^2.$$

Proof [Proof of Lemma 14]

We use mathematical induction for the proof. The proof in the following uses Lemma 15, which is a property of strongly convex functions. For the base case $T = 0$, it holds because $0 \leq 0$.

Assume it holds for $T = \tau - 1$, that is

$$\sum_{t=1}^{\tau-1} m_t(x_t) - m_t(z_{t+1}) + \sum_{t=1}^{\tau-1} \alpha_t \ell_t(z_{t+1}) - \alpha_t \ell_t(w) \leq D_{\tau-1} \quad (31)$$

Then,

$$\begin{aligned}
 & \sum_{t=1}^{\tau} m_t(x_t) - m_t(z_{t+1}) + \sum_{t=1}^{\tau} \alpha_t \ell_t(z_{t+1}) \\
 & \stackrel{(a)}{\leq} D_{\tau-1} + m_{\tau}(x_{\tau}) - m_{\tau}(z_{\tau+1}) + \alpha_{\tau} \ell_{\tau}(z_{\tau+1}) + \sum_{t=1}^{\tau-1} \alpha_t \ell_t(z_{\tau}) \\
 & \stackrel{(b)}{\leq} D_{\tau-1} - \frac{\sigma_{\tau-1}}{2} \|z_{\tau} - x_{\tau}\|^2 + m_{\tau}(x_{\tau}) - m_{\tau}(z_{\tau+1}) + \alpha_{\tau} \ell_{\tau}(z_{\tau+1}) + \sum_{t=1}^{\tau-1} \alpha_t \ell_t(x_{\tau}) \\
 & = D_{\tau-1} - \frac{\sigma_{\tau-1}}{2} \|z_{\tau} - x_{\tau}\|^2 - m_{\tau}(z_{\tau+1}) + \alpha_{\tau} \ell_{\tau}(z_{\tau+1}) + \left(\sum_{t=1}^{\tau-1} \alpha_t \ell_t(x_{\tau}) + m_{\tau}(x_{\tau}) \right) \\
 & \stackrel{(c)}{\leq} D_{\tau-1} - \frac{\sigma_{\tau-1}}{2} \|z_{\tau} - x_{\tau}\|^2 - \frac{\sigma'_{\tau-1}}{2} \|z_{\tau+1} - x_{\tau}\|^2 - m_{\tau}(z_{\tau+1}) + \alpha_{\tau} \ell_{\tau}(z_{\tau+1}) \\
 & \quad + \left(\sum_{t=1}^{\tau-1} \alpha_t \ell_t(z_{\tau+1}) + m_{\tau}(z_{\tau+1}) \right) \\
 & = D_{\tau-1} - \frac{\sigma_{\tau-1}}{2} \|z_{\tau} - x_{\tau}\|^2 - \frac{\sigma'_{\tau-1}}{2} \|z_{\tau+1} - x_{\tau}\|^2 + \sum_{t=1}^{\tau} \alpha_t \ell_t(z_{\tau+1}) \\
 & \leq D_{\tau} + \sum_{t=1}^{\tau} \alpha_t \ell_t(w)
 \end{aligned} \tag{32}$$

where (a) is by the assumption that it holds at $\tau - 1$ and by setting $w = z_{\tau}$, (b) is by using the strong convexity lemma (Lemma 15) to get the following,

$$\sum_{t=1}^{\tau-1} \alpha_t \ell_t(z_{\tau}) \leq \sum_{t=1}^{\tau-1} \alpha_t \ell_t(x_{\tau}) - \frac{\sigma_{\tau-1}}{2} \|z_{\tau} - x_{\tau}\|^2, \tag{33}$$

as z_{τ} is the minimizer of the $\sigma_{\tau-1}$ -strongly convex function $\sum_{t=1}^{\tau-1} \alpha_t \ell_t(\cdot)$ for $\sigma_{\tau-1} := \sum_{s=1}^{\tau-1} \frac{\alpha_s}{L} = \frac{A_{\tau-1}}{L}$, (c) is by using Lemma 15 again to get

$$\sum_{t=1}^{\tau-1} \alpha_t \ell_t(x_{\tau}) + m_{\tau}(x_{\tau}) \leq \sum_{s=1}^{\tau-1} \alpha_s \ell_s(z_{\tau+1}) + m_{\tau}(z_{\tau+1}) - \frac{\sigma'_{\tau-1}}{2} \|z_{\tau+1} - x_{\tau}\|^2, \tag{34}$$

as x_{τ} is the minimizer of the $\sigma'_{\tau-1}$ -strongly convex function $\sum_{t=1}^{\tau-1} \alpha_t \ell_t(\cdot) + m_{\tau}(\cdot)$ for $\sigma'_{\tau-1} := \sum_{s=1}^{\tau-1} \frac{\alpha_s}{L} + \frac{\alpha_{\tau}}{L} = \frac{A_{\tau}}{L}$, and the last inequality is by the fact that $z_{\tau+1}$ is the minimizer of $\sum_{t=1}^{\tau} \alpha_t \ell_t(\cdot)$. \blacksquare

Appendix E. Optimistic FTRL

In this subsection, we analyze the regret bound for OPTIMISTIC-FTRL. Let θ_t be the loss function in round t and let the cumulative loss vector be $L_t = \sum_{s=1}^t \theta_s$. The update of OPTIMISTIC-FTRL is

$$y_t = \arg \min_{y \in \mathcal{Y}} \langle y, L_{t-1} + m_t \rangle + \frac{1}{\eta} R(y), \tag{35}$$

where m_t is the learner's guess of the loss vector θ_t , $R(\cdot)$ is a β -strong convex function with respect to a norm ($\|\cdot\|$) on the constraint set \mathcal{Y} and η is a parameter. The analysis of [Syrkanis et al. \(2015\)](#) is based on the assumption that \mathcal{Y} is a simplex. Here, we want to extend their results for \mathcal{Y} being an arbitrary convex set. Define the regret to be $\text{Regret} := \sum_{t=1}^T \langle y_t - y^*, \theta_t \rangle$. The proof basically replicates the one of [Luo \(2017\)](#) except that we consider any convex set and any strongly convex regularization term instead of negative entropy as in his note.

