

On Tractable Φ -Equilibria in Non-Concave Games

Yang Cai* Constantinos Daskalakis[†] Haipeng Luo[‡] Chen-Yu Wei[§]

Weiqliang Zheng[¶]

July 4, 2024

Abstract

While Online Gradient Descent and other no-regret learning procedures are known to efficiently converge to a coarse correlated equilibrium in games where each agent’s utility is concave in their own strategy, this is not the case when utilities are non-concave – a common scenario in machine learning applications involving strategies parameterized by deep neural networks, or when agents’ utilities are computed by neural networks, or both. Non-concave games introduce significant game-theoretic and optimization challenges: (i) Nash equilibria may not exist; (ii) local Nash equilibria, though existing, are intractable; and (iii) mixed Nash, correlated, and coarse correlated equilibria generally have infinite support and are intractable. To sidestep these challenges, we revisit the classical solution concept of Φ -equilibria introduced by [Greenwald and Jafari \[2003\]](#), which is guaranteed to exist for an arbitrary set of strategy modifications Φ even in non-concave games [[Stoltz and Lugosi, 2007](#)]. However, the tractability of Φ -equilibria in such games remains elusive. In this paper, we initiate the study of tractable Φ -equilibria in non-concave games and examine several natural families of strategy modifications. We show that when Φ is finite, there exists an efficient uncoupled learning algorithm that converges to the corresponding Φ -equilibria. Additionally, we explore cases where Φ is infinite but consists of local modifications, showing that Online Gradient Descent can efficiently approximate Φ -equilibria in non-trivial regimes.

*Yale University. Email: yang.cai@yale.edu

[†]MIT CSAIL. Email: costis@csail.mit.edu

[‡]University of Southern California. Email: haipengl@usc.edu

[§]University of Virginia. Email: chenyu.wei@virginia.edu

[¶]Yale University. Email: weiqiang.zheng@yale.edu

1 Introduction

Von Neumann’s celebrated minimax theorem establishes the existence of Nash equilibrium in all two-player zero-sum games where the players’ utilities are continuous as well as *concave* in their own strategy [v. Neumann, 1928].¹ This assumption that players’ utilities are concave, or quasi-concave, in their own strategies has been a cornerstone for the development of equilibrium theory in Economics, Game Theory, and a host of other theoretical and applied fields that make use of equilibrium concepts. In particular, (quasi-)concavity is key for showing the existence of many types of equilibrium, from generalizations of min-max equilibrium [Fan, 1953, Sion, 1958] to competitive equilibrium in exchange economies [Arrow and Debreu, 1954, McKenzie, 1954], mixed Nash equilibrium in finite normal-form games [Nash Jr, 1950], and, more generally, Nash equilibrium in (quasi-)concave games [Debreu, 1952, Rosen, 1965].

Not only are equilibria guaranteed to exist in concave games, but it is also well-established—thanks to a long line of work at the interface of game theory, learning and optimization whose origins can be traced to Dantzig’s work on linear programming [George B. Dantzig, 1963], Brown and Robinson’s work on fictitious play [Brown, 1951, Robinson, 1951], Blackwell’s approachability theorem [Blackwell, 1956] and Hannan’s consistency theory [Hannan, 1957]—that several solution concepts are efficiently computable both centrally and via decentralized learning dynamics. For instance, it is well-known that the learning dynamics produced when the players of a game iteratively update their strategies using no-regret learning algorithms, such as online gradient descent, is guaranteed to converge to Nash equilibrium in two-player zero-sum concave games, and to coarse correlated equilibrium in multi-player general-sum concave games [Cesa-Bianchi and Lugosi, 2006]. The existence of such simple decentralized dynamics further justifies using these solution concepts to predict the outcome of real-life multi-agent interactions where agents deploy strategies, obtain feedback, and use that feedback to update their strategies.

While (quasi-)concave utilities have been instrumental in the development of equilibrium theory, as described above, they are also too restrictive an assumption. Several modern applications and outstanding challenges in Machine Learning, from training Generative Adversarial Networks (GANs) to Multi-Agent Reinforcement Learning (MARL) as well as generic multi-agent Deep Learning settings where the agents’ strategies are parameterized by deep neural networks or their utilities are computed by deep neural networks, or both, give rise to games where the agents’ utilities are *non-concave* in their own strategies. We call these games *non-concave*, following Daskalakis [2022].

Unfortunately, classical equilibrium theory quickly hits a wall in non-concave games. First, Nash equilibria are no longer guaranteed to exist. Second, while mixed Nash, correlated and coarse correlated equilibria do exist—under convexity and compactness of the strategy sets [Glicksberg, 1952], which we have been assuming all along in our discussion so far, they have infinite support, in general [Karlin, 2014]. Finally, they are computationally intractable; so, a fortiori, they are also intractable to attain via decentralized learning dynamics.

In view of the importance of non-concave games in emerging ML applications and the afore-described state-of-affairs, our investigation is motivated by the following broad and largely open question:

Question from [Daskalakis, 2022]: *Is there a theory of non-concave games? What solution concepts are meaningful, universal, and tractable?*

¹Throughout this paper, we model games using the standard convention in Game Theory that each player has a utility function that they want to maximize. This is, of course, equivalent to modeling the players as loss minimizers, a convention more common in learning. When we say that a player’s utility is concave (respectively non-concave) in their strategy, this is the same as saying that the player’s loss is convex (respectively non-convex) in their strategy.

Table 1: A comparison between different solution concepts in multi-player non-concave games. We include definitions of Nash equilibrium, mixed Nash equilibrium, (coarse) correlated equilibrium, strict local Nash equilibrium, and second-order local Nash equilibrium in [Appendix B](#). We also give a detailed discussion on the existence and complexity of these solution concepts in [Appendix B](#).

Solution Concept	Incentive Guarantee	Existence	Complexity (to compute or check existence)
Nash equilibrium	Global stability	✗	NP-hard
Mixed Nash equilibrium		✓	NP-hard
(Coarse) Correlated equilibrium		✓	NP-hard
Strict local Nash equilibrium	Local stability	✗	NP-hard
Second-order local Nash equilibrium	Second-order stability	✗	NP-hard
Local Nash equilibrium	First-order stability	✓	PPAD-hard
Φ -equilibrium (finite $ \Phi $)	Stability against finite deviations	✓	Efficient ε -approximation for any $\varepsilon > 0$
Conv($\Phi(\delta)$)-equilibrium (finite $ \Phi(\delta) $)	First-order stability when $\varepsilon = \Omega(\delta^2)$	✓	Efficient ε -approximation for $\varepsilon = \Omega(\delta^2)$
$\Phi_{\text{proj}}(\delta)$ -equilibrium		✓	Efficient ε -approximation via GD/OG for $\varepsilon = \Omega(\delta^2)$
$\Phi_{\text{Int}}(\delta)$ -equilibrium		✓	Efficient ε -approximation via no-regret learning for $\varepsilon = \Omega(\delta^2)$
$\Phi_{\text{Int}^+}(\delta)$ -equilibrium		High-order stability when $\varepsilon = o(\delta^2)$	✓

1.1 Contributions

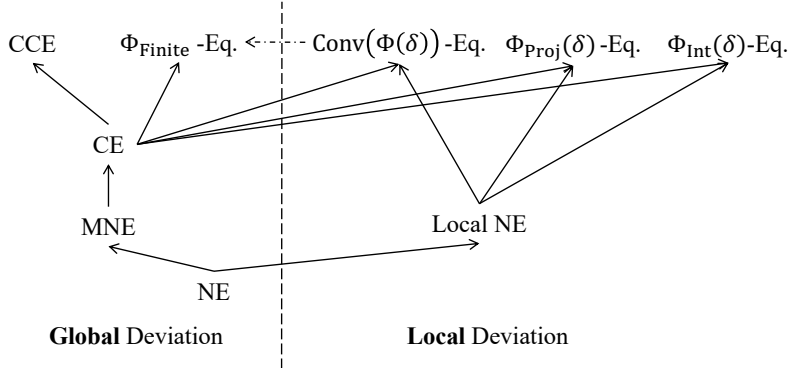
We study Daskalakis’ question through the lens of the classical solution concept of Φ -equilibria introduced by [Greenwald and Jafari \[2003\]](#). This concept is guaranteed to exist for virtually any set of strategy modifications Φ , even in non-concave games, as demonstrated by [Stoltz and Lugosi \[2007\]](#).² However, the tractability of Φ -equilibria in such games remains elusive. In this paper, we initiate the study of tractable Φ -equilibria in non-concave games and examine several natural families of strategy modifications.

Φ -Equilibrium. The concept of Φ -equilibrium generalizes (coarse) correlated equilibrium. A Φ -equilibrium is a joint distribution over $\Pi_{i=1}^n \mathcal{X}_i$, the Cartesian product of all players’ strategy sets, and is defined in terms of a set, $\Phi^{\mathcal{X}_i}$, of *strategy modifications*, for each player i . The set $\Phi^{\mathcal{X}_i}$ contains functions mapping \mathcal{X}_i to itself. A joint distribution over strategy profiles qualifies as a $\Phi = \Pi_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium if no player i can increase their expected utility by using any strategy modification function, $\phi_i \in \Phi^{\mathcal{X}_i}$, on the strategy sampled from the joint distribution. The larger the set Φ , the stronger the incentive guarantee offered by the Φ -equilibrium. For example, if $\Phi^{\mathcal{X}_i}$ contains all constant functions, the corresponding Φ -equilibrium coincides with the notion of coarse correlated equilibrium. Throughout the paper, we also consider ε -approximate Φ -equilibria, where no player can gain more than ε by deviating using any function from $\Phi^{\mathcal{X}_i}$. We study several families of Φ and illustrate their relationships in [Figure 1](#).

Finite Set of Global Deviations. The first case we consider is when each player i ’s set of strategy modifications, $\Phi^{\mathcal{X}_i}$, contains a finite number of arbitrary functions mapping \mathcal{X}_i to itself. As shown in [[Greenwald and Jafari, 2003](#)], if there exists an online learning algorithm where each player i is guaranteed to have sublinear $\Phi^{\mathcal{X}_i}$ -regret, the empirical distribution of joint strategies played converges to a $\Phi = \Pi_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium. [Gordon et al. \[2008\]](#) consider Φ -regret minimization but for concave reward functions, and

²[Stoltz and Lugosi \[2007\]](#) only require the elements of Φ to be measurable functions.

Figure 1: The relationship between different solution concepts in non-concave games. An arrow from one solution concept to another means the former is contained in the latter. The dashed arrow from $\text{Conv}(\Phi(\delta))$ -equilibria to Φ_{Finite} -equilibria means the former is contained in the latter when $\Phi(\delta) = \Phi_{\text{Finite}}$.



their results, therefore, do not apply to non-concave games. [Stoltz and Lugosi \[2007\]](#) provide an algorithm that achieves no Φ^{x_i} -regret in non-concave games; however, their algorithm requires a fixed-point computation per step, making it computationally inefficient.³ Our first contribution is to provide an efficient randomized algorithm that achieves no Φ^{x_i} -regret for each player i with high probability.

Contribution 1: Let \mathcal{X} be a convex and compact set, and Ψ an arbitrary finite set of strategy modification functions for \mathcal{X} . We design a randomized online learning algorithm that achieves $O\left(\sqrt{T \log |\Psi|}\right)$ Ψ -regret, with high probability, for *arbitrary* bounded reward functions on \mathcal{X} ([Theorem 2](#)). The algorithm operates in time $\sqrt{T}|\Psi|$ per iteration. If every player in a *non-concave* game adopts this algorithm, the empirical distribution of strategy profiles played forms an ε -approximate $\Phi = \prod_{i=1}^n \Phi^{x_i}$ -equilibrium, with high probability, for any $\varepsilon > 0$, after $\text{poly}\left(\frac{1}{\varepsilon}, \log(\max_i |\Phi^{x_i}|), \log n\right)$ iterations.

If players have infinitely many global strategy modifications, we can extend [Algorithm 1](#) by discretizing the set of strategy modifications under mild assumptions, such as the modifications being Lipschitz ([Corollary 1](#)). The empirical distribution of the strategy profiles still converges to the corresponding Φ -equilibrium, but at a much slower rate of $O(T^{-\frac{1}{d+2}})$, where d is the dimension of the set of strategies. Additionally, the algorithm requires exponential time in the dimension per iteration, making it inefficient. This inefficiency is unavoidable, as the problem remains intractable even when Φ contains only constant functions.

To address the limitations associated with infinitely large global strategy modifications, a natural approach is to focus on local deviations instead. The corresponding Φ -equilibrium will guarantee local stability. The study of local equilibrium concepts in non-concave games has received significant attention in recent years—see e.g., [[Ratliff et al., 2016](#), [Hazan et al., 2017](#), [Daskalakis and Panageas, 2018](#), [Jin et al., 2020](#), [Daskalakis et al., 2021b](#)]. However, these solution concepts either are not guaranteed to exist, are restricted to sequential two-player zero-sum games [[Mangoubi and Vishnoi, 2021](#)], only establish local convergence guarantees for learning dynamics—see e.g., [[Daskalakis and Panageas, 2018](#), [Wang et al., 2020](#), [Fiez et al., 2020](#)], only establish asymptotic convergence guarantees—see e.g., [[Daskalakis et al., 2023](#)], or involve non-standard solution concepts where local stability is not with respect to a distribution over strategy profiles [[Hazan et al., 2017](#)].

We study the tractability of Φ -equilibrium with infinitely large Φ sets that consist solely of local strategy

³The existence of the fixed point is guaranteed by the Schauder-Cauty fixed-point theorem [[Cauty, 2001](#)], a generalization of the Brouwer fixed-point theorem. Hence, it's unlikely such fixed points are tractable.

modifications. These local solution concepts are guaranteed to exist in general multi-player non-concave games. Specifically, we focus on the following three families of natural deviations.

- **Projection based Local Deviations:** Each player i 's set of strategy modifications, denoted by $\Phi_{\text{Proj}}^{\mathcal{X}_i}(\delta)$, contains all deviations that attempt a small step from their input in a fixed direction and project if necessary, namely are of the form $\phi_v(x) = \Pi_{\mathcal{X}_i}[x - v]$, where $\|v\| \leq \delta$ and $\Pi_{\mathcal{X}_i}$ stands for the ℓ_2 -projection onto \mathcal{X}_i .
- **Convex Combination of Finitely Many Local Deviations:** Each player i 's set of strategy modifications, denoted by $\text{Conv}(\Phi^{\mathcal{X}_i}(\delta))$, contains all deviations that can be represented as a convex combination of a finite set of δ -local strategy modifications, i.e., $\|\phi(x) - x\| \leq \delta$ for all $\phi \in \Phi^{\mathcal{X}_i}(\delta)$.
- **Interpolation based Local Deviations:** each player i 's set of local strategy modifications, denoted by $\Phi_{\text{Int}}^{\mathcal{X}_i}(\delta)$, that contains all deviations that *interpolate* between the input strategy and another strategy in \mathcal{X}_i . Formally, each element $\phi_{\lambda, x'}(x)$ of $\Phi_{\text{Int}}^{\mathcal{X}_i}(\delta)$ can be represented as $(1 - \lambda)x + \lambda x'$ for some $x' \in \mathcal{X}_i$ and $\lambda \leq \delta/D_{\mathcal{X}_i}$ ($D_{\mathcal{X}_i}$ is the diameter of \mathcal{X}_i).

For our three families of local strategy modifications, we explore the tractability of Φ -equilibrium within a regime we term the *first-order stationary regime*, where $\varepsilon = \Omega(\delta^2)$, with δ representing the maximum deviation allowed for a player. An ε -approximate Φ -equilibrium in this regime ensures first-order stability. This regime is particularly interesting for two reasons: (i) [Daskalakis et al. \[2021b\]](#) have demonstrated that computing an ε -approximate δ -local Nash equilibrium in this regime is intractable.⁴ This poses an intriguing question: can correlating the players' strategies, as in a Φ -equilibrium, potentially make the problem tractable? (ii) Extending our algorithm, initially designed for finite sets of strategy modifications, to these three sets of local deviations results in inefficiency; specifically, the running time becomes exponential in one of the problem's natural parameters. Designing efficient algorithms for this regime thus presents challenges. Despite these, we show the following:

Contribution 2: For any $\delta > 0$, for each of the three families of infinite δ -local strategy modifications mentioned above, there exists an efficient uncoupled learning algorithm that converges to an ε -approximate Φ -equilibrium of the non-concave game in the first-order stationary regime, i.e., $\varepsilon = \Omega(\delta^2)$.

We present our results for the projection-based local deviation in [Theorem 3](#) and [Theorem 4](#). Our result for the convex combination of local deviations can be found in [Theorem 5](#). [Theorem 6](#) contains our result for the interpolation-based local deviations. Similar to the finite case, our algorithms build on the connection between Φ -regret minimization and Φ -equilibrium. Given that our strategy modifications are non-standard, it is a priori unclear how to minimize the corresponding Φ -regret. For instance, to our knowledge, no algorithm is known to minimize $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret even when the reward functions are concave, and provably $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is incomparable to external regret ([Examples 3](#) and [4](#)). However, via a novel analysis, we show that Online Gradient Descent ([GD](#)) and Optimistic Gradient ([OG](#)) achieve a near-optimal $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret guarantee ([Theorem 3](#) and [Theorem 9](#)).

Our results provide efficient uncoupled algorithms to compute ε -approximate $\Phi(\delta)$ -equilibria in the first-order stationary regime $\varepsilon = \Omega(\delta^2)$. A natural question arises: can we approximate Φ -equilibria within $\varepsilon = o(\delta^2)$ efficiently? We answer this question in the negative: it is in fact NP-hard to achieve

⁴A strategy profile is considered an ε -approximate δ -local Nash equilibrium if no player can gain more than ε by deviating within a δ distance.

$\varepsilon = o(\delta^2)$. This hardness result further illustrates the necessity to consider the $\varepsilon = \Omega(\delta^2)$ regime.

Contribution 3: We show that the first-order stationary regime, i.e., $\varepsilon = \Omega(\delta^2)$, is the best one can hope for in the efficient computation of an ε -approximate $\Phi(\delta)$ -equilibrium when $\Phi(\delta)$ contains all δ -local strategy modifications. Specifically, in [Theorem 11](#), we prove that, unless $P = NP$, for any $\varepsilon \leq \text{poly}(d, n, G, L) \cdot \delta^{2+c}$ with any constant $c > 0$, there is no algorithm with running time $\text{poly}(d, n, G, L, 1/\varepsilon, 1/\delta)$ that can find an ε -approximate $\Phi(\delta)$ -equilibrium. Here, G is the Lipschitzness, and L is the smoothness of the players' utilities. This result holds even for a single-player game over $\mathcal{X} = [0, 1]^d$. Moreover, we show in [Theorem 12](#) that a similar hardness result holds even for a very restricted set of strategy modifications $\Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$ that is the union of $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ and one additional strategy modification.

We remark that existing hardness results for local maximum [[Daskalakis et al., 2021b](#)] could be extended to approximate $\Phi(\delta)$ -equilibrium. However, these results only apply to the “global regime”, where δ equals $D_{\mathcal{X}}$, the diameter of the action set.⁵ In contrast, our hardness results hold for a range of δ and rule out efficient approximation for any $\varepsilon = o(\delta^2)$, which cannot be directly derived from results that concern the “global regime” [[Daskalakis et al., 2021b](#)]. Our analysis is also novel and different from [[Daskalakis et al., 2021b](#)]: we construct a reduction from the NP-hard maximum clique problem using the Motzkin-Straus Theorem [[Motzkin and Straus, 1965](#)]. As a byproduct of our analysis, we also obtain new hardness results for computing approximate local maximum and approximate local Nash equilibrium ([Corollary 2](#)). See [Figure 2](#) for a summary.

Further related work is discussed in [Appendix A](#).

2 Preliminaries

A ball of radius $r > 0$ centered at $x \in \mathbb{R}^d$ is denoted by $B_d(x, r) := \{x' \in \mathbb{R}^d : \|x - x'\| \leq r\}$. We use $\|\cdot\|$ for ℓ_2 norm throughout. We also write $B_d(\delta)$ for a ball centered at the origin with radius δ . For $a \in \mathbb{R}$, we use $[a]^+$ to denote $\max\{0, a\}$. We denote $D_{\mathcal{X}}$ as the diameter of a set \mathcal{X} .

Continuous / Smooth Games. An n -player *continuous game* has a set of n players $[n] := \{1, 2, \dots, n\}$. Each player $i \in [n]$ has a nonempty convex and compact strategy set $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$. For a joint strategy profile $x = (x_i, x_{-i}) \in \prod_{j=1}^n \mathcal{X}_j$, the reward of player i is determined by a utility function $u_i : \prod_{j=1}^n \mathcal{X}_j \rightarrow [0, 1]$. We denote by $d = \sum_{i=1}^n d_i$ the dimensionality of the game and assume $\max_{i \in [n]} \{D_{\mathcal{X}_i}\} \leq D$. A *smooth game* is a continuous game whose utility functions further satisfy the following assumption.

Assumption 1 (Smooth Games). *The utility function $u_i(x_i, x_{-i})$ for any player $i \in [n]$ is differentiable and satisfies*

1. (*G-Lipschitzness*) $\|\nabla_{x_i} u_i(x)\| \leq G$ for all i and $x \in \prod_{j=1}^n \mathcal{X}_j$;
2. (*L-smoothness*) there exists $L_i > 0$ such that $\|\nabla_{x_i} u_i(x_i, x_{-i}) - \nabla_{x_i} u_i(x'_i, x_{-i})\| \leq L_i \|x_i - x'_i\|$ for all $x_i, x'_i \in \mathcal{X}_i$ and $x_{-i} \in \prod_{j \neq i} \mathcal{X}_j$. We denote $L = \max_i L_i$ as the smoothness of the game.

Crucially, we make no assumption on the concavity of $u_i(x_i, x_{-i})$.

⁵In [Appendix I](#), we also prove NP-hardness of ε -approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium ($\Phi_{\text{Int}}(\delta)$ -equilibrium) when $\delta = D_{\mathcal{X}}$ and $\varepsilon = 1$.

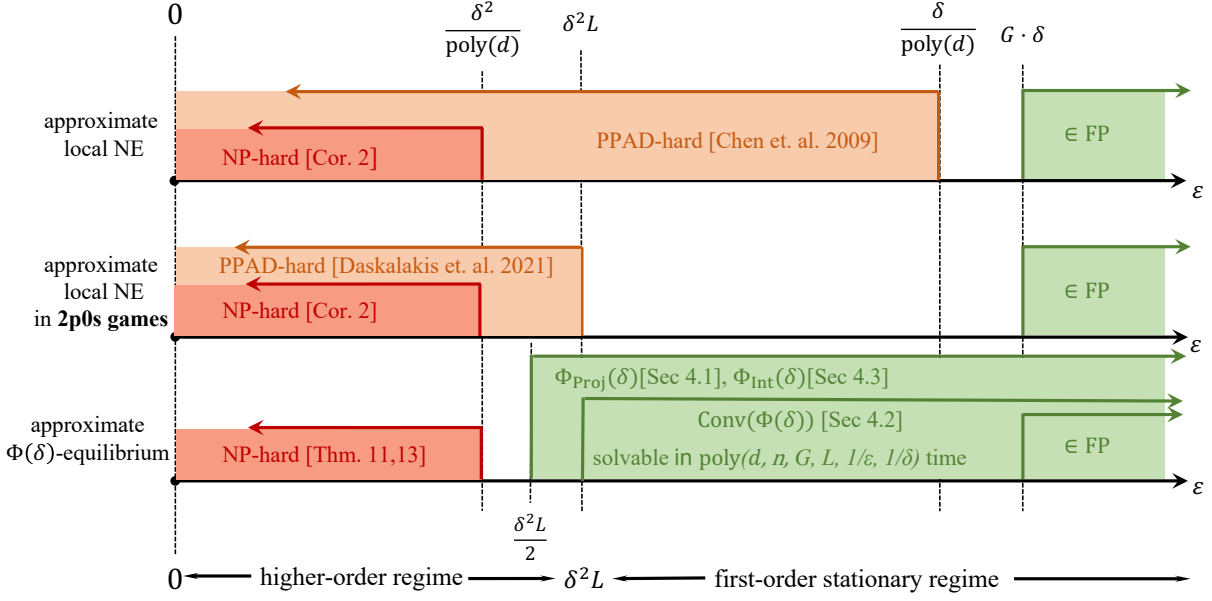


Figure 2: Complexity of computing an ε -approximate δ -local Nash equilibrium and ε -approximate $\Phi(\delta)$ -equilibrium in G -Lipschitz and L -smooth d -dimensional games. We consider cases where $G, L = O(\text{poly}(d))$. The regime $\varepsilon \geq G\delta$ is trivial since the game is G -Lipschitz. The PPAD-hardness of approximate local Nash equilibrium follows from approximate (global) Nash equilibrium in bimatrix games due to linearity of the utility function [Chen et al., 2009]. The PPAD-hardness of approximate local Nash equilibrium in two-player zero-sum games is proved in [Daskalakis et al., 2021b]. The NP-hardness of ε -approximate $\Phi(\delta)$ -equilibrium is proven for $\Phi_{\text{All}}(\delta)$ (Theorem 11) and $\Phi_{\text{Int}^+}(\delta)$ (Theorem 12) in Appendix H. The NP-hardness of ε -approximate δ -local Nash equilibrium is implied by Corollary 2. The positive results for ε -approximate $\Phi(\delta)$ -equilibrium in the first-order stationary regime hold for $\Phi_{\text{Proj}}(\delta)$ (Section 4.1), $\Phi_{\text{Int}}(\delta)$ (Section 4.3), and $\text{Conv}(\Phi(\delta))$ when $|\Phi(\delta)|$ is finite (Section 4.2).

Φ -equilibrium and Φ -regret Below we formally introduce the concept of Φ -equilibrium and its relationship with online learning and Φ -regret minimization.

Definition 1 (Φ -equilibrium [Greenwald and Jafari, 2003, Stoltz and Lugosi, 2007]). *In a continuous game, a distribution σ over joint strategy profiles $\prod_{i=1}^n \mathcal{X}_i$ is an ε -approximate Φ -equilibrium for some $\varepsilon \geq 0$ and a profile of strategy modification sets $\Phi = \prod_{i=1}^n \Phi_i$ if and only if for all player $i \in [n]$, $\max_{\phi \in \Phi_i} \mathbb{E}_{x \sim \sigma} [u_i(\phi(x_i), x_{-i})] \leq \mathbb{E}_{x \sim \sigma} [u_i(x)] + \varepsilon$. When $\varepsilon = 0$, we call σ a Φ -equilibrium.*

We consider the standard online learning setting: at each day $t \in [T]$, the learner chooses an action x^t from a nonempty convex compact set $\mathcal{X} \subseteq \mathbb{R}^m$ and the adversary chooses a possibly non-convex loss function $f^t : \mathcal{X} \rightarrow \mathbb{R}$, then the learner suffers a loss $f^t(x^t)$ and receives feedback. In this paper, we focus on two feedback models: (1) the player receives an oracle for $f^t(\cdot)$; (2) the player receives only the gradient $\nabla f^t(x^t)$. The classic goal of an online learning algorithm is to minimize the *external regret* defined as $\text{Reg}^T := \max_{x \in \mathcal{X}} \sum_{t=1}^T (f^t(x^t) - f^t(x))$. An algorithm is called *no-regret* if its external regret is sublinear in T . The notion of Φ -regret generalizes external regret by allowing more general strategy modifications.

Definition 2 (Φ -regret). *Let Φ be a set of strategy modification functions $\{\phi : \mathcal{X} \rightarrow \mathcal{X}\}$. For $T \geq 1$, the*

Φ -regret of an online learning algorithm is $\text{Reg}_{\Phi}^T := \max_{\phi \in \Phi} \sum_{t=1}^T (f^t(x^t) - f^t(\phi(x^t)))$. An algorithm is called no Φ -regret if its Φ -regret is sublinear in T .

Many classic notions of regret can be interpreted as Φ -regret. For example, the external regret is Φ_{ext} -regret where Φ_{ext} contains all constant strategy modifications $\phi_{x^*}(x) = x^*$ for all $x^* \in \mathcal{X}$. The swap regret on simplex Δ^m is Φ_{swap} -regret where Φ_{swap} contains all linear transformations $\phi : \Delta^m \rightarrow \Delta^m$. A fundamental result for learning in games is that no- Φ -regret learning dynamics in games converge to an approximate Φ -equilibrium [Greenwald and Jafari, 2003].

Theorem 1. [Greenwald and Jafari, 2003] *If each player $i \in [n]$ has Φ_i -regret that is upper bounded by $\text{Reg}_{\Phi_i}^T$, then their empirical distribution of strategy profiles played is an $(\max_{i \in [n]} \text{Reg}_{\Phi_i}^T / T)$ -approximate Φ -equilibrium.*

3 Tractable Φ -Equilibrium for Finite Φ via Sampling

In this section, we revisit the problem of computing and learning an Φ -equilibrium in non-concave games when each player's set of strategy modifications $\Phi^{\mathcal{X}_i}$ is finite.

The pioneering work of Stoltz and Lugosi [2007] gives a no- Φ -regret algorithm for this case where each player chooses a distribution over strategies in each round. This result also implies convergence to Φ -equilibrium. However, the algorithm by Stoltz and Lugosi [2007] is not computationally efficient. In each iteration, their algorithm requires computing a distribution that is stationary under a transformation that can be represented as a mixture of the modifications in Φ . The existence of such a stationary distribution is guaranteed by the Schauder-Cauty fixed-point theorem Cauty [2001], but the distribution might require exponential support and be intractable to find.

Our main result in this section is an efficient Φ -regret minimization algorithm (Algorithm 1) that circumvents the step of the exact computation of a stationary distribution. Consequently, our algorithm also ensures efficient convergence to a Φ -equilibrium when adopted by all players.

Algorithm 1: Φ -regret minimization for non-concave reward via sampling

Input: $x_{\text{root}} \in \mathcal{X}$, $h \geq 2$, an external regret minimization algorithm \mathfrak{R}_{Φ} over Φ

Output: A Φ -regret minimization algorithm for \mathcal{X}

```

1 function NEXTSTRATEGY()
2    $p^t \leftarrow \mathfrak{R}_{\Phi}.\text{NEXTSTRATEGY}()$ . Note that  $p_t$  is a distribution over  $\Phi$ .
3   return  $x^t \leftarrow \text{SAMPLESTRATEGY}(x_{\text{root}}, h, p^t)$ .
4 function OBSERVEREWARD( $u^t(\cdot)$ )
5   Set  $u_{\Phi}^t(\phi) = u^t(\phi(x^t))$  for all  $\phi \in \Phi$ .
6    $\mathfrak{R}_{\Phi}.\text{OBSERVEREWARD}(u_{\Phi}^t(\cdot))$ .
```

Theorem 2. *Let \mathcal{X} be a convex and compact set, Φ be an arbitrary finite set of strategy modification functions for \mathcal{X} , and $u^1(\cdot), \dots, u^T(\cdot)$ be an arbitrary sequence of possibly non-concave reward functions from \mathcal{X} to $[0, 1]$. If we instantiate Algorithm 1 with the Hedge algorithm as the regret minimization algorithm \mathfrak{R}_{Φ} over Φ and $d = \sqrt{T}$, the algorithm guarantees that, with probability at least $1 - \beta$, it produces a sequence of strategies x^1, \dots, x^T with Φ -regret at most $\max_{\phi \in \Phi} \sum_{t=1}^T u^t(\phi(x^t)) - \sum_{t=1}^T u^t(x^t) \leq 8\sqrt{T}(\log |\Phi| + \log(1/\beta))$. Moreover, the algorithm runs in time $O(\sqrt{T}|\Phi|)$ per iteration.*

Algorithm 2: SAMPLESTRATEGY

Input: $x_{\text{root}} \in \mathcal{X}$, $h \geq 2$, $p^t \in \Delta(\Phi)$

Output: $x \in \mathcal{X}$

- 1 $x_1 \leftarrow x_{\text{root}}$.
 - 2 **for** $2 \leq k \leq h$ **do**
 - 3 $\phi \leftarrow$ sample from Φ according to p^t .
 - 4 $x_k = \phi(x_{k-1})$.
 - 5 **return** x from $\{x_1, \dots, x_h\}$ uniformly at random.
-

If all players in a non-concave continuous game employ [Algorithm 1](#), then with probability at least $1 - \beta$, for any $\varepsilon > 0$, the empirical distribution of strategy profiles played forms an ε -approximate $\Phi = \prod_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium, after $\text{poly}\left(\frac{1}{\varepsilon}, \log(\max_i |\Phi^{\mathcal{X}_i}|), \log \frac{n}{\beta}\right)$ iterations.

High-level ideas We adopt the framework in [Stoltz and Lugosi, 2007]. The framework contains two steps in each iteration t : (1) the learner runs a no-external-regret algorithm over Φ which outputs $p^t \in \Delta(\Phi)$ in each iteration t ; (2) the learner chooses a stationary distribution $\mu^t = \sum_{\phi \in \Phi} p^t \phi(\mu^t)$, where we slightly abuse notation to use $\phi(\mu^t)$ to denote the image measure of μ by ϕ . However, how to compute the stationary distribution μ^t efficiently is unclear. We essentially provide a computationally efficient way to carry out step (2) without computing this stationary distribution.

- We first construct an ε -approximate stationary distribution by recursively applying strategy modifications from Φ . The constructed distribution can be viewed as a tree. Our construction is inspired by the recent work of [Zhang et al. \[2024\]](#) for concave games. The main difference here is that for non-concave games, the distribution needs to be approximately stationary with respect to a *mixture* of strategy modifications rather than a single one as in concave games. Consequently, this leads to an approximate stationary distribution with prohibitively high support size $(|\Phi|)^{\sqrt{T}}$, as opposed to \sqrt{T} in [Zhang et al., 2024] for concave games.
- Despite the exponentially large support size of the distribution, we utilize its tree structure to design a simple and efficient sampling procedure that runs in time \sqrt{T} . Equipped with such a sampling procedure, we provide an efficient randomized algorithm that generates a sequence of strategies so that, with high probability, the Φ -regret for this sequence of strategies is at most $O(\sqrt{T \log |\Phi|})$.

We defer the full proof of [Theorem 2](#) to [Section 3.1](#). An extension of [Theorem 2](#) to infinite Φ holds when the rewards $\{u^t\}_{t \in [T]}$ are G -Lipschitz and Φ admits an α -cover with size $N(\alpha)$. In particular, when Φ is the set of all M -Lipschitz functions over $[0, 1]^d$, Φ admits an α -cover with $\log N(\alpha)$ of the order $(1/\alpha)^d$. In this case, we have

Corollary 1. *There is a randomized algorithm such that, with probability at least $1 - \beta$, the Φ -regret is bounded by $c \cdot T^{\frac{d+1}{d+2}} \cdot \log(1/\beta)$, where c only depends on G and M . The algorithm runs in time $\text{poly}(T, N(T^{-1/(d+2)}))$.*

3.1 Proof of Theorem 2

For a distribution $\mu \in \Delta(\mathcal{X})$ over strategy space \mathcal{X} , we slightly abuse notation and define its expected utility as

$$u^t(\mu) := \mathbb{E}_{x \sim \mu} [u^t(x)] \in [0, 1].$$

We define $\phi(\mu)$ the image of μ under transformation ϕ . In each iteration t , the learner chooses their strategy $x^t \in \mathcal{X}$ according to the distribution μ^t . For a sequence of strategies $\{x^t\}_{t \in [T]}$, the Φ -regret is

$$\text{Reg}_{\Phi}^T := \max_{\phi \in \Phi} \left\{ \sum_{t=1}^T (u^t(\phi(x^t)) - u^t(x^t)) \right\}.$$

Algorithm 1 uses an external regret minimization algorithm R_{Φ} over Φ which outputs a distribution $p^t \in \Delta(\Phi)$. We can then decompose the Φ -regret into two parts.

$$\text{Reg}_{\Phi}^T = \underbrace{\max_{\phi \in \Phi} \left\{ \sum_{t=1}^T u^t(\phi(x^t)) - \mathbb{E}_{\phi' \sim p^t} [u^t(\phi'(x^t))] \right\}}_{\text{I: external regret over } \Phi} + \underbrace{\sum_{t=1}^T \mathbb{E}_{\phi' \sim p^t} [u^t(\phi'(x^t))] - u^t(x^t)}_{\text{II: approximation error of stationary distribution}}.$$

I: Bounding the external regret over Φ . The external regret over Φ can be bounded directly. This is equivalent to an online expert problem: in each iteration t , the external regret minimizer \mathfrak{R}_{Φ} chooses $p^t \in \Delta(\Phi)$ and the adversary then determines the utility of each expert $\phi \in \Phi$ as $u^t(\phi(x^t))$. We choose the external regret minimizer \mathfrak{R}_{Φ} to be the Hedge algorithm [Freund and Schapire, 1999], which gives

$$\max_{\phi \in \Phi} \left\{ \sum_{t=1}^T u^t(\phi(x^t)) - \mathbb{E}_{\phi' \sim p^t} [u^t(\phi'(x^t))] \right\} \leq 2\sqrt{T \log |\Phi|}, \quad (1)$$

where we use the fact that the utility function u^t is bounded in $[0, 1]$.

II: Bounding error due to sampling from an approximate stationary distribution. We first define a distribution μ^t using a complete $|\Phi|$ -ary tree with depth h . The root of this tree is an arbitrary strategy $x_{\text{root}} \in \mathcal{X}$. Each internal node x has exactly $|\Phi|$ children, denoted as $\{\phi(x)\}_{\phi \in \Phi}$. The distribution μ^t is supported on the nodes of this tree. Next, we define the probability for each node under the distribution μ^t . The root node x_{root} receives probability $\frac{1}{h}$. The probability of other nodes is defined in a recursive manner. For every node $x = \phi(x_p)$ where x_p is its parent, x receives probability $\Pr_{\mu^t}[x] = \Pr_{\mu^t}[x_p] \cdot p^t(\phi)$. It is then clear that the total probability of the children of a node x_p is exactly $\Pr_{\mu^t}[x \text{ is } x_p\text{'s child}] = \sum_{\phi} \Pr[x_p] \cdot p^t(\phi) = \Pr[x_p]$. Denote the set of nodes in depth k as N_k . We have $\Pr_{\mu^t}[x \in N_k] = \frac{1}{h}$ for every depth $1 \leq k \leq h$. Thus the distribution μ^t supports on $\frac{|\Phi|^d - 1}{|\Phi| - 1}$ points and is well-defined. By the construction above, we know x^t output by Algorithm 2 is a sample from μ^t .

Claim 1. Algorithm 2 generates a sample from the distribution μ^t .

Now we show that the approximation error of μ^t is bounded by $O(\frac{1}{h})$. We can evaluate the approximation error of μ^t :

$$\begin{aligned} & \mathbb{E}_{\phi \sim p^t} [u^t(\phi(\mu^t))] - u^t(\mu^t) \\ &= \mathbb{E}_{\phi \sim p^t} \left[\sum_{k=1}^h \sum_{x \in N_k} \Pr_{\mu^t}[x] u^t(\phi(x)) \right] - \left[\sum_{k=1}^h \sum_{x \in N_k} \Pr_{\mu^t}[x] u^t(x) \right] \\ &= \sum_{k=1}^h \sum_{x \in N_k} \left(\mathbb{E}_{\phi \sim p^t} \left[\Pr_{\mu^t}[x] u^t(\phi(x)) \right] - \Pr_{\mu^t}[x] u^t(x) \right). \end{aligned}$$

We recall that for a node $x = \phi(x_p)$ with x_p being its parent, we have $\Pr_{\mu^t}[x] = \Pr_{\mu^t}[x_p] \cdot p^t(\phi)$. Thus for any $1 \leq k \leq h-1$, we have

$$\begin{aligned} & \sum_{x \in N_k} \left(\mathbb{E}_{\phi \sim p^t} \left[\Pr_{\mu^t}[x] u^t(\phi(x)) \right] - \Pr_{\mu^t}[x] u^t(x) \right) \\ &= \sum_{x \in N_k} \left(\sum_{\phi \in \Phi} p^t(\phi) \Pr_{\mu^t}[x] u^t(\phi(x)) - \Pr_{\mu^t}[x] u^t(x) \right) \\ &= \sum_{x \in N_{k+1}} \Pr_{\mu^t}[x] u^t(x) - \sum_{x \in N_k} \Pr_{\mu^t}[x] u^t(x). \end{aligned}$$

Using the above equality, we get

$$\begin{aligned} & \mathbb{E}_{\phi \sim p^t} [u^t(\phi(\mu^t))] - u^t(\mu^t) \\ &= \sum_{k=1}^{h-1} \sum_{x \in N_{k+1}} \Pr_{\mu^t}[x] u^t(x) + \sum_{x \in N_h} \sum_{\phi \in \Phi} p^t(\phi) \Pr_{\mu^t}[x] u^t(\phi(x)) - \sum_{k=2}^h \sum_{x \in N_k} \Pr_{\mu^t}[x] u^t(x) - \Pr_{\mu^t}[x_{\text{root}}] u^t(x_{\text{root}}) \\ &= \sum_{x \in N_h} \sum_{\phi \in \Phi} p^t(\phi) \Pr_{\mu^t}[x] u^t(\phi(x)) - \Pr_{\mu^t}[x_{\text{root}}] u^t(x_{\text{root}}) \\ &\leq \frac{1}{h}, \end{aligned}$$

where in the last inequality we use the fact that $\sum_{x \in N_k} \Pr_{\mu^t}[x] = \frac{1}{h}$ for all $1 \leq k \leq h$ and the utility function u^t is bounded in $[0, 1]$. Therefore, for $x^t \sim \mu^t$, the sequence of random variables

$$\sum_{t=1}^{\tau} \left(\mathbb{E}_{\phi \sim p^t} [u^t(\phi(x^t))] - u^t(x^t) - \frac{1}{h} \right), \text{ for } \tau \geq 1.$$

is a super-martingale. Thanks to the boundedness of the utility function, we can apply the Hoeffding-Azuma Inequality and get for any $\varepsilon > 0$.

$$\Pr \left[\sum_{t=1}^T \left(\mathbb{E}_{\phi \sim p^t} [u^t(\phi(x^t))] - u^t(x^t) - \frac{1}{h} \right) \geq \varepsilon \right] \leq \exp \left(-\frac{\varepsilon^2}{8T} \right). \quad (2)$$

Combining (1) and (2) with $\varepsilon = \sqrt{8T \log(1/\beta)}$ and $h = \sqrt{T}$, we have that, with probability $1 - \beta$, that

$$\begin{aligned} \text{Reg}_{\Phi}^T &= \max_{\phi \in \Phi} \left\{ \sum_{t=1}^T u^t(\phi(x^t)) - \mathbb{E}_{\phi' \sim p^t} [u^t(\phi'(x^t))] \right\} + \sum_{t=1}^T \mathbb{E}_{\phi' \sim p^t} [u^t(\phi'(x^t))] - u^t(x^t) \\ &\leq 2\sqrt{T \log |\Phi|} + \frac{T}{h} + \sqrt{8T \log(1/\beta)} \\ &\leq 8\sqrt{T(\log |\Phi| + \log(1/\beta))}. \end{aligned}$$

Convergence to Φ -equilibrium If all players in a non-concave continuous game employ [Algorithm 1](#), then we know for each player i , with probability $1 - \frac{\beta}{n}$, its $\Phi^{\mathcal{X}_i}$ -regret is upper bounded by

$$8\sqrt{T(\log |\Phi^{\mathcal{X}_i}| + \log(n/\beta))}$$

By a union bound over all n players, we get with probability $1 - \beta$, every player i 's $\Phi^{\mathcal{X}_i}$ -regret is upper bounded by $8\sqrt{T(\log |\Phi^{\mathcal{X}_i}| + \log(n/\beta))}$. Now by [Theorem 1](#), we know the empirical distribution of strategy profiles played forms an ε -approximate $\Phi = \prod_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium, as long as $T \geq \frac{64(\log |\Phi^{\mathcal{X}_i}| + \log(n/\beta))}{\varepsilon^2}$ iterations. This completes the proof of [Theorem 2](#)

4 Approximate Φ -Equilibria under Infinite Local Strategy Modifications

This section studies Φ -equilibrium when $|\Phi|$ is infinite. It is, in general, computationally hard to compute a Φ -equilibrium even if Φ contains all constant deviations. Instead, we focus on Φ that consists solely of local strategy modifications. We introduce several natural classes of local strategy modifications and provide efficient online learning algorithms that converge to ε -approximate Φ -equilibrium in the first-order stationary regime where $\varepsilon = \Omega(\delta^2 L)$. These approximate Φ -equilibria guarantee first-order stability.

Definition 3 (δ -local strategy modification). *For each agent i , we call a set of strategy modifications $\Phi^{\mathcal{X}_i}$ δ -local if for all $x \in \mathcal{X}_i$ and $\phi_i \in \Phi^{\mathcal{X}_i}$, $\|\phi_i(x) - x\| \leq \delta$. We use notation $\Phi^{\mathcal{X}_i}(\delta)$ to denote a δ -local strategy modification set for agent i . We also use $\Phi(\delta) = \prod_{i=1}^n \Phi^{\mathcal{X}_i}(\delta)$ to denote a profile of δ -local strategy modification sets.*

Below we present a useful reduction from computing an ε -approximate $\Phi(\delta)$ -equilibrium in *non-concave* smooth games to $\Phi^{\mathcal{X}_i}(\delta)$ -regret minimization against *convex* losses for any $\varepsilon \geq \frac{\delta^2 L}{2}$. The key observation here is that the L -smoothness of the utility function permits the approximation of a non-concave function with a linear function within a δ -neighborhood. This approximation yields an error of at most $\frac{\delta^2 L}{2}$ that lies in the first-order stationary regime. We defer the proof to [Appendix C](#).

Lemma 1 (No $\Phi(\delta)$ -Regret for Convex Losses to Approximate $\Phi(\delta)$ -Equilibrium in Non-Concave Games). *For any $T \geq 1$ and $\delta > 0$, let \mathcal{A} be an algorithm that guarantees to achieve no more than $\text{Reg}_{\Phi^{\mathcal{X}_i}(\delta)}^T$ $\Phi^{\mathcal{X}_i}(\delta)$ -regret for convex loss functions for each agent $i \in [n]$. Then*

1. *The $\Phi^{\mathcal{X}_i}(\delta)$ -regret of \mathcal{A} for non-convex and L -smooth loss functions is at most $\text{Reg}_{\Phi^{\mathcal{X}_i}(\delta)}^T + \frac{\delta^2 L T}{2}$ for each agent i .*
2. *When every agent employs \mathcal{A} in a non-concave L -smooth game, their empirical distribution of the joint strategies played converges to a $(\max_{i \in [n]} \{\text{Reg}_{\Phi^{\mathcal{X}_i}(\delta)}^T\} / T + \frac{\delta^2 L}{2})$ -approximate $\Phi(\delta)$ -equilibrium.*

Computing Φ -Equilibrium via Generic Φ -Regret Minimization. By [Lemma 1](#), it suffices to design no Φ -regret algorithms against convex losses for efficient equilibrium computation. Although Φ -regret minimization is extensively studied [[Greenwald and Jafari, 2003](#), [Hazan and Kale, 2007](#), [Stoltz and Lugosi, 2007](#), [Gordon et al., 2008](#), [Dagan et al., 2023](#), [Peng and Rubinstein, 2023](#)], to our knowledge, all of these algorithms, when applied to compute a $(\varepsilon, \Phi(\delta))$ -equilibrium for a general δ -local strategy modification set $\Phi(\delta)$ (using [Lemma 1](#)), require running time exponential in either $1/\varepsilon$ or the dimension d . In the following sections, we show that for several natural choices of $\Phi(\delta)$, ε -approximate $\Phi(\delta)$ -equilibrium can be computed efficiently, i.e. polynomial in $1/\varepsilon$ and d , using simple algorithms.

4.1 Projection-Based Local Strategy Modifications

In this section, we study a set of local strategy modifications based on projection. Specifically, the set $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ encompasses all deviations that essentially add a fixed displacement vector v to the input strategy and project back to the feasible set:

$$\Phi_{\text{Proj}}^{\mathcal{X}}(\delta) := \{\phi_{\text{Proj},v}(x) = \Pi_{\mathcal{X}}[x - v] : v \in B_d(\delta)\}.$$

It is clear that $\|\phi_v(x) - x\| \leq \|v\| \leq \delta$. The induced $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is

$$\text{Reg}_{\text{Proj},\delta}^T := \max_{v \in B_d(\delta)} \sum_{t=1}^T (f^t(x^t) - f^t(\Pi_{\mathcal{X}}[x^t - v])).$$

We also define $\Phi_{\text{Proj}}(\delta) = \prod_{i=1}^n \Phi_{\text{Proj}}^{\mathcal{X}_i}(\delta)$.

By [Lemma 1](#), to achieve convergence to an approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium, it suffices to minimize $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret against convex losses. However, to our knowledge, there does not exist an algorithm that minimizes $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret even in the convex case. In fact, external regret and $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret are provably incomparable: a sequence of actions may suffer high Reg^T but low $\text{Reg}_{\text{Proj},\delta}^T$ ([Example 3](#)) or vice versa ([Example 4](#)). At a high level, the external regret competes against a fixed action, whereas $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is more akin to the notion of *dynamic regret*, competing with a sequence of varying actions. Despite this, surprisingly, we show that classical algorithms like Online Gradient Descent ([GD](#)) and Optimistic Gradient Descent ([OG](#)), known for minimizing external regret, also attain near-optimal $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret. We defer the examples and missing proofs to [Appendix D](#).

4.1.1 $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret Minimization in the Adversarial Setting

We show that ([GD](#)) enjoys an $O(G\sqrt{\delta D_{\mathcal{X}}T})$ $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret despite the difference between the external regret and $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret. First, let us recall the update rule of [GD](#): given initial point $x^1 \in \mathcal{X}$ and step size $\eta > 0$, [GD](#) updates in each iteration t :

$$x^{t+1} = \Pi_{\mathcal{X}}[x - \eta \nabla f^t(x^t)]. \tag{GD}$$

The key step in our analysis for [GD](#) is simple but novel and general. We extend the analysis to the Optimistic Gradient ([OG](#)) algorithm in [Section 4.1.2](#).

Theorem 3. *Let $\delta > 0$ and $T \in \mathbb{N}$. For any convex and G -Lipschitz loss functions $\{f^t : \mathcal{X} \rightarrow \mathbb{R}\}_{t \in [T]}$, the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of [GD](#) with step size $\eta > 0$ is $\text{Reg}_{\text{Proj},\delta}^T \leq \frac{\delta^2}{2\eta} + \frac{\eta}{2}G^2T + \frac{\delta D_{\mathcal{X}}}{\eta}$. We can choose η*

optimally as $\frac{\sqrt{\delta(\delta+D_{\mathcal{X}})}}{G\sqrt{T}}$ and attain $\text{Reg}_{\text{Proj},\delta}^T \leq 2G\sqrt{\delta(\delta+D_{\mathcal{X}})T}$. For any $\delta > 0$ and any $\varepsilon > 0$, when all players employ **GD** in a smooth game, their empirical distribution of played strategy profiles converges to an $(\varepsilon + \frac{\delta^2 L}{2})$ -approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium in $O(1/\varepsilon^2)$ iterations.

Remark 1. Note that $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret can also be viewed as the dynamic regret [Zinkevich, 2003] with changing comparators $\{p^t := \Pi_{\mathcal{X}}[x - v]\}$. However, we remark that our analysis does not follow from standard $O(\frac{(1+P_T)}{\eta} + \eta T)$ dynamic regret bound of **GD** [Zinkevich, 2003] since P_T , defined as $\sum_{t=2}^T \|p^t - p^{t-1}\|$, can be $\Omega(\eta T)$.

Proof. Let us denote $v \in B_d(\delta)$ a fixed deviation and define $p^t = \Pi_{\mathcal{X}}[x^t - v]$. By standard analysis of **GD** [Zinkevich, 2003] (see also the proof of [Bubeck et al., 2015, Theorem 3.2]), we have

$$\begin{aligned} \sum_{t=1}^T (f^t(x^t) - f^t(p^t)) &\leq \sum_{t=1}^T \frac{1}{2\eta} \left(\|x^t - p^t\|^2 - \|x^{t+1} - p^t\|^2 + \eta^2 \|\nabla f^t(x^t)\|^2 \right) \\ &\leq \sum_{t=1}^{T-1} \frac{1}{2\eta} \left(\|x^{t+1} - p^{t+1}\|^2 - \|x^{t+1} - p^t\|^2 \right) + \frac{\delta^2}{2\eta} + \frac{\eta}{2} G^2 T, \end{aligned}$$

where the last step uses $\|x^1 - p^1\| \leq \delta$ and $\|\nabla f^t(x^t)\| \leq G$. Here the terms $\|x^{t+1} - p^{t+1}\|^2 - \|x^{t+1} - p^t\|^2$ do not telescope, and we further relax them in the following key step.

Key Step: We relax the first term as:

$$\begin{aligned} \|x^{t+1} - p^{t+1}\|^2 - \|x^{t+1} - p^t\|^2 &= \langle p^t - p^{t+1}, 2x^{t+1} - p^t - p^{t+1} \rangle \\ &= \langle p^t - p^{t+1}, 2x^{t+1} - 2p^{t+1} \rangle - \|p^t - p^{t+1}\|^2 \\ &= 2\langle p^t - p^{t+1}, v \rangle + 2\langle p^t - p^{t+1}, x^{t+1} - v - p^{t+1} \rangle - \|p^t - p^{t+1}\|^2 \\ &\leq 2\langle p^t - p^{t+1}, v \rangle - \|p^t - p^{t+1}\|^2, \end{aligned}$$

where in the last inequality we use the fact that p^{t+1} is the projection of $x^{t+1} - v$ onto \mathcal{X} and p^t is in \mathcal{X} . Now we get a telescoping term $2\langle p^t - p^{t+1}, v \rangle$ and a negative term $-\|p^t - p^{t+1}\|^2$. The negative term is useful for improving the regret analysis in the game setting, but we ignore it for now. Combining the two inequalities above, we have

$$\begin{aligned} \sum_{t=1}^T (f^t(x^t) - f^t(p^t)) &\leq \frac{\delta^2}{2\eta} + \frac{\eta}{2} G^2 T + \frac{1}{\eta} \sum_{t=1}^{T-1} \langle p^t - p^{t+1}, v \rangle \\ &= \frac{\delta^2}{2\eta} + \frac{\eta}{2} G^2 T + \frac{1}{\eta} \langle p^1 - p^T, v \rangle \leq \frac{\delta^2}{2\eta} + \frac{\eta}{2} G^2 T + \frac{\delta D_{\mathcal{X}}}{\eta}. \end{aligned}$$

Since the above holds for any v with $\|v\| \leq \delta$, it also upper bounds $\text{Reg}_{\text{Proj},\delta}^T$. \square

Lower bounds for $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret We complement our upper bound with two lower bounds for $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret minimization. The first one is an $\Omega(\delta G\sqrt{T})$ lower bound for any online learning algorithms against linear loss functions (Theorem 8). The second one is an $\Omega(\delta^2 LT)$ lower bound for any algorithm that satisfies the *linear span* assumption, which holds for **GD** and **OG** against L -smooth non-convex losses. Combining with Lemma 1, this lower bound suggests that **GD** attains nearly optimal $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret, even in the non-convex setting, among a natural family of gradient-based algorithms. We defer the theorem statements and detailed discussion to Appendix D.

4.1.2 Improved $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret in the Game Setting

Any online algorithm suffers an $\Omega(\sqrt{T}) \Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret even against linear loss functions by [Theorem 8](#). This lower bound, however, holds only in the *adversarial* setting. In this section, we show an improved $O(T^{\frac{1}{4}})$ individual $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret bound under a slightly stronger smoothness assumption ([Assumption 2](#)) in the *game* setting, where players interact with each other using the same algorithm, previous results show improved external regret [[Syrkkanis et al., 2015](#), [Chen and Peng, 2020](#), [Daskalakis et al., 2021a](#), [Anagnostides et al., 2022a,b](#), [Farina et al., 2022a](#)]. This assumption is naturally satisfied by finite normal-form games and is also made for results about concave games [[Farina et al., 2022a](#)].

Assumption 2. For any player $i \in [n]$, the utility $u_i(x)$ satisfies $\|\nabla_{x_i} u_i(x) - \nabla_{x_i} u_i(x')\| \leq L\|x - x'\|$ for all $x, x' \in \mathcal{X}$.

We study the Optimistic Gradient (OG) algorithm [[Rakhlin and Sridharan, 2013](#)], an optimistic variant of GD that has been shown to have improved individual *external* regret guarantee in the game setting [[Syrkkanis et al., 2015](#)]. The OG algorithm initializes $w^0 \in \mathcal{X}$ arbitrarily and $g^0 = 0$. In each step $t \geq 1$, the algorithm plays x^t , receives feedback $g^t := \nabla f^t(x^t)$, and updates w^t , as follows:

$$x^t = \Pi_{\mathcal{X}}[w^{t-1} - \eta g^{t-1}], \quad w^t = \Pi_{\mathcal{X}}[w^{t-1} - \eta g^t]. \quad (\text{OG})$$

We show that OG has $O(\sqrt{T}) \Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret in the adversarial setting ([Theorem 9](#) in [Appendix D](#)) and fast $O(T^{1/4}) \Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret and convergence to approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium in games.

Theorem 4 (Improved Individual $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret of OG in the Game Setting). *In a G -Lipschitz L -smooth (in the sense of [Assumption 2](#)) game, when all players employ OG with step size $\eta > 0$, then for each player i , $\delta > 0$, and $T \geq 1$, their individual $\Phi_{\text{Proj}}^{\mathcal{X}_i}(\delta)$ -regret denoted as $\text{Reg}_{\text{Proj},\delta}^{T,i}$ is $\text{Reg}_{\text{Proj},\delta}^{T,i} \leq \frac{\delta D}{\eta} + \eta G^2 + 3nL^2G^2\eta^3T$. Choosing $\eta = \min\{(\delta D/(nL^2G^2T))^{\frac{1}{4}}, (\delta D)^{\frac{1}{2}}/G\}$, we have $\text{Reg}_{\text{Proj},\delta}^{T,i} \leq 4(\delta D)^{\frac{3}{4}}(nL^2G^2T)^{\frac{1}{4}} + 2\sqrt{\delta DG}$. Furthermore, for any $\delta > 0$ and any $\varepsilon > 0$, their empirical distribution of played strategy profiles converges to an $(\varepsilon + \frac{\delta^2 L}{2})$ -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium in $O(1/\varepsilon^{\frac{4}{3}})$ iterations.*

4.2 Convex Combination of Finite Local Strategy Modifications

This section considers $\text{Conv}(\Phi)$ where Φ is a finite set of local strategy modifications. The set of infinite strategy modifications $\text{Conv}(\Phi)$ is defined as $\text{Conv}(\Phi) = \{\phi_p(x) = \sum_{\phi \in \Phi} p(\phi)\phi(x) : p \in \Delta(\Phi)\}$. Our main result is an efficient algorithm ([Algorithm 3](#)) that guarantees convergence to an ε -approximate $\text{Conv}(\Phi)$ -equilibrium in a smooth game satisfying [Assumption 1](#) for any $\varepsilon > \delta^2 L$. Due to space constraints, we defer [Algorithm 3](#) and the proof to [Appendix E](#).

Theorem 5. *Let \mathcal{X} be a convex and compact set, Φ be an arbitrary finite set of δ -local strategy modification functions for \mathcal{X} , and $u^1(\cdot), \dots, u^T(\cdot)$ be a sequence of G -Lipschitz and L -smooth but possibly non-concave reward functions from \mathcal{X} to $[0, 1]$. If we instantiate [Algorithm 3](#) with the Hedge algorithm as the regret minimization algorithm \mathfrak{R}_{Φ} over $\Delta(\Phi)$ and $K = \sqrt{T}$, the algorithm guarantees that, with probability at least $1 - \beta$, it produces a sequence of strategies x^1, \dots, x^T with $\text{Conv}(\Phi)$ -regret at most $8\sqrt{T}(G\delta\sqrt{\log|\Phi|} + \sqrt{\log(1/\beta)}) + \delta^2 LT$. The algorithm runs in time $\sqrt{T}|\Phi|$ per iteration.*

If all players in a non-concave smooth game employ [Algorithm 3](#), then with probability $1 - \beta$, for any $\varepsilon > 0$, the empirical distribution of strategy profiles played forms an $(\varepsilon + \delta^2 L)$ -approximate $\Phi = \Pi_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium, after $\text{poly}\left(\frac{1}{\varepsilon}, G, \log(\max_i |\Phi^{\mathcal{X}_i}|), \log \frac{n}{\beta}\right)$ iterations.

Proof Sketch We adopt the framework in [Stoltz and Lugosi, 2007, Gordon et al., 2008] (as described in Section 3) with two main modifications. First, we utilize the L -smoothness of the utilities to transform the problem of external regret over $\Delta(\Phi)$ against non-concave rewards into a linear optimization problem. Second, we use the technique of “fixed point in expectation” [Zhang et al., 2024] to circumvent the intractable problem of finding a fixed point.

4.3 Interpolation-Based Local Strategy Modifications

We introduce a natural set of local strategy modifications and the corresponding local equilibrium notion. Given any set of (possibly non-local) strategy modifications $\Psi = \{\psi : \mathcal{X} \rightarrow \mathcal{X}\}$, we define a set of *local* strategy modifications as follows: for $\delta \leq D_{\mathcal{X}}$ and $\lambda \in [0, 1]$, each strategy modification $\phi_{\lambda, \psi}$ interpolates the input strategy x with the modified strategy $\psi(x)$: formally,

$$\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta) := \{\phi_{\lambda, \psi}(x) := (1 - \lambda)x + \lambda\psi(x) : \psi \in \Psi, \lambda \leq \delta/D_{\mathcal{X}}\}.$$

Note that for any $\psi \in \Psi$ and $\lambda \leq \frac{\delta}{D_{\mathcal{X}}}$, we have $\|\phi_{\lambda, \psi}(x) - x\| = \lambda\|x - \psi(x)\| \leq \delta$, respecting the locality constraint. The induced $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret can be written as

$$\text{Reg}_{\text{Int}, \Psi, \delta}^T := \max_{\psi \in \Psi, \lambda \leq \frac{\delta}{D_{\mathcal{X}}}} \sum_{t=1}^T (f^t(x^t) - f^t((1 - \lambda)x^t + \lambda\psi(x^t))).$$

To guarantee convergence to the corresponding Φ -equilibrium, it suffices to minimize $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret against convex losses, which we show further reduces to Ψ -regret minimization against convex losses (Theorem 10 in Appendix F).

CCE-like Instantiation In the special case where Ψ contains only *constant* strategy modifications (i.e., $\psi(x) = x^*$ for all x), we get a coarse correlated equilibrium (CCE)-like instantiation of local equilibrium, which limits the gain by interpolating with any *fixed* strategy. We denote the resulting set of local strategy modification simply as $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$. We can apply any no-external regret algorithm for efficient $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -regret minimization and computation of ε -approximate $\Phi_{\text{Int}}(\delta)$ -equilibrium in the first-order stationary regime as summarized in Theorem 6. We also discuss faster convergence rates in the game setting in Appendix F.

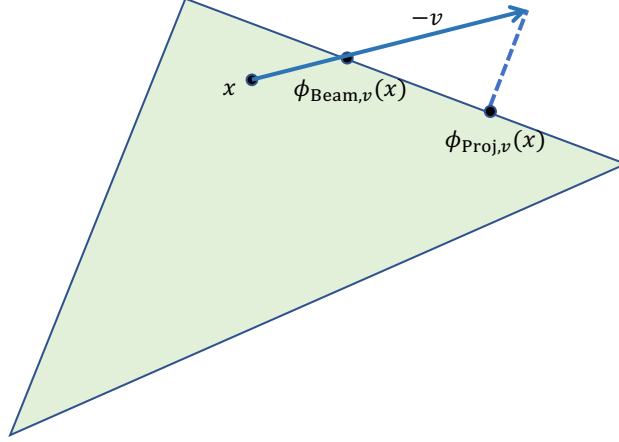
Theorem 6. *For the Online Gradient Descent algorithm (GD) [Zinkevich, 2003] with step size $\eta = \frac{D_{\mathcal{X}}}{G\sqrt{T}}$, its $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -regret is at most $2\delta G\sqrt{T}$. Furthermore, for any $\delta > 0$ and any $\varepsilon > \frac{\delta^2 L}{2}$, when all players employ the GD algorithm in a smooth game, their empirical distribution of played strategy profiles converges to an $(\varepsilon + \frac{\delta^2 L}{2})$ -approximate $\Phi_{\text{Int}}(\delta)$ -equilibrium in $O(1/\varepsilon^2)$ iterations.*

4.4 Beam-Search Local Strategy Modifications and Local Equilibria

In Section 4.1 and Section 4.3, we have shown that GD achieves near-optimal performance for both $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -regret and $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret. In this section, we introduce another natural set of local strategy modifications, $\Phi_{\text{Beam}}^{\mathcal{X}}(\delta)$, which is similar to $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$. Specifically, the set $\Phi_{\text{Beam}}^{\mathcal{X}}(\delta)$ contains deviations that try to move as far as possible in a fixed direction (see Figure 3 for an illustration of the difference between $\phi_{\text{Beam}, v}(x)$ and $\phi_{\text{Proj}, v}(x)$):

$$\Phi_{\text{Beam}}^{\mathcal{X}}(\delta) := \{\phi_{\text{Beam}, v}(x) = x - \lambda^* v : v \in B_d(\delta), \lambda^* = \max\{\lambda : x - \lambda v \in \mathcal{X}, \lambda \in [0, 1]\}\}.$$

Figure 3: Illustration of $\phi_{\text{Proj},v}(x)$ and $\phi_{\text{Beam},v}(x)$



It is clear that $\|\phi_{\text{Beam},v}(x) - x\| \leq \|v\| \leq \delta$. We can similarly derive the notion of $\Phi_{\text{Beam}}^{\mathcal{X}}$ -regret and $(\varepsilon, \Phi_{\text{Beam}}(\delta))$ -equilibrium. Surprisingly, we show that **GD** suffers linear $\Phi_{\text{Beam}}^{\mathcal{X}}(\delta)$ -regret (proof deferred to [Appendix G](#)).

Theorem 7. *For any $\delta, \eta < \frac{1}{2}$ and $T \geq 1$, there exists a sequence of linear loss functions $\{f^t : \mathcal{X} \subseteq [0, 1]^2 \rightarrow \mathbb{R}\}_{t \in [T]}$ such that **GD** with step size η suffers $\Omega(\delta T) \Phi_{\text{Beam}}^{\mathcal{X}}(\delta)$ -regret.*

Our results show that even for simple local strategy modification sets $\Phi(\delta)$, the landscape of efficient local $\Phi(\delta)$ -regret minimization is already quite rich, and many basic and interesting questions remain open.

5 Discussion and Future Directions

In this paper, we initiate the study of tractable Φ -equilibria in non-concave games and examine several natural families of strategy modifications. For any Φ that contains only a finite number of strategy modifications, we design an efficient randomized Φ -regret minimization algorithm, which provides efficient uncoupled learning dynamics that converge to the corresponding Φ -equilibria. Additionally, we study several classes of $\Phi(\delta)$ that contain an infinite number of δ -local strategy modifications and show efficient uncoupled learning dynamics that converge to an ε -approximate $\Phi(\delta)$ -equilibrium in the first-order stationary regime, where $\varepsilon = \Omega(\delta^2)$. We justify our focus on the first-order stationary regime by proving an NP-hardness result for achieving an approximation error $\varepsilon = o(\delta^2)$, even when the set of strategy modifications is simple, such as $\Phi(\delta) = \Phi_{\text{Int}^+}(\delta)$, which includes only Φ_{Int} and one additional strategy modification. These results are presented in [Appendix H](#).

Below, we discuss several future directions on Φ -regret minimization and Φ -equilibria computation.

Other efficiently minimizable Φ -regret. In this paper, we propose several natural sets of Φ 's that admit efficient Φ -regret minimization. It will be interesting to investigate for which other strategy modifications Φ , the corresponding Φ -regret can be minimized efficiently. One natural candidate is $\Phi_{\text{Beam}}^{\mathcal{X}}(\delta)$.

Improved $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret in games. We show in [Theorem 4](#) that the optimistic gradient (**OG**) dynamics guarantees that an improved individual $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of $O(T^{1/4})$. Could we design uncoupled learning

dynamics with better individual regret guarantees, consequently leading to faster convergence to an approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium?

References

- Jacob Abernethy, Kevin A Lai, and Andre Wibisono. Last-iterate convergence rates for min-max optimization: Convergence of hamiltonian gradient descent and consensus optimization. In *Algorithmic Learning Theory*, pages 3–47. PMLR, 2021.
- Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization oracle. In *Conference on Learning Theory*, pages 18–29. PMLR, 2019.
- Ioannis Anagnostides, Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Noah Golowich, and Tuomas Sandholm. Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 2022a.
- Ioannis Anagnostides, Gabriele Farina, Christian Kroer, Chung-Wei Lee, Haipeng Luo, and Tuomas Sandholm. Uncoupled learning dynamics with $o(\log t)$ swap regret in multiplayer games. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022b.
- Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Near-optimal phi-regret learning in extensive-form games. In *International Conference on Machine Learning (ICML)*, 2023.
- Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, pages 265–290, 1954.
- Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.
- Sergul Aydore, Tianhao Zhu, and Dean P Foster. Dynamic local regret for non-convex online forecasting. *Advances in neural information processing systems*, 32, 2019.
- Yu Bai, Chi Jin, Song Mei, Ziang Song, and Tiancheng Yu. Efficient phi-regret minimization in extensive-form games via online mirror descent. *Advances in Neural Information Processing Systems*, 35:22313–22325, 2022.
- Martino Bernasconi, Matteo Castiglioni, Alberto Marchesi, Francesco Trovò, and Nicola Gatti. Constrained phi-equilibria. In *International Conference on Machine Learning*, 2023.
- David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, January 1956. ISSN 0030-8730. URL <https://projecteuclid.org/journals/pacific-journal-of-mathematics/volume-6/issue-1/An-analog-of-the-minimax-theorem-for-vector-payoffs/pjm/1103044235.full>. Publisher: Pacific Journal of Mathematics, A Non-profit Corporation.
- George W Brown. Iterative solution of games by fictitious play. *Act. Anal. Prod Allocation*, 13(1):374, 1951.

- Sébastien Bubeck et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.
- Yang Cai and Weiqiang Zheng. Accelerated single-call methods for constrained min-max optimization. *International Conference on Learning Representations (ICLR)*, 2023.
- Robert Cauty. Solution du problème de point fixe de Schauder. *Fundamenta Mathematicae*, 170:231–246, 2001. ISSN 0016-2736, 1730-6329. doi: 10.4064/fm170-3-2. URL <https://www.impan.pl/en/publishing-house/journals-and-series/fundamenta-mathematicae/all/170/3/88631/solution-du-probleme-de-point-fixe-de-schauder>. Publisher: Instytut Matematyczny Polskiej Akademii Nauk.
- Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- Xi Chen and Binghui Peng. Hedging in games: Faster convergence of external and swap regrets. *Advances in Neural Information Processing Systems (NeurIPS)*, 33:18990–18999, 2020.
- Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the ACM (JACM)*, 56(3):1–57, 2009.
- Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From External to Swap Regret 2.0: An Efficient Reduction and Oblivious Adversary for Large Action Spaces, December 2023. URL <http://arxiv.org/abs/2310.19786>. arXiv:2310.19786 [cs].
- Constantinos Daskalakis. Non-concave games: A challenge for game theory’s next 100 years. *Cowles Preprints*, 2022.
- Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. In *the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2018.
- Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.
- Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. Near-optimal no-regret learning in general games. *Advances in Neural Information Processing Systems (NeurIPS)*, 2021a.
- Constantinos Daskalakis, Stratis Skoulakis, and Manolis Zampetakis. The complexity of constrained min-max optimization. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 2021b.
- Constantinos Daskalakis, Noah Golowich, Stratis Skoulakis, and Emmanouil Zampetakis. Stay-on-the-ridge: Guaranteed convergence to local minimax equilibrium in nonconvex-nonconcave games. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 5146–5198. PMLR, 2023.
- Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893, 1952.

- Jelena Diakonikolas, Constantinos Daskalakis, and Michael Jordan. Efficient methods for structured nonconvex-nonconcave min-max optimization. *International Conference on Artificial Intelligence and Statistics*, 2021.
- Ky Fan. Minimax theorems. *Proceedings of the National Academy of Sciences*, 39(1):42–47, 1953.
- Gabriele Farina, Ioannis Anagnostides, Haipeng Luo, Chung-Wei Lee, Christian Kroer, and Tuomas Sandholm. Near-optimal no-regret learning dynamics for general convex games. *Advances in Neural Information Processing Systems*, 35:39076–39089, 2022a.
- Gabriele Farina, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Simple uncoupled no-regret learning dynamics for extensive-form correlated equilibrium. *Journal of the ACM*, 69(6), 2022b.
- Tanner Fiez and Lillian J Ratliff. Local convergence analysis of gradient descent ascent with finite timescale separation. In *Proceedings of the International Conference on Learning Representation*, 2021.
- Tanner Fiez, Benjamin Chasnov, and Lillian Ratliff. Implicit learning dynamics in stackelberg games: Equilibria characterization, convergence analysis, and empirical study. In *International Conference on Machine Learning*, pages 3133–3144. PMLR, 2020.
- Yoav Freund and Robert E. Schapire. Adaptive Game Playing Using Multiplicative Weights. *Games and Economic Behavior*, 29:79–103, 1999.
- Xiand Gao, Xiaobo Li, and Shuzhong Zhang. Online learning with non-convex losses and non-stationary regret. In *International Conference on Artificial Intelligence and Statistics*, pages 235–243. PMLR, 2018.
- George B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.
- Irving L Glicksberg. A further generalization of the kakutani fixed theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.
- Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *Proceedings of the 25th international conference on Machine learning*, pages 360–367, 2008.
- Amy Greenwald and Amir Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003. Proceedings*, pages 2–12. Springer, 2003.
- Ziwei Guan, Yi Zhou, and Yingbin Liang. Online nonconvex optimization with limited instantaneous oracle feedback. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 3328–3355. PMLR, 2023.
- Nadav Hallak, Panayotis Mertikopoulos, and Volkan Cevher. Regret minimization in stochastic non-convex learning via a proximal-gradient approach. In *International Conference on Machine Learning*, pages 4008–4017. PMLR, 2021.
- James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3: 97–139, 1957.

- Elad Hazan and Satyen Kale. Computational Equivalence of Fixed Points and No Regret Algorithms, and Convergence to Equilibria. In *Advances in Neural Information Processing Systems*, volume 20. Curran Associates, Inc., 2007. URL <https://proceedings.neurips.cc/paper/2007/hash/e4bb4c5173c2ce17fd8fcd40041c068f-Abstract.html>.
- Elad Hazan, Karan Singh, and Cyril Zhang. Efficient regret minimization in non-convex games. In *International Conference on Machine Learning*, pages 1433–1441. PMLR, 2017.
- Amélie Héliou, Matthieu Martin, Panayotis Mertikopoulos, and Thibaud Rahier. Online non-convex optimization with imperfect feedback. *Advances in Neural Information Processing Systems*, 33:17224–17235, 2020.
- Ya-Ping Hsieh, Panayotis Mertikopoulos, and Volkan Cevher. The limits of min-max optimization algorithms: Convergence to spurious non-critical sets. In *International Conference on Machine Learning*, pages 4337–4348. PMLR, 2021.
- Chi Jin, Praneeth Netrapalli, and Michael Jordan. What is local optimality in nonconvex-nonconcave min-max optimization? In *International conference on machine learning (ICML)*, pages 4880–4889. PMLR, 2020.
- Samuel Karlin. *Mathematical methods and theory in games, programming, and economics: Volume II: the theory of infinite games*, volume 2. Addison-Wesley, 1959.
- Samuel Karlin. *Mathematical Methods and Theory in Games, Programming, and Economics: Volume 2: The Theory of Infinite Games*. Elsevier, 2014.
- Walid Krichene, Maximilian Balandat, Claire Tomlin, and Alexandre Bayen. The hedge algorithm on a continuum. In *International Conference on Machine Learning*, pages 824–832. PMLR, 2015.
- Odalric-Ambrym Maillard and Rémi Munos. Online learning in adversarial lipschitz environments. In *Joint european conference on machine learning and knowledge discovery in databases*, pages 305–320. Springer, 2010.
- Oren Mangoubi and Nisheeth K Vishnoi. Greedy adversarial equilibrium: an efficient alternative to nonconvex-nonconcave min-max optimization. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 896–909, 2021.
- Eric Mazumdar, Lillian J Ratliff, and S Shankar Sastry. On gradient-based learning in continuous games. *SIAM Journal on Mathematics of Data Science*, 2(1):103–131, 2020.
- Lionel McKenzie. On equilibrium in graham’s model of world trade and other competitive systems. *Econometrica*, pages 147–161, 1954.
- Panayotis Mertikopoulos and Zhengyuan Zhou. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173:465–507, 2019.
- Dustin Morrill, Ryan D’Orazio, Reza Sarfati, Marc Lanctot, James R Wright, Amy R Greenwald, and Michael Bowling. Hindsight and sequential rationality of correlated play. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021a.

- Dustin Morrill, Ryan D’Orazio, Marc Lanctot, James R Wright, Michael Bowling, and Amy R Greenwald. Efficient deviation types and learning for hindsight rationality in extensive-form games. In *International Conference on Machine Learning*, pages 7818–7828. PMLR, 2021b.
- T. S. Motzkin and E. G. Straus. Maxima for Graphs and a New Proof of a Theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965. ISSN 0008-414X, 1496-4279. doi: 10.4153/CJM-1965-053-6. URL https://www.cambridge.org/core/product/identifier/S0008414X00039493/type/journal_article.
- Katta G. Murty and Santosh N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, June 1987. ISSN 1436-4646. doi: 10.1007/BF02592948. URL <https://doi.org/10.1007/BF02592948>.
- John F Nash Jr. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.
- Binghui Peng and Aviad Rubinstein. Fast swap regret minimization and applications to approximate correlated equilibria, November 2023. URL <http://arxiv.org/abs/2310.19647>. arXiv:2310.19647 [cs].
- Thomas Pethick, Puya Latafat, Panagiotis Patrinos, Olivier Fercoq, and Volkan Cevherå. Escaping limit cycles: Global convergence for constrained nonconvex-nonconcave minimax problems. In *International Conference on Learning Representations (ICLR)*, 2022.
- Georgios Piliouras, Mark Rowland, Shayegan Omidshafiei, Romuald Elie, Daniel Hennes, Jerome Connor, and Karl Tuyls. Evolutionary dynamics and phi-regret minimization in games. *Journal of Artificial Intelligence Research*, 74:1125–1158, 2022.
- Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Beyond regret. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 559–594. JMLR Workshop and Conference Proceedings, 2011.
- Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. *Advances in Neural Information Processing Systems*, 2013.
- Lillian J Ratliff, Samuel A Burden, and S Shankar Sastry. On the characterization of local nash equilibria in continuous games. *IEEE transactions on automatic control*, 61(8):2301–2307, 2016.
- Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pages 296–301, 1951.
- J Ben Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, pages 520–534, 1965.
- Dravyansh Sharma. No internal regret with non-convex loss functions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.
- Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8(4):171–176, 1958.
- Ziang Song, Song Mei, and Yu Bai. Sample-efficient learning of correlated equilibria in extensive-form games. *Advances in Neural Information Processing Systems*, 35:4099–4110, 2022.

- Gilles Stoltz and Gábor Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.
- Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. In *Algorithmic Learning Theory*, pages 845–861. PMLR, 2020.
- Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. Fast convergence of regularized learning in games. *Advances in Neural Information Processing Systems (NeurIPS)*, 2015.
- J v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.
- Bernhard Von Stengel and Françoise Forges. Extensive-form correlated equilibrium: Definition and computational complexity. *Mathematics of Operations Research*, 33(4):1002–1022, 2008.
- Yuanhao Wang, Guodong Zhang, and Jimmy Ba. On solving minimax optimization locally: A follow-the-ridge approach. In *International Conference on Learning Representations (ICLR)*, 2020.
- Brian Hu Zhang, Ioannis Anagnostides, Gabriele Farina, and Tuomas Sandholm. Efficient ϕ -regret minimization with low-degree swap deviations in extensive-form games, 2024.
- Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (ICML)*, 2003.
- David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 681–690, 2006.

Contents

1	Introduction	2
1.1	Contributions	3
2	Preliminaries	6
3	Tractable Φ-Equilibrium for Finite Φ via Sampling	8
3.1	Proof of Theorem 2	10
4	Approximate Φ-Equilibria under Infinite Local Strategy Modifications	12
4.1	Projection-Based Local Strategy Modifications	13
4.1.1	$\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret Minimization in the Adversarial Setting	13
4.1.2	Improved $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret in the Game Setting	15
4.2	Convex Combination of Finite Local Strategy Modifications	15
4.3	Interpolation-Based Local Strategy Modifications	16
4.4	Beam-Search Local Strategy Modifications and Local Equilibria	16
5	Discussion and Future Directions	17
A	Related Work	24
B	Additional Preliminaries: Solution Concepts in Non-Concave Games	25
C	Proof of Lemma 1	27
D	Missing Details in Section 4.1	28
D.1	Differences between External Regret and $\Phi_{\text{Proj}}^{\mathcal{X}}$ -regret	28
D.2	Lower bounds for $\Phi_{\text{Proj}}^{\mathcal{X}}$ -Regret	29
D.2.1	Proof of Theorem 8	29
D.2.2	Proof of Proposition 1	30
D.3	Improved $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret in the Game Setting and Proof of Theorem 4	31
D.3.1	Proof of Theorem 4	32
E	Missing Details in Section 4.2	32
E.1	Proof of Theorem 5	32
F	Missing details in Section 4.3	34
G	Proof of Theorem 7	36
H	First-Order Stationary Regime	37
H.1	Hardness when $\Phi(\delta)$ contains all local strategy modifications	37
H.2	Restricted Deviations	41
I	Hardness for Approximate $\Phi_{\text{Proj}}(\delta)/\Phi_{\text{Int}}(\delta)$-Equilibrium when $\delta = D$	44
I.1	Proof of Theorem 14	45
I.2	Proof of Corollary 3	46

J	Removing the D dependence for $\Phi_{\text{Proj}}^{\mathcal{X}}$-regret	46
J.1	One-Dimensional Case	46
J.2	d -Dimensional Box Case	47

A Related Work

Non-Concave Games. An important special case of multi-player games are two-player zero-sum games, which are defined in terms of some function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ that one of the two players say the one choosing $x \in \mathcal{X}$, wants to minimize, while the other player, the one choosing $y \in \mathcal{Y}$, wants to maximize. Finding Nash equilibrium in such games is tractable in the *convex-concave* setting, i.e. when $f(x, y)$ is convex with respect to the minimizing player’s strategy, x , and concave with respect to the maximizing player’s strategy, y , but it is computationally intractable in the general *nonconvex-nonconcave* setting. Namely, a Nash equilibrium may not exist, and it is NP-hard to determine if one exists and, if so, find it. Moreover, in this case, stable limit points of gradient-based dynamics are not necessarily Nash equilibria, not even local Nash equilibria [Daskalakis and Panageas, 2018, Mazumdar et al., 2020]. Moreover, there are examples including the “Polar Game” [Pethick et al., 2022] and the “Forsaken Matching Pennies” [Hsieh et al., 2021] showing that for GD / OG and many other no-regret learning algorithms in nonconvex-nonconcave min-max optimization, the last-iterate does not converge and even the average-iterate fails to be a stationary point. We emphasize that the convergence guarantees we provide for GD / OG in Section 4.1 and Section 4.3 holds for the empirical distribution of play, not the average-iterate or the last-iterate.

A line of work focuses on computing Nash equilibrium under additional structure in the game. This encompasses settings where the game satisfies the (weak) Minty variational inequality [Mertikopoulos and Zhou, 2019, Diakonikolas et al., 2021, Pethick et al., 2022, Cai and Zheng, 2023], or is sufficiently close to being bilinear [Abernethy et al., 2021]. However, the study of universal solution concepts in the nonconvex-nonconcave setting is sparse. Daskalakis et al. [2021b] proved the existence and computational hardness of local Nash equilibrium. In a more recent work, [Daskalakis et al., 2023] proposes second-order algorithms with asymptotic convergence to local Nash equilibrium. Several works study sequential two-player zero-sum games with additional assumptions about the player who goes second. They propose equilibrium concepts such as *local minimax points* [Jin et al., 2020], *differentiable Stackelberg equilibrium* [Fiez et al., 2020], and *greedy adversarial equilibrium* [Mangoubi and Vishnoi, 2021]. Notably, local minimax points are stable limit points of Gradient-Descent-Ascent (GDA) dynamics [Jin et al., 2020, Wang et al., 2020, Fiez and Ratliff, 2021] while greedy adversarial equilibrium can be computed efficiently using second-order algorithms in the unconstrained setting [Mangoubi and Vishnoi, 2021]. In contrast to these studies, we focus on the more general case of multi-player non-concave games.

Local Equilibrium. To address the limitations associated with classical, global equilibrium concepts, a natural approach is to focus on developing equilibrium concepts that guarantee local stability instead. One definition of interest is the strict local Nash equilibrium, wherein each player’s strategy corresponds to a local maximum of their utility function, given the other players’ strategies. Unfortunately, a strict local Nash equilibrium may not always exist, as demonstrated in Example 1. Furthermore, a weaker notion—the second-order local Nash equilibrium, where each player has no incentive to deviate based on the second-order Taylor expