Optimization, Learning, and Games with Predictable Sequences

Alexander Rakhlin
University of Pennsylvania

Karthik Sridharan
University of Pennsylvania

Abstract

We provide several applications of Optimistic Mirror Descent, an online learning algorithm based on the idea of predictable sequences. First, we recover the Mirror Prox algorithm, prove an extension to Hölder-smooth functions, and apply the results to saddle-point type problems. Second, we prove that a version of Optimistic Mirror Descent (which has a close relationship to the Exponential Weights algorithm) can be used by two strongly-uncoupled players in a finite zero-sum matrix game to converge to the minimax equilibrium at the rate of $O((\log T)/T)$. This addresses a question of Daskalakis et al [5]. Further, we consider a partial information version of the problem. We then apply the results to approximate convex programming and show a simple algorithm for the approximate Max-Flow problem.

1 Introduction

Recently, no-regret algorithms have received increasing attention in a variety of communities, including theoretical computer science, optimization, and game theory [3, 1]. The wide applicability of these algorithms is arguably due to the black-box regret guarantees that hold for arbitrary sequences. However, such regret guarantees can be loose if the sequence being encountered is not "worst-case". Such a reduction in "arbitrariness" of the sequence can arise from the particular structure of the problem at hand, and should be exploited. For instance, in some applications of online methods, the sequence comes from an additional computation done by the learner, thus being far from arbitrary.

One way to formally capture the partially benign nature of data is through a notion of predictable sequences [9]. We exhibit applications of this idea in several domains. First, we show that the Mirror Prox method [7], designed for optimizing non-smooth structured saddle-point problems, can be viewed as an instance of the predictable sequence approach. Predictability in this case is due precisely to smoothness of the inner optimization part and the saddle-point structure of the problem. We extend the results to Hölder-smooth functions, interpolating between the case of well-predictable gradients and "unpredictable" gradients.

Second, we address the question raised in [5] about existence of "simple" algorithms that converge at the rate of $O(T^{-1})$ when employed in an uncoupled manner by players in a zero-sum finite matrix game, yet maintain the usual $O(T^{-1/2})$ rate against arbitrary sequences. Moreover, the prior knowledge of whether the other player is collaborating is not required, as the algorithm is fully adaptive. Here, the additional predictability comes from the fact that both players attempt to converge to the minimax value. We also tackle a partial information version of the problem where the player only has access to the payoffs of the mixed actions played by the two players on each round rather than the entire vector.

Our third application is to convex programming: optimization of a linear function subject to convex constraints. This problem often arises in theoretical computer science, and we show that the idea of predictable sequences can be used here too. We exhibit a simple algorithm for ϵ-approximate max-flow for a graph with d edges in time $\tilde{O}(d^{3/2}/\epsilon)$, a performance previously obtained through a relatively involved procedure [6].

2 Online Learning with Predictable Gradient Sequences

Let us describe the online convex optimization (OCO) problem and the basic algorithm studied in [9, 4]. Let $\mathcal{F}$ be a convex set of moves of the learner. On round $t = 1, \ldots, T$, the learner makes a prediction $f_t \in \mathcal{F}$ and observes a convex function G_t on $\mathcal{F}$. The objective is to keep regret $\frac{1}{T}\sum_{t=1}^{T} G_t(f_t) - G_t(f^*)$ small for any $f^* \in \mathcal{F}$. Let $\mathcal{R}$ be a 1-strongly convex function w.r.t. some norm $\|\cdot\|$ on $\mathcal{F}$, and let $g_0 = \arg\min_{g\in\mathcal{F}} \mathcal{R}(g)$. Suppose that at the beginning of every round t, the learner has access to M_t, a vector computable based on the past observations or side information. In this paper we study the **Optimistic Mirror Descent** algorithm, defined by the interleaved sequence

$$f_t = \underset{f\in\mathcal{F}}{\operatorname{argmin}}\ \eta_t \langle f, M_t\rangle + D_{\mathcal{R}}(f, g_{t-1})\ ,\ g_t = \underset{g\in\mathcal{F}}{\operatorname{argmin}}\ \eta_t \langle g, \nabla G_t(f_t)\rangle + D_{\mathcal{R}}(g, g_{t-1}) \tag{1}$$

where $D_{\mathcal{R}}$ is the Bregman Divergence with respect to $\mathcal{R}$ and $\{\eta_t\}$ is a sequence of step sizes that can be chosen adaptively based on the sequence observed so far. The method adheres to the OCO protocol since M_t is available at the beginning of round t, and $\nabla G_t(f_t)$ becomes available after the prediction f_t is made. The sequence $\{f_t\}$ will be called primary, while $\{g_t\}$ – secondary. This method was proposed in [4] for $M_t = \nabla G_{t-1}(f_{t-1})$, and the following lemma is a straightforward extension of the result in [9] for general M_t:

Lemma 1. *Let $\mathcal{F}$ be a convex set in a Banach space $\mathcal{B}$. Let $\mathcal{R} : \mathcal{B} \mapsto \mathbb{R}$ be a 1-strongly convex function on $\mathcal{F}$ with respect to some norm $\|\cdot\|$, and let $\|\cdot\|_*$ denote the dual norm. For any fixed step-size η, the Optimistic Mirror Descent Algorithm yields, for any $f^* \in \mathcal{F}$,*

$$\begin{aligned}\sum_{t=1}^{T} G_t(f_t) - G_t(f^*) &\le \sum_{t=1}^{T} \langle f_t - f^*, \nabla_t\rangle \\ &\le \eta^{-1}R^2 + \sum_{t=1}^{T} \|\nabla_t - M_t\|_* \|g_t - f_t\| - \frac{1}{2\eta}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g_{t-1} - f_t\|^2\right) \end{aligned} \tag{2}$$

where $R \ge 0$ is such that $D_{\mathcal{R}}(f^*, g_0) \le R^2$ and $\nabla_t = \nabla G_t(f_t)$.

When applying the lemma, we will often use the simple fact that

$$\|\nabla_t - M_t\|_* \|g_t - f_t\| = \inf_{\rho>0}\left\{\frac{\rho}{2}\|\nabla_t - M_t\|_*^2 + \frac{1}{2\rho}\|g_t - f_t\|^2\right\}\ . \tag{3}$$

In particular, by setting $\rho = \eta$, we obtain the (unnormalized) regret bound of $\eta^{-1}R^2 + (\eta/2)\sum_{t=1}^{T}\|\nabla_t - M_t\|_*^2$, which is $R\sqrt{2\sum_{t=1}^{T}\|\nabla_t - M_t\|_*^2}$ by choosing η optimally. Since this choice is not known ahead of time, one may either employ the doubling trick, or choose the step size adaptively:

Corollary 2. *Consider step size $\eta_t = R_{\max}\min\left\{\left(\sqrt{\sum_{i=1}^{t-1}\|\nabla_i - M_i\|_*^2} + \sqrt{\sum_{i=1}^{t-2}\|\nabla_i - M_i\|_*^2}\right)^{-1}, 1\right\}$ with $R_{\max}^2 = \sup_{f,g\in\mathcal{F}} D_{\mathcal{R}}(f,g)$. Then regret of the Optimistic Mirror Descent algorithm is upper bounded by $3.5R_{\max}\left(\sqrt{\sum_{t=1}^{T}\|\nabla_t - M_t\|_*^2} + 1\right)/T$.*

These results indicate that tighter regret bounds are possible if one can guess the next gradient ∇_t by computing M_t. One such case arises in *offline* optimization of a smooth function, whereby the previous gradient turns out to be a good proxy for the next one. More precisely, suppose we aim to optimize a function $G(f)$ whose gradients are Lipschitz continuous: $\|\nabla G(f) - \nabla G(g)\|_* \le H\|f - g\|$ for some $H > 0$. In this optimization setting, no guessing of M_t is needed: we may simply query the oracle for the gradient and set $M_t = \nabla G(g_{t-1})$. The Optimistic Mirror Descent then becomes

$$f_t = \underset{f\in\mathcal{F}}{\operatorname{argmin}}\ \eta_t \langle f, \nabla G(g_{t-1})\rangle + D_{\mathcal{R}}(f, g_{t-1})\ ,\ g_t = \underset{g\in\mathcal{F}}{\operatorname{argmin}}\ \eta_t \langle g, \nabla G(f_t)\rangle + D_{\mathcal{R}}(g, g_{t-1})$$

which can be recognized as the Mirror Prox method, due to Nemirovski [7]. By smoothness, $\|\nabla G(f_t) - M_t\|_* = \|\nabla G(f_t) - \nabla G(g_{t-1})\|_* \le H\|f_t - g_{t-1}\|$, and Lemma 1 with Eq. (3) and $\rho = \eta = 1/H$ immediately yields a bound

$$\sum_{t=1}^{T} G(f_t) - G(f^*) \le HR^2,$$

which implies that the average $\bar{f}_T = \frac{1}{T}\sum_{t=1}^{T} f_t$ satisfies $G(\bar{f}_T) - G(f^*) \le HR^2/T$, a known bound for Mirror Prox. We now extend this result to arbitrary α-Hölder smooth functions, that is convex functions G such that $\|\nabla G(f) - \nabla G(g)\|_* \le H\|f - g\|^{\alpha}$ for all $f, g \in \mathcal{F}$.

Lemma 3. *Let $\mathcal{F}$ be a convex set in a Banach space $\mathcal{B}$ and let $\mathcal{R} : \mathcal{B} \mapsto \mathbb{R}$ be a 1-strongly convex function on $\mathcal{F}$ with respect to some norm $\|\cdot\|$. Let G be a convex α-Hölder smooth function with constant $H > 0$ and $\alpha \in [0,1]$. Then the average $\bar{f}_T = \frac{1}{T}\sum_{t=1}^{T} f_t$ of the trajectory given by Optimistic Mirror Descent Algorithm enjoys*

$$G(\bar{f}_T) - \inf_{f\in\mathcal{F}} G(f) \le \frac{8HR^{1+\alpha}}{T^{\frac{1+\alpha}{2}}}$$

where $R \ge 0$ is such that $\sup_{f\in\mathcal{F}} D_{\mathcal{R}}(f, g_0) \le R$.

This result provides a smooth interpolation between the $T^{-1/2}$ rate at $\alpha = 0$ (that is, no predictability of the gradient is possible) to the T^{-1} rate when the smoothness structure allows for a dramatic speed up with a very simple modification of the original Mirror Descent.

3 Structured Optimization

In this section we consider the structured optimization problem

$$\underset{f\in\mathcal{F}}{\operatorname{argmin}}\ G(f)$$

where $G(f)$ is of the form $G(f) = \sup_{x\in\mathcal{X}} \phi(f, x)$ with $\phi(\cdot, x)$ convex for every $x \in \mathcal{X}$ and $\phi(f, \cdot)$ is concave for every $f \in \mathcal{F}$. Both $\mathcal{F}$ and $\mathcal{X}$ are assumed to be convex sets. While G itself need not be smooth, it has been recognized that the structure can be exploited to improve rates of optimization if the function ϕ is smooth [8]. From the point of view of online learning, we will see that the optimization problem of the saddle point type can be solved by playing two online convex optimization algorithms against each other (henceforth called Players I and II).

