A Complete Proofs of Section 4

A.1 Proof of Theorem 1

A.1.1 Highlight

The key idea to proving Theorem 1 is that the gap function $\rho(x) \coloneqq f_x(x) - \min_{x' \in X} f_x(x')$ can be used as a residual function for the above EP/VI/FP in Theorem 1. That is, $\rho(x)$ is non-negative, computable in polynomial time (it is a convex program), and $\rho(x) = 0$ if and only if $x \in X^\star$ (because $f_x(\cdot)$ is convex $\forall x \in \mathcal{X}$). Therefore, to show Theorem 1, we only need to prove that solving one of these problems is equivalent to achieving sublinear dynamic regret.

First, suppose an algorithm generates a sequence $\{x_n \in \mathcal{X}\}$ such that $\lim_{n\to\infty} x_n = x^\star$, for some $x^\star \in X^\star$. To show this implies $\{x_n \in \mathcal{X}\}$ has sublinear dynamic regret, we first show $\lim_{x \to x^\star \in X^\star} \rho(x) = 0$. Then define $\rho_n = \rho(x_n)$. Because $\lim_{n\to\infty} \rho_n = 0$, we have $\text{Regret}_N^d = \sum_{n=1}^N \rho_n = o(N)$.

Next, we prove the opposite direction. Suppose an algorithm generates a sequence $\{x_n \in \mathcal{X}\}$ with sublinear dynamic regret. This implies that $\hat{\rho}_N \coloneqq \min_n \rho_n \le \frac{1}{N} \sum_{n=1}^N \rho_n$ is in $o(1)$ and non-increasing. Thus, $\lim_{N\to\infty} \hat{\rho}_N = 0$. As ρ is a proper residual, the algorithm solves the EP/VI/FP problem by returning the decision associated with $\hat{\rho}_N$.

The proof of PPAD-completeness is based on converting the residual of a Brouwer's fixed-point problem to a bifunction, and use the solution along with $\hat{\rho}_N$ above as the approximate solution.

Note that the gap function ρ, despite motivated by dynamic regret here, corresponds to a natural gap function $r_{ep}(x) \coloneqq \max_{x' \in \mathcal{X}} -F(x, x')$ used in the EP literature, showing again a close connection between the dynamic regret and the EP in Theorem 1. Nonetheless, $\rho(x)$ is not conventional for VIs and FPs. Below we relate $\rho(x)$ to some standard residuals of VIs and FPs under a stronger assumption on f.

Proposition 11. *For $\epsilon > 0$, consider some $x_\epsilon \in \mathcal{X}$ such that $\rho(x_\epsilon) \le \epsilon$. If $f_{x_\epsilon}(\cdot)$ is α-strongly convex, then $\lim_{\epsilon\to 0} \langle \nabla f_{x_\epsilon}(x_\epsilon), x - x_\epsilon \rangle \ge 0$, $\forall x \in \mathcal{X}$, and $\lim_{\epsilon \to 0} \|x_\epsilon - T(x_\epsilon)\| = 0$.*

A.1.2 Full proof

Now we give the details of the steps above.

We first show the solutions sets of the EP, the VI, and the FP are identical.

- 2. $\implies$ 3.
  Let $x_{\text{VI}}^\star \in \mathcal{X}$ be a solution to $\text{VI}(\mathcal{X}, F)$ where $F(x) = \nabla f_x(x)$. That is, for all $x \in \mathcal{X}$, $\langle \nabla f_{x_{\text{VI}}^\star}(x_{\text{VI}}^\star), x - x_{\text{VI}}^\star \rangle \ge 0$. The sufficient first-order condition for optimality implies that $x_{\text{VI}}^\star$ is optimal for $f_{x_{\text{VI}}^\star}$. Therefore, $f_{x_{\text{VI}}^\star}(x_{\text{VI}}^\star) \le f_{x_{\text{VI}}^\star}(x)$ for all $x \in \mathcal{X}$, meaning that $x_{\text{VI}}^\star$ is also a solution to $\text{EP}(\mathcal{X}, \Phi)$ where $\Phi(x, x') = f_x(x') - f_x(x)$.

- 3. $\implies$ 4.
  Let $x_{\text{EP}}^\star \in \mathcal{X}$ be a solution to $\text{EP}(\mathcal{X}, \Phi)$. By definition, it satisfies $f_{x_{\text{EP}}^\star}(x_{\text{EP}}^\star) \le f_{x_{\text{EP}}^\star}(x)$ for all $x \in \mathcal{X}$, which implies $x_{\text{EP}}^\star = \arg\min_{x \in \mathcal{X}} f_{x_{\text{EP}}^\star}(x) = T(x_{\text{EP}}^\star)$. Therefore, $x_{\text{EP}}^\star$ is a also solution to $\text{FP}(\mathcal{X}, T)$, where $T(x') = \arg\min_{x \in \mathcal{X}} f_{x'}(x)$.

- 4. $\implies$ 2.
  If $x_{\text{FP}}^\star$ is a solution to $\text{FP}(\mathcal{X}, T)$, then $x_{\text{FP}}^\star = \arg\min_{x \in \mathcal{X}} f_{x_{\text{FP}}^\star}(x)$. By the necessary first-order condition for optimality, we have $\langle \nabla f_{x_{\text{FP}}^\star}(x_{\text{FP}}^\star) x - x_{\text{FP}}^\star \rangle \ge 0$ for all $x \in \mathcal{X}$. Therefore $x_{\text{FP}}^\star$ is also a solution to $\text{VI}(\mathcal{X}, F)$ where $F(x) = \nabla f_x(x)$.

Let $X^\star$ denote their common solution sets. To finish the proof of equivalence in Theorem 1, we only need to show that converging to $X^\star$ is equivalent to achieving sublinear dynamic regret.

- Suppose there is an algorithm that generates a sequence $\{x_n \in \mathcal{X}\}$ such that $\lim_{n\to\infty} x_n = x^\star$, for some $x^\star \in X^\star$. To show this implies $\{x_n \in \mathcal{X}\}$ has sublinear dynamic regret, we need a continuity lemma.

Lemma 1. $\lim_{x \to x^\star \in X^\star} \rho(x) = 0$.

Proof. Let $\bar{x} \in T(x)$. Using convexity, we can derive that

$$\begin{aligned}
&\rho(x) = f_x(x) - f_x(\bar{x}) \leq \langle \nabla f_x(x), x - \bar{x} \rangle \\
&\leq \langle \nabla f_{x^\star}(x^\star), x - \bar{x} \rangle + \|\nabla f_{x^\star}(x^\star) - \nabla f_x(x)\|_* \|x - \bar{x}\| \\
&\leq \langle \nabla f_{x^\star}(x^\star), x^\star - \bar{x} \rangle + \|\nabla f_{x^\star}(x^\star)\|_* \|x - x^\star\| + \|\nabla f_{x^\star}(x^\star) - \nabla f_x(x)\|_* \|x - \bar{x}\| \\
&\leq \|\nabla f_{x^\star}(x^\star)\|_* \|x - x^\star\| + \|\nabla f_{x^\star}(x^\star) - \nabla f_x(x)\|_* \|x - \bar{x}\|
\end{aligned}$$

where the second and the third inequalities are due to Cauchy-Schwarz inequality, and the last inequality is due to that $x^\star$ solves $\mathrm{VI}(\mathcal{X}, \nabla f)$. By continuity of ∇f, the above upper bound vanishes as $x \to x^\star$. □

For short hand, let us define $\rho_n = \rho(x_n)$; we can then write $\mathrm{Regret}_N^d = \sum_{n=1}^N \rho_n$. By Lemma 1, $\lim_{n\to\infty} x = x^\star$ implies that $\lim_{n\to\infty} \rho_n = 0$. Finally, we show by contradiction that $\lim_{n\to\infty} \rho_n = 0$ implies $\mathrm{Regret}_N^d = o(N)$. Suppose the dynamic regret is linear. Then $c > 0$ exists such that there is a subsequence $\{\rho_{n_i}\}$ satisfying $\rho_{n_i} \geq c$ for all n_i. However, this contradicts with $\lim_{n\to\infty} \rho_n = 0$.

- We can also prove the opposite direction. Suppose an algorithm generates a sequence $\{x_n \in \mathcal{X}\}$ with sublinear dynamic regret. This implies that $\hat{\rho}_N \coloneqq \min_n \rho_n \leq \frac{1}{N} \sum_{n=1}^N \rho_n$ is in $o(N)$ and non-increasing. Thus, $\lim_{N\to\infty} \hat{\rho}_N = 0$ and the algorithm solves the VI/EP/FP problem because ρ is a residual. Alternatively, we may view $\hat{\rho}$ as a Lyapunov-like function. The sequence of minimizers $\hat{x}_N = \arg\min_{x_n} \rho(x_n)$ are confined to the level sets of ρ, which converge to the zero-level set. Since $\mathcal{X}$ is compact, $\hat{x}_N$ converges to this set.

Finally, we show the PPAD-completeness by proving that achieving sublinear dynamic regret with polynomial dependency on d implies solving a Brouwer's problem (finding a fixed point of a continuous point-to-point map on a convex compact set). Because Brouwer's problem is known to be PPAD-complete Daskalakis et al. (2009), we can use this algorithm to solve all PPAD problems.

Given a Brouwer's problem on $\mathcal{X}$ with some continuous map $\hat{T}$. We can define the bifunction f as $f_{x'}(x) = \frac{1}{2}\|x - \hat{T}(x')\|_2^2$, where $\|\cdot\|_2$ is Euclidean. Obviously, this f satisfies Definition 1, and its gap function is zero at $x^\star$ if and only $x^\star$ is a solution to the Brouwer's problem. Suppose we have an algorithm that achieves sublinear dynamic regret for continuous online learning. We can use the definition $\hat{\rho}_N$ in the proof above to return a solution whose gap function is less than $\frac{1}{2}\epsilon^2$, which implies an ϵ-approximate solution to Brouwer's problem (i.e. $\|x - \hat{T}(x)\| \leq \epsilon$). If the dynamic regret depends polynomially on d, we have such an N in $poly(d)$, which implies solving any Brouwer's problem in polynomial time.

A.1.3 Proof of Proposition 11

For the VI problem, let $x_\epsilon^* = T(x_\epsilon)$ and notice that

$$\frac{\alpha}{2}\|x_\epsilon - x_\epsilon^*\|^2 \leq f_{x_\epsilon}(x_\epsilon) - f_{x_\epsilon}(x_\epsilon^*) \leq \epsilon \tag{5}$$

for some $\alpha > 0$. Therefore, for any $x \in \mathcal{X}$,

$$\begin{aligned}
\langle \nabla f_{x_\epsilon}(x_\epsilon), x - x_\epsilon \rangle &\geq \langle \nabla f_{x_\epsilon}(x_\epsilon^*), x - x_\epsilon \rangle - \|\nabla f_{x_\epsilon}(x_\epsilon^*) - \nabla f_{x_\epsilon}(x_\epsilon)\|_* \|x - x_\epsilon\| \\
&\geq \langle \nabla f_{x_\epsilon}(x_\epsilon^*), x - x_\epsilon^* \rangle - \|\nabla f_{x_\epsilon}(x_\epsilon^*)\|_* \|x_\epsilon - x_\epsilon^*\| - \|\nabla f_{x_\epsilon}(x_\epsilon^*) - \nabla f_{x_\epsilon}(x_\epsilon)\|_* \|x - x_\epsilon\| \\
&\geq -\|\nabla f_{x_\epsilon}(x_\epsilon^*)\|_* \|x_\epsilon - x_\epsilon^*\| - \|\nabla f_{x_\epsilon}(x_\epsilon^*) - \nabla f_{x_\epsilon}(x_\epsilon)\|_* \|x - x_\epsilon\|
\end{aligned}$$

Since $\|x_\epsilon - x_\epsilon^*\|^2 \leq \frac{2\epsilon}{\alpha}$, by continuity of ∇f_{x_ϵ}, it satisfies that $\lim_{\epsilon\to 0} \langle \nabla f_{x_\epsilon}(x_\epsilon), x - x_\epsilon \rangle \geq 0$, $\forall x \in \mathcal{X}$.