Theorem 16 OPTIMISTIC-FTRL (35) has regret

$$\begin{aligned} \text{Regret} := \sum_{t=1}^T \langle y_t - y^*, \theta_t \rangle &\leq \frac{R(y^*) - R(z) - \frac{\beta}{2} (\sum_{t=1}^T \|y_t - z_t\|^2 + \sum_{t=1}^T \|y_t - z_{t+1}\|^2)}{\eta} \\ &\quad + \sum_{t=1}^T \frac{\eta}{\beta} \|\theta_t - m_t\|_*^2, \end{aligned} \quad (36)$$

where $z_t = \arg \min_{y \in \mathcal{Y}} \langle y, L_{t-1} \rangle + \frac{1}{\eta} R(y)$ is the update of the standard FOLLOWTHEREGULARIZEDLEADER, and $z = \arg \min_{y \in \mathcal{Y}} R(y)$.

Proof We can re-write the regret as

$$\text{Regret} := \sum_{t=1}^T \langle y_t - y^*, \theta_t \rangle = \sum_{t=1}^T \langle y_t - z_{t+1}, \theta_t - m_t \rangle + \sum_{t=1}^T \langle y_t - z_{t+1}, m_t \rangle + \langle z_{t+1} - y^*, \theta_t \rangle \quad (37)$$

Let us analyze the first sum

$$\sum_{t=1}^T \langle y_t - z_{t+1}, \theta_t - m_t \rangle. \quad (38)$$

Now using Lemma 17 with $y_1 = y_t$, $u_1 = \sum_{s=1}^{t-1} \theta_s + m_t$ and $y_2 = z_{t+1}$, $u_2 = \sum_{s=1}^t \theta_s$ in the lemma, we have

$$\sum_{t=1}^T \langle y_t - z_{t+1}, \theta_t - m_t \rangle \leq \sum_{t=1}^T \|y_t - z_{t+1}\| \|\theta_t - m_t\|_* \leq \sum_{t=1}^T \frac{\eta}{\beta} \|\theta_t - m_t\|_*^2. \quad (39)$$

For the other sum,

$$\sum_{t=1}^T \langle y_t - z_{t+1}, m_t \rangle + \langle z_{t+1} - y^*, \theta_t \rangle, \quad (40)$$

we are going to show that it is upper-bounded by

$$\frac{R(y^*) - R(z) - D_T}{\eta}, \quad (41)$$

for any $y^* \in \mathcal{Y}$, where $D_T = \sum_{t=1}^T \frac{\beta}{2} \|y_t - z_t\|^2 + \frac{\beta}{2} \|y_t - z_{t+1}\|^2$.

Using induction, we see that the base case $T = 0$ clearly is true. Assume that it also holds for $T - 1$, for a $T \geq 1$. Then, we have

$$\begin{aligned}
 & \sum_{t=1}^T \langle y_t - z_{t+1}, m_t \rangle + \langle z_{t+1}, \theta_t \rangle \\
 & \leq \langle y_T - z_{T+1}, m_T \rangle + \langle z_{T+1}, \theta_T \rangle + \frac{R(z_T) - R(z) - D_{T-1}}{\eta} + \langle z_T, L_{T-1} \rangle \\
 & \stackrel{(a)}{\leq} \langle y_T - z_{T+1}, m_T \rangle + \langle z_{T+1}, \theta_T \rangle + \frac{R(y_T) - R(z) - D_{T-1} - \frac{\beta}{2} \|y_T - z_T\|^2}{\eta} \\
 & \quad + \langle y_T, L_{T-1} \rangle \\
 & = \langle z_{T+1}, \theta_T - m_T \rangle + \frac{R(y_T) - R(z) - D_{T-1} - \frac{\beta}{2} \|y_T - z_T\|^2}{\eta} + \langle y_T, L_{T-1} + m_T \rangle \quad (42) \\
 & \stackrel{(b)}{\leq} \langle z_{T+1}, \theta_T - m_T \rangle + \frac{R(z_{T+1}) - R(z) - D_{T-1} - \frac{\beta}{2} \|y_T - z_T\|^2 - \frac{\beta}{2} \|y_T - z_{T+1}\|^2}{\eta} \\
 & \quad + \langle z_{T+1}, L_{T-1} + m_T \rangle \\
 & = \langle z_{T+1}, L_T \rangle + \frac{R(z_{T+1}) - R(z) - D_T}{\eta} \\
 & \stackrel{(c)}{\leq} \langle y^*, L_T \rangle + \frac{R(y^*) - R(z) - D_T}{\eta},
 \end{aligned}$$

where (a) and (b) are by strong convexity so that

$$\langle z_T, L_{T-1} \rangle + \frac{R(z_T)}{\eta} \leq \langle y_T, L_{T-1} \rangle + \frac{R(y_T)}{\eta} - \frac{\beta}{2\eta} \|y_T - z_T\|^2, \quad (43)$$

and

$$\langle y_T, L_{T-1} + m_T \rangle + \frac{R(y_T)}{\eta} \leq \langle z_{T+1}, L_{T-1} + m_T \rangle + \frac{R(z_{T+1})}{\eta} - \frac{\beta}{2\eta} \|y_T - z_{T+1}\|^2. \quad (44)$$

(c) is because z_{T+1} is the optimal point of $\arg \min_y \langle y, L_T \rangle + \frac{R(y)}{\eta}$.

Now we have shown that

$$\sum_{t=1}^T \langle y_t - z_{t+1}, m_t \rangle + \langle z_{t+1} - y^*, \theta_t \rangle \leq \frac{R(y^*) - R(z) - D_T}{\eta}, \quad (45)$$

for any $y^* \in \mathcal{Y}$.

Combining (39) and (45) completes the theorem (36). ■

Lemma 17 Denote $y_1 = \arg \min_y \langle y, u_1 \rangle + \frac{1}{\eta} R(y)$ and $y_2 = \arg \min_y \langle y, u_2 \rangle + \frac{1}{\eta} R(y)$ for a β -strongly convex function $R(\cdot)$ with respect to a norm $\|\cdot\|$. We have $\|y_1 - y_2\| \leq \frac{\eta}{\beta} \|u_1 - u_2\|_*$.