Specifically, assume that Player I produces a sequence $f_1, \dots, f_T$ by using a regret-minimization algorithm, such that

$$\frac{1}{T}\sum_{t=1}^{T} \phi(f_t, x_t) - \inf_{f\in\mathcal{F}} \frac{1}{T}\sum_{t=1}^{T} \phi(f, x_t) \le \mathrm{Rate}^1(x_1, \dots, x_T) \tag{4}$$

and Player II produces $x_1, \dots, x_T$ with

$$\frac{1}{T}\sum_{t=1}^{T} (-\phi(f_t, x_t)) - \inf_{x\in\mathcal{X}} \frac{1}{T}\sum_{t=1}^{T} (-\phi(f_t, x)) \le \mathrm{Rate}^2(f_1, \dots, f_T)\ . \tag{5}$$

By a standard argument (see e.g. [?]),

$$\inf_f \frac{1}{T}\sum_{t=1}^{T} \phi(f, x_t) \le \inf_f \phi(f, \bar{x}_T) \le \sup_x \inf_f \phi(f, x) \le \inf_f \sup_x \phi(f, x) \le \sup_x \phi\left(\bar{f}_T, x\right) \le \sup_x \frac{1}{T}\sum_{t=1}^{T} \phi(f_t, x)$$

where $\bar{f}_T = \frac{1}{T}\sum_{t=1}^{T} f_t$ and $\bar{x}_T = \frac{1}{T}\sum_{t=1}^{T} x_t$. By adding (4) and (5), we have

$$\sup_{x\in\mathcal{X}} \frac{1}{T}\sum_{t=1}^{T} \phi(f_t, x) - \inf_{f\in\mathcal{F}} \frac{1}{T}\sum_{t=1}^{T} \phi(f, x_t) \le \mathrm{Rate}^1(x_1, \dots, x_T) + \mathrm{Rate}^2(f_1, \dots, f_T) \tag{6}$$

which sandwiches the previous sequence of inequalities up to the sum of regret rates and implies near-optimality of $\bar{f}_T$ and $\bar{x}_T$.

Lemma 4. *Suppose both players employ the Optimistic Mirror Descent algorithm with, respectively, predictable sequences M_t^1 and M_t^2, 1-strongly convex functions $\mathcal{R}_1$ on $\mathcal{F}$ (w.r.t. $\|\cdot\|_{\mathcal{F}}$) and $\mathcal{R}_2$ on $\mathcal{X}$ (w.r.t. $\|\cdot\|_{\mathcal{X}}$), and fixed learning rates η and η'. Let $\{f_t\}$ and $\{x_t\}$ denote the primary sequences of the players while let $\{g_t\},\{y_t\}$ denote the secondary. Then for any $\alpha,\beta>0$,*

$$
\begin{aligned}
&\sup_{x\in\mathcal{X}}\phi\left(\bar{f}_T,x\right)-\inf_{f\in\mathcal{F}}\sup_{x\in\mathcal{X}}\phi(f,x) \qquad (7)\\
&\le \frac{R_1^2}{\eta}+\frac{\alpha}{2}\sum_{t=1}^{T}\|\nabla_f\phi(f_t,x_t)-M_t^1\|_{\mathcal{F}^*}^2+\frac{1}{2\alpha}\sum_{t=1}^{T}\|g_t-f_t\|_{\mathcal{F}}^2-\frac{1}{2\eta}\sum_{t=1}^{T}\left(\|g_t-f_t\|_{\mathcal{F}}^2+\|g_{t-1}-f_t\|_{\mathcal{F}}^2\right)\\
&\quad+\frac{R_2^2}{\eta'}+\frac{\beta}{2}\sum_{t=1}^{T}\|\nabla_x\phi(f_t,x_t)-M_t^2\|_{\mathcal{X}^*}^2+\frac{1}{2\beta}\sum_{t=1}^{T}\|y_t-x_t\|_{\mathcal{X}}^2-\frac{1}{2\eta'}\sum_{t=1}^{T}\left(\|y_t-x_t\|_{\mathcal{X}}^2+\|y_{t-1}-x_t\|_{\mathcal{X}}^2\right)
\end{aligned}
$$

where $R_1^2=\sup_f D_{\mathcal{R}_1}(f,g_0)$, $R_2^2=\sup_x D_{\mathcal{R}_1}(x,y_0)$, and $\bar{f}_T=\frac{1}{T}\sum_{t=1}^T f_t$.

The proof of Lemma 7 is immediate from Lemma 1. We obtain the following corollary:

Corollary 5. *Suppose $\phi:\mathcal{F}\times\mathcal{X}\mapsto\mathbb{R}$ is Hölder smooth in the following sense:*

$$
\begin{aligned}
&\|\nabla_f\phi(f,x)-\nabla_f\phi(g,x)\|_{\mathcal{F}^*}\le H_1\|f-g\|_{\mathcal{F}}^{\alpha},\quad \|\nabla_f\phi(f,x)-\nabla_f\phi(f,y)\|_{\mathcal{F}^*}\le H_2\|x-y\|_{\mathcal{X}}^{\alpha'}\\
\text{and}\quad &\|\nabla_x\phi(f,x)-\nabla_x\phi(g,x)\|_{\mathcal{X}^*}\le H_4\|f-g\|_{\mathcal{F}}^{\beta},\quad \|\nabla_x\phi(f,x)-\nabla_x\phi(f,y)\|_{\mathcal{X}^*}\le H_3\|x-y\|_{\mathcal{X}}^{\beta'}.
\end{aligned}
$$

Let $\gamma=\min\{\alpha,\alpha',\beta,\beta'\}$, $H=\max\{H_1,H_2,H_3,H_4\}$. Suppose both players employ Optimistic Mirror Descent with $M_t^1=\nabla_f\phi(g_{t-1},y_{t-1})$ and $M_t^2=\nabla_x\phi(g_{t-1},y_{t-1})$, where $\{g_t\}$ and $\{y_t\}$ are the secondary sequences updated by the two algorithms, and with step sizes $\eta=\eta'=(R_1^2+R_2^2)^{\frac{1-\gamma}{2}}(2H)^{-1}\left(\frac{T}{2}\right)^{\frac{\gamma-1}{2}}$. Then

$$
\sup_{x\in\mathcal{X}}\phi\left(\bar{f}_T,x\right)-\inf_{f\in\mathcal{F}}\sup_{x\in\mathcal{X}}\phi(f,x)\le\frac{4H(R_1^2+R_2^2)^{\frac{1+\gamma}{2}}}{T^{\frac{1+\gamma}{2}}} \qquad (8)
$$

As revealed in the proof of this corollary, the negative terms in (7), that come from an upper bound on regret of Player I, in fact contribute to cancellations with positive terms in regret of Player II, and vice versa. Such coupling of the two regret upper bounds for the players can be seen as leading to faster rates under the appropriate assumptions, and this idea will be exploited to a great extent in the proofs of the next section.

4 Zero-sum Game and Uncoupled Dynamics

The notions of a zero-sum matrix game and a minimax equilibrium are arguably the most basic and important notions of game theory. The tight connection between linear programming and minimax equilibrium suggests that there might be simple dynamics that can lead the two players of the game to eventually converge to the equilibrium value. Existence of such simple or natural dynamics is of interest in behavioral economics, where one asks how agents can discover static solution concepts of the game iteratively and without extensive communication.

More formally, let $A\in[-1,1]^{n\times m}$ be a matrix with bounded entries. The two players aim to find a pair of near-optimal mixed strategies $(\bar{f},\bar{x})\in\Delta_n\times\Delta_m$ such that $\bar{f}^\top A\bar{x}$ is close to the minimax value $\min_{f\in\Delta_n}\max_{x\in\Delta_m}f^\top Ax$, where Δ_n is the probability simplex over n actions. Of course, this is a particular form of the saddle point problem considered in the previous section, with $\phi(f,x)=f^\top Ax$. It is well-known (and follows immediately from (6)) that the players can achieve the goal of computing near-optimal strategies by simply playing no-regret algorithms [**?**]. More precisely, on round t, the players I and II "predict" the mixed strategies f_t and x_t and observe Ax_t and $f_t^\top A$, respectively. While black-box regret minimization algorithms, such as Exponential Weights, immediately yield $O(T^{-1/2})$ convergence rates, Daskalakis et al [5] asked whether faster methods exist. To make the problem well-posed, it is required that the two players are *strongly uncoupled*: neither A nor the number of available actions of the opponent is known to each player, no "funny bit arithmetic" is allowed, and memory storage of each player allows only for constant number of payoff vectors. The authors of [5] exhibited a near-optimal algorithm that, if used by both players, yields a pair of mixed strategies that constitutes an $O\left(\frac{\log(m+n)(\log T+(\log(m+n))^{3/2})}{T}\right)$-approximate minimax equilibrium. Furthermore, the method has a regret bound of the same order as Exponential Weights when faced with an arbitrary sequence. The method in [5] is an application of the excessive gap technique of Nesterov, and requires careful choreography and interleaving of rounds between the two non-communicating players. The authors, therefore, asked whether a simple algorithm (e.g. a modification of Exponential Weights) can in fact achieve the same result. We answer this in the affirmative. While a direct application of Mirror Prox does not yield the result (and also does not provide strong decoupling), below we show that a modification of Optimistic Mirror Descent achieves the goal. Furthermore, by choosing the step size adaptively, the same method guarantees the typical $O(T^{-1/2})$ regret if not faced with a compliant player.

In Section 4.1, we analyze the "first-order information" version of the problem, as described above: upon playing the respective mixed strategies f_t and x_t on round t, Player I observes Ax_t and Player II observes $f_t^\intercal A$. Then, in Section 4.2, we consider an interesting extension to partial information, whereby the players submit their moves f_t, x_t but only observe the real value $f_t^\intercal A x_t$. Recall that in both cases the matrix A is not known to the players.

4.1 First-Order Information

Consider the following simple algorithm. Initialize $g'_0 \in \Delta_n$ and $y'_0 \in \Delta_m$ to be uniform distributions and proceed as follows:

```
On round t, Player I performs
   Play      f_t and observe Ax_t
   Update    g_t(i) ∝ g'_{t-1}(i) exp{-η_t [Ax_t]_i},     g'_t = (1 - 1/T^2) g_t + (1/(nT^2)) 1_n
             f_{t+1}(i) ∝ g'_t(i) exp{-η_{t+1} [Ax_t]_i}
while simultaneously Player II performs
   Play      x_t and observe f_t^T A
   Update    y_t(i) ∝ y'_{t-1}(i) exp{-η'_t [f_t^T A]_i},     y'_t = (1 - 1/T^2) y_t + (1/(nT^2)) 1_m
             x_{t+1}(i) ∝ y'_t(i) exp{-η'_{t+1} [f_t^T A]_i}
```

Here, $\mathbf{1}_n \in \mathbb{R}^n$ is a vector of all ones. Other than the "mixing in" of the uniform distribution, the algorithm for both players is simply the Optimistic Mirror Descent with the (negative) entropy function. In fact, the step of mixing in the uniform distribution is only needed when some coordinate of g_t (resp., y_t) is smaller than $1/(nT)$. Furthermore, this step is also not needed if none of the players deviate from the prescribed method. In such a case, the resulting algorithm is simply the constant step-size Exponential Weights $f_t(i) \propto \exp\{-\eta \sum_{s=1}^{t-2}[Ax_{s-1}]_i + 2\eta[Ax_{t-1}]_i\}$, but with a factor 2 in front of the *latest* loss vector!