For the fixed-point problem, similarly by (5), we see that $\lim_{\epsilon\to 0} \|x_\epsilon - T(x_\epsilon)\| = 0$

A.2 Proofs of Proposition 4

Proof of Proposition 4. Let $x_\star \in X_{\star\star}$. It holds that $\forall x \in \mathcal{X}$, $0 \geq \Phi(x, x_\star) = f_x(x_\star) - f_x(x) \geq \langle \nabla f_x(x), x_\star - x \rangle$, which implies $x_\star \in X_\star$. The condition for the converse case is obvious. □

A.3 Proof of Proposition 5

Because ∇f_x is α-strongly monotone, we can derive

$$\begin{aligned}\langle \nabla f_x(x) - \nabla f_y(y), x - y\rangle &= \langle \nabla f_x(x) - \nabla f_x(y), x - y\rangle + \langle \nabla f_x(y) - \nabla f_y(y), x - y\rangle \\ &\geq \alpha \|x-y\|^2 - \|\nabla f_x(y) - \nabla f_y(y)\|_* \|x-y\| \\ &\geq (\alpha - \beta)\|x-y\|^2\end{aligned}$$

$\forall x, y \in \mathcal{X}$, where the last step is due to β-regularity.

A.4 Proof of Proposition 6

The result follows immediately from the following lemma.

Lemma 2. *Suppose f is (α,β)-regular with $\alpha > 0$. Then F in Theorem 1 is point-valued and $\frac{\beta}{\alpha}$-Lipschitz.*

Proof. Let $x^* = F(x)$ and $y^* = F(y)$ for some $x, y \in \mathcal{X}$. By strong convexity, x^* and y^* are unique, and $\nabla f_x(\cdot)$ is α-strongly monotone; therefore it holds that

$$\begin{aligned}\langle \nabla f_x(y^*), y^* - x^*\rangle &\geq \langle \nabla f_x(x^*), y^* - x^*\rangle + \alpha\|x^* - y^*\|^2 \\ &\geq \alpha \|x^* - y^*\|^2\end{aligned}$$

Since y^* satisfies $\langle \nabla f_y(y^*), x^* - y^*\rangle \geq 0$, the above inequality implies that

$$\begin{aligned}\alpha\|x^* - y^*\|^2 &\leq \langle \nabla f_x(y^*), y^* - x^*\rangle \\ &\leq \langle \nabla f_x(y^*) - \nabla f_y(y^*), y^* - x^*\rangle \\ &\leq \|\nabla f_x(y^*) - \nabla f_y(y^*)\|_* \|y^* - x^*\| \\ &\leq \beta\|x-y\|\|y^* - x^*\|\end{aligned}$$

Rearranging the inequality gives the statement. □

B Dual Solution and Strongly Convex Sets

We show when the strong convexity property of $\mathcal{X}$ implies the existence of dual solution for VIs. We first recall the definition of strongly convex sets.

Definition 4. Let $\alpha_\mathcal{X} \geq 0$. A set $\mathcal{X}$ is called *$\alpha_\mathcal{X}$-strongly convex* if, for any $x, x' \in \mathcal{X}$ and $\lambda \in [0,1]$, it holds that, for all unit vector v, $\lambda x + (1-\lambda)x' + \frac{\alpha_\mathcal{X}\lambda(1-\lambda)}{2}\|x - x'\|^2 v \in \mathcal{X}$.

When $\alpha_\mathcal{X} = 0$, the definition reduces to usual convexity. Also, we see that this definition implies $\alpha_\mathcal{X} \leq \frac{4}{D_\mathcal{X}}$. In other words, larger sets are less strongly convex. This can also be seen from the lemma below.

Lemma 3. (Journée et al., 2010, Theorem 12) *Let f be non-negative, α-strongly convex, and β-smooth on a Euclidean space. Then the set $\{x | f(x) \leq r\}$ is $\frac{\alpha}{\sqrt{2r\beta}}$-strongly convex.*

Here we present the existence result.

Proposition 12. *Let $x^\star \in X^\star$. If $\mathcal{X}$ is $\alpha_\mathcal{X}$-strongly convex $\forall x \in \mathcal{X}$, it holds that $\langle F(x^*), x - x^*\rangle \geq \frac{\alpha_\mathcal{X}}{2}\|x - x^*\|^2 \|F(x^*)\|_*$. If further F is L-Lipschitz, this implies $\langle F(x), x - x^*\rangle \geq (\frac{\alpha_\mathcal{X}}{2}\|F(x^*)\|_* - L)\|x - x^*\|^2$, i.e. when $\alpha_\mathcal{X} \geq \frac{2L}{\|F(x^*)\|_*}$, $x^\star \in X_\star$.*

Proof of Proposition 12. Let $g = F(x^\star)$. Let $y = \lambda x + (1-\lambda)x^\star$ and $d = -\lambda(1-\lambda)\frac{\alpha_\mathcal{X}}{2}\|x - y\|^2 v$, for some $\lambda \in [0,1]$ and some unit vector v to be decided later. By $\alpha_\mathcal{X}$-strongly convexity of $\mathcal{X}$, we have $y + d \in \mathcal{X}$. We can derive

$$\begin{aligned}\langle g, x - x^\star\rangle &= \langle g, x - y - d\rangle + \langle g, y + d - x^\star\rangle \\ &\geq \langle g, x - y\rangle - \langle g, d\rangle \\ &= (1-\lambda)\langle g, x - x^\star\rangle - \langle g, d\rangle\end{aligned}$$

which implies $\langle g, x - x^\star\rangle \geq \frac{-\langle g, d\rangle}{\lambda} = (1-\lambda)\frac{\alpha_\mathcal{X}}{2}\|x - x^\star\|^2 \langle g, v\rangle$. Since we are free to choose λ and v, we can set $\lambda = 0$ and $v = \arg\max_{v:\|v\|\leq 1} \langle g, v\rangle$, which yields the inequality in the statement. □

C Complete Proofs of Section 5

In this section, we describe a general strategy to reduce monotone equilibrium problems (EPs) to continuous online learning (COL) problems. This reduction can be viewed as refinement and generalization of the classic reduction from convex optimization to adversarial online learning and that from saddle-point problem to two-player adversarial online learning. In comparison, our reduction

1. results in a single-player online learning problem, which allows for unified algorithm design
2. considers potential continuous relationship of the losses between different rounds through the setup of COL, which leads to a predictable online problem amenable to acceleration techniques, such as (Rakhlin and Sridharan, 2013; Juditsky et al., 2011; Cheng et al., 2019a).
3. and extends the concept to general convex problems, namely, monotone EPs, which includes of course convex optimization and convex-concave saddle-point problems but also fixed-point problems (FPs), variational inequalities (VIs), etc.

The results here are summarized as Theorem 2 and Theorem 3.

Here we further suppose $\Phi(x, x) = 0$ in the definition of EP. This is not a strong condition. First all the common source problems in introduced below in Appendix C.1.1 satisfy this condition. Generally, suppose we have some EP problem with $\Phi'(x, x) > 0$ for some x. We can define $\Phi(x, x) = \Phi'(x, x') - \Phi'(x, x')$. Then the solution of $\mathrm{EP}(\mathcal{X}, \Phi)$ are subset of the solution $\mathrm{EP}(\mathcal{X}, \Phi')$. In other words, allowing $\Phi(x, x) > 0$ only makes problem easier. We note that the below reduction and discussion can easily be extended to work directly with EPs with $\Phi(x, x) > 0$ by defining instead $f_x(x') = \Phi(x, x') - \Phi(x, x)$, but this will make the presentation less clean.

C.1 Background: Equilibrium Problems (EPs)

Let $\mathcal{X}$ be a compact and convex set in a finite dimensional space. Let $F : x \times x' \mapsto \Phi(x, x')$ be a bifunction[7] that is continuous in the first argument, convex in the second argument, and satisfies $\Phi(x, x) = 0$.[8] The problem $\mathrm{EP}(\mathcal{X}, F)$ aims to find $x^\star \in \mathcal{X}$ such that

$$\Phi(x^\star, x) \geq 0, \qquad \forall x \in \mathcal{X}$$

Its dual problem $\mathrm{DEP}(\mathcal{X}, F)$ finds $x_{\star\star} \in \mathcal{X}$ such that

$$\Phi(x, x_{\star\star}) \leq 0, \qquad \forall x \in \mathcal{X}$$

Based on the problem's definition, a natural residual (or gap function) of $\mathrm{EP}(\mathcal{X}, F)$ is

$$r_{ep}(x) \coloneqq - \min_{x' \in \mathcal{X}} \Phi(x, x')$$

which says the degree that the inequality in the EP definition is violated. A residual for $\mathrm{DEP}(\mathcal{X}, F)$ can be defined similarly as

$$r_{dep}(x') \coloneqq \max_{x \in \mathcal{X}} \Phi(x, x')$$

Sometimes EPs are called maxInf (or minSup) problems (Jofré and Wets, 2014), because

$$x^\star \in \arg\min_{x \in \mathcal{X}} r_{ep}(x) = \arg\max_{x \in \mathcal{X}} \min_{x' \in \mathcal{X}} \Phi(x, x')$$

In a special case, when $\Phi(\cdot, x)$ is concave. It reduces to a saddle-point problem.

We say a bifunction F is *monotone* if it satisfies

$$\Phi(x, x') + \Phi(x', x) \leq 0,$$

[7] We impose convexity and continuity to simplify the setup; similar results hold for subdifferentials and Lipschitz continuity defined based on hemi-continuity.

[8] As discussed, we concern only EP with $\Phi(x, x) = 0$ here

and we say F is skew-symmetric if

$$\Phi(x, x') = -\Phi(x, x'),$$

which implies F is monotone. Finally, we say the problem $\mathrm{EP}(\mathcal{X}, F)$ is monotone, if its bifunction F is monotone.

C.1.1 Examples

We review some source problems of EPs. Please refer to e.g. (Jofré and Wets, 2014; Konnov and Schaible, 2000) for a more complete survey.

Convex Optimization Consider $\min_{x\in\mathcal{X}} h(x)$ where h is convex. We can simply define

$$\Phi(x, x') = h(x') - h(x)$$

which is a skew-symmetric (and therefore monotone) bifunction.

We can also define (following the VI given by its optimality condition)

$$\Phi(x, x') = \langle \nabla h(x), x' - x \rangle .$$

We can easily verify that this construction is also monotone

$$\Phi(x, x') + \Phi(x', x) = \langle \nabla h(x), x' - x \rangle + \langle \nabla h(x'), x - x' \rangle = \langle \nabla h(x) - \nabla h(x'), x' - x \rangle \leq 0.$$

Suppose h is μ-strongly convex. We can also consider

$$\Phi(x, x') = \langle \nabla h(x), x' - x \rangle + \frac{\mu'}{2} \|x' - x\|^2$$

where $\mu' \leq \mu$. Such F is still monotone:

$$\Phi(x, x') + \Phi(x', x) = \langle \nabla h(x) - \nabla h(x'), x' - x \rangle + \mu' \|x' - x\|^2 \leq 0.$$

Saddle-Point Problem Let $\mathcal{U}$ and $\mathcal{V}$ to convex and compact sets in a finite dimensional space. Consider a convex-concave saddle point problem

$$\min_{u\in\mathcal{U}} \max_{v\in\mathcal{V}} \phi(u, v) \tag{6}$$

in which ϕ is continuous, $\phi(\cdot, y)$ is convex, and $\phi(x, \cdot)$ is concave. It is well known that in this case

$$\min_{u\in\mathcal{U}} \max_{v\in\mathcal{V}} \phi(u, v) = \max_{v\in\mathcal{V}} \min_{u\in\mathcal{U}} \phi(u, v) =: \phi^\star.$$

We can define a EP by the bifunction

$$\Phi(x, x') := -\phi(u, v') + \phi(u', v). \tag{7}$$

By definition we have the skew symmetry property, which implies monotonicity.

Variational Inequality A VI with a vector-valued map F finds $x^\star \in \mathcal{X}$ such that

$$\langle F(x^\star), x - x^\star \rangle \geq 0, \qquad \forall x \in \mathcal{X}.$$

To turn that into a EP, we can simply define

$$\Phi(x, x') = \langle F(x), x' - x \rangle .$$

Mixed Variational Inequality (MVI) MVI considers problems that finds $x^\star \in \mathcal{X}$ such that

$$h(x) - h(x^\star) + \langle F(x^\star), x - x^\star \rangle \geq 0, \qquad \forall x \in \mathcal{X}.$$

Following the previous idea, we can define its equivalent EP through the bifunction

$$\Phi(x, x') = h(x') - h(x) + \langle F(x), x' - x \rangle$$

C.2 More insights into residuals of primal and dual EPs

We derive further relationships between primal and dual EPs. These properties will be useful for understanding the reduction introduced in the next section.

C.2.1 Monotonicity

By the definition of monotonicity, $\Phi(x, x') + \Phi(x', x) \leq 0$, we can relate the primal and the dual residuals: for $\hat{x} \in \mathcal{X}$,

$$r_{dep}(\hat{x}) = \max_{x \in \mathcal{X}} \Phi(x, \hat{x}) \leq \max_{x \in \mathcal{X}} -\Phi(\hat{x}, x) = r_{ep}(\hat{x})$$

Let $X^\star$ and $X_{\star\star}$ be the solution sets of the EP and DEP, respectively. In other words, for monotone EPs, $X^\star \subseteq X_{\star\star}$.

C.2.2 Continuity

When $\Phi(\cdot, x)$ is continuous, it can be shown that $X^\star \subseteq X_{\star\star}$ (Konnov and Schaible, 2000) (this can be relaxed to hemi-continuity). Below we relate the primal and the dual residuals in this case. It implies that the convergence rate of the primal residual is slower than the dual residual.

Proposition 13. *Suppose $\Phi(\cdot, x)$ is L-Lipschitz continuous for any $x \in \mathcal{X}$ and $\max_{x,x' \in \mathcal{X}} \|x - x'\| \leq D$. If $r_{dep}(x) \leq 2LD$, the $r_{ep}(x) \leq 2\sqrt{2LD}\sqrt{r_{dep}(x)}$.*

Suppose in addition $\Phi(x, \cdot)$ is μ-strongly convex with $\mu > 0$. If $r_{dep}(x) \leq \frac{L^2}{\mu}$, we can remove the dependency on D and show $r_{ep}(x) \leq 2.8(\frac{L^2}{\mu})^{1/3} r_{dep}(x)^{2/3}$.

Proof. Let $y \in \mathcal{X}$ be arbitrary. Define $z = \tau x + (1 - \tau) y$, where $\tau \in [0, 1]$. Suppose x is an ϵ-approximate dual solution, i.e.,

$$r_{dep}(x) = \max_{x' \in \mathcal{X}} \Phi(x', x) = \epsilon$$

By convexity and $\Phi(z, z) = 0$, we can write

$$\begin{aligned} \epsilon \geq \Phi(z, x) &= \Phi(z, x) - \Phi(z, z) \\ &\geq \Phi(z, x) - \tau \Phi(z, x) - (1 - \tau)\Phi(z, y) = (1 - \tau)(\Phi(z, x) - \Phi(z, y)) \end{aligned}$$

Using this, we can then show

$$\begin{aligned} -\Phi(x, y) &= -\Phi(x, y) + \Phi(z, y) + (\Phi(z, x) - \Phi(z, y)) - \Phi(z, x) + \Phi(x, x) \\ &\leq |\Phi(z, y) - \Phi(x, y)| + |\Phi(x, x) - \Phi(z, x)| + \Phi(z, x) - \Phi(z, y) \\ &\leq 2(1 - \tau) L \|x - y\| + \Phi(z, x) - \Phi(z, y) && (\because \text{Lipschitz condition}) \\ &\leq 2(1 - \tau) L \|x - y\| + \frac{\epsilon}{1 - \tau} && (\because \text{The inequality above}) \\ &\leq 2(1 - \tau) LD + \frac{\epsilon}{1 - \tau} \end{aligned}$$

Assume $\epsilon \leq 2LD$ and let $(1 - \tau) = \sqrt{\frac{\epsilon}{2LD}}$, which satisfies $\tau \in [0, 1]$. Then

$$-\Phi(x, y) \leq 2\sqrt{2LD\epsilon}$$

When we have μ-strong convexity, we have a tighter bound

$$\epsilon \geq \Phi(z,x) = \Phi(z,x) - \Phi(z,z) \geq \Phi(z,x) - \tau\Phi(z,x) - (1-\tau)\Phi(z,y) + \frac{\mu\tau(1-\tau)}{2}\|x-y\|^2$$
$$= (1-\tau)(\Phi(z,x) - \Phi(z,y)) + \frac{\mu\tau(1-\tau)}{2}\|x-y\|^2$$

Using this, we can instead show (following similar steps as above)

$$\begin{aligned} -\Phi(x,y) &\leq 2(1-\tau)L\|x-y\| + \Phi(z,x) - \Phi(z,y) \\ &\leq 2(1-\tau)L\|x-y\| + \frac{\epsilon}{1-\tau} - \frac{\mu\tau}{2}\|x-y\|^2 \\ &\leq \frac{\epsilon}{1-\tau} + \frac{2L^2(1-\tau)^2}{\mu\tau} \end{aligned}$$

where the last inequality is simply $bx - \frac{a}{2}x^2 \leq \frac{b^2}{2a}$ for $a > 0$. Assume $\epsilon \leq \frac{L^2}{\mu} =: \frac{K}{2}$ and let $(1-\tau) = (\frac{\epsilon}{K})^{1/3} \in [0,1]$. We have the following inequality, where the last step uses $\epsilon \leq \frac{K}{2}$.

$$-\Phi(x,y) \leq \frac{\epsilon}{1-\tau} + \frac{2L^2(1-\tau)^2}{\mu\tau} = \epsilon^{2/3}K^{1/3}\left(1 + \frac{1}{1-(\frac{\epsilon}{K})^{1/3}}\right) \leq 2.2\epsilon^{2/3}K^{1/3}$$

□

C.2.3 Equivalence between primal and dual EPs.

An interesting special case of EP is those with *skew-symmetric* bifunctions, i.e.

$$\Phi(x,x') = -\Phi(x',x)$$

In this case, the EP and the DEP become identical

$$(DEP) \quad \Phi(x, x_{\star\star}) \leq 0 \qquad \iff \qquad -\Phi(x_{\star\star}, x) \leq 0 \qquad \iff \qquad \Phi(x_{\star\star}, x) \geq 0 \quad (EP)$$

and we have $X^\star = X_{\star\star}$ and naturally matching residuals

$$r_{dep}(\hat{x}) = r_{ep}(\hat{x}).$$

Recall from the results of the previous two subsections, generally, when $\Phi(\cdot, x)$ is Lipschitz and F is monotone (but not skew-symmetric), we have $X^\star = X_{\star\star}$ (as known before) but only ($\Phi(x,\cdot)$ is convex)

$$r_{dep}(x) \leq r_{ep}(x) \leq \sqrt{2LD}\sqrt{r_{dep}(x)} \tag{8}$$

or ($\Phi(x,\cdot)$ is μ-strongly convex)

$$r_{dep}(x) \leq r_{ep}(x) \leq 2.8(\frac{L^2}{\mu})^{1/3} r_{dep}(x)^{2/3}$$

C.2.4 Relationship with VIs

We can reduce a EP into a VI problem. We observe that if a point $x^\star \in \mathcal{X}$ satisfies

$$\Phi(x^\star, x) \geq 0, \qquad \forall x \in \mathcal{X}$$

if only if

$$\nabla_2 \Phi(x^\star, x^\star)^\top (x - x^\star) \geq 0, \qquad \forall x \in \mathcal{X}$$

(i.e. $x^\star$ is a global minimum of the function $\Phi(x^\star, \cdot)$), where ∇_2 denotes the partial derivative with respect to the second argument. Therefore, $\text{EP}(\mathcal{X}, \Phi)$ is equivalent to $\text{VI}(\mathcal{X}, F)$

$$\text{find } x^\star \in \mathcal{X} \quad \text{s.t.} \qquad \langle F(x), x' - x \rangle \geq 0, \qquad \forall x' \in \mathcal{X}$$

if we define F as

$$F : x \in \mathcal{X} \mapsto F(x) = \nabla_2 \Phi(x, x) \tag{9}$$

In a sense, this VI problem is a linearization of the EP problem. In other words, VIs are EPs whose bifunction satisfies that $\Phi(x, \cdot)$ is linear.

By the definition in (9), we can show that

$$r_{dvi}(\hat{x}) \leq r_{dep}(\hat{x}) \qquad \text{and} \qquad r_{ep}(\hat{x}) \leq r_{vi}(\hat{x})$$

And if Φ is monotone, then $F = \nabla_2 \Phi(x, x)$ is monotone (though the opposite is not true), because

$$\begin{aligned} \langle F(x), x' - x \rangle = \langle \nabla_2 \Phi(x, x), x' - x \rangle &\leq \Phi(x, x') && (\because \text{Convexity}) \\ &\leq -\Phi(x', x) && (\because \text{Monotonicity}) \\ &\leq \langle \nabla_2 \Phi(x', x'), x' - x \rangle = \langle F(x'), x' - x \rangle && (\because \text{Convexity}) \end{aligned}$$

Note the converse is not true, unless $\Phi(x, \cdot)$ is linear.

C.3 Reduction from Equilibrium Problems to Continuous Online Learning

Now we present the general reduction strategy. Given a EP $(\mathcal{X}, \Phi)$, we propose to define a COL problem by identifying

$$f_x(x') = \Phi(x, x')$$

We can see that this definition is consistent with Theorem 1: due to $\Phi(x, x) = 0$, it satisfies

$$f_x(x') - f_x(x) = \Phi(x, x') - \Phi(x, x) = \Phi(x, x')$$

Therefore, we can say a COL is *normalized* if $f_x(x) = 0$. In this case, f and Φ are interchangeable.

Below we relate the dynamic regret $\text{Regret}_N^d := \sum_{n=1}^N f_{x_n}(x_n) - \min_{x \in \mathcal{X}} f_{x_n}(x)$ and the static regret $\text{Regret}_N^s := \sum_{n=1}^N f_{x_n}(x_n) - \min_{x \in \mathcal{X}} \sum_{n=1}^N f_{x_n}(x)$ of this problem to the convergence to the EP's solution; note that the above definitions use the fact that in COL $l_n(x) = f_{x_n}(x)$.

C.3.1 Dynamic Regret and Primal Residual

We first observe that each instant term in the dynamic regret of this COL problem is exactly the residual function:

$$f_{x_n}(x_n) - \min_{x \in \mathcal{X}} f_{x_n}(x) = -\min_{x \in \mathcal{X}} \Phi(x_n, x) = r_{ep}(x_n)$$

Therefore, the average dynamic regret describes the rate the gap function converges to zero:

$$\sum_{n=1}^N r_{ep}(x_n) = \sum_{n=1}^N f_{x_n}(x_n) - \min_{x \in \mathcal{X}} f_{x_n}(x) = \text{Regret}_N^d$$

Note that the above relationship holds also for weighted dynamic regret. In general, it means that if the average dynamic regret converges, then the last iterate must converge to the solution set of the EP (since the residual is non-negative.)

C.3.2 Static Regret and Dual Residual of Monotone EPs

Next we relate the weighted static regret to the dual residual of the EP. Let $\{w_n\}$ be such that $w_n > 0$. Let $\hat{x}_N = \frac{1}{x_{1:N}} \sum_{n=1}^N w_n x_n$ for some $\{x_n \in \mathcal{X}\}_{n=1}^N$, where we define $w_{1:N} \coloneqq \sum_{n=1}^N w_n$. We can derive

$$\begin{aligned}
r_{dep}(\hat{x}_N) &= \max_{x\in\mathcal{X}} \Phi(x, \hat{x}_N) \\
&\leq \max_{x\in\mathcal{X}} \frac{1}{w_{1:N}} \sum_{n=1}^N w_n \Phi(x, x_n) && (\because \text{Convexity}) \\
&\leq \max_{x\in\mathcal{X}} \frac{1}{w_{1:N}} \sum_{n=1}^N -w_n \Phi(x_n, x) && (\because \text{Monotonocity}) \\
&= -\min_{x\in\mathcal{X}} \frac{1}{w_{1:N}} \sum_{n=1}^N w_n \Phi(x_n, x) \\
&= \frac{1}{w_{1:N}} \sum_{n=1}^N w_n \Phi(x_n, x_n) - \min_{x\in\mathcal{X}} \frac{1}{w_{1:N}} \sum_{n=1}^N w_n \Phi(x_n, x) && (\because \Phi(x_n, x_n) = 0) \\
&= \frac{1}{w_{1:N}} \left(\sum_{n=1}^N w_n f_n(x_n) - \min_{x\in\mathcal{X}} \sum_{n=1}^N w_n f_n(x) \right) \\
&=: \frac{\text{Regret}_N^s(w)}{w_{1:N}}
\end{aligned}$$

Note that the inequality $r_{dep}(\hat{x}_N) \leq \frac{\text{Regret}_N^s(w)}{w_{1:N}}$ holds for *any* sequence $\{x_n\}$ and $\{w_n\}$. Interestingly, by (8), we see that by the definition of regrets and the property of monotonicity and local Lipschitz continuity, it holds that

$$\frac{r_{ep}(\hat{x}_N)^2}{2LD} \leq r_{dep}(\hat{x}_N) \leq \frac{\text{Regret}_N^s(w)}{w_{1:N}} \leq \frac{\text{Regret}_N^d(w)}{w_{1:N}} =: \frac{\sum_{n=1}^N w_n r_{ep}(x_n)}{w_{1:N}}$$

where L is the Lipschitz constant of $\Phi(\cdot, x)$ and D is the size of $\mathcal{X}$.

C.4 Summary

Let us summarize the insights gained from the above discussions.

1. We can reduce EP($\mathcal{X}, \Phi$) with monotone Φ to the COL problem with $l_n(x) = \Phi(x_n, x)$
2. In this COL, the convergence in (weighted) average dynamic regret implies the convergence of the last iterate to the primal solution set. The convergence in (weighted) average static regret implies the convergence of the (weighted) average decision to the dual solution set.
3. Because any dual solution is a primal solution when $\Phi(\cdot, x)$ is continuous, this implies the (weighted) average solution above also converges to the primal solution set. Particularly, if the problem is Lipschitz, we can show $r_{ep} \leq O(\sqrt{r_{dep}})$ and therefore we can also quantify the exact quality of $\hat{x}_N$ in terms of the primal EP (though it results in a slower rate).
4. When the problem is skew-symmetric (as in the case of common reductions from optimization and saddle-point problems), we have exactly $r_{ep} = r_{dep}$. This means the average static regret rate directly implies the quality of $\hat{x}_N$ in terms of the primal residual, *without* rate degradation.

D Complete Proofs of Section 6

D.1 Proof of Theorem 4

The main idea is based on the decomposition that

$$\text{Regret}_N^d = \sum_{n=1}^N f_{x_n}(x_n) - f_{x_n}(x^\star) + \sum_{n=1}^N f_{x_n}(x^\star) - f_{x_n}(x_n^*) \tag{10}$$

For the first term, $\sum_{n=1}^{N} f_{x_n}(x_n) - f_{x_n}(x^\star) = \text{Regret}_N^s(x^\star) \leq \text{Regret}_N^s$ and $f_{x_n}(x_n) - f_{x_n}(x^\star) \leq \langle \nabla f_{x_n}(x_n), x_n - x^\star \rangle \leq G\Delta_n$. For the second term, we derive

$$\begin{aligned}
&f_{x_n}(x^\star) - f_{x_n}(x_n^*) \\
&\leq \langle \nabla f_{x_n}(x^\star), x^\star - x_n^* \rangle - \frac{\alpha}{2}\|x^\star - x_n^*\|^2 \\
&\leq \langle \nabla f_{x_n}(x^\star) - \nabla_{x^\star} f(x^\star), x^\star - x_n^* \rangle - \frac{\alpha}{2}\|x^\star - x_n^*\|^2 \\
&\leq \|\nabla f_{x_n}(x^\star) - \nabla_{x^\star} f(x^\star)\|_* \|x^\star - x_n^*\| - \frac{\alpha}{2}\|x^\star - x_n^*\|^2 \\
&\leq \beta \|x_n - x^\star\| \|x^\star - x_n^*\| - \frac{\alpha}{2}\|x^\star - x_n^*\|^2 \\
&\leq \min\{\beta D_{\mathcal{X}} \|x_n - x^\star\|, \frac{\beta^2}{2\alpha}\|x_n - x^\star\|^2\}
\end{aligned}$$

in which the second inequality is due to that $x^\star \in X^\star$ and the fourth inequality is due to β-regularity. Combining the two terms gives the upper bound. For the lower bound, we notice that when $x_\star \in X_\star$, we have $f_{x_n}(x_n) - f_{x_n}(x_\star) \geq 0$. Since by Proposition 1 $x_\star \in X^\star$ is also true, we can use (10) and the fact that $f_{x_n}(x_\star) - f_{x_n}(x_n^*) \geq \frac{\alpha}{2}\|x_\star - x_n^*\|^2$ to derive the lower bound.

D.2 Proof of Corollary 1

By Proposition 5, ∇f is $(\alpha - \beta)$-strongly monotone, implying $\langle \nabla f_{x_n}(x_n), x_n - x^\star \rangle \geq (\alpha - \beta)\Delta_n^2$, where we recall that $\Delta_n = \|x_n - x^\star\|$ and $x^\star \in X^\star$. Because $\sum_{n=1}^{N} \langle \nabla f_{x_n}(x_n), x_n - x^\star \rangle = \text{Regret}_N^s(x^\star) \leq \text{Regret}_N^s$, we have by Theorem 4 the inequality in the statement.

D.3 Proof of Proposition 7

In this case, by Proposition 6, T is non-expansive. We know that, e.g., Mann iteration (Mann, 1953), i.e., for $\eta_n \in (0, 1)$ we set

$$x_{n+1} = \eta_n x_n + (1 - \eta_n) x_n^*, \tag{11}$$

converges to some $x^\star \in X^\star$; in view of (11), the greedy is update is equivalent to Mann iteration with $\eta_n = 1$. As Mann iteration converges in general Hilbert space, by Theorem 1, it has sublinear dynamic regret with some constant that is polynomial in d.

D.4 Proof of Proposition 8

We first establish a simple lemma related to the smoothness of $\nabla f_x(x)$ and then a result on the convergence of the Bregman divergence $B_R(x_n \| x^\star)$. The purpose of the second lemma is to establish essentially a contraction showing that the distance between the equilibrium point $x^\star$ and x_n strictly decreases.

Lemma 4. *If, $\forall x \in \mathcal{X}$, $\nabla f_\cdot(x)$ is β-Lipschitz continuous and $f_x(\cdot)$ is γ-smooth, then, for any $x, y \in \mathcal{X}$,*

$$\|\nabla f_x(x) - \nabla f_y(y)\|_* \leq (\gamma + \beta)\|x - y\|.$$

Proof. For any $x, y \in \mathcal{X}$, it holds that

$$\begin{aligned}
\|\nabla f_x(x) - \nabla f_y(y)\|_* &\leq \|\nabla f_x(x) - \nabla f_y(x) + \nabla f_y(x) - \nabla f_y(y)\|_* \\
&\leq \|\nabla f_x(x) - \nabla f_y(x)\|_* + \|\nabla f_y(x) - \nabla f_y(y)\|_* \\
&\leq \beta\|x - y\| + \gamma\|x - y\|.
\end{aligned}$$

The last inequality uses β-regularity and γ-smoothness of $\nabla f_\cdot(x)$ and $f_y(\cdot)$, respectively. □

Lemma 5. *If f is (α, β)-regular, $f_x(\cdot)$ is γ-smooth for all $x \in \mathcal{X}$, and R is 1-strongly convex and L-smooth, then for the online mirror descent algorithm it holds that*

$$B_R(x^\star \| x_n) \leq \left(1 - 2\eta(\alpha - \beta)L^{-1} + \eta^2(\gamma + \beta)^2\right)^{n-1} B_R(x^\star \| x_1).$$

Proof. By the mirror descent update rule in (4), $\langle \eta \nabla f_{x_n}(x_n) + \nabla R(x_{n+1}) - \nabla R(x_n), x^\star - x_{n+1} \rangle \geq 0$. Since $x^\star \in X_\star$, $\langle \eta \nabla f_{x^\star}(x^\star), x_{n+1} - x^\star \rangle \geq 0$. Combining these inequalities yields $\eta \langle \nabla f_{x_n}(x_n) - \nabla f_{x^\star}(x^\star), x_{n+1} - x^\star \rangle \leq \langle \nabla R(x_{n+1}) - \nabla R(x_n), x^\star - x_{n+1} \rangle$. Then by the three-point equality of the Bregman divergence, we have

$$B_R(x^\star \| x_{n+1}) \leq B_R(x^\star \| x_n) - B_R(x_{n+1} \| x_n) - \eta \langle \nabla f_{x_n}(x_n) - \nabla f_{x^\star}(x^\star), x_{n+1} - x^\star \rangle.$$