Proof From strong convexity, we have

$$\langle y_1, u_1 \rangle + \frac{1}{\eta} R(y_1) \leq \langle y_2, u_1 \rangle + \frac{1}{\eta} R(y_2) - \frac{\beta}{2\eta} \|y_1 - y_2\|^2,$$

and

$$\langle y_2, u_2 \rangle + \frac{1}{\eta} R(y_2) \leq \langle y_1, u_2 \rangle + \frac{1}{\eta} R(y_1) - \frac{\beta}{2\eta} \|y_1 - y_2\|^2.$$

Adding the above inequalities together, we get

$$\langle y_2 - y_1, u_2 - u_1 \rangle \leq -\frac{\beta}{\eta} \|y_1 - y_2\|^2,$$

or

$$\langle y_1 - y_2, u_2 - u_1 \rangle \geq \frac{\beta}{\eta} \|y_1 - y_2\|^2.$$

So, by using Hölder's inequality,

$$\frac{\eta}{\beta} \|u_1 - u_2\|_* \geq \|y_1 - y_2\|.$$

■

Appendix F. Proof of Corollary 11

Proof [Proof of Corollary 11] Let $\epsilon = O\left(\frac{L(1+\frac{L}{\sigma})}{\beta T^2}\right)$. We can see from (8) that the y iterates are exactly the actions of the y -player in Theorem 10 when $R(y) = \gamma_{\mathcal{Y}}^2(y)$. Likewise, the x -player's iterates are exactly those of the x -player in Theorem 10, as shown in Equation 18. As the weight vector α is the same as in Theorem 10, we know that $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an ϵ -approximate saddle point to the FW game $g(x, y) = f^*(x) - \langle x, \bar{y}_\alpha \rangle$. Then we have:

$$f(\bar{y}_\alpha) = \sup_{x \in \mathcal{X}} \langle x, \bar{y}_\alpha \rangle - f^*(x) = - \inf_{x \in \mathcal{X}} g(x, \bar{y}_\alpha) \leq -V^* + \epsilon$$

where V^* is the value of the FW game, which we showed in (10) is equal to $-\inf_{y \in \mathcal{Y}} f(y)$.

Also note that Algorithm 2 does not need to store all the points $\{y_t\}$, as it can instead maintain a variable recording the weighted sum of the points. ■

Appendix G. The y -player plays BETHEPERTURBEDLEADER in the FW game

Suppose now the y -player plays in round t plays BETHEPERTURBEDLEADER,

$$y_t = E_{\xi \sim B_u(\cdot)} [\arg \min_{y \in \mathcal{Y}} \langle y, \sum_{s=1}^t x_s + \xi \rangle], \quad (46)$$

where $B_u(\cdot)$ is a uniform sample on the ℓ_2 ball of radius u . We are going to analyze the distance term $\|y_t - y_{t-1}\|^2$ in the regret of the x -player (22). Define $\Psi(x) = E_{\xi \sim B_u(\cdot)} [\max_{y \in \mathcal{Y}} \langle y, x \rangle]$. If

the max expression inside the expectation has a unique maximizer we can swap the expectation and gradient (Proposition 2.2 in (Bertsekas (1973))). So, the gradient can be written as $\nabla\Psi(x) = E_{\xi \sim B_u(\cdot)}[\arg \max_{y \in \mathcal{Y}} \langle y, x \rangle]$. Now we see that the update of Be-the-Perturbed-Leader (46) can be connected to the gradient of $\Psi(\cdot)$. That is,

$$y_t = \nabla\Psi\left(-\sum_{s=1}^t x_s\right) = E_{\xi \sim B_u(\cdot)}[\arg \max_{y \in \mathcal{Y}} \langle y, -\sum_{s=1}^t x_s + \xi \rangle]. \quad (47)$$

Using Lemma E.2 in (Duchi et al. (2012)), we know that Φ is a $\frac{L_0\sqrt{d}}{u}$ smooth function, where L_0 is the bound of the norm $\|\nabla\Psi(\cdot)\|$ and d is the dimension of \mathcal{Y} .

Consequently,

$$\|y_t - y_{t-1}\|^2 = \left\| \nabla\Psi\left(-\sum_{s=1}^t x_s\right) - \nabla\Psi\left(-\sum_{s=1}^{t-1} x_s\right) \right\|^2 \leq \frac{L_0^2 d}{u^2} \left\| -\sum_{s=1}^t x_s + \sum_{s=1}^{t-1} x_s \right\|^2 \leq \frac{LL_0^2 d}{u^2}, \quad (48)$$

where the last inequality is by the fact that x_t is a gradient vector and that the norm of gradient is bounded by the smooth constant L by the assumption. One can also show that regret of Be-the-perturbed-leader is bounded by $O(u)$ (e.g. Kalai and Vempala (2005a); Cesa-Bianchi and Lugosi (2006)).

Combining α -REG^x of (22) and α -REG^y,

$$\begin{aligned} & \frac{\alpha\text{-REG}^x}{A_T} + \frac{\alpha\text{-REG}^y}{A_T} \\ & \leq \frac{\sum_{t=1}^T \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right) + O(u)}{\sum_{t=1}^T \alpha_t} \\ & \leq O\left(\frac{\left(\sum_{t=1}^T \frac{L^2 L_0^2 d}{\sigma u^2} \right) + u}{T^2} \right), \end{aligned} \quad (49)$$

where the last inequality is due to that $\alpha_t = t$ and so that $A_T = O(T^2)$. Setting $u = O(T^{1/3})$ leads to $\frac{\alpha\text{-REG}^x}{A_T} + \frac{\alpha\text{-REG}^y}{A_T} = O\left(\frac{L^2 L_0^2 d}{\sigma T^{5/3}}\right)$. The rate is better than $O(1/T)$ when the dimension d is small. Note that we do not make any assumption on the constraint set \mathcal{Y} , so the rate holds for any convex set \mathcal{Y} . Yet, to get the fast rate, one needs to have a efficient way to compute (46), which seems to involve large number of calls to the linear oracle in each iteration. We leave it as an open problem.

Appendix H. The y -player plays BESTRESPONSE in the FW game

If the constraint set is strongly convex, $\|y_t - y_{t-1}\|^2$ is indeed shrinking even if the y -player just plays `BestResponse`, assuming the gradient is lower bounded by a constant B . To see this, let us use Lemma 18 below,

$$\begin{aligned} \|y_t - y_{t-1}\| & \leq \frac{2\|x_t - x_{t-1}\|}{\lambda(\|x_t\| + \|x_{t-1}\|)} = \frac{2\|\nabla f(\bar{y}_{t-1}) - \nabla f(\bar{y}_{t-2})\|}{\lambda(\|\nabla f(\bar{y}_{t-1})\| + \|\nabla f(\bar{y}_{t-2})\|)} \\ & \leq \frac{2L\|\bar{y}_{t-1} - \bar{y}_{t-2}\|}{\lambda(\|\nabla f(\bar{y}_{t-1})\| + \|\nabla f(\bar{y}_{t-2})\|)} = O\left(\frac{L}{\lambda t B}\right), \end{aligned} \quad (50)$$

where the first inequality is due to the lemma and the last equality is by assuming $\|\nabla f(\cdot)\| \geq B$.