Proposition 6. *Let $A \in [-1,1]^{n\times m}$, $\mathcal{F} = \Delta_n$, $\mathcal{X} = \Delta_m$. If both players use above algorithm with, respectively, $M_t^1 = Ax_{t-1}$ and $M_t^2 = f_{t-1}^\intercal A$, and the adaptive step sizes*

$$\eta_t = \min\left\{\log(nT)\left(\sqrt{\textstyle\sum_{i=1}^{t-1}\|Ax_i - Ax_{i-1}\|_*^2} + \sqrt{\textstyle\sum_{i=1}^{t-2}\|Ax_i - Ax_{i-1}\|_*^2}\right)^{-1}, \frac{1}{11}\right\}$$

and

$$\eta'_t = \min\left\{\log(mT)\left(\sqrt{\textstyle\sum_{i=1}^{t-1}\|f_i^\intercal A - f_{i-1}^\intercal A\|_*^2} + \sqrt{\textstyle\sum_{i=1}^{t-2}\|f_i^\intercal A - f_{i-1}^\intercal A\|_*^2}\right)^{-1}, \frac{1}{11}\right\}$$

respectively, then the pair $(\bar{f}_T, \bar{x}_T)$ is an $O\left(\frac{\log m+\log n+\log T}{T}\right)$-approximate minimax equilibrium. Furthermore, if only one player (say, Player I) follows the above algorithm, her regret against any sequence $x_1,\dots,x_T$ of plays is

$$O\left(\frac{\log(nT)}{T}\left(\sqrt{\sum_{t=1}^{T}\|Ax_t - Ax_{t-1}\|_*^2} + 1\right)\right). \tag{9}$$

In particular, this implies the worst-case regret of $O\left(\frac{\log(nT)}{\sqrt{T}}\right)$ *in the general setting of online linear optimization.*

We remark that (9) can give intermediate rates for regret in the case that the second player deviates from the prescribed strategy but produces "stable" moves. For instance, if the second player follows a mirror descent algorithm (or Follow the Regularized Leader / Exponential Weights method) with step size η, one can typically show stability $\|x_t - x_{t-1}\| = O(\eta)$. In this case, (9) yields the rate $O\left(\frac{\eta \log T}{\sqrt{T}}\right)$ for the first player. A typical setting of $\eta = O(T^{-1/2})$ for the second player still ensures the $O(\log T/T)$ regret for the first player.

Let us finish with a technical remark. The reason for the extra step of "mixing in" the uniform distribution stems from the goal of having an adaptive method that still attains $O(T^{-1/2})$ regret if the other player deviated from using the algorithm. If one is only interested in the dynamics when both players cooperate, this step is not necessary, and in this case the extraneous $\log T$ factor disappears from the above bound, leading to the $O\left(\frac{\log n + \log m}{T}\right)$ convergence. On the technical side, the need for the extra step is the following. The adaptive step size result of Corollary 2 involves the term $R^2_{\max} = \sup_g D_{\mathcal{R}_1}(f^*, g)$ which is potentially infinite for the negative entropy function $\mathcal{R}_1$. It is possible that the doubling trick or the analysis of Auer et al [2] (who encountered the same problem for the Exponential Weights algorithm) can remove the extra $\log T$ factor while still preserving the regret minimization property. Finally, $R_{\max}$ is small for the p-norm function $\mathcal{R}_1$, and so the use of this regularizer avoids the extraneous logarithmic in T factor while still preserving the logarithmic dependence on n and m. However, projection onto simplex under the p-norm is not as elegant as the Exponential Weights update.

4.2 Partial Information

We now turn to the partial (or, zero-th order) information model. Recall that the matrix A is not known to the players, yet we are interested in finding ϵ-optimal minimax strategies. On each round, the two players choose mixed strategies $f_t \in \Delta_n$ and $x_t \in \Delta_m$, respectively, and observe $f_t^\top A x$. Now the question is, how many such observations do we need to get to an ϵ-optimal minimax strategy? Can this be done while still ensuring the usual no-regret rate?

The specific setting we consider below requires that on each round t, the two players play four times, and that these four plays are δ-close to each other (that is, $\|f_t^i - f_t^j\|_1 \le \delta$ for $i, j \in \{1, \dots, 4\}$). Interestingly, up to logarithmic factors, the fast rate of the previous section is possible even in this scenario, yet we do require the knowledge of the number of actions of the opposing player (or, an upper bound on this number). We leave it as an open problem the question of whether one can attain the $1/T$-type rate with only one play per round.

Player I
$u_1, \dots, u_{n-1}$: orthonormal basis of Δ_n
Initialize $g_1, f_1 = (\frac{1}{n}, \dots, \frac{1}{n})$
Draw $i_0 \sim \mathrm{Unif}([n-1])$
At time $t = 1$ to T
  Play f_t
  Draw $i_t \sim \mathrm{Unif}([n-1])$
  Observe :
    $r_t^+ = (f_t + \delta u_{i_{t-1}})^\top A x_t$
    $r_t^- = (f_t - \delta u_{i_{t-1}})^\top A x_t$
    $\bar{r}_t^+ = (f_t + \delta u_{i_t})^\top A x_t$
    $\bar{r}_t^- = (f_t - \delta u_{i_t})^\top A x_t$
  Build estimates :
    $\hat{a}_t = \frac{n}{2\delta}(r_t^+ - r_t^-) u_{i_{t-1}}$
    $\bar{a}_t = \frac{n}{2\delta}(\bar{r}_t^+ - \bar{r}_t^-) u_{i_t}$
  Update :
    $g_t(i) \propto g'_{t-1}(i) \exp\{-\eta_t \hat{a}_t(i)\}$
    $g'_t = \left(1 - \frac{1}{T^2}\right) g_t + \frac{\mathbf{1}}{nT^2}$
    $f_{t+1}(i) \propto g'_t(i) \exp\{-\eta_{t+1} \bar{a}_t(i)\}$
End

Player II
$v_1, \dots, v_{m-1}$: orthonormal basis of Δ_m
Initialize $y_1, x_1 = (\frac{1}{m}, \dots, \frac{1}{m})$
Draw $j_0 \sim \mathrm{Unif}([m-1])$
At time $t = 1$ to T
  Play x_t
  Draw $j_t \sim \mathrm{Unif}([m-1])$
  Observe :
    $s_t^+ = -f_t^\top A(x_t + \delta v_{j_{t-1}})$
    $s_t^- = -f_t^\top A(x_t - \delta v_{j_{t-1}})$
    $\bar{s}_t^+ = -f_t^\top A(x_t + \delta v_{j_t})$
    $\bar{s}_t^- = -f_t^\top A(x_t - \delta v_{j_t})$
  Build estimates :
    $\hat{b}_t = \frac{m}{2\delta}(s_t^+ - s_t^-) v_{j_{t-1}}$
    $\bar{b}_t = \frac{m}{2\delta}(\bar{s}_t^+ - \bar{s}_t^-) v_{j_t}$
  Update :
    $y_t(i) \propto y'_{t-1}(i) \exp\{-\eta'_t \hat{b}_t(i)\}$
    $y'_t = \left(1 - \frac{1}{T^2}\right) y_t + \frac{\mathbf{1}}{nT^2}$
    $x_{t+1}(i) \propto y'_t(i) \exp\{-\eta'_{t+1} \bar{b}_t(i)\}$
End

Lemma 7. *Let $A \in [-1,1]^{n\times m}$, $\mathcal{F} = \Delta_n$, $\mathcal{X} = \Delta_m$, and let δ be small enough (e.g. exponentially small in m, n, T). If both players use above algorithms with the adaptive step sizes*

$$\eta_t = \min\left\{\sqrt{\log(nT)}\frac{\sqrt{\sum_{i=1}^{t-1}\|\hat{a}_i-\bar{a}_{i-1}\|_*^2}-\sqrt{\sum_{i=1}^{t-2}\|\hat{a}_i-\bar{a}_{i-1}\|_*^2}}{\|\hat{a}_{t-1}-\bar{a}_{t-2}\|_*^2}, \frac{1}{28m\sqrt{\log(mT)}}\right\}$$

and

$$\eta'_t = \min\left\{\sqrt{\log(mT)}\frac{\sqrt{\sum_{i=1}^{t-1}\left\|\hat{b}_i-\bar{b}_{i-1}\right\|_*^2}-\sqrt{\sum_{i=1}^{t-2}\left\|\hat{b}_i-\bar{b}_{i-1}\right\|_*^2}}{\left\|\hat{b}_{t-1}-\bar{b}_{t-2}\right\|_*^2}, \frac{1}{28n\sqrt{\log(nT)}}\right\}$$

respectively, then the pair $(\bar{f}_T, \bar{x}_T)$ is an

$$O\left(\frac{\left(m\log(nT)\sqrt{\log(mT)} + n\log(mT)\sqrt{\log(nT)}\right)}{T}\right)$$

-approximate minimax equilibrium. Furthermore, if only one player (say, Player I) follows the above algorithm, her (averaged over T) regret against any sequence $x_1, \dots, x_T$ of plays is bounded by

$$O\left(\frac{m\sqrt{\log(mT)}\log(nT) + n\sqrt{\log(nT)\sum_{t=1}^{T}\|x_t - x_{t-1}\|^2}}{T}\right)$$

We leave it as an open problem to provide an algorithm that attains the $1/T$-type rate when both players only observe the value $e_i^\top A e_j = A_{i,j}$ upon drawing *pure* actions i, j from their respective mixed strategies f_t, x_t. We hypothesize a rate better than $T^{-1/2}$ is not possible in this scenario.

5 Approximate Smooth Convex Programming

In this section we show how one can use the structured optimization results from Section 3 for approximate convex programming problems. Specifically consider the optimization problem,

$$\begin{aligned} \underset{f\in\mathcal{G}}{\operatorname{argmax}} \quad & c^\top f \\ \text{s.t.} \quad & \forall i \in [d],\ \ G_i(f) \le 1 \end{aligned} \tag{10}$$

where $\mathcal{G}$ is a convex set and each G_i is an H-smooth convex function. Assume that the optimal value of the above optimization problem is given by $F^* > 0$. Without loss of generality we assume F^* is known (one typically performs binary search if it is not known). Define the sets $\mathcal{F} = \{f : f \in \mathcal{G}, c^\top f = F^*\}$ and $\mathcal{X} = \Delta_d$. Now note that the convex programming problem in (10) can be reformulated as the alternative optimization problem of ,

$$\underset{f\in\mathcal{F}}{\operatorname{argmin}}\ \max_{i\in[d]} G_i(f) = \underset{f\in\mathcal{F}}{\operatorname{argmin}}\ \sup_{x\in\mathcal{X}} \sum_{i=1}^{d} x[i] G_i(f)\ . \tag{11}$$

Observe that it is of the saddle-point form studied earlier in the paper in Section 3. We may think of the first player as aiming to minimize the above expression over $\mathcal{F}$ (and thus satisfy the constraints), while the second player maximizes over a mixture of constraints over Δ_d.

Lemma 8. *Assume there exists $f_0 \in \mathcal{G}$ such that $c^\top f_0 \ge 0$ and for every $i \in [d]$, $G_i(f_0) \le 1 - \gamma$. Fix $\epsilon > 0$ and consider the solution*

$$\hat{f}_T = (1-\alpha)\bar{f}_T + \alpha f_0$$

where $\alpha = \frac{\epsilon}{\epsilon+\gamma}$ and $\bar{f}_T = \frac{1}{T}\sum_{t=1}^{T} f_t \in \mathcal{F}$ is the average of the trajectory returned by employing the optimization procedure as in Lemma 4 to the optimization problem specified in Eq. (11) with $\mathcal{R}_1(\cdot) = \frac{1}{2}\|\cdot\|_2^2$, $\mathcal{R}_2$ the entropy function, $\eta = \frac{1}{\eta'} = \frac{\|f^*-g_0\|_2}{H\sqrt{\log d}}$ $M_t^1 = \sum_{i=1}^{d} y_{t-1}[i]\nabla G_i(g_{t-1})$ and $M_t^2 = (G_1(g_{t-1}), \dots, G_d(f_t))$. Let number of iterations T be such that

$$T > \frac{4H\|f^* - g_0\|_2\sqrt{\log d}}{\epsilon}$$

where $g_0 \in \mathcal{F}$ is some initialization and $f^* \in \mathcal{F}$ is the (unknown) solution to the optimization problem. We have that $\hat{f}_T \in \mathcal{G}$ satisfies all d constraints and is $\frac{\epsilon}{\gamma}$-approximate, that is

$$c^\top \hat{f}_T \geq \left(1 - \frac{\epsilon}{\gamma}\right) F^*$$

The above lemma tells us that using the predictable sequences for the two players approach, one can obtain an $\frac{\epsilon}{\gamma}$ approximate solution to smooth convex programming problem in number of iterations at most order $1/\epsilon$. If T_1 is the time complexity for single update of the predictable sequence algorithm of Player I and T_2 is the time complexity for single update of the predictable sequence algorithm of Player II then we can conclude that the time complexity of the overall procedure is $O\left(\frac{(T_1+T_2)}{\epsilon}\right)$

5.1 Application to Max-Flow

We now apply the result to the problem of finding max-flow between a source and a sink in a network. Specifically consider the max-flow problem on a network where each edge has capacity 1 (the method can easily be extended to varying capacity on each edge). The max flow problem consists of finding the maximal flow between a source and sink in the network such that the capacity constraint on each edge is satisfied. Let us consider the case when the number of edges d in the network is the same order as number of vertices in the network. The max-flow problem can be considered as an instance of a convex (linear) programming problem. We therefore apply the proposed algorithm for structured optimization to obtain an approximate solution to this problem.