Because of the $(\alpha - \beta)$-strong monotonicity of $\nabla f_x(x)$, the above inequality implies

$$\begin{aligned}
B_R(x^\star \| x_{n+1}) &\leq B_R(x^\star \| x_n) - B_R(x_{n+1} \| x_n) - \eta \langle \nabla f_{x_n}(x_n) - \nabla f_{x^\star}(x^\star), x_{n+1} - x_n \rangle \\
&\quad - \eta \langle \nabla f_{x_n}(x_n) - \nabla f_{x^\star}(x^\star), x_n - x^\star \rangle \\
&\leq B_R(x^\star \| x_n) - B_R(x_{n+1} \| x_n) - \eta \langle \nabla f_{x_n}(x_n) - \nabla f_{x^\star}(x^\star), x_{n+1} - x_n \rangle - \eta(\alpha - \beta) \| x^\star - x_n \|^2 \\
&\leq B_R(x^\star \| x_n) + \frac{\eta^2 (\gamma + \beta)^2}{2} \| x^\star - x_n \|^2 - \eta(\alpha - \beta) \| x^\star - x_n \|^2 \\
&\leq \left(1 + \eta^2 (\gamma + \beta)^2 - 2\eta(\alpha - \beta) L^{-1}\right) B_R(x^\star \| x_n).
\end{aligned}$$

The third inequality results from the Cauchy-Scwharz inequality followed by maximizing over $\|x_{n+1} - x_n\|$ and then applying Lemma 4. The last inequality uses the fact that R is 1-strongly convex and L-smooth. □

If $\alpha > \beta$ and η is chosen such that $\eta < \frac{2(\alpha-\beta)}{L(\gamma+\beta)^2}$, we can see that the online mirror descent algorithm guarantees linear convergence of $B_R(x^\star \| x_n)$ to zero with rate $(1 - 2\eta(\alpha - \beta)L^{-1} + \eta^2(\gamma + \beta)^2) \in (0, 1)$. By strong convexity, we have,

$$\begin{aligned}
\Delta_n = \| x^\star - x_n \| &\leq \sqrt{2 B_R(x^\star \| x_n)} \\
&\leq \sqrt{2} \left(1 + \eta^2 (\gamma + \beta)^2 - 2\eta(\alpha - \beta) L^{-1}\right)^{\frac{n-1}{2}} B_R(x^\star \| x_0)^{1/2}.
\end{aligned}$$

The proposition follows immediately from combining this result and Theorem 4.

D.5 Proof of Proposition 9

Recall that $g_n = \nabla l_n(x_n) + \epsilon_n + \xi_n$. As discussed previously, we assume there exist constants $0 \leq \sigma, \kappa < \infty$ such that $\mathbb{E}\left[\|\epsilon_n\|_*^2\right] \leq \sigma^2$ and $\|\xi_n\|_*^2 \leq \kappa^2$ for all n. The mirror descent update rule is given by

$$x_{n+1} = \arg\min_{x \in \mathcal{X}} \langle \eta_n g_n, x \rangle + B_R(x \| x_n). \tag{12}$$

We use Corollary 1 along with known results for the static regret to bound the dynamic regret in the stochastic case. The main idea of the proof is to show the result for the linearized losses. By convexity, this can be used to bound both terms in Corollary 1.

Let u be any fixed vector in $\mathcal{X}$, chosen independent of the learner's decisions $x_1, \ldots, x_n$. The first-order condition for optimality of (12) yields $\langle \eta_n g_n, x_{n+1} - u \rangle \leq \langle u - x_{n+1}, \nabla R(x_{n+1}) - \nabla R(x_n) \rangle$. We use this condition to bound the linearized losses as in the proof of Proposition 8. We can bound the linearized losses by the magnitude of the stochastic gradients and Bregman divergences between u and the learner's decisions:

$$\begin{aligned}
\langle g_n, x_n - u \rangle &\leq \frac{1}{\eta_n} \langle u - x_{n+1}, \nabla R(x_{n+1}) - \nabla R(x_n) \rangle + \langle g_n, x_n - x_{n+1} \rangle \\
&= \frac{1}{\eta_n} B_R(u \| x_n) - \frac{1}{\eta_n} B_R(u \| x_{n+1}) - \frac{1}{\eta_n} B_R(x_{n+1} \| x_n) + \langle g_n, x_n - x_{n+1} \rangle \\
&\leq \frac{1}{\eta_n} B_R(u \| x_n) - \frac{1}{\eta_n} B_R(u \| x_{n+1}) - \frac{1}{2\eta_n} \| x_n - x_{n+1} \|^2 + \| g_n \|_* \| x_n - x_{n+1} \| \\
&\leq \frac{1}{\eta_n} B_R(u \| x_n) - \frac{1}{\eta_n} B_R(u \| x_{n+1}) + \frac{\eta_n}{2} \| g_n \|_*^2.
\end{aligned}$$

The first inequality follows from adding $\langle g_n, x_n - x_{n+1} \rangle$ to both sides of the inequality from the first-order condition for optimality. The equality uses the three-point equality of the Bregman divergence. The second inequality follows from the Cauchy-Schwarz inequality and the fact that $\frac{1}{2}\|x_n - x_{n+1}\|^2 \leq B_R(x_{n+1}\|x_n)$ due to the 1-strong convexity of R. The last inequality maximizes over $\|x_n - x_{n+1}\|$.

Define $\mathcal{R} = \sup_{w_1,w_2\in\mathcal{X}} B_R(w_1\|w_2)$, which is bounded. Note that $\mathbb{E}\left[\|g_n\|_*^2\right] \leq 3(G^2+\sigma^2+\kappa^2)$. Therefore, summing from $n=1$ to N, it holds for any $u \in \mathcal{X}$ selected before learning,

$$\mathbb{E}\left[\sum_{n=1}^{N}\langle g_n, x_n - u\rangle\right] \leq \mathbb{E}\left[\sum_{n=1}^{N}\left(\frac{1}{\eta_n} - \frac{1}{\eta_{n-1}}\right)\mathcal{R} + \frac{3}{2}(G^2+\sigma^2+\kappa^2)\eta_n\right]$$

After rearrangement, we have

$$\mathbb{E}\left[\sum_{n=1}^{N}\langle \nabla l_n(x_n) + \epsilon_n, x_n - u\rangle\right] \leq \mathbb{E}\left[\sum_{n=1}^{N}\left(\frac{1}{\eta_n} - \frac{1}{\eta_{n-1}}\right)\mathcal{R} + \frac{3}{2}(G^2+\sigma^2+\kappa^2)\eta_n + D_{\mathcal{X}}\|\xi_n\|_*\right].$$

Choosing $\eta_n = \frac{1}{\sqrt{n}}$, $\eta_n = \eta_1$, and $u = x^\star$ (because $x^\star$ is fixed for a fixed f selected before learning) yields $\mathbb{E}\left[\sum_{n=1}^{N}\langle \nabla l_n(x_n) + \epsilon_n, x_n - x^\star\rangle\right] = O(\sqrt{N} + \Xi)$. Because of the law of total expectation and that x_n does not depend on ϵ_n, we have $\mathbb{E}[\widetilde{\text{Regret}}_N^s(x^\star)] = \mathbb{E}\left[\sum_{n=1}^{N}\langle \nabla l_n(x_n) + \epsilon_n, x_n - x^\star\rangle\right]$. Further, by convexity, it follows $\mathbb{E}[\text{Regret}_N^s(x^\star)] \leq \mathbb{E}[\widetilde{\text{Regret}}_N^s(x^\star)]$. Then, we may apply Corollary 1 to obtain the result. Note that there is no requirement that R is smooth.

E Complete Proofs of Section 7

E.1 Proof of Proposition 10

Because $\nabla l_n(\cdot)$ is α-strongly monotone, it holds

$$\left\langle \nabla l_n(x_{n-1}^*), x_{n-1}^* - x_n^*\right\rangle \geq \alpha\|x_{n-1}^* - x_n^*\|^2$$

Since y^* satisfies $\left\langle \nabla l_{n-1}(x_{n-1}^*), x_n^* - x_{n-1}^*\right\rangle \geq 0$, the above inequality implies that

$$\begin{aligned}\alpha\|x_n^* - x_{n-1}^*\|^2 &\leq \left\langle \nabla l_n(x_{n-1}^*) - \nabla l_{n-1}(x_{n-1}^*), x_{n-1}^* - x_n^*\right\rangle \\ &\leq (\beta\|x_n - x_{n-1}\| + a_n)\|x_{n-1}^* - x_n^*\|\end{aligned}$$

Rearranging the inequality gives the statement.

E.2 Proof of Theorem 5

For convenience, define $\lambda := \frac{\beta}{\alpha}$. Recall that, by the mirror descent update rule, the first-order conditions for optimality of both x_{x+1} and x_n^* yield, for all $x \in \mathcal{X}$,

$$\begin{aligned}\langle \eta\nabla l_n(x_n), x - x_{n+1}\rangle &\geq \langle \nabla R(x_n) - \nabla R(x_{n+1}), x - x_{n+1}\rangle \\ \langle \nabla l_n(x_n^*), x - x_n^*\rangle &\geq 0.\end{aligned}$$

The proof requires many intermediate steps, which we arrange in a series of lemmas that typically follow from each other in order. Ultimately, we aim to achieve a result that resembles a contraction as done in Proposition 8 but with additional terms due to the adversarial component of the predictable problem. We begin with general bounds on the Bregman divergence beteween the learner's decisions and the optimal decisions.

Lemma 6. *At round n, for an (α,β)-predictable problem under the mirror descent algorithm, if l_n is γ-smooth and R is 1-strongly convex and L-smooth, then it holds that*

$$\begin{aligned}B_R(x_{n+1}^*\|x_{n+1}) \leq\ & B_R(x_{n+1}^*\|x_n^*) + B_R(x_n^*\|x_{n+1}) \\ & + \lambda\|x_{n+1} - x_n\|\|\nabla R(x_n^*) - \nabla R(x_{n+1})\|_* + \frac{a_n}{\alpha}\|\nabla R(x_n^*) - \nabla R(x_{n+1})\|_*\end{aligned}$$

and, in the next round,

$$B_R(x_n^*\|x_{n+1}) \leq B_R(x_n^*\|x_n) - B_R(x_{n+1}\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \eta\gamma\|x_n - x_n^*\|\|x_{n+1} - x_n\|.$$

Proof. The first result uses the basic three-point equality of the Bregman divergence followed by the Cauchy-Schwarz inequality and Proposition 10. Note that this first part of the lemma does not require that x_n is generated from a mirror descent algorithm:

$$\begin{aligned} B_R(x^*_{n+1}\|x_{n+1}) &= B_R(x^*_{n+1}\|x^*_n) + B_R(x^*_n\|x_{n+1}) + \langle x^*_{n+1} - x^*_n, \nabla R(x^*_n) - \nabla R(x_{n+1})\rangle \\ &\leq B_R(x^*_{n+1}\|x^*_n) + B_R(x^*_n\|x_{n+1}) + \|x^*_{n+1} - x^*_n\|\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* \\ &\leq B_R(x^*_{n+1}\|x^*_n) + B_R(x^*_n\|x_{n+1}) \\ &\quad + \lambda\|x_{n+1} - x_n\|\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* + \frac{a_n}{\alpha}\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_*. \end{aligned}$$

For the second part of the lemma, we require using the first-order conditions of optimality of both x_{n+1} for the mirror descent update and x^*_n for l_n:

$$\begin{aligned} B_R(x^*_n\|x_{n+1}) &= B_R(x^*_n\|x_n) - B_R(x_{n+1}\|x_n) + \langle x^*_n - x_{n+1}, \nabla R(x_n) - \nabla R(x_{n+1})\rangle \\ &\leq B_R(x^*_n\|x_n) - B(x_{n+1}\|x_n) + \eta\langle \nabla l_n(x^*_n) - \nabla l_n(x_n), x_n - x^*_n\rangle \\ &\quad + \eta\langle \nabla l_n(x^*_n) - \nabla l_n(x_n), x_{n+1} - x_n\rangle \\ &\leq B_R(x^*_n\|x_n) - B_R(x_{n+1}\|x_n) - \alpha\eta\|x_n - x^*_n\|^2 + \eta\gamma\|x_n - x^*_n\|\|x_{n+1} - x_n\|. \end{aligned}$$

The first line again applies the three-point equality of the Bregman divergence. The second line combines both first-order optimality conditions to bound the inner product. The last inequality uses the strong convexity of l_n to bound $\eta\langle \nabla l_n(x^*_n) - \nabla l_n(x_n), x_n - x^*_n\rangle \leq -\alpha\eta\|x_n - x^*_n\|^2$ and the Cauchy-Schwarz inequality along with the smoothness of l_n to bound the other inner product. □

The second result also leads to a natural corollary that will be useful later in the full proof.

Corollary 2. *Under the same conditions as Lemma 6, it holds that*

$$B_R(x^*_n\|x_{n+1}) = \left(1 - 2\alpha\eta L^{-1} + \eta^2\gamma^2\right) B_R(x^*_n\|x_n).$$

Proof. We start with the first inequality of Lemma 6 and then maximize over $\|x_{n+1} - x_n\|^2$. Finally, we applying the strong convexity and smoothness of R to achieve the result:

$$\begin{aligned} B_R(x^*_n\|x_{n+1}) &\leq B_R(x^*_n\|x_n) - B_R(x_{n+1}\|x_n) - \alpha\eta\|x_n - x^*_n\|^2 + \eta\gamma\|x_n - x^*_n\|\|x_{n+1} - x_n\| \\ &\leq (1 - 2\alpha\eta L^{-1})B_R(x^*_n\|x_n) - \frac{1}{2}\|x_{n+1} - x_n\|^2 + \eta\gamma\|x_n - x^*_n\|\|x_{n+1} - x_n\| \\ &\leq (1 - 2\alpha\eta L^{-1})B_R(x^*_n\|x_n) + \eta^2\gamma^2 B_R(x^*_n\|x_n) = \left(1 - 2\alpha\eta L^{-1} + \eta^2\gamma^2\right) B_R(x^*_n\|x_n). \end{aligned}$$ □

We can combine both results of Lemma 6 in order to show

$$\begin{aligned} B_R(x^*_{n+1}\|x_{n+1}) &\leq B_R(x^*_{n+1}\|x^*_n) + \lambda\|x_{n+1} - x_n\|\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* + \frac{a_n}{\alpha}\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* \\ &\quad + B_R(x^*_n\|x_n) - B(x_{n+1}\|x_n) - \alpha\eta\|x_n - x^*_n\|^2 + \eta\gamma\|x_n - x^*_n\|\|x_{n+1} - x_n\|. \end{aligned}$$

Some of the terms in the above inequality can be grouped and bounded above. By L-smoothness of R, we have $B_R(x^*_{n+1}\|x^*_n) \leq \frac{L}{2}\|x^*_{n+1} - x^*_n\|^2 \leq \frac{L}{2}\left(\lambda\|x_n - x_{n+1}\| + \frac{a_n}{\alpha}\right)^2 = \frac{L}{2}\left(\lambda^2\|x_n - x_{n+1}\|^2 + \frac{a_n^2}{\alpha^2} + \frac{2a_n\lambda}{\alpha}\|x_n - x_{n+1}\|\right)$. Because, R is 1-strongly convex, $L \geq 1$; therefore, the previous inequality can be bounded from above using L^2 instead of L. While this artificially worsens the bound, it will be useful for simplifying the conditions sufficient for sublinear dynamic regret. 1-strong convexity of R also gives us $-B_R(x_{n+1}, x_n) \leq -\frac{1}{2}\|x_{n+1} - x_n\|^2$. Applying these upper bounds and then aggregating terms yields

$$\begin{aligned} B_R(x^*_{n+1}\|x_{n+1}) &\leq -\frac{(1 - L^2\lambda^2)}{2}\|x_n - x_{n+1}\|^2 + \left(\lambda\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* + \eta\gamma\|x_n - x^*_n\|\right)\|x_n - x_{n+1}\| \\ &\quad + B_R(x^*_n\|x_n) - \alpha\eta\|x_n - x^*_n\|^2 + \frac{a_n}{\alpha}\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* + \frac{a_n^2 L}{2\alpha^2} + \frac{a_n L\lambda}{\alpha}\|x_n - x_{n+1}\| \\ &\leq -\frac{(1 - L^2\lambda^2)}{2}\|x_n - x_{n+1}\|^2 + \left(\lambda\|\nabla R(x^*_n) - \nabla R(x_{n+1})\|_* + \eta\gamma\|x_n - x^*_n\|\right)\|x_n - x_{n+1}\| \end{aligned}$$

$$
\begin{aligned}
&+ B_R(x_n^*\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \frac{a_n L}{\alpha} D_{\mathcal{X}} + \frac{a_n^2 L}{2\alpha^2} + \frac{a_n L \lambda}{\alpha} D_{\mathcal{X}} \\
\leq\ & \frac{\lambda^2 \|\nabla R(x_n^*) - \nabla R(x_{n+1})\|_*^2 + \eta^2\gamma^2\|x_n - x_n^*\|^2}{1 - L^2\lambda^2} + B_R(x_n^*\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \zeta_n \\
\leq\ & \frac{\lambda^2 L^2 \|x_n^* - x_{n+1}\|^2 + \eta^2\gamma^2\|x_n - x_n^*\|^2}{1 - L^2\lambda^2} + B_R(x_n^*\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \zeta_n \\
\leq\ & \frac{2\lambda^2 L^2 B_R(x_n^*\|x_{n+1}) + 2\eta^2\gamma^2 B_R(x_n^*\|x_n)}{1 - L^2\lambda^2} + B_R(x_n^*\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \zeta_n,
\end{aligned}
$$

where $\zeta_n = \frac{a_n L D_{\mathcal{X}}}{\alpha}(1+\lambda) + \frac{a_n^2 L}{2\alpha^2}$. The third inequality follows from maximizing over $\|x_n - x_{n+1}\|$ and then applying $(a+b)^2 \leq 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$. For this operation, we require that $L^2\lambda^2 < 1$. The fourth inequality uses L-smoothness of R. The last inequality uses the fact that R is 1-strongly convex to bound the squared normed differences by the Bregman divergence.

We then use Corollary 2 to bound this result on $B_R(x_{n+1}^*\|x_{n+1})$ in terms of only $B_R(x_n^*\|x_n)$ and the appropriate constants:

$$
\begin{aligned}
B_R(x_{n+1}^*\|x_{n+1}) &\leq \frac{2L^2\lambda^2 B_R(x_n^*\|x_{n+1}) + 2\eta^2\gamma^2 B_R(x_n^*\|x_n)}{1 - L^2\lambda^2} + B_R(x_n^*\|x_n) - \alpha\eta\|x_n - x_n^*\|^2 + \zeta_n \\
&\leq \frac{2L^2\lambda^2}{1 - L^2\lambda^2}\left(1 - 2\alpha\eta L^{-1} + \eta^2\gamma^2\right) B_R(x_n^*\|x_n) + \frac{2\eta^2\gamma^2}{1 - L^2\lambda^2} B_R(x_n^*\|x_n) \\
&\quad + B_R(x_n^*\|x_n) - 2\alpha\eta L^{-1} B_R(x_n^*\|x_n) + \zeta_n \\
&= \left(1 - 2\alpha\eta L^{-1} + \frac{2\eta^2\gamma^2}{1 - L^2\lambda^2} + \frac{2L^2\lambda^2}{1 - L^2\lambda^2} - \frac{4L\lambda^2\alpha\eta}{1 - L^2\lambda^2} + \frac{2L^2\lambda^2\eta^2\gamma^2}{1 - L^2\lambda^2}\right) B_R(x_n^*\|x_n) + \zeta_n \\
&= \left(\frac{1 + L^2\lambda^2}{1 - L^2\lambda^2}\right)\left(1 - 2\alpha\eta L^{-1} + 2\eta^2\gamma^2\right) B_R(x_n^*\|x_n) + \zeta_n.
\end{aligned}
$$

Thus, we have arrived at an inequality that resembles a contraction. However, the stepsize $\eta > 0$ may be chosen such that it minimizes the factor in front of the Bregman divergence. This can be achieved, but it requires that additional constraints are put on the value of λ.

Lemma 7. *If $\lambda < \frac{\alpha}{2L^2\gamma}$ and $\eta = \frac{\alpha}{2L\gamma^2}$, then*

$$
\left(\frac{1 + L^2\lambda^2}{1 - L^2\lambda^2}\right)\left(1 - 2\alpha\eta L^{-1} + 2\eta^2\gamma^2\right) < 1
$$

Proof. By optimizing over choices of η, it can be seen that

$$
1 - 2\alpha\eta L^{-1} + 2\eta^2\gamma^2 \geq 1 - \frac{\alpha^2}{2L^2\gamma^2},
$$

where η is chosen to be $\frac{\alpha}{2L\gamma^2}$. Therefore, in order to realize a contraction, we must have

$$
1 > \left(\frac{1 + L^2\lambda^2}{1 - L^2\lambda^2}\right)\left(1 - \frac{\alpha^2}{2L^2\gamma^2}\right).
$$

Alternatively,

$$
0 > 2L^2\lambda^2 - \frac{\alpha^2}{2L^2\gamma^2} - \frac{\lambda^2\alpha^2}{2\gamma^2}.
$$

The quantity on the right hand size of the above inequality is in fact smaller than $2L^2\lambda^2 - \frac{\alpha^2}{2L^2\gamma^2}$, meaning that it is sufficient to have the condition for a contraction be: $\frac{\alpha}{2L^2\gamma} > \lambda$.. □

Note that $\frac{\alpha}{2L^2\gamma} < 1$ since $L \geq 1$ and $\gamma \geq \alpha$ by the definitions of smoothness of R and l_n, respectively. Thus, this condition required to guarantee the contraction is stricter than requiring that $\lambda < 1$. If this condition is satisfied and if we set $\eta = \frac{\alpha}{2L\gamma^2}$, then we can further examine the contraction in terms of constants that depend only on the properties of l_n and R:

$$\begin{aligned} B_R(x_{n+1}^* \| x_{n+1}) &\leq \left(\frac{1+L^2\lambda^2}{1-L^2\lambda^2}\right)\left(1 - 2\alpha\eta L^{-1} + 2\eta^2\gamma^2\right) B_R(x_n^* \| x_n) + \zeta_n \\ &< \left(\frac{1+\frac{\alpha^2}{4L^2\gamma^2}}{1-\frac{\alpha^2}{4L^2\gamma^2}}\right)\left(1 - \frac{\alpha^2}{2L^2\gamma^2}\right) B_R(x_n^* \| x_n) + \zeta_n \\ &= \left(1 - \frac{\frac{\alpha^4}{8L^4\gamma^4}}{1-\frac{\alpha^2}{4L^2\gamma^2}}\right) B_R(x_n^* \| x_n) + \zeta_n. \end{aligned}$$

It is easily verified that the factor in front of the Bregman divergence on the right side is less than 1 and greater than $\frac{5}{6}$.

By applying the above inequality recursively, we can derive the inequality below

$$\frac{1}{2}\|x_n - x_n^*\|^2 \leq B_R(x_n^* \| x_n) \leq \rho^{n-1} B_R(x_1^* \| x_1) + \sum_{k=1}^{n-1} \rho^{n-k-1}\zeta_k,$$

where $\rho = \left(\frac{1+L^2\lambda^2}{1-L^2\lambda^2}\right)\left(1 - 2\alpha\eta L^{-1} + 2\eta^2\gamma^2\right) < 1$. Therefore the dynamic regret can be bounded as

$$\begin{aligned} \text{Regret}_N^d &= \sum_{n=1}^N f_n(x_n) - f_n(x_n^*) \leq G\sum_{n=1}^N \|x_n - x_n^*\| \\ &\leq \sqrt{2} G B_R(x_1^* \| x_1)^{1/2} \sum_{n=1}^N \rho^{\frac{n-1}{2}} + \sqrt{2} G \sum_{n=2}^N \left(\sum_{k=1}^{n-1} \rho^{n-k-1}\zeta_k\right)^{1/2} \\ &\leq \sqrt{2} G B_R(x_1^* \| x_1)^{1/2} \sum_{n=1}^N \rho^{\frac{n-1}{2}} + \sqrt{2} G \sum_{n=2}^N \sum_{k=1}^{n-1} \rho^{\frac{n-k-1}{2}}\zeta_k^{1/2}, \end{aligned}$$

where both inequalities use the fact that for $a, b > 0$, $a + b \leq a + b + 2\sqrt{ab} = (\sqrt{a} + \sqrt{b})^2$. The left-hand term is clearly bounded above by a constant since $\sqrt{\rho} < 1$. Analysis of the right-hand term is not as obvious, so we establish the following lemma independently.

Lemma 8. *If $\rho < 1$ and $\zeta_n = \frac{a_n L D_{\mathcal{X}}}{\alpha}(1+\lambda) + \frac{a_n^2 L}{2\alpha^2}$, then it holds that*

$$\sqrt{2}\sum_{n=2}^N \sum_{k=1}^{n-1} \rho^{\frac{n-k-1}{2}}\zeta_k^{1/2} = O(A_N + \sqrt{N A_N}).$$

Proof.

$$\sum_{n=2}^N \sum_{k=1}^{n-1} \rho^{\frac{n-k-1}{2}}\zeta_k^{1/2} = \sum_{n=1}^{N-1} \zeta_n^{1/2}\left(1 + \rho^{\frac{1}{2}} + \ldots + \rho^{\frac{N-1-n}{2}}\right) \leq \frac{1}{1-\sqrt{\rho}} \sum_{n=1}^{N-1} \sqrt{\zeta_n}.$$

The last inequality upper bounds the finite geometric series with the value of the infinite geometric series since again $\sqrt{\rho} < 1$ for each k. Recall that ζ_n was defined as

$$\zeta_n = \frac{a_n L D_{\mathcal{X}}}{\alpha}(1+\lambda) + \frac{a_n^2 L}{2\alpha^2}.$$

Therefore, the over the square roots can be bounded:

$$\sum_{n=1}^{N-1} \sqrt{\zeta_n} \leq \sqrt{\frac{L D_{\mathcal{X}}}{\alpha}(1+\lambda)} \sum_{n=1}^{N-1} \sqrt{a_n} + \alpha^{-1}\sqrt{\frac{L}{2}} \sum_{n=1}^{N-1} a_n.$$

While the right-hand summation is simply the definition of A_{N-1}, the left-hand summation yields $\sum_{n=1}^{N-1} \sqrt{a_n} \leq \sqrt{(N-1)A_{N-1}}$. □

Then the total dynamic regret has order $\text{Regret}_N^d = O(1 + A_N + \sqrt{N A_N})$.

E.3 Proof of Theorem 6

E.3.1 Euclidean Space with $\frac{\beta}{\alpha} = 1$

The proof first requires a result from analysis on the convergence of sequences that are nearly monotonic.

Lemma 9. *Let $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and $(b_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be two sequences satisfying $b_n \geq 0$ and $\sum_{k=1}^{n} a_k < \infty$ $\forall n \in \mathbb{N}$. If $b_{n+1} \leq b_n + a_n$, then the sequence b_n converges.*

Proof. Define $u_1 := b_1$ and $u_n := b_n - \sum_{k=1}^{n-1} a_k$. Note that $u_1 = b_1 \geq b_2 - a_1 = u_2$. Recursively, $b_n - a_{n-1} \leq b_{n-1} \implies b_n - \sum_{k=1}^{n-1} a_k \leq b_{n-1} - \sum_{k=1}^{n-2}$. Therefore, $u_n \leq u_{n+1}$. Note that $(u_n)_{n\in\mathbb{N}})$ is bounded below because $b_n \geq 0$ and $\sum_{k=1}^{n} a_k < \infty$. This implies that $(u_n)_{n\in\mathbb{N}}$ converges. Because $(\sum_{k=1}^{n} a_k)_{n\in\mathbb{N}}$, also converges, $(b_n)_{n\in\mathbb{N}}$ must converge. □

The majority of the proof follows a similar line of reasoning as a standard result in the field of discrete-time pursuit-evasion games Alexander et al. (2006). Let $\|\cdot\|$ denote the Euclidean norm. We aim to show that if the distance between the learner's decision x_n and the optimal decision x_n^* does not converge to zero, then they travel unbounded in a straight line, which is a contradiction.

Consider the following update rule which essentially amounts to a constrained greedy update:

$$x_{n+1} = \frac{x_n + x_n^*}{2}$$

x_{n+1} is well defined at each round because $\mathcal{X}$ is convex. Define $c_n := \|x_n - x_n^*\|$. Then we have

$$\begin{aligned}
0 \leq c_{n+1} &= \|x_{n+1} - x_{n+1}^*\| \\
&\leq \|x_{n+1} - x_n^*\| + \|x_{n+1}^* - x_n^*\| \\
&= \frac{1}{2}\|x_n - x_n^*\| + \|x_{n+1}^* - x_n^*\| \\
&\leq \frac{1}{2}\|x_n - x_n^*\| + \|x_{n+1} - x_n\| + \frac{a_n}{\alpha} \qquad (\because \text{Proposition 10}) \\
&= \|x_n - x_n^*\| + \frac{a_n}{\alpha} = c_n + \frac{a_n}{\alpha}
\end{aligned}$$

Because it is assumed that $\sum_{n=1}^{\infty} a_n < \infty$, the sequences $(c_n)_{n\in\mathbb{N}}$ and $(a_n)_{n\in\mathbb{N}}$ satisfy the sufficient conditions of Lemma 9. Thus the sequence $(c_n)_{n\in\mathbb{N}}$ converges, so there exists a limit point $C := \lim_{n\to\infty} c_n \geq 0$. Towards a contradiction, consider the case where $C > 0$. We will prove that this leads the points to follow a straight line in the following lemma.

Lemma 10. *Let θ_n denote the angle between the vectors from x_n^* to x_{n+1}^* and from x_n^* to x_{n+1}. If $\lim_{n\to\infty} c_n > 0$, then $\lim_{n\to\infty} \cos\theta_n = -1$.*

Proof. At round $n+1$ we can write the distance between the learner's decision and the optimal decision in terms of the previous round:

$$\begin{aligned}
C^2 &= \lim_{n\to\infty} \|x_{n+1} - x_{n+1}^*\|^2 \\
&= \lim_{n\to\infty} \left(\|x_{n+1} - x_n^*\|^2 + \|x_{n+1}^* - x_n^*\|^2 - 2\|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\|\cos\theta_n\right) \\
&\leq \lim_{n\to\infty} \left(\frac{1}{4}\|x_n - x_n^*\|^2 + \|x_n - x_{n+1}\|^2 + \frac{a_n^2}{\alpha^2} + \frac{2a_n}{\alpha}\|x_n - x_{n+1}\| - 2\|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\|\cos\theta_n\right) \\
&= \lim_{n\to\infty} \left(\frac{1}{2}\|x_n - x_n^*\|^2 + \frac{a_n^2}{\alpha^2} + \frac{2a_n}{\alpha}\|x_n - x_{n+1}\| - 2\|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\|\cos\theta_n\right) \\
&= \lim_{n\to\infty} \frac{1}{2}\|x_n - x_n^*\|^2 - 2\lim_{n\to\infty} \|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\|\cos\theta_n \\
&= \frac{1}{2}C^2 - 2\lim_{n\to\infty} \|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\|\cos\theta_n
\end{aligned}$$

The first inequality follows because $\|x_{n+1} - x_n^*\| = \frac{1}{2}\|x_n - x_n^*\|$ and $\|x_{n+1}^* - x_n^*\| \leq \|x_{n+1} - x_n\| + \frac{a_n}{\alpha}$ due to Proposition 10. The next equality again uses $\|x_{n+1} - x_n^*\| = \frac{1}{2}\|x_n - x_n^*\|$. The second to last line follows from passing the limit through the sum, where we have $\lim_{n\to\infty} a_n = 0$ because $A_\infty < \infty$. That is, the inequality above implies

$$2 \lim_{n\to\infty} \|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\| \cos\theta_n = -\frac{C^2}{2} < 0$$

which in turn implies $\lim_{n\to\infty} \cos\theta_n < 0$. This leads to an upper bound

$$\begin{aligned}
-2 \lim_{n\to\infty} \|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\| \cos\theta_n &= \left(-2 \lim_{n\to\infty} \cos\theta_n\right) \lim_{n\to\infty} \|x_{n+1} - x_n^*\|\|x_{n+1}^* - x_n^*\| \\
&\leq \left(-2 \lim_{n\to\infty} \cos\theta_n\right) \lim_{n\to\infty} \frac{1}{2}\|x_n - x_n^*\| \left(\|x_{n+1} - x_n\| + \frac{a_n}{\alpha}\right) \\
&= \frac{-C^2}{2} \lim_{n\to\infty} \cos\theta_n
\end{aligned}$$

Combining these two inequalities, we can then conclude $C^2 \leq \frac{C^2}{2} - \frac{C^2}{2}\cos\theta \leq C^2$. A necessary condition in order for the bounds to be satisfied is $\cos\theta = -1$. □

When $C > 0$, Lemma 10 therefore implies the points $x_n, x_{n+1}, x_n^*, x_{n+1}^*$ are colinear in the limit. Thus, $\|x_n - x_{n+m}\|$ grows unbounded in m, which contradicts the compactness of $\mathcal{X}$. The alternative case must then be true: $C = \lim_{n\to\infty} \|x_n - x_n^*\| = 0$. The dynamic regret can then be bounded as:

$$\text{Regret}_N^d = \sum_{n=1}^{N} l_n(x_n) - l_n(x_n^*) \leq G \sum_{n=1}^{N} \|x_n - x_n^*\|$$

Since $\|x_N - x_N^*\| \to 0$, we know $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \|x_n - x_n^*\| = 0$. Therefore, the dynamic regret is sublinear.

Note that this result does not reveal a rate of convergence, only that $\|x_n - x_n^*\|$ converges to zero, which is enough for sublinear dynamic regret.

E.3.2 One-dimensional Space with arbitrary $\frac{\beta}{\alpha}$

In the case where $d = 1$, we aim to prove sublinear dynamic regret regardless of α and β by showing that x_n essentially traps x_n^* by taking conservative steps as before. Rather than the constraint being $|x_n - x_{n+1}| \leq \frac{1}{2}|x_n - x_n^*|$, we choose x_{n+1} in the direction of x_n^* subject to $|x_n - x_{n+1}| \leq \frac{1}{1+\lambda}|x_n - x_n^*|$. Specifically, we will use the following update rule:

$$x_{n+1} = \frac{\lambda x_n + x_n^*}{1 + \lambda} \tag{13}$$

Recall that sublinear dynamic regret is implied by $c_n := |x_n - x_n^*|$ converging to zero as $n \to \infty$. Therefore, below we will prove the above update rule results in $\lim_{n\to\infty} c_n = 0$. Like our discussions above, this implies achieving sublinear dynamic regret but not directly its rate.

Suppose at any time $|x_n - x_n^*| = 0$. Then we are done since the learner can repeated play the same decision without x_n^* changing. Below we consider the case $|x_n - x_n^*| \neq 0$. We prove this by contradiction. First we observe that the update in (13) makes sure that, at any round, x_{n+1}^* cannot switch to the opposite side of x_n^* with respect to x_{n+1} and x_n; namely it is guaranteed that $(x_{n+1}^* - x_{n+1})(x_n^* - x_{n+1}) \geq 0$ and $(x_{n+1}^* - x_n)(x_n^* - x_n) \geq 0$.

Towards a contradiction, suppose that there is some $C > 0$ such that $|x_n - x_n^*| \geq C$ for infinitely many n. Then x_n at every round moves a distance of at least $\frac{C}{1+\lambda}$ in the same direction infinitely since x_{n+1}^* always lies the same side of x_{n+1} as x_n^*. This contradicts the compactness of $\mathcal{X}$. Therefore $|x_n - x_n^*|$ must converge to zero.

F New Insights into Imitation Learning

In this section, we investigate an application of the COL framework in the sequential decision problem of online IL (Ross et al., 2011). We consider an episodic MDP with state space $\mathcal{S}$, action space $\mathcal{A}$, and finite horizon H.

For any $s, s' \in \mathcal{S}$ and $a \in \mathcal{A}$, the transition dynamics $\mathcal{P}$ gives the conditional density, denoted by $\mathcal{P}(s'|s,a)$, of transitioning to s' starting from state s and applying action a. The reward of state s and action a is denoted as $r(s,a)$. A deterministic policy π is a mapping from $\mathcal{S}$ to a density over $\mathcal{A}$. We suppose the MDP starts from some fixed initial state distribution. We denote the probability of being in state s at time t under policy π as $d_t^\pi(s)$, and we define the average state distribution under π as $d^\pi(s) = \frac{1}{T}\sum_{t=1}^T d_t^\pi(s)$.

In IL, we assume that $\mathcal{P}$ and r are unknown to the learner, but, during training time, the learner is given access to an expert policy $\pi^\star$ and full knowledge of a supervised learning loss function $c(s, \pi; \pi^\star)$, defined for each state $s \in \mathcal{S}$. The objective of IL is to solve

$$\min_{\pi \in \Pi} \quad \mathbb{E}_{s \sim d^\pi}\left[c(s, \pi; \pi^\star)\right], \tag{14}$$

where Π is the set of allowable parametric policies, which will be assumed to be convex. Note that it is often the case that $\pi^\star \notin \Pi$.

As d^π is not known analytically, optimizing (14) directly leads to a reinforcement learning problem and therefore can be sample inefficient. *Online IL*, such as the popular DAgger algorithm Ross et al. (2011), bypasses this difficulty by reducing (14) into a sequence of supervised learning problems. Below we describe a general construction of online IL: at the nth iteration (1) execute the learner's current policy π_n in the MDP to collect state-action samples; (2) update π_{n+1} with information of the stochastic approximation of $l_n(\pi) = \mathbb{E}_{d^{\pi_n}}\left[c(s,\pi;\pi^\star)\right]$ based the samples collected in the first step. Importantly, we remark that in these empirical risks, the states are sampled according to d^{π_n} of the learner's policy.

The use of online learning to analyze online IL is well established (Ross et al., 2011). As studied in Cheng and Boots (2018); Lee et al. (2018), these online losses can be formulated as a bifunction, $l_n(\pi) = f_{\pi_n}(\pi) = \mathbb{E}_{s \sim d^{\pi_n}}\left[c(s,\pi;\pi^\star)\right]$, and the policy class Π can be viewed as the decision set $\mathcal{X}$. Naturally, this online learning formulation results in many online IL algorithms resembling standard online learning algorithms, such as follow-the-leader (FTL), which uses full information feedback $l_n(\cdot) = \mathbb{E}_{s \sim d^{\pi_n}}\left[c(s,\cdot;\pi^\star)\right]$ at each round (Ross et al., 2011), and mirror descent (Sun et al., 2017), which uses the first-order feedback $\nabla l_n(\pi_n) = \mathbb{E}_{d^{\pi_n}}\left[\nabla_{\pi_n} c(s,\pi_n;\pi^\star)\right]$. This feedback can also be approximated by unbiased samples. The original work by Ross et al. (2011) analyzed FTL in the static regret case by immediate reductions to known static regret bounds of FTL. However, a crucial objective is understanding when these algorithms converge to useful solutions in terms of policy performance, which more recent work has attempted to address (Cheng and Boots, 2018; Lee et al., 2018; Cheng et al., 2019b). According to these refined analyses, dynamic regret is a more appropriate solution concept to online IL when $\pi^\star \notin \Pi$, which is the common case in practice.

Below we frame online IL in the proposed COL framework and study its properties based on the properties of COL that we obtained in the previous sections. We have already shown that the per-round loss $l_n(\cdot)$ can be written as the evaluation of a bifunction $f_{\pi_n}(\cdot)$. This COL problem is actually an (α,β)-regular COL problem when the expected supervised learning loss $\mathbb{E}_{s\sim d^{\pi_n}}[c(s,\pi;\pi^\star)]$ is strongly convex in π and the state distribution d^π is Lipschitz continuous (see Ross et al. (2011); Cheng and Boots (2018); Lee et al. (2018)). We can then leverage our results in the COL framework to immediately answer an interesting question in the online IL problem.

Proposition 14. *When $\alpha > \beta$, there exists a* unique *policy $\widehat{\pi}$ that is optimal on its own distribution:*

$$\mathbb{E}_{s \sim d_{\widehat{\pi}_n}}\left[c(s,\widehat{\pi};\pi^\star)\right] = \min_{\pi \in \Pi} \mathbb{E}_{s \sim d_{\widehat{\pi}_n}}\left[c(s,\pi;\pi^\star)\right].$$

This result is immediate from the fact that $\alpha > \beta$ implies that $\nabla f_\pi(\pi)$ is a μ-strongly monotone VI with $\mu = \beta - \alpha$ by Proposition 5. The VI is therefore guaranteed to have a unique solution (Facchinei and Pang, 2007).

Furthermore, we can improve upon the known conditions sufficient to find this policy through online gradient descent and give a non-asymptotic convergence guarantee through a reduction to strongly monotone VIs. We will additionally assume that f is γ-smooth in π, satisfying $\|\nabla f_{\pi'}(\pi_1) - \nabla f_{\pi'}(\pi_2)\| \leq \gamma \|\pi_1 - \pi_2\|$ for any fixed query argument π'.

We then apply our results from Section 6.1. Specifically, we consider mirror descent with $B_R(\pi\|\pi') = \frac{1}{2}\|\pi - \pi'\|_2^2$, which is equivalent to online gradient descent studied in Sun et al. (2017); Lee et al. (2018). Note that $R = \frac{1}{2}\|\pi\|_2^2$, which is 1-strongly convex and 1-smooth. Then, we apply Lemma 5.

Corollary 3. *If $\alpha > \beta$ and the stepsize is chosen such that $\eta = \frac{\alpha-\beta}{(\gamma+\beta)^2}$, then, under the online gradient descent algorithm with deterministic feedback $g_n = \nabla l_n(\pi_n)$, it holds that*

$$\|\pi_n - \widehat{\pi}\|^2 \leq \left(1 - \left(\frac{\alpha - \beta}{\gamma + \beta}\right)^2\right)^{n-1} \|\pi_1 - \widehat{\pi}\|^2$$

By Proposition 8, Regret_N^d will therefore be sublinear (in fact, $\text{Regret}_N^d = O(1)$) and the policy converges linearly to the policy that is optimal on its own distribution, $\widehat{\pi}$. The only condition required on the problem itself is $\alpha > \beta$ while the state-of-the-art sufficient condition of Lee et al. (2018) additionally requires $\frac{\alpha}{\gamma} > \frac{2\beta}{\alpha}$. The result also gives a new non-asymptotic convergence rate to $\widehat{\pi}$.

The above result only considers the case when the feedback is deterministic; i.e., there is no sampling error due to executing the policy on the MDP, and the risk $\mathbb{E}_{d^{\pi_n}}\left[c(s, \pi; \pi^\star)\right]$ is known exactly at each round. While this is a standard starting point in analysis of online IL algorithms (Ross et al., 2011), we are also interested in the more realistic stochastic case, which has so far not been analyzed for the online gradient descent algorithm in online IL. It turns out that the COL framework can be easily leveraged here too to provide a sublinear dynamic regret bound.

At round n, we consider observing the empirical risk $\tilde{l}_n(\pi) = \frac{1}{T}\sum_{t=1}^{T} c(s_t, \pi; \pi^\star)$ where $s_t \sim d_t^{\pi_n}$. Note that $\mathbb{E}[\tilde{l}_n(\pi)|\pi_n] = l_n(\pi)$ and it is easy to show that the first-order feedback $\nabla \tilde{l}_n(\pi_n)$ can be modeled as the expected gradient with an additive zero-mean noise: $g_n = \nabla l_n(\pi_n) + \epsilon_n$. For simplicity, we assume $\mathbb{E}\left[\|\epsilon_n\|^2\right] < \infty$.

Corollary 4. *If $\alpha > \beta$ and the stepsize is chosen as $\eta_n = \frac{1}{\sqrt{n}}$, then, under online gradient descent with stochastic feedback, it holds that $\mathbb{E}[\text{Regret}_N^d] = O(\sqrt{N})$.*

This corollary follows from Proposition 9, which in turn leverages the reduction to static regret in Corollary 1.