Since the y -player plays **BESTRESPONSE**, its regret is 0. Combining this with α -REG^{*x*} of (22) and using (50) we have

$$\begin{aligned} & \frac{\alpha\text{-REG}^x}{A_T} + \frac{\alpha\text{-REG}^y}{A_T} \\ & \leq \frac{\sum_{t=1}^T \alpha_t \left(\left(\frac{\alpha_t L}{A_t} \right) \|y_t - y_{t-1}\|^2 + \frac{1}{\sigma} \left(\frac{\alpha_t L}{A_t} \right)^2 \|y_t - y_{t-1}\|^2 \right)}{\sum_{t=1}^T \alpha_t} \\ & = O\left(\frac{L \log T}{\lambda B T^2}\right). \end{aligned} \tag{51}$$

Lemma 18 ³

Denote $x_p = \arg \max_{x \in \mathcal{K}} \langle p, x \rangle$ and $x_q = \arg \max_{x \in \mathcal{K}} \langle q, x \rangle$, where $p, q \in \mathbb{R}^d$ are any nonzero vectors. If a compact set \mathcal{K} is a λ -strongly convex set, then

$$\|x_p - x_q\| \leq \frac{2 \|p - q\|}{\lambda(\|p\| + \|q\|)}. \tag{52}$$

Proof Recall that a strongly convex set \mathcal{K} can be written as intersection of some Euclidean balls (c.f. definition 3 in Section 1.1). Namely,

$$\mathcal{K} = \bigcap_{u \in \mathcal{S}} B_{\frac{1}{\lambda}} \left(x_u - \frac{u}{\lambda} \right).$$

Let $x_p = \arg \max_{x \in \mathcal{K}} \langle \frac{p}{\|p\|}, x \rangle$ and $x_q = \arg \max_{x \in \mathcal{K}} \langle \frac{q}{\|q\|}, x \rangle$. Based on the definition of strongly convex sets, we can see that $x_q \in B_{\frac{1}{\lambda}}(x_p - \frac{p}{\lambda\|p\|})$ and $x_p \in B_{\frac{1}{\lambda}}(x_q - \frac{q}{\lambda\|q\|})$. Therefore,

$$\|x_q - x_p - \frac{p}{\lambda\|p\|}\|^2 \leq \frac{1}{\lambda^2},$$

which leads to

$$\|p\| \cdot \|x_p - x_q\|^2 \leq \frac{2}{\lambda} \langle x_p - x_q, p \rangle. \tag{53}$$

Similarly,

$$\|x_p - x_q - \frac{q}{\lambda\|q\|}\|^2 \leq \frac{1}{\lambda^2},$$

which results in

$$\|q\| \cdot \|x_p - x_q\|^2 \leq \frac{2}{\lambda} \langle x_q - x_p, q \rangle. \tag{54}$$

Summing (53) and (54), one gets $(\|p\| + \|q\|)\|x_p - x_q\|^2 \leq \frac{2}{\lambda} \langle x_p - x_q, p - q \rangle$. Applying the Cauchy-Schwarz inequality completes the proof. \blacksquare

3. Polovinkin (1996) has discussed the smoothness of the support function on strongly convex sets. Here, we state the more general result and give a new proof.

Appendix I.

Proposition 19 *For arbitrary x , let $\ell(\cdot) := g(\cdot, y_x)$. Then $\nabla_x \ell(\cdot) \in \partial_x s(\cdot)$.*

Proof Consider any point $w \in \mathcal{X}$,

$$\begin{aligned} s(w) - s(x) &= g(w, y_w) - g(x, y_x) \\ &= g(w, y_w) - g(w, y_x) + g(w, y_x) - g(x, y_x) \geq 0 + g(w, y_x) - g(x, y_x) \\ &\geq \langle \partial_x g(\cdot, y_x), w - x \rangle = \langle \nabla_x \ell(\cdot), w - x \rangle \end{aligned} \quad (55)$$

where the first inequality is because that y_w is the best response to w , the second inequality is due to the convexity. The overall statement implies that $\nabla_x \ell(\cdot)$ is a subgradient of s at x . \blacksquare

We assume that

$$g(x, y) = a(x) + x^\top M y - b(y), \quad (56)$$

where $a(x)$ is any σ_x -strongly convex function of x , M is a matrix, and $b(y)$ is any σ_y -strongly convex function of y ⁴.

Proposition 20 *Assume the payoff function is in the form of (56) and that the norm is ℓ_2 norm (i.e. $\|\cdot\| = \|\cdot\|_2$). Given any points $w, z \in \mathcal{X}$, the gradient of $s(\cdot)$ satisfies $\|\nabla_w s(w) - \nabla_z s(z)\| \leq (\frac{\|M\|^2}{\sigma_y} + L)\|w - z\|$. This shows that $s(\cdot)$ is a $(\frac{\|M\|^2}{\sigma_y} + L)$ -smooth function.*

Proof

$$\begin{aligned} \|\nabla_w s(\cdot) - \nabla_z s(\cdot)\| &\stackrel{(a)}{=} \|\nabla_w g(\cdot, y_w) - \nabla_z g(\cdot, y_z)\| \\ &\stackrel{(b)}{\leq} \|\nabla_w g(\cdot, y_w) - \nabla_w g(\cdot, y_z)\| + \|\nabla_w g(\cdot, y_z) - \nabla_z g(\cdot, y_z)\| \\ &\stackrel{(c)}{\leq} \|\nabla_w g(\cdot, y_w) - \nabla_w g(\cdot, y_z)\| + L\|w - z\| \\ &= \|M(y_w - y_z)\| + L\|w - z\| \leq \|M\| \cdot \|y_w - y_z\| + L\|w - z\| \end{aligned} \quad (57)$$

where (a) is because of the gradient coincidence (Proposition 19), (b) is due to triangle inequality, and (c) is the result of the L -smoothness of $g(\cdot, \cdot)$, and the last inequality is a standard matrix inequality.

Now we are going to bound $\|y_w - y_z\|$. Recall that the y -player plays BESTRESPONSE, so $y_w = \arg \min_{y \in \mathcal{Y}} -\langle My, w \rangle + b(y)$ and $y_z = \arg \min_{y \in \mathcal{Y}} -\langle My, z \rangle + b(y)$. Since $b(\cdot)$ is a σ_y -strongly convex function with respect to $\|\cdot\|_2$, its conjugate $b^*(\cdot)$ is $\frac{1}{\sigma_y}$ -strongly smooth with respect to $\|\cdot\|_2$.

In the following, denote function $\psi_w(y) = -\langle My, w \rangle + b(y)$ so that $y_w = \arg \min_y \psi_w(y)$ and $y_z = \arg \min_y \psi_w(y) + \langle M^\top(w - z), y \rangle := \arg \min_y \psi_z(y)$. By definition of Fenchel-conjugate (c.f. Rockafellar (1996); Borwein and Lewis (2006)),

$$\psi_w^*(0) = \max_y \langle 0, y \rangle - \psi_w(y) = -\min_y \psi_w(y) = -\psi_w(y_w), \quad (58)$$

4. We note that some machine learning applications can be written into this form (see, e.g. Xu et al. (2005); Zhang and Xiao (2015)).

and

$$\nabla \psi_w^*(0) = y_w. \quad (59)$$

Similarly,

$$\psi_w^*(M^\top(z-w)) = \max_y \langle M^\top(z-w), y \rangle - \psi_w(y) = -\min_y (\langle M^\top(w-z), y \rangle + \psi_w(y)) = -\psi_z(y_z), \quad (60)$$

and

$$\nabla \psi_w^*(M^\top(z-w)) = y_z. \quad (61)$$

Therefore,

$$\|y_w - y_z\| = \|\nabla \psi^*(0) - \nabla \psi^*(M^\top(z-w))\| \leq \frac{1}{\sigma_b} \|0 - M^\top(z-w)\| = \frac{\|M\|}{\sigma_b} \|w - z\|. \quad (62)$$

Combining the above results, we have

$$\|\nabla_w s(\cdot) - \nabla_z s(\cdot)\| \leq \|M\| \|y_w - y_z\| + L \|w - z\| \leq \left(\frac{\|M\|}{\sigma_b} + L \right) \|w - z\|. \quad (63)$$

■

Appendix J.

Proposition 21 Assume that for all $x \in \mathcal{X}$, $\arg \min_{y \in \mathcal{Y}} g(x, y) \in \mathcal{Y}$. Also assume that the function $g(x, y)$ is σ_y -strongly concave in y and that the following smoothness condition holds for any given points $w, z \in \mathcal{X}$, $\|\nabla_{y_z} g(w, \cdot) - \nabla_{y_z} g(z, \cdot)\| \leq L \|w - z\|$. Then, $s(\cdot)$ is a $L(1 + \frac{2L}{\sigma_y})$ -smooth function.

Proof Following the same proof in (57), we have

$$\|\nabla_w s(\cdot) - \nabla_z s(\cdot)\| \leq L \|y_w - y_z\| + L \|w - z\|. \quad (64)$$

Now we need to bound $\|y_w - y_z\|$. Note that $-g(x, \cdot)$ is a σ_y -strongly convex function. Therefore, we have

$$(-g(w, y_z)) - (-g(w, y_w)) \geq \langle -\nabla_{y_w} g(w, \cdot), y_z - y_w \rangle + \frac{\sigma_y}{2} \|y_w - y_z\|^2 \geq \frac{\sigma_y}{2} \|y_w - y_z\|^2. \quad (65)$$

This means that

$$g(w, y_w) - g(w, y_z) \geq \frac{\sigma_y}{2} \|y_w - y_z\|^2.$$

From this, we can also conclude that ⁵

$$\|\nabla_{y_z} g(w, \cdot)\|^2 \geq \frac{\sigma_y}{2} (g(w, y_w) - g(w, y_z)). \quad (66)$$

5. $\sqrt{\frac{2}{\sigma_y} (g(w, y_w) - g(w, y_z))} \|\nabla_{y_z} g(w, \cdot)\| \geq \|y_w - y_z\| \|\nabla_{y_z} g(w, \cdot)\| \geq \langle y_z - y_w, -\nabla_{y_z} g(w, \cdot) \rangle \geq (-g(w, y_z)) - (-g(w, y_w)).$

Since y is unconstrained, $\nabla_{y_z} g(z, \cdot) = 0$, from the assumption that $\|\nabla_{y_z} g(w, \cdot) - \nabla_{y_z} g(z, \cdot)\| \leq L\|w - z\|$, we have

$$\|\nabla_{y_z} g(w, \cdot)\| \leq L\|w - z\|. \quad (67)$$

Putting it all together, we get:

$$\frac{\sigma_y}{2} \|y_w - y_z\|^2 \stackrel{(65)}{\leq} g(w, y_w) - g(w, y_z) \stackrel{(66)}{\leq} \frac{2}{\sigma_y} \|\nabla_{y_z} g(w, \cdot)\|^2 \stackrel{(67)}{\leq} \frac{2L^2}{\sigma_y} \|w - z\|^2 \quad (68)$$

So, $\|y_w - y_z\| \leq \frac{2L}{\sigma_y} \|w - z\|$. Substituting it back to (64) completes the proof. \blacksquare

Appendix K. Proof of Theorem 12

Theorem 12 Assume g satisfies any conditions such that the resulting $s(\cdot)$ is L -smooth and σ -strongly convex function for constants L and σ and moreover that $(\arg \min_{x \in \mathbb{R}^d} s(x)) \in \mathcal{X}$. Suppose we instantiate Algorithm 1, where x -player uses SC-AdaGrad (Algorithm 3) for loss functions $\alpha_t \ell_t(\cdot) = \alpha_t g(\cdot, y_t)$, y -player uses BESTRESPONSE, and the sequence α is defined as $\alpha_t := \|\nabla \ell_t(x_t)\|^{-2}$. Let $T \geq \frac{L}{\sigma} \log 3$, and let G be a constant such that $\|\nabla \ell_t\| \leq G$ for all t . Then the output $(\bar{x}_\alpha, \bar{y}_\alpha)$ is an ϵ -equilibrium of g where $\epsilon = O\left(\frac{G^2 T}{L} e^{-\frac{\sigma}{L} T}\right)$.

The proof of Theorem 12 is a slight modification of the proof of Theorem 3.2 in Levy (2017), combined with Proposition 19. As before, we let $y_x = \arg \max_{y \in \mathcal{Y}} g(x, y)$. Recall that $\ell_t(\cdot) = g(\cdot, y_{x_t})$ and note $\ell_t(x_t) = g(x_t, y_{x_t}) = s(x_t)$. Before proving the theorem, let us introduce some lemmas.

Lemma 22 For any L -smooth convex function $\ell(\cdot) : \mathbb{R}^d \mapsto \mathbb{R}$, if $x^* = \arg \min_{x \in \mathbb{R}^d} \ell(x)$, then

$$\|\nabla \ell(x)\|^2 \leq 2L(\ell(x) - \ell(x^*)), \quad \forall x \in \mathbb{R}^d.$$

Lemma 23 (Levy (2017)) For any non-negative real numbers $a_1, \dots, a_n \geq 1$,

$$\sum_{i=1}^n \frac{a_i}{\sum_{j=1}^i a_j} \leq 1 + \log \left(\sum_{i=1}^n a_i \right).$$

Lemma 24 (Hazan et al. (2007); Levy (2017)) Assume the loss function $\ell_t(\cdot) : \mathcal{K} \mapsto \mathbb{R}$ is θ_t -strongly convex. Setting $\eta_t = \left(\sum_{s=1}^t \theta_s\right)^{-1}$ in Algorithm 3, we get the following regret for any $x \in \mathcal{K}$:

$$\sum_{t=1}^T \ell_t(x_t) - \sum_{t=1}^T \ell_t(x) \leq \frac{1}{2} \sum_{t=1}^T \eta_t \|\nabla \ell_t(x_t)\|^2 \quad (69)$$

where $x_{t+1} = x_t - \eta_t \nabla \ell_t(x_t)$.

Lemma 25 Under the conditions in Theorem 12, $G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2} \geq \frac{1}{3} e^{\frac{\sigma}{L} T}$

Algorithm 3 Strongly-Convex Adaptive Gradient Descent (SC-AdaGrad) (Levy (2017))

```

1: for  $t = 1, 2, \dots, T$  do
2:   Play  $x_t \in \mathcal{K}$ 
3:   Receive a  $\theta_t$ -strongly convex loss function  $\ell_t(\cdot)$ 
4:   Set  $\eta_t = \frac{1}{\sum_{\tau=1}^T \theta_\tau}$ 
5:   Update  $x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta_t \nabla \ell_t(x_t))$ 
6: end for
    
```

Proof

$$T = \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^2}{\|\nabla \ell_t(x_t)\|^2} \stackrel{\text{Proposition 19}}{=} \sum_{t=1}^T \frac{\|\nabla s(x_t)\|^2}{\|\nabla \ell_t(x_t)\|^2} \stackrel{\text{Lemma 22}}{\leq} \sum_{t=1}^T \frac{2L}{\|\nabla \ell_t(x_t)\|^2} (s(x_t) - s(x^*)) \quad (70)$$

$$\leq \sum_{t=1}^T \frac{2L}{\|\nabla \ell_t(x_t)\|^2} (\ell_t(x_t) - \ell_t(x^*)), \quad (71)$$

where the last inequality follows from the fact that $s(x_t) = \ell_t(x_t)$ and $\ell_t(x) = g(x, y_t) \leq g(x, y_x) = s(x)$ for any x .

Now we define the following function, which is $\frac{\sigma}{\|\nabla \ell_t(x_t)\|^2}$ -strongly convex.

$$\tilde{\ell}_t(x) = \frac{1}{\|\nabla \ell_t(x_t)\|^2} \nabla \ell_t(x_t)^T x + \frac{\sigma}{2 \|\nabla \ell_t(x_t)\|^2} \|x - x_t\|^2 \quad (72)$$

Then we have

$$\begin{aligned}
 T &\leq \sum_{t=1}^T \frac{2L}{\|\nabla \ell_t(x_t)\|^2} (\ell_t(x_t) - \ell_t(x^*)) \stackrel{(a)}{\leq} \sum_{t=1}^T \frac{2L}{\|\nabla \ell_t(x_t)\|^2} (\nabla \ell_t(x_t)^T (x_t - x^*) - \frac{\sigma}{2} \|x_t - x^*\|^2) \\
 &= 2L \sum_{t=1}^T (\tilde{\ell}_t(x_t) - \tilde{\ell}_t(x^*)) \\
 &\stackrel{(b)}{\leq} \frac{L}{\sigma} \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^{-2}}{\sum_{\tau=1}^t \|\nabla \ell_\tau(x_\tau)\|^{-2}} \\
 &\stackrel{(c)}{\leq} \frac{L}{\sigma} \left(1 + \log \left(G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2} \right) \right)
 \end{aligned} \quad (73)$$

where (a) is from strong convexity, (b) is by applying Lemma 24 to the function $\tilde{\ell}_t(\cdot)$, and (c) is due to Lemma 23. Thus, we get

$$G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2} \geq \frac{1}{3} e^{\frac{\sigma}{L} T} \quad (74)$$

■

Proof [Proof of Theorem 12]

We consider the weighed loss setting in Algorithm 1: in each round, the loss function of the x player is $\alpha_t \ell_t(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$, where $\ell_t(\cdot) := g(\cdot, y_t)$. The y -learner, on the other hand, observes her own sequence of loss functions $\alpha_t h_t(\cdot) : \mathcal{Y} \rightarrow \mathbb{R}$, where $h_t(\cdot) := -g(x_t, \cdot)$. To continue, let us “re-define” the regret of the x -player in the weighted loss game: $\alpha\text{-REG}^x := \sum_{t=1}^T (\alpha_t \ell_t(x_t) - \alpha_t \ell_t(x^*))$, where $x^* = \arg \min_{x \in \mathcal{X}} s(x)$.

$$\begin{aligned}
 \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t g(x_t, y_t) &= \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t \ell_t(x_t) \\
 &= \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t \ell_t(x^*) + \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t (\ell_t(x_t) - \ell_t(x^*)) \\
 &= \frac{1}{\sum_{s=1}^T \alpha_s} \sum_{t=1}^T \alpha_t \ell_t(x^*) + \frac{\alpha\text{-REG}^x}{\sum_{s=1}^T \alpha_s} \\
 &\leq \sum_{t=1}^T \frac{1}{\sum_{s=1}^T \alpha_s} \alpha_t g(x^*, y_t) + \epsilon_T \\
 &\stackrel{(a)}{\leq} g\left(x^*, \sum_{t=1}^T \frac{1}{\sum_{s=1}^T \alpha_s} \alpha_t y_t\right) + \epsilon_T \tag{75} \\
 &\leq \max_{y \in \mathcal{Y}} g(x^*, y) + \epsilon_T \tag{76} \\
 &\stackrel{(b)}{=} s(x^*) + \epsilon_T \tag{77} \\
 &= \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} g(x, y) + \epsilon_T,
 \end{aligned}$$

where (a) is by Jensen, (b) is by the definition of the function $s(x)$. Then, by combining the above result with (28), one get the following result

Proposition 26 Denote $A_T = \sum_{s=1}^T \alpha_s$, ϵ_T the upper bound of $\frac{\alpha\text{-REG}^x}{A_T}$, and δ_T the upper bound of $\frac{\alpha\text{-REG}^y}{A_T}$. Then, $\bar{x}_T := \frac{\sum_{s=1}^T \alpha_s x_s}{A_T}$ satisfies

$$\max_{y \in \mathcal{Y}} g(\bar{x}_T, y) \leq V^* + \epsilon_T + \delta_T.$$

However, since the y -player plays best response, $\alpha\text{-REG}^y = 0$, we only need to consider $\frac{\alpha\text{-REG}^x}{A_T}$.

Here, we consider the weight to be $\alpha_t = \|\nabla \ell_t(x_t)\|^{-2}$.

$$\begin{aligned}
 \frac{\alpha\text{-REG}^x}{A_T} &:= \frac{\sum_{t=1}^T \alpha_t (\ell_t(x_t) - \ell_t(x^*))}{A_T} \stackrel{(a)}{\leq} \frac{1}{2\sigma A_T} \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^{-2}}{\sum_{\tau=1}^t \|\nabla \ell_t(x_\tau)\|^{-2}} \\
 &\stackrel{(b)}{\leq} \frac{G^2 \left(1 + \log \left(G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2}\right)\right)}{2\sigma G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2}} \stackrel{(c)}{\leq} \frac{3G^2 \left(1 + \left(\frac{\sigma T}{L}\right)\right)}{2\sigma e^{\frac{\sigma}{L}T}} = O\left(\frac{G^2 T}{L} e^{-\frac{\sigma}{L}T}\right) \tag{78}
 \end{aligned}$$

where (a) has be shown in (73), (b) is by Lemma 23, (c) is by (74) and the fact that $\frac{1+\log z}{z}$ is monotonically decreasing for $z \geq 1$.

Next, we relate the weighted regret to the game

$$\begin{aligned} \frac{\alpha\text{-REG}^x}{A_T} &:= \frac{\sum_{t=1}^T \alpha_t (\ell_t(x_t) - \ell_t(x^*))}{A_T} = \frac{\sum_{t=1}^T \alpha_t (s(x_t) - g(x^*, y_t))}{A_T} \\ &\stackrel{(a)}{\geq} h(\bar{x}_T) - g(x^*, \bar{y}_T) = g(\bar{x}_T, y_{\bar{x}_T}) - g(x^*, \bar{y}_T), \end{aligned} \quad (79)$$

where in (a) we use Jensen for both sums. Thus, our final result is the following,

$$g(\bar{x}_T, y_{\bar{x}_T}) - g(x^*, \bar{y}_T) \leq O\left(\frac{G^2 T}{L} e^{-\frac{\sigma}{L} T}\right), \quad (80)$$

which also implies

$$\begin{aligned} g(\bar{x}_T, y_{\bar{x}_T}) - \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} g(x, y) &= g(\bar{x}_T, y_{\bar{x}_T}) - g(x^*, y_{x^*}) \stackrel{(a)}{\leq} g(\bar{x}_T, y_{\bar{x}_T}) - g(x^*, \bar{y}_T) \\ &\stackrel{(80)}{\leq} O\left(\frac{G^2 T}{L} e^{-\frac{\sigma}{L} T}\right), \end{aligned} \quad (81)$$

where (a) is because of $g(x^*, \bar{y}_T) \geq g(x^*, y_{x^*})$ by definition of best response.

But, we also note that $g(\bar{x}_T, y_{\bar{x}_T}) = s(\bar{x}_T) \geq \min_{x \in \mathcal{X}} s(x) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} g(x, y)$. This shows that $g(\bar{x}_T, y_{\bar{x}_T})$ is an $O(\exp(-T))$ approximate saddle point solution, so we have completed the proof. ■

Appendix L. Proof of Theorem 13

Theorem 13 Consider the FW game in which $g(x, y) = -x^\top y + f^*(x)$. Suppose that $f(\cdot)$ is a L -smooth convex function and that \mathcal{Y} is a λ -strongly convex set. Also assume that the gradients of the f in \mathcal{Y} are bounded away from 0, i.e., $\max_{y \in \mathcal{Y}} \|\nabla f(y)\| \geq B$. Let $T \geq \frac{L}{\lambda B} \log 3$. Then, there exists a FW-like algorithm that has $O(\exp(-\frac{\lambda B}{L} T))$ rate.

In this section, we give a proof of Theorem 13. Let us first restate the theorem with the explicit algorithm. Note that in the algorithm that we describe below the weights α_t are not predefined but rather depend on the queries of the algorithm. These adaptive weights are explicitly defined in Algorithm 4 which is used by the x -player. The y -player plays best response.

Algorithm 4 Strongly-Convex Adaptive Follow-the-Leader (SC-AFTL)

- 1: **for** $t = 1, 2, \dots, T$ **do**
 - 2: Play $x_t \in \mathcal{K}$
 - 3: Receive a strongly convex loss function $\alpha_t \ell_t(\cdot)$ with $\alpha_t = \frac{1}{\|\nabla \ell_t(x_t)\|^2}$.
 - 4: Update $x_{t+1} = \min_{x \in \mathcal{X}} \sum_{s=1}^t \alpha_s \ell_s(x)$
 - 5: **end for**
-

Before proving the theorem, let us focus on the strategy of the x -player which we describe in Algorithm 4. This algorithm is equivalent to performing FTL updates over the following loss sequence:

$$\left\{ \tilde{\ell}_t(x) = \alpha_t \ell_t(x) \right\}_{t=1}^T$$

And in the case of FW game, the L -smoothness of f implies that each $\tilde{\ell}_t(x)$ is $\frac{1/L}{\|\nabla \ell_t(x_t)\|^2}$ -strongly-convex. Thus Lemma 8 implies the following holds for any $x \in \mathcal{X}$:

$$\sum_{t=1}^T \tilde{\ell}_t(x_t) - \sum_{t=1}^T \tilde{\ell}_t(x) \leq \frac{L}{2} \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^{-2}}{\sum_{s=1}^t \|\nabla \ell_s(x_s)\|^{-2}}. \quad (82)$$

Proof [Proof of Theorem 13] Since the y -player plays best response, $\alpha\text{-REG}^y = 0$, we only need to show that $\alpha\text{-REG}^x \leq O(\exp(-\frac{\lambda B}{L}T))$, which we do next.

We start by showing that $s(x) = \max_{y \in \mathcal{Y}} -x^\top y + f^*(x)$ is a smooth function. We have that

$$\begin{aligned} \|\nabla_w s(\cdot) - \nabla_z s(\cdot)\| &= \|\arg \max_{y \in \mathcal{Y}} (-w^\top y + f^*(w)) - \arg \max_{y \in \mathcal{Y}} (-z^\top y + f^*(z))\| \\ &= \|\arg \max_{y \in \mathcal{Y}} (-w^\top y) - (\arg \max_{y' \in \mathcal{Y}} -z^\top y')\| \leq \frac{2\|w - z\|}{\lambda(\|w\| + \|z\|)} \leq \frac{\|w - z\|}{\lambda B}, \end{aligned} \quad (83)$$

where the second to last inequality uses Lemma 18 regarding λ -strongly convex sets, and the last inequality is by assuming the gradient of $\|\nabla f(\cdot)\| \geq B$ and the fact that $w, z \in \mathcal{X}$ are gradients of f . This shows that $s(\cdot)$ is a smooth function with smoothness constant $L' = \frac{1}{\lambda B}$.

$$T = \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^2}{\|\nabla \ell_t(x_t)\|^2} \stackrel{\text{Proposition 19}}{=} \sum_{t=1}^T \frac{\|\nabla s(x_t)\|^2}{\|\nabla \ell_t(x_t)\|^2} \stackrel{\text{Lemma 22}}{\leq} \sum_{t=1}^T \frac{L'}{\|\nabla \ell_t(x_t)\|^2} (s(x_t) - s(x^*)) \quad (84)$$

$$\leq \sum_{t=1}^T \frac{L'}{\|\nabla \ell_t(x_t)\|^2} (\ell_t(x_t) - \ell_t(x^*)), \quad (85)$$

where the last inequality follows from the fact that $s(x_t) = \ell_t(x_t)$ and $\ell_t(x) = g(x, y_t) \leq g(x, y_x) = s(x)$ for any x . We can apply Lemma 22 because $\arg \min_{x \in \mathbb{R}^d} \{\max_{y \in \mathcal{Y}} -x^\top y + f^*(x)\} \in \mathcal{X}$, as \mathcal{X} is the gradient space.

To continue, we have

$$\begin{aligned} T &\leq \sum_{t=1}^T \frac{L'}{\|\ell_t(x_t)\|^2} (\ell_t(x_t) - \ell_t(x^*)) \\ &\stackrel{(a)}{=} \sum_{t=1}^T L' (\tilde{\ell}_t(x_t) - \tilde{\ell}_t(x^*)) \\ &\stackrel{(b)}{\leq} \frac{L \cdot L'}{2} \sum_{t=1}^T \frac{\|\nabla \ell_t(x_t)\|^{-2}}{\sum_{s=1}^t \|\nabla \ell_s(x_t)\|^{-2}} \\ &\stackrel{(c)}{\leq} \frac{L \cdot L'}{2} \left(1 + \log(G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2}) \right), \end{aligned}$$

where (a) is by the definition of $\tilde{\ell}_t(\cdot)$, and (b) is shown using (82) with strong convexity parameter of $\ell_t(\cdot)$ being $\frac{1}{L}$, and (c) is by Lemma 23,

$$\sum_{t=1}^T \frac{\|\ell_t(x_t)\|^{-2}}{\sum_{s=1}^t \|\ell_s(x_s)\|^{-2}} = \sum_{t=1}^T \frac{G^2 \|\ell_t(x_t)\|^{-2}}{\sum_{s=1}^t G^2 \|\ell_s(x_s)\|^{-2}} \leq 1 + \log(G^2 \sum_{t=1}^T \|\ell_t(x_t)\|^{-2}).$$

Thus, we get

$$G^2 \sum_{t=1}^T \|\nabla \ell_t(x_t)\|^{-2} = O(e^{\frac{1}{L \cdot L'} T}) = O(e^{\frac{\lambda B}{L} T}). \quad (86)$$

The rest of the proof is the same as in (79). So, we can get that $g(\bar{x}_T, y_{\bar{x}_T})$ is an $O(\exp(-\frac{\lambda B}{L} T))$ approximate saddle point solution. This completes the proof. ■

Appendix M. Supremum of strongly convex functions is strongly convex

Lemma 27 $s(x) = \sup_{y \in \mathcal{Y}} g(x, y)$ is σ -strongly convex if $g(\cdot, y)$ is σ -strongly convex for any $y \in \mathcal{Y}$.

Proof For any $\theta \in [0, 1]$ and any $w, z \in X$,

$$\begin{aligned} s(\theta w + (1 - \theta)z) &= g(\theta w + (1 - \theta)z, \hat{y}) \\ &\leq \theta g(w, \hat{y}) + (1 - \theta)g(z, \hat{y}) - \frac{\sigma}{2}\theta(1 - \theta)\|w - z\|^2 \\ &\leq \theta s(w) + (1 - \theta)s(z) - \frac{\sigma}{2}\theta(1 - \theta)\|w - z\|^2, \end{aligned} \quad (87)$$

where the first equality holds for a specific $\hat{y} \in \mathcal{Y}$. ■