In the max-flow problem $\mathcal{G}$ is given by a set of linear equalities and so is $\mathcal{F}$. Further, if we use Euclidean norm squared as regularizer for flow player, then projection step can be performed $O(d)$ time using conjugate gradient method. This is because we are simply minimizing Euclidean norm squared subject to equality constraints which is well conditioned. Hence $T_1 = O(d)$. Similarly the Exponential Weights update also has time complexity $O(d)$ as there are order d constraints, and so overall time complexity to produce ϵ approximate solution is given by $O(nd)$, where n is the number of iterations of the proposed procedure. Specifically we shall assume that we know the value of the maximum flow F^* (if not we can use binary search to obtain it) and knowing this we have the following corollary.

Corollary 9. *Applying the procedure for smooth convex programming from Lemma 8 to the max-flow problem with $f_0 = \mathbf{0} \in \mathcal{G}$ the 0 flow, we can conclude that the time complexity to compute an ϵ-approximate max-flow is bounded by*

$$O\left(\frac{d^{3/2}\sqrt{\log d}}{\epsilon}\right)$$

This time complexity matches the known result from [6], but with a much simpler procedure (gradient descent for the flow player and Exponential Weights for the constraints). It would be interesting to see whether the techniques presented here can be used to improve the dependence on d to $d^{4/3}$ or better while maintaining the $1/\epsilon$ dependence. While the result of [**?**] has the improved $d^{4/3}$ dependence, the complexity in terms of ϵ is much worse.

6 Discussion

We close this paper with a discussion. As we showed, the notion of using extra information about the sequence is a powerful tool with applications in optimization, convex programming, game theory, to name a few. All the applications considered in this paper, however, used some notion of smoothness for constructing the predictable process M_t. An interesting direction of further research is to isolate more general conditions under which the next gradient is predictable, perhaps even when the functions are not smooth in any sense. For instance one could use techniques from bundle methods to further restrict the set of possible gradients the function being optimized can have at various points in the feasible set. This could then be used to solve for the right predictable sequence to use so as to optimize the bounds. Using this notion of selecting predictable sequences one can hope to derive adaptive optimization procedures that in practice can provide rapid convergence.

References

[1] S. Arora, E. Hazan, and S. Kale. The multiplicative weights update method: A meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[2] P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002.

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

[4] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In *COLT*, 2012.

[5] C. Daskalakis, A. Deckelbaum, and A. Kim. Near-optimal no-regret algorithms for zero-sum games. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 235–254. SIAM, 2011.

[6] Andrew V Goldberg and Satish Rao. Beyond the flow decomposition barrier. *Journal of the ACM (JACM)*, 45(5):783–797, 1998.

[7] A. Nemirovski. Prox-method with rate of convergence O(1/t) for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1):229–251, 2004.

[8] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–152, 2005.

[9] A. Rakhlin and K. Sridharan. Online learning with predictable sequences. In *Proceedings of the 26th Annual Conference on Learning Theory (COLT)*, 2013.

Proofs

Proof of Lemma 1. For any $f^* \in \mathcal{F}$,

$$\langle f_t - f^*, \nabla_t \rangle = \langle f_t - g_t, \nabla_t - M_t \rangle + \langle f_t - g_t, M_t \rangle + \langle g_t - f^*, \nabla_t \rangle \tag{12}$$

First observe that

$$\langle f_t - g_t, \nabla_t - M_t \rangle \le \| f_t - g_t \| \, \| \nabla_t - M_t \|_* \ . \tag{13}$$

Any update of the form $a^* = \arg\min_{a \in A} \langle a, x \rangle + D_{\mathcal{R}}(a, c)$ satisfies for any $d \in A$

$$\langle a^* - d, x \rangle \le D_{\mathcal{R}}(d, c) - D_{\mathcal{R}}(d, a^*) - D_{\mathcal{R}}(a^*, c) \ . \tag{14}$$

This yields

$$\langle f_t - g_t, M_t \rangle \le \frac{1}{\eta} \left(D_{\mathcal{R}}(g_t, g_{t-1}) - D_{\mathcal{R}}(g_t, f_t) - D_{\mathcal{R}}(f_t, g_{t-1}) \right) \tag{15}$$

and

$$\langle g_t - f^*, \nabla_t \rangle \le \frac{1}{\eta} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - D_{\mathcal{R}}(g_t, g_{t-1}) \right) \ . \tag{16}$$

Combining, $\langle f_t - f^*, \nabla_t \rangle$ is upper bounded by

$$\begin{aligned}
&\| \nabla_t - M_t \|_* \| f_t - g_t \| + \frac{1}{\eta} \left(D_{\mathcal{R}}(g_t, g_{t-1}) - D_{\mathcal{R}}(g_t, f_t) - D_{\mathcal{R}}(f_t, g_{t-1}) \right) \\
&+ \frac{1}{\eta} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - D_{\mathcal{R}}(g_t, g_{t-1}) \right)) \\
&= \| \nabla_t - M_t \|_* \| f_t - g_t \| + \frac{1}{\eta} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - D_{\mathcal{R}}(g_t, f_t) - D_{\mathcal{R}}(f_t, g_{t-1}) \right) \\
&\le \| \nabla_t - M_t \|_* \| f_t - g_t \| + \frac{1}{\eta} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - \frac{1}{2} \| g_t - f_t \|^2 - \frac{1}{2} \| g_{t-1} - f_t \|^2 \right)
\end{aligned} \tag{17}$$

where in the last step we used strong convexity: for any f, f', $D_{\mathcal{R}}(f, f') \ge \frac{1}{2} \| f - f' \|^2$. Summing over $t = 1, \ldots, T$ yields, for any $f^* \in \mathcal{F}$,

$$\sum_{t=1}^T \langle f_t - f^*, \nabla_t \rangle \le \eta^{-1} D_{\mathcal{R}}(f^*, g_0) + \sum_{t=1}^T \| \nabla_t - M_t \|_* \| g_t - f_t \| - \frac{1}{2\eta} \sum_{t=1}^T \left(\| g_t - f_t \|^2 + \| g_{t-1} - f_t \|^2 \right) \ .$$

Appealing to convexity of G_t's completes the proof. □

Proof of Corollary 2. Let us re-work the proof of Lemma 1 for the case of a changing η_t. Eq. (15) and (16) are now replaced by

$$\langle f_t - g_t, M_t \rangle \le \frac{1}{\eta_t} \left(D_{\mathcal{R}}(g_t, g_{t-1}) - D_{\mathcal{R}}(g_t, f_t) - D_{\mathcal{R}}(f_t, g_{t-1}) \right) \tag{18}$$

and

$$\langle g_t - f^*, \nabla_t \rangle \le \frac{1}{\eta_t} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - D_{\mathcal{R}}(g_t, g_{t-1}) \right) \ . \tag{19}$$

The upper bound of Eq. (17) becomes

$$\| \nabla_t - M_t \|_* \| f_t - g_t \| + \frac{1}{\eta_t} \left(D_{\mathcal{R}}(f^*, g_{t-1}) - D_{\mathcal{R}}(f^*, g_t) - \frac{1}{2} \| g_t - f_t \|^2 - \frac{1}{2} \| g_{t-1} - f_t \|^2 \right) \ .$$

Summing over $t = 1, \ldots, T$ yields, for any $f^* \in \mathcal{F}$,

$$\begin{aligned}
\sum_{t=1}^T \langle f_t - f^*, \nabla_t \rangle &\le \eta_1^{-1} D_{\mathcal{R}}(f^*, g_0) + \sum_{t=2}^T D_{\mathcal{R}}(f^*, g_{t-1}) \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) + \sum_{t=1}^T \| \nabla_t - M_t \|_* \| g_t - f_t \| \\
&\quad - \sum_{t=1}^T \frac{1}{2\eta_t} \left(\| g_t - f_t \|^2 + \| g_{t-1} - f_t \|^2 \right) \\
&\le (\eta_1^{-1} + \eta_T^{-1}) R_{\max}^2 + \sum_{t=1}^T \| \nabla_t - M_t \|_* \| g_t - f_t \| - \frac{1}{2} \sum_{t=1}^T \eta_t^{-1} \left(\| g_t - f_t \|^2 + \| g_{t-1} - f_t \|^2 \right)
\end{aligned} \tag{20}$$

Observe that

$$\eta_t = R_{\max} \min\left\{ \frac{1}{\sqrt{\sum_{i=1}^{t-1} \|\nabla_i - M_i\|_*^2} + \sqrt{\sum_{i=1}^{t-2} \|\nabla_i - M_i\|_*^2}}, 1 \right\} \tag{21}$$

$$= R_{\max} \min\left\{ \frac{\sqrt{\sum_{i=1}^{t-1} \|\nabla_i - M_i\|_*^2} - \sqrt{\sum_{i=1}^{t-2} \|\nabla_i - M_i\|_*^2}}{\|\nabla_{t-1} - M_{t-1}\|_*^2}, 1 \right\} \tag{22}$$

From (21),

$$\eta_t^{-1} \le R_{\max}^{-1} \max\left\{ 2\sqrt{\sum_{i=1}^{t-1} \|\nabla_i - M_i\|_*^2}, 1 \right\} .$$

Using this step size in Equation (20) and defining $\eta_1 = 1$, $\sum_{t=1}^{T} \langle f_t - f^*, \nabla_t \rangle$ is upper bounded by

$$R_{\max}\left(2\sqrt{\sum_{t=1}^{T-1} \|\nabla_t - M_t\|_*^2} + 2 \right) + \sum_{t=1}^{T} \|\nabla_t - M_t\|_* \|g_t - f_t\| - \frac{1}{2}\sum_{t=1}^{T} \eta_t^{-1} \left(\|g_t - f_t\|^2 + \|g_{t-1} - f_t\|^2 \right)$$

$$\le R_{\max}\left(2\sqrt{\sum_{t=1}^{T-1} \|\nabla_t - M_t\|_*^2} + 2 \right) + \frac{1}{2}\sum_{t=1}^{T} \eta_{t+1} \|\nabla_t - M_t\|_*^2 + \frac{1}{2}\sum_{t=1}^{T} \eta_{t+1}^{-1} \|f_t - g_t\|^2 - \frac{1}{2}\sum_{t=1}^{T} \eta_t^{-1} \|g_t - f_t\|^2$$

where we used (3) with $\rho = \eta_{t+1}$ and dropped one of the positive terms. The last two terms can be upper bounded as

$$\frac{1}{2}\sum_{t=1}^{T} \eta_{t+1}^{-1} \|f_t - g_t\|^2 - \frac{1}{2}\sum_{t=1}^{T} \eta_t^{-1} \|g_t - f_t\|^2 \le \frac{R_{\max}^2}{2} \sum_{t=1}^{T} \left(\eta_{t+1}^{-1} - \eta_t^{-1} \right) \le \frac{R_{\max}^2}{2} \eta_{T+1}^{-1},$$

yielding an upper bound

$$R_{\max}\left(2\sqrt{\sum_{t=1}^{T} \|\nabla_t - M_t\|_*^2} + 2 \right) + \frac{1}{2}\sum_{t=1}^{T} \eta_{t+1} \|\nabla_t - M_t\|_*^2 + \frac{R_{\max}^2}{2} \eta_{T+1}^{-1}$$

$$\le 3R_{\max}\left(\sqrt{\sum_{t=1}^{T} \|\nabla_t - M_t\|_*^2} + 1 \right) + \frac{1}{2}\sum_{t=1}^{T} \eta_{t+1} \|\nabla_t - M_t\|_*^2 .$$

In view of (21), we arrive at

$$3R_{\max}\left(\sqrt{\sum_{t=1}^{T} \|\nabla_t - M_t\|_*^2} + 1 \right) + \frac{R_{\max}}{2}\sum_{t=1}^{T} \left(\sqrt{\sum_{i=1}^{t} \|\nabla_i - M_i\|_*^2} - \sqrt{\sum_{i=1}^{t-1} \|\nabla_i - M_i\|_*^2} \right)$$

$$\le 3R_{\max}\left(\sqrt{\sum_{t=1}^{T} \|\nabla_t - M_t\|_*^2} + 1 \right) + \frac{R_{\max}}{2}\sqrt{\sum_{i=1}^{T} \|\nabla_i - M_i\|_*^2}$$

$$\le 3.5\, R_{\max}\left(\sqrt{\sum_{t=1}^{T} \|\nabla_t - M_t\|_*^2} + 1 \right)$$

□

Proof of Lemma 3. Let $\nabla_t = \nabla G(f_t)$ and $M_t = \nabla G(g_{t-1})$. Then by Lemma 1 and by Hölder smoothness,

$$\sum_{t=1}^{T} \langle f_t - f^*, \nabla_t \rangle \le \frac{R^2}{\eta} + H\sum_{t=1}^{T} \|g_t - f_t\|^{1+\alpha} - \frac{1}{2\eta}\sum_{t=1}^{T} \|g_t - f_t\|^2 . \tag{23}$$

We can re-write the middle term in the upper bound as

$$H\sum_{t=1}^{T}\|g_t - f_t\|^{1+\alpha} = \sum_{t=1}^{T} H\left((1+\alpha)\eta\right)^{\frac{1+\alpha}{2}}\left(\frac{\|g_t - f_t\|}{\sqrt{(1+\alpha)\eta}}\right)^{1+\alpha}$$

$$\le \left(\sum_{t=1}^{T} H^{\frac{2}{1-\alpha}}\left((1+\alpha)\eta\right)^{\frac{1+\alpha}{1-\alpha}}\right)^{\frac{1-\alpha}{2}}\left(\sum_{t=1}^{T}\frac{\|g_t - f_t\|^2}{(1+\alpha)\eta}\right)^{\frac{1+\alpha}{2}}$$

by Hölder's inequality with conjugate powers $1/p = (1-\alpha)/2$ and $1/q = (1+\alpha)/2$. We further upper bound the last term using AM-GM inequality as

$$\frac{1-\alpha}{2}\left(TH^{\frac{2}{1-\alpha}}(1+\alpha)^{\frac{1+\alpha}{1-\alpha}}\eta^{\frac{1+\alpha}{1-\alpha}}\right) + \frac{1}{2\eta}\sum_{t=1}^{T}\|g_t - f_t\|^2 \ .$$

Plugging into (23),

$$\sum_{t=1}^{T}\langle f_t - f^*, \nabla_t\rangle \le \frac{R^2}{\eta} + \frac{1-\alpha}{2}\left(TH^{\frac{2}{1-\alpha}}(1+\alpha)^{\frac{1+\alpha}{1-\alpha}}\eta^{\frac{1+\alpha}{1-\alpha}}\right) \ .$$

Setting $\eta = R^{1-\alpha}H^{-1}(1+\alpha)^{-\frac{1+\alpha}{2}}(1-\alpha)^{-\frac{1-\alpha}{2}}T^{-\frac{1-\alpha}{2}}$ yields an upper bound of

$$\sum_{t=1}^{T}\langle f_t - f^*, \nabla_t\rangle \le HR^{1+\alpha}(1+\alpha)^{\frac{1+\alpha}{2}}(1-\alpha)^{\frac{1-\alpha}{2}}T^{\frac{1-\alpha}{2}} \le 8HR^{1+\alpha}T^{\frac{1-\alpha}{2}} \ .$$

□

Proof of Corollary 5. Using Lemma 4,

$$\sup_{x\in\mathcal{X}}\phi\left(\frac{1}{T}\sum_{t=1}^{T}f_t, x\right) - \inf_{f\in\mathcal{F}}\sup_{x\in\mathcal{X}}\phi(f,x)$$

$$\le \frac{R_1^2}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\|\nabla_f\phi(f_t,x_t) - \nabla_f\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{F}^*} - \frac{1}{2\eta}\sum_{t=1}^{T}\|g_{t-1} - f_t\|^2_{\mathcal{F}}$$

$$+ \frac{R_2^2}{\eta'} + \frac{\eta'}{2}\sum_{t=1}^{T}\|\nabla_x\phi(f_t,x_t) - \nabla_x\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{X}^*} - \frac{1}{2\eta'}\sum_{t=1}^{T}\|y_{t-1} - x_t\|^2_{\mathcal{X}}$$

Using $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$ and the smoothness assumption yields

$$\frac{\eta}{2}\sum_{t=1}^{T}\|\nabla_f\phi(f_t,x_t) - \nabla_f\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{F}^*}$$

$$\le \eta\sum_{t=1}^{T}\|\nabla_f\phi(f_t,x_t) - \nabla_f\phi(g_{t-1},x_t)\|^2_{\mathcal{F}^*} + \eta\sum_{t=1}^{T}\|\nabla_f\phi(g_{t-1},x_t) - \nabla_f\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{F}^*}$$

$$\le \eta H_1^2\sum_{t=1}^{T}\|f_t - g_{t-1}\|^{2\alpha}_{\mathcal{F}} + \eta H_2^2\sum_{t=1}^{T}\|x_t - y_{t-1}\|^{2\alpha'}_{\mathcal{X}}$$

and similarly

$$\frac{\eta'}{2}\sum_{t=1}^{T}\|\nabla_x\phi(f_t,x_t) - \nabla_x\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{X}^*}$$

$$\le \eta'\sum_{t=1}^{T}\|\nabla_x\phi(f_t,x_t) - \nabla_x\phi(f_t,y_{t-1})\|^2_{\mathcal{X}^*} + \eta'\sum_{t=1}^{T}\|\nabla_x\phi(f_t,y_{t-1}) - \nabla_x\phi(g_{t-1},y_{t-1})\|^2_{\mathcal{X}^*}$$

$$\le \eta' H_3^2\sum_{t=1}^{T}\|x_t - y_{t-1}\|^{2\beta'}_{\mathcal{X}} + \eta' H_4^2\sum_{t=1}^{T}\|f_t - g_{t-1}\|^{2\beta}_{\mathcal{F}} \ .$$

Combining, we get the upper bound of

$$\frac{R_1^2}{\eta} + \sum_{t=1}^{T}(4\alpha\eta)^{\alpha}\eta H_1^2\left(\frac{\|f_t - g_{t-1}\|_{\mathcal{F}}}{\sqrt{4\alpha\eta}}\right)^{2\alpha} + \sum_{t=1}^{T}(4\alpha'\eta')^{\alpha'}\eta H_2^2\left(\frac{\|x_t - y_{t-1}\|_{\mathcal{X}}}{\sqrt{4\alpha'\eta'}}\right)^{2\alpha'} - \frac{1}{2\eta}\sum_{t=1}^{T}\|g_{t-1} - f_t\|^2_{\mathcal{F}}$$

$$+ \frac{R_2^2}{\eta'} + \sum_{t=1}^{T}(4\beta'\eta')^{\beta'}\eta' H_3^2\left(\frac{\|x_t - y_{t-1}\|_{\mathcal{X}}}{\sqrt{4\beta'\eta'}}\right)^{2\beta'} + \sum_{t=1}^{T}(4\beta\eta)^{\beta}\eta' H_4^2\left(\frac{\|f_t - g_{t-1}\|_{\mathcal{F}}}{\sqrt{4\beta\eta}}\right)^{2\beta} - \frac{1}{2\eta'}\sum_{t=1}^{T}\|y_{t-1} - x_t\|^2_{\mathcal{X}}$$

As in the proof of Lemma 3, we use Hölder inequality to further upper bound by

$$
\begin{aligned}
&\frac{R_1^2}{\eta} + \frac{R_2^2}{\eta'} - \frac{1}{2\eta}\sum_{t=1}^{T}\|g_{t-1} - f_t\|_{\mathcal{F}}^2 - \frac{1}{2\eta'}\sum_{t=1}^{T}\|y_{t-1} - x_t\|_{\mathcal{X}}^2 \qquad (24)\\
&\quad + \left(T(4\alpha\eta)^{\frac{\alpha}{1-\alpha}}\eta^{\frac{1}{1-\alpha}}H_1^{\frac{2}{1-\alpha}}\right)^{1-\alpha}\left(\sum_{t=1}^{T}\frac{\|f_t - g_{t-1}\|_{\mathcal{F}}^2}{4\alpha\eta}\right)^{\alpha} + \left(T(4\alpha'\eta')^{\frac{\alpha'}{1-\alpha'}}\eta^{\frac{1}{1-\alpha'}}H_2^{\frac{2}{1-\alpha'}}\right)^{1-\alpha'}\left(\sum_{t=1}^{T}\frac{\|x_t - y_{t-1}\|_{\mathcal{X}}^2}{4\alpha'\eta'}\right)^{\alpha'}\\
&\quad + \left(T(4\beta'\eta')^{\frac{\beta'}{1-\beta'}}\eta'^{\frac{1}{1-\beta'}}H_3^{\frac{2}{1-\beta'}}\right)^{1-\beta'}\left(\sum_{t=1}^{T}\frac{\|x_t - y_{t-1}\|_{\mathcal{X}}^2}{4\beta'\eta'}\right)^{\beta'} + \left(T(4\beta\eta)^{\frac{\beta}{1-\beta}}\eta'^{\frac{1}{1-\beta}}H_4^{\frac{2}{1-\beta}}\right)^{1-\beta}\left(\sum_{t=1}^{T}\frac{\|f_t - g_{t-1}\|_{\mathcal{F}}^2}{4\beta\eta}\right)^{\beta}\\
&\le \frac{R_1^2}{\eta} + \frac{R_2^2}{\eta'} + \left((1-\alpha)(4\alpha\eta)^{\frac{\alpha}{1-\alpha}}\eta^{\frac{1}{1-\alpha}}H_1^{\frac{2}{1-\alpha}}\right)T + \left((1-\alpha')(4\alpha'\eta')^{\frac{\alpha'}{1-\alpha'}}\eta^{\frac{1}{1-\alpha'}}H_2^{\frac{2}{1-\alpha'}}\right)T\\
&\quad + \left((1-\beta')(4\beta'\eta')^{\frac{\beta'}{1-\beta'}}\eta'^{\frac{1}{1-\beta'}}H_3^{\frac{2}{1-\beta'}}\right)T + \left((1-\beta)(4\beta\eta)^{\frac{\beta}{1-\beta}}\eta'^{\frac{1}{1-\beta}}H_4^{\frac{2}{1-\beta}}\right)T
\end{aligned}
$$

Setting $\eta = \eta'$ we get an upper bound of

$$
\begin{aligned}
&\frac{R_1^2 + R_2^2}{\eta} + \left((1-\alpha)(4\alpha)^{\frac{\alpha}{1-\alpha}}\eta^{\frac{1+\alpha}{1-\alpha}}H_1^{\frac{2}{1-\alpha}}\right)T + \left((1-\alpha')(4\alpha')^{\frac{\alpha'}{1-\alpha'}}\eta^{\frac{1+\alpha'}{1-\alpha'}}H_2^{\frac{2}{1-\alpha'}}\right)T\\
&\quad + \left((1-\beta')(4\beta')^{\frac{\beta'}{1-\beta'}}\eta^{\frac{1+\beta'}{1-\beta'}}H_3^{\frac{2}{1-\beta'}}\right)T + \left((1-\beta)(4\beta)^{\frac{\beta}{1-\beta}}\eta^{\frac{1+\beta}{1-\beta}}H_4^{\frac{2}{1-\beta}}\right)T\\
&\le \frac{R_1^2 + R_2^2}{\eta} + \left((1-\alpha)(\alpha)^{\frac{\alpha}{1-\alpha}}\eta^{\frac{1+\alpha}{1-\alpha}}(2H_1)^{\frac{2}{1-\alpha}}\right)T + \left((1-\alpha')(\alpha')^{\frac{\alpha'}{1-\alpha'}}\eta^{\frac{1+\alpha'}{1-\alpha'}}(2H_2)^{\frac{2}{1-\alpha'}}\right)T\\
&\quad + \left((1-\beta')(\beta')^{\frac{\beta'}{1-\beta'}}\eta^{\frac{1+\beta'}{1-\beta'}}(2H_3)^{\frac{2}{1-\beta'}}\right)T + \left((1-\beta)(\beta)^{\frac{\beta}{1-\beta}}\eta^{\frac{1+\beta}{1-\beta}}(2H_4)^{\frac{2}{1-\beta}}\right)T\\
&\le \frac{R_1^2 + R_2^2}{\eta} + \left(\eta^{\frac{1+\alpha}{1-\alpha}}(2H_1)^{\frac{2}{1-\alpha}}\right)T + \left(\eta^{\frac{1+\alpha'}{1-\alpha'}}(2H_2)^{\frac{2}{1-\alpha'}}\right)T\\
&\quad + \left(\eta^{\frac{1+\beta'}{1-\beta'}}(2H_3)^{\frac{2}{1-\beta'}}\right)T + \left(\eta^{\frac{1+\beta}{1-\beta}}(2H_4)^{\frac{2}{1-\beta}}\right)T\\
&\le \frac{R_1^2 + R_2^2}{\eta} + (2H\eta)^{\frac{1+\gamma}{1-\gamma}}HT
\end{aligned}
$$

Finally picking step size as $\eta = (R_1^2 + R_2^2)^{\frac{1-\gamma}{2}}(2H)^{-1}\left(\frac{T}{2}\right)^{\frac{\gamma-1}{2}}$ we conclude that

$$
\sup_{x\in\mathcal{X}}\phi\left(\frac{1}{T}\sum_{t=1}^{T}f_t, x\right) - \inf_{f\in\mathcal{F}}\sup_{x\in\mathcal{X}}\phi(f,x) \le \frac{4H(R_1^2 + R_2^2)^{\frac{1+\gamma}{2}}}{T^{\frac{1+\gamma}{2}}} \qquad (25)
$$

□

Proof of Proposition 6. Let $\mathcal{R}_1(f) = \sum_{i=1}^{n} f(i)\ln f(i)$ and, respectively, $\mathcal{R}_2(x) = \sum_{i=1}^{m} x(i)\ln x(i)$. These functions are strongly convex with respect to $\|\cdot\|_1$ norm on the respective flat simplex. We first upper bound regret of Player I, writing ∇_t as a generic observation vector, later to be chosen as Ax_t, and M_t as a generic predictable sequence, later chosen to be Ax_{t-1}. Observe that $\|g'_t - g_t\|_1 \le 1/T$. Then

$$
\langle f_t - f^*, \nabla_t\rangle = \langle f_t - g_t, \nabla_t - M_t\rangle + \langle f_t - g_t, M_t\rangle + \langle g_t - f^*, \nabla_t\rangle
$$

By the update rule,

$$
\langle f_t - g_t, M_t\rangle \le \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(g_t, g'_{t-1}) - D_{\mathcal{R}_1}(g_t, f_t) - D_{\mathcal{R}_1}(f_t, g'_{t-1})\right)
$$

and, assuming $\|\nabla_t\|_\infty \le 1$,

$$
\langle g_t - f^*, \nabla_t\rangle \le \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g_t) - D_{\mathcal{R}_1}(g_t, g'_{t-1})\right)
$$

We conclude that $\langle f_t - f^*, \nabla_t \rangle$ is upper bounded by

$$
\begin{aligned}
&\|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(g_t, g'_{t-1}) - D_{\mathcal{R}_1}(g_t, f_t) - D_{\mathcal{R}_1}(f_t, g'_{t-1})\right) \\
&+ \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g_t) - D_{\mathcal{R}_1}(g_t, g'_{t-1}))\right) \\
&= \|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g_t) - D_{\mathcal{R}_1}(g_t, f_t) - D_{\mathcal{R}_1}(f_t, g'_{t-1})\right) \\
&\le \|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g_t) - \frac{1}{2}\|g_t - f_t\|^2 - \frac{1}{2}\|g'_{t-1} - f_t\|^2\right) \\
&= \|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g'_t) - \frac{1}{2}\|g_t - f_t\|^2 - \frac{1}{2}\|g'_{t-1} - f_t\|^2\right) \\
&\qquad + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_t) - D_{\mathcal{R}_1}(f^*, g_t)\right) \\
&= \|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g'_t) - \frac{1}{2}\|g_t - f_t\|^2 - \frac{1}{2}\|g'_{t-1} - f_t\|^2\right) \\
&\qquad + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_t) - D_{\mathcal{R}_1}(f^*, g_t)\right) \\
&= \|\nabla_t - M_t\|_* \|f_t - g_t\| + \frac{1}{\eta_t}\left(D_{\mathcal{R}_1}(f^*, g'_{t-1}) - D_{\mathcal{R}_1}(f^*, g'_t) - \frac{1}{2}\|g_t - f_t\|^2 - \frac{1}{2}\|g'_{t-1} - f_t\|^2\right) \\
&\qquad + \frac{1}{\eta_t}\ln\frac{g_t(i^*)}{g'_t(i^*)} .
\end{aligned}
$$

Now let us bound the term $\frac{1}{\eta_t}\ln\frac{g_t(i^*)}{g'_t(i^*)}$. First note that whenever $g'_t(i^*) \ge g_t(i^*)$ then the term is negative. Since $g'_t(i^*) = (1 - 1/T^2)g_t(i^*) + 1/nT^2$ we see that whenever $g_t(i^*) \le 1/n$ this term is negative. On the other hand, for $g_t(i^*) > 1/n$ we can bound

$$\ln\frac{g_t(i^*)}{g'_t(i^*)} = \ln\frac{g_t(i)}{(1 - 1/T^2)g_t(i) + 1/(nT^2)} \le \frac{2}{T^2},$$

Using the above in the bound on $\langle f_t - f^*, \nabla_t \rangle$ and summing over $t = 1, \ldots, T$, and using the fact that the step size are non-increasing, we conclude that

$$
\begin{aligned}
\sum_{t=1}^T \langle f_t - f^*, \nabla_t \rangle &\le \eta_1^{-1} D_{\mathcal{R}_1}(f^*, g_0) + \sum_{t=2}^T D_{\mathcal{R}_1}(f^*, g'_{t-1})\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) + \sum_{t=1}^T \|\nabla_t - M_t\|_* \|g_t - f_t\| \\
&\quad - \sum_{t=1}^T \frac{1}{2\eta_t}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) + \frac{2}{T^2}\sum_{t=1}^T \frac{1}{\eta_t} \\
&\le (\eta_1^{-1} + \eta_T^{-1}) R_{1,\max}^2 + \sum_{t=1}^T \|\nabla_t - M_t\|_* \|g_t - f_t\| - \frac{1}{2}\sum_{t=1}^T \eta_t^{-1}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) \\
&\quad + \frac{2}{T^2}\sum_{t=1}^T \frac{1}{\eta_t} .
\end{aligned} \tag{26}
$$

where i^* is the coordinate on which f^* is 1 (the best action in hind-sight) and $R_{1,\max}^2$ is an upper bound on the largest KL divergence between f^* and any g' that has all coordinates at least $1/(nT^2)$. Since f^* is a vertex of the flat simplex, we may take $R_{1,\max}^2 \triangleq \log(nT^2)$. Also note that $1/\eta_t \le \sqrt{T}$ and so $\frac{2}{T^2}\sum_{t=1}^T \frac{1}{\eta_t} \le \frac{1}{T^{1/2}} \le 1$. Hence we conclude that for our matrix game, from the above result, bound on regret of Player I is given by,

$$\sum_{t=1}^T \langle f_t - f^*, \nabla_t \rangle \le (\eta_1^{-1} + \eta_T^{-1}) R_{1,\max}^2 + \sum_{t=1}^T \|Ax_t - Ax_{t-1}\|_* \|g_t - f_t\| - \frac{1}{2}\sum_{t=1}^T \eta_t^{-1}\left(\|g'_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) + 1 \tag{27}$$

Observe that

$$\eta_t = \min\left\{ R_{1,\max}^2 \frac{\sqrt{\sum_{i=1}^{t-1}\|Ax_i - Ax_{i-1}\|_*^2} - \sqrt{\sum_{i=1}^{t-2}\|Ax_i - Ax_{i-1}\|_*^2}}{\|Ax_{t-1} - Ax_{t-2}\|_*^2}, \frac{1}{11} \right\}$$

and

$$11 \le \eta_t^{-1} \le \max\left\{2R_{1,\max}^{-2}\sqrt{\sum_{i=1}^{t-1}\|Ax_i - Ax_{i-1}\|_*^2}, 11\right\}$$

With this, the upper bound on Player I's unnormalized regret is

$$1 + 22R_{1,\max}^2 + 2\sqrt{\sum_{t=1}^{T-1}\|Ax_t - Ax_{t-1}\|_*^2} + \sum_{t=1}^{T}\|Ax_t - Ax_{t-1}\|_*\|g_t - f_t\| - \frac{11}{2}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right)$$

Adding the regret of the second player who uses step size η'_t, the overall bound on the suboptimality, as in Eq.(6), is

$$\begin{aligned}
&2 + 22R_{1,\max}^2 + 2\sqrt{\sum_{t=1}^{T-1}\|Ax_t - Ax_{t-1}\|_*^2} + \sum_{t=1}^{T}\|Ax_t - Ax_{t-1}\|_*\|g_t - f_t\| \\
&+ 22R_{2,\max}^2 + 2\sqrt{\sum_{t=1}^{T-1}\|f_t^\intercal A - f_{t-1}^\intercal A\|_*^2} + \sum_{t=1}^{T}\|f_t^\intercal A - f_{t-1}^\intercal A\|_*\|y_t - x_t\| \\
&- \frac{11}{2}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) - \frac{11}{2}\sum_{t=1}^{T}\left(\|y_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right)
\end{aligned}$$

By over-bounding with $\sqrt{c} \le c + 1$ for $c \ge 0$, we obtain an upper bound

$$\begin{aligned}
\sum_{t=1}^{T}\langle f_t - f^*, \nabla_t\rangle \le\ & 6 + 22R_{1,\max}^2 + \sum_{t=1}^{T-1}\|Ax_t - Ax_{t-1}\|_*^2 + \sum_{t=1}^{T}\|Ax_t - Ax_{t-1}\|_*\|g'_t - f_t\| \\
&+ 22R_{2,\max}^2 + \sum_{t=1}^{T-1}\|f_t^\intercal A - f_{t-1}^\intercal A\|_*^2 + \sum_{t=1}^{T}\|f_t^\intercal A - f_{t-1}^\intercal A\|_*\|y'_t - x_t\| \\
&- \frac{11}{2}\sum_{t=1}^{T}\left(\|g'_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) - \frac{11}{2}\sum_{t=1}^{T}\left(\|y'_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right) \\
&\le 6 + 22R_{1,\max}^2 + \frac{5}{2}\sum_{t=1}^{T}\|Ax_t - Ax_{t-1}\|_*^2 + \frac{1}{2}\sum_{t=1}^{T}\|g_t - f_t\|^2 \\
&+ 22R_{2,\max}^2 + \frac{5}{2}\sum_{t=1}^{T}\|f_t^\intercal A - f_{t-1}^\intercal A\|_*^2 + \frac{1}{2}\sum_{t=1}^{T}\|y_t - x_t\|^2 \\
&- \frac{11}{2}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) - \frac{11}{2}\sum_{t=1}^{T}\left(\|y_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right)
\end{aligned}$$

Since each entry of the matrix is bounded by 1,

$$\|Ax_t - Ax_{t-1}\|_*^2 \le \|x_t - x_{t-1}\|^2 \le 2\|x_t - y'_{t-1}\|^2 + 2\|x_{t-1} - y'_{t-1}\|^2$$

and similar inequality holds for the other player too. This leads to an upper bound of

$$\begin{aligned}
&6 + 22R_{1,\max}^2 + 22R_{2,\max}^2 + \frac{1}{2}\sum_{t=1}^{T}\|y_t - x_t\|^2 + \frac{1}{2}\sum_{t=1}^{T}\|g_t - f_t\|^2 \\
&+ 5\sum_{t=1}^{T}\left(\|g'_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) + 5\sum_{t=1}^{T}\left(\|y'_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right) \\
&- \frac{11}{2}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right) - \frac{11}{2}\sum_{t=1}^{T}\left(\|y_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right) \\
&\le 6 + 22R_{1,\max}^2 + 22R_{2,\max}^2 \\
&+ 5\sum_{t=1}^{T}\left(\|g'_t - f_t\|^2 - \|g_t - f_t\|^2\right) + 5\sum_{t=1}^{T}\left(\|y'_t - x_t\|^2 - \|y_t - x_t\|^2\right)\ . \qquad (28)
\end{aligned}$$

Now note that

$$\begin{aligned}
\|g'_t - f_t\|^2 - \|g_t - f_t\|^2 &= \left(\|g'_t - f_t\| + \|g_t - f_t\|\right)\left(\|g'_t - f_t\| - \|g_t - f_t\|\right) \\
&\le \left(\|g'_t - f_t\| + \|g_t - f_t\|\right)\left(\|g'_t - g'_t\|\right) \le \frac{4}{T^2}
\end{aligned}$$

Similarly we also have that $\|y_t' - x_t\|^2 - \|y_t - x_t\|^2 \le \frac{4}{T^2}$. Using these in Eq **??** we conclude that the overall bound on the suboptimality, as in Eq.(6), is

$$\begin{aligned} 6 + 22R_{1,\max}^2 + 22R_{2,\max}^2 + \frac{40}{T} &= 6 + 22\log(nT^2) + 22\log(mT^2) + \frac{40}{T} \\ &= 6 + 22\log(nmT^4) + \frac{40}{T} \ . \end{aligned}$$

This proves the result for the case when both players adhere to the prescribed algorithm. Now, consider the case when Player I adheres, but we do not make any assumption about Player II. Then, from Eq. (27) and Eq. (3) with $\rho = \eta_t$, the upper bound on, $\sum_{t=1}^T \langle f_t - f^*, \nabla_t \rangle$, the unnormalized regret of Player I's is

$$\begin{aligned}
&(\eta_1^{-1} + \eta_T^{-1}) R_{1,\max}^2 + 1 + \sum_{t=1}^T \|Ax_t - Ax_{t-1}\|_* \|g_t - f_t\| - \frac{1}{2}\sum_{t=1}^T \eta_t^{-1} \left(\|g_t - f_t\|^2 + \|g_{t-1}' - f_t\|^2 \right) \\
&\le 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \sum_{t=1}^T \|Ax_t - Ax_{t-1}\|_* \|g_t - f_t\| - \frac{1}{2}\sum_{t=1}^T \eta_t^{-1} \|g_t - f_t\|^2 \\
&\le 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{1}{2}\sum_{t=1}^T \eta_{t+1} \|Ax_t - Ax_{t-1}\|_*^2 + \frac{1}{2}\sum_{t=1}^T (\eta_{t+1}^{-1} - \eta_t^{-1}) \|g_t - f_t\|^2 \\
&= 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{R_{1,\max}^2}{2}\sum_{t=1}^T \left(\sqrt{\sum_{i=1}^t \|Ax_i - Ax_{i-1}\|_*^2} - \sqrt{\sum_{i=1}^{t-1} \|Ax_i - Ax_{i-1}\|_*^2} \right) \\
&\quad + \frac{1}{2}\sum_{t=1}^T (\eta_{t+1}^{-1} - \eta_t^{-1}) \|g_t - f_t\|^2 \\
&\le 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{R_{1,\max}^2}{2}\sqrt{\sum_{i=1}^T \|Ax_i - Ax_{i-1}\|_*^2} + 2\sum_{t=1}^T (\eta_{t+1}^{-1} - \eta_t^{-1}) \\
&\le 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{R_{1,\max}^2}{2}\sqrt{\sum_{i=1}^T \|Ax_i - Ax_{i-1}\|_*^2} + 4\eta_{T+1}^{-1} \\
&\le 22R_{1,\max}^2 + 1 + 2\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{R_{1,\max}^2}{2}\sqrt{\sum_{i=1}^T \|Ax_i - Ax_{i-1}\|_*^2} \\
&\qquad + 4\max\left\{ 2R_{1,\max}^{-2}\sqrt{\sum_{i=1}^T \|Ax_i - Ax_{i-1}\|_*^2}, 11 \right\} \\
&\le 22R_{1,\max}^2 + 45 + 10\sqrt{\sum_{t=1}^{T-1} \|Ax_t - Ax_{t-1}\|_*^2} + \frac{R_{1,\max}^2}{2}\sqrt{\sum_{i=1}^T \|Ax_i - Ax_{i-1}\|_*^2} \\
&\le 22R_{1,\max}^2 + 45 + \frac{20 + R_{1,\max}^2}{2}\sqrt{\sum_{t=1}^T \|Ax_t - Ax_{t-1}\|_*^2}
\end{aligned}$$

concluding the proof.

□

Proof of Lemma 7. We start with the observation that $\hat{a}_t$ and $\bar{a}_{t-1}$ are unbiased estimates of Ax_t and Ax_{t-1} respectively. Thats is $\mathbb{E}_{i_{t-1}}[\hat{a}_t] = Ax_t$ and $\mathbb{E}_{i_{t-1}}[\bar{a}_{t-1}] = Ax_{t-1}$. Hence we have

$$\mathbb{E}\left[\sum_{t=1}^T f_t^\top Ax_t - \inf_{f \in \Delta_n} \sum_{t=1}^T f^\top Ax_t\right] \le \mathbb{E}\left[\sum_{t=1}^T \langle f_t, \hat{a}_t \rangle - \inf_{f \in \Delta_n} \sum_{t=1}^T \langle f, \hat{a}_t \rangle\right]$$

Using the predictable sequences result, specifically using the same line of proof as the one used to arrive at Equation 26 in Proposition 6 we get that the unnormalized regret for Player I can be upper bounded as,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} f_t^\top A x_t - \inf_{f\in\Delta_n}\sum_{t=1}^{T} f^\top A x_t\right] \\
&\le \mathbb{E}\left[(\eta_1^{-1}+\eta_T^{-1})R_{1,\max}^2 + 1 + \sum_{t=1}^{T}\|\hat{a}_t-\bar{a}_{t-1}\|_*\|g_t-f_t\| - \frac{1}{2}\sum_{t=1}^{T}\eta_t^{-1}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right)\right] \\
&\le \mathbb{E}\Bigg[(\eta_1^{-1}+\eta_T^{-1})R_{1,\max}^2 + 1 + \sum_{t=1}^{T}\eta_{t+1}\|\hat{a}_t-\bar{a}_{t-1}\|_*^2 + \frac{1}{4}\sum_{t=1}^{T}\eta_{t+1}^{-1}\|g_t-f_t\|^2 \\
&\qquad - \frac{1}{2}\sum_{t=1}^{T}\eta_t^{-1}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right)\Bigg] \\
&\le \mathbb{E}\Bigg[(\eta_1^{-1}+\eta_T^{-1})R_{1,\max}^2 + 1 + \sum_{t=1}^{T}\eta_{t+1}\|\hat{a}_t-\bar{a}_{t-1}\|_*^2 + \sum_{t=1}^{T}(\eta_{t+1}^{-1}-\eta_t^{-1}) \\
&\qquad - \frac{1}{4}\sum_{t=1}^{T}\eta_t^{-1}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right)\Bigg] \\
&\le \mathbb{E}\left[2R_{1,\max}^2(\eta_1^{-1}+\eta_T^{-1}) + 1 + \sum_{t=1}^{T}\eta_{t+1}\|\hat{a}_t-\bar{a}_{t-1}\|_*^2 - \frac{1}{4}\sum_{t=1}^{T}\eta_t^{-1}\left(\|g'_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right)\right]
\end{aligned}$$

Since

$$\eta_t = \min\left\{R_{1,\max}\frac{\sqrt{\sum_{i=1}^{t-1}\|\hat{a}_i-\bar{a}_{i-1}\|_*^2}-\sqrt{\sum_{i=1}^{t-2}\|\hat{a}_i-\bar{a}_{i-1}\|_*^2}}{\|\hat{a}_{t-1}-\bar{a}_{t-2}\|_*^2}, \frac{1}{28m\, R_{2,\max}}\right\}$$

and

$$28mR_{2,\max} \le \eta_t^{-1} \le \max\left\{2R_{1,\max}^{-1}\sqrt{\sum_{i=1}^{t-1}\|Ax_i-Ax_{i-1}\|_*^2}, 28m\, R_{2,\max}\right\}$$

With this, the upper bound on Player I's unnormalized regret is

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} f_t^\top A x_t - \inf_{f\in\Delta_n}\sum_{t=1}^{T} f^\top A x_t\right] &\le 56mR_{2,\max}R_{1,\max}^2 + 1 + \frac{7}{2}R_{1,\max}\sqrt{\sum_{t=1}^{T}\|\hat{a}_t-\bar{a}_{t-1}\|_*^2} \qquad (29) \\
&\quad - \frac{28mR_{2,\max}}{4}\sum_{t=1}^{T}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right)
\end{aligned}$$

Both players are honest : We first consider the case when both players play the prescribed algorithm. In this case, a similar regret bound holds for Player II. Adding the regret of the second player who uses step size η'_t, the overall bound on the suboptimality, as in Eq.(6), is

$$\begin{aligned}
&2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max}+nR_{2,\max}\right) + \frac{7}{2}R_{1,\max}\sqrt{\sum_{t=1}^{T}\|\hat{a}_t-\bar{a}_{t-1}\|_*^2} + \frac{7}{2}R_{2,\max}\sqrt{\sum_{t=1}^{T}\left\|\hat{b}_t-\bar{b}_{t-1}\right\|_*^2} \\
&\quad - 7mR_{2,\max}\sum_{t=1}^{T}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right) - 7nR_{1,\max}\sum_{t=1}^{T}\left(\|y_t-x_t\|^2+\|y'_{t-1}-x_t\|^2\right)
\end{aligned}$$

Now note that

$$\begin{aligned}
\|\hat{a}_t-\bar{a}_{t-1}\|_* &\le \frac{n}{2\delta}\|u_{i_{t-1}}\|_*\left|r_t^+-r_t^-+\bar{r}_t^--\bar{r}_t^+\right| = \frac{n}{2\delta}\|u_{i_{t-1}}\|_*\left|2\delta u_{i_{t-1}}^\top A(x_t-x_{t-1})\right| \le n\left|A(x_t-x_{t-1})\right| \\
&\le n\|x_t-x_{t-1}\|
\end{aligned}$$

Similarly we have $\left\|\hat{b}_t-\bar{b}_{t-1}\right\|_* \le m\|f_t-f_{t-1}\|$. Hence using this, we can bound the sub-optimality as

$$\begin{aligned}
&2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max}+nR_{2,\max}\right) + \frac{7}{2}nR_{1,\max}\sqrt{\sum_{t=1}^{T}\|x_t-x_{t-1}\|^2} + \frac{7}{2}mR_{2,\max}\sqrt{\sum_{t=1}^{T}\|f_t-f_{t-1}\|^2} \\
&\quad - 7mR_{2,\max}\sum_{t=1}^{T}\left(\|g_t-f_t\|^2+\|g'_{t-1}-f_t\|^2\right) - 7nR_{1,\max}\sum_{t=1}^{T}\left(\|y_t-x_t\|^2+\|y'_{t-1}-x_t\|^2\right)
\end{aligned}$$

Using the fact that $\sqrt{c^2} \le c^2 + 1$ we further bound sub-optimality by

$$2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max} + nR_{2,\max}\right) + \frac{7}{2}mR_{2,\max} + \frac{7}{2}nR_{1,\max} + \frac{7}{2}nR_{1,\max}\sum_{t=1}^{T}\|x_t - x_{t-1}\|^2$$
$$+ \frac{7}{2}mR_{2,\max}\sum_{t=1}^{T}\|f_t - f_{t-1}\|^2 - 7mR_{2,\max}\sum_{t=1}^{T}\left(\|g_t - f_t\|^2 + \|g'_{t-1} - f_t\|^2\right)$$
$$- 7nR_{1,\max}\sum_{t=1}^{T}\left(\|y_t - x_t\|^2 + \|y'_{t-1} - x_t\|^2\right)$$

Now note that
$$\|x_t - x_{t-1}\|^2 \le 2\|x_t - y'_{t-1}\|^2 + 2\|x_{t-1} - y'_{t-1}\|^2$$
and similarly
$$\|f_t - f_{t-1}\|^2 \le 2\|f_t - g'_{t-1}\|^2 + 2\|f_{t-1} - g'_{t-1}\|^2$$
Hence we can conclude that sub-optimality is bounded by

$$\begin{aligned}
&2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max} + nR_{2,\max}\right) + \frac{7}{2}mR_{2,\max} + \frac{7}{2}nR_{1,\max} \\
&\quad + 7mR_{2,\max}\sum_{t=1}^{T}\left(\|g'_t - f_t\|^2 - \|g_t - f_t\|^2\right) + 7nR_{1,\max}\sum_{t=1}^{T}\left(\|y'_t - x_t\|^2 - \|y_t - x_t\|^2\right) \\
&\quad \le 2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max} + nR_{2,\max}\right) + \frac{7}{2}mR_{2,\max} + \frac{7}{2}nR_{1,\max} \\
&\qquad + 28mR_{2,\max}\sum_{t=1}^{T}\|g'_t - g_t\| + 28nR_{1,\max}\sum_{t=1}^{T}\|y'_t - y_t\| \\
&\quad \le 2 + 56R_{1,\max}R_{2,\max}\left(mR_{1,\max} + nR_{2,\max}\right) + \frac{7}{2}mR_{2,\max} + \frac{7}{2}nR_{1,\max} \\
&\qquad + \frac{28(mR_{2,\max} + nR_{1,\max})}{T}
\end{aligned}$$

Just as in the proof of Proposition 6 we have $R_{1,\max} \le \sqrt{\log(nT^2)}$ and $R_{2,\max} \le \sqrt{\log(mT^2)}$ and so overall we get the bound on sub-optimality :

$$2 + 56\sqrt{\log(mT^2)\log(nT^2)}\left(m\sqrt{\log(nT^2)} + n\sqrt{\log(mT^2)}\right) + \frac{7}{2}m\sqrt{\log(mT^2)} + \frac{7}{2}n\sqrt{\log(nT^2)}$$
$$+ \frac{28(m\sqrt{\log(mT^2)} + n\sqrt{\log(nT^2)})}{T}$$

Player II deviates from algorithm : Now let us consider the case when the Player 2 deviates from the prescribed algorithm. In this case, note that starting from Eq. (29) and simply dropping the negative term we get,

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t^\top A x_t - \inf_{f\in\Delta_n}\sum_{t=1}^{T} f^\top A x_t\right] \le 56mR_{2,\max}R_{1,\max}^2 + 1 + \frac{7}{2}R_{1,\max}\sqrt{\sum_{t=1}^{T}\|\hat{a}_t - \bar{a}_{t-1}\|_*^2}$$

As we noted earlier, $\|\hat{a}_t - \bar{a}_{t-1}\|_* \le n\|x_t - x_{t-1}\|$ and so,

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t^\top A x_t - \inf_{f\in\Delta_n}\sum_{t=1}^{T} f^\top A x_t\right] \le 56mR_{2,\max}R_{1,\max}^2 + 1 + \frac{7}{2}nR_{1,\max}\sqrt{\sum_{t=1}^{T}\|x_t - x_{t-1}\|^2}$$

Further noting that $R_{1,\max} \le \sqrt{\log(nT)}$ and $R_{2,\max} \le \sqrt{\log(mT)}$ we conclude that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t^\top A x_t - \inf_{f\in\Delta_n}\sum_{t=1}^{T} f^\top A x_t\right] \le 56m\sqrt{\log(mT)}\log(nT) + 1 + \frac{7}{2}n\sqrt{\log(nT)\sum_{t=1}^{T}\|x_t - x_{t-1}\|^2}$$

This concludes the proof.

□

***Proof of Lemma 8*.** Noting that the constraints are all H-strongly smooth and that the objective w.r.t. constraint player is linear (weighted average over constraints), we can apply Lemma 4 to the optimization problem with $\mathcal{R}_1(\cdot) = \frac{1}{2}\|\cdot\|_2^2$ and $\mathcal{R}_2$ the entropy function to obtain that

$$\max_{i\in[d]} G_i(\bar{f}_T) - \operatorname*{argmin}_{f\in\mathcal{F}} \max_{i\in[d]} G_i(f)$$
$$\le \frac{\|f^* - g_0\|_2^2}{\eta} + \frac{\alpha}{2}\sum_{t=1}^{T}\|\sum_{i=1}^{d} x_t[i]\nabla G_i(f_t) - \sum_{i=1}^{d} y_{t-1}[i]\nabla G_i(g_{t-1})\|_2^2 + \frac{1}{2\alpha}\sum_{t=1}^{T}\|g_t - f_t\|_2^2 - \frac{1}{2\eta}\sum_{t=1}^{T}\left(\|g_t - f_t\|_2^2 + \|g_{t-1} - f_t\|_2^2\right)$$
$$+ \frac{\log d}{\eta'} + \frac{\beta}{2}\sum_{t=1}^{T}\|\mathbf{G}(f_t) - \mathbf{G}(f_t)\|_\infty^2 + \frac{1}{2\beta}\sum_{t=1}^{T}\|y_t - x_t\|_1^2 - \frac{1}{2\eta'}\sum_{t=1}^{T}\left(\|y_t - x_t\|_1^2 + \|y_{t-1} - x_t\|_1^2\right)$$
$$\le \frac{\|f^* - g_0\|_2^2}{\eta} + \frac{H+1}{2}\sum_{t=1}^{T}\|f_t - g_{t-1}\|_2^2 - \frac{1}{2\eta}\sum_{t=1}^{T}\left(\|g_t - f_t\|_2^2 + \|g_{t-1} - f_t\|_2^2\right) + \frac{\log d}{\eta'}$$

where last step we picked $\beta = \eta'$ and $\alpha = 1$. Piciking $\eta' = 1/\eta$ appropriately we conclude that

$$\max_{i\in[d]} G_i(\bar{f}_T) - \operatorname*{argmin}_{f\in\mathcal{F}} \max_{i\in[d]} G_i(f) \le \frac{4H(\|f^* - g_0\|_2\sqrt{\log d})}{T}$$

Now since T is such that $T \ge \frac{4H(\|f^* - g_0\|_2\sqrt{\log d})}{\epsilon}$ we can conclude that

$$\max_{i\in[d]} G_i(\bar{f}_T) - \operatorname*{argmin}_{f\in\mathcal{F}} \max_{i\in[d]} G_i(f) \le \epsilon$$

Observe that for an optimal solution $f^* \in \mathcal{G}$ to the original optimization problem (10) we have that $f^* \in \mathcal{F}$ and $\forall i, G_i(f^*) \le 1$. Thus,

$$\max_{i\in[d]} G_i\left(\bar{f}_T\right) \le 1 + \epsilon$$

Hence, $\bar{f}_T \in \mathcal{F}$ is a solution that attains the optimum value F^* and almost satisfies the constraints (violates by at most ϵ). Now we have from the lemma statement that $f_0 \in \mathcal{G}$ is such that $c^\top f_0 \ge 0$ and for every $i \in [d]$, $G_i(f_0) \le 1 - \gamma$. Hence by convexity of G_i, we have that for every $i \in [d]$,

$$G_i\left(\alpha f_0 + (1-\alpha)\bar{f}_T\right) \le \alpha G_i(f_0) + (1-\alpha)G_i(\bar{f}_T) \le \alpha(1-\gamma) + (1-\alpha)(1+\epsilon) \le 1$$

Thus for $\alpha = \frac{\epsilon}{\epsilon+\gamma}$ and $\hat{f}_T = (1-\alpha)\bar{f}_T + \alpha f_0$ we can conclude that $\hat{f}_T \in \mathcal{G}$ and that all the constraints are satisfied. That is for every $i \in [d]$, $G_i(\hat{f}_T) \le 1$. Also note that

$$c^\top \hat{f}_T = (1-\alpha)c^\top \bar{f}_T + \alpha c^\top f_0 = (1-\alpha)F^* = \frac{\gamma F^*}{\epsilon + \gamma}$$

and, hence, $\hat{f}_T$ is an approximate maximizer, that is

$$c^\top(f^* - \hat{f}_T) \le F^* - \frac{\gamma F^*}{\epsilon+\gamma} = \frac{F^*\epsilon + F^*\gamma - \gamma F^*}{\epsilon + \gamma} = \epsilon\left(\frac{F^*}{\epsilon+\gamma}\right) \le \frac{\epsilon}{\gamma}F^*$$

Thus we obtain a $(1+\frac{\epsilon}{\gamma})$-optimal solution in the multiplicative sense which concludes the proof. □

***Proof of Corollary 9*.** As mentioned, for both players, the time to perform each step of the optimistic mirror descent in the max-flow problem is $O(d)$. Now further note that max-flow is a linear programming problem and so we are ready tp apply Lemma 8. Specifically for f_0 we use the $\mathbf{0}$ flow which is in $\mathcal{G}$ (though not in $\mathcal{F}$) and note that for f_0 we have that $\gamma = 1$. Applying Lemma 8 we get that number of iterations T we need to reach an ϵ approximate solution is given by

$$T \le \frac{4H\|f^* - g_0\|_2\sqrt{\log d}}{\epsilon}$$

Now we can use $g_0 = \operatorname*{argmin}_{g\in\mathcal{F}} \|g\|^2$ which can be computed in $O(d)$ time. Now note that $\|f^* - g_0\|_2 \le \|f^*\|_2 + \|g_0\|_2 \le 2\sqrt{d}$. Hence number of iterations is at most

$$T \le \frac{8H\sqrt{d\log d}}{\epsilon}$$

Since each iteration has time complexity $O(d)$, the overall complexity of the algorithm is given by

$$O\left(\frac{d^{3/2}\sqrt{\log d}}{\epsilon}\right)$$

this concludes the proof. $\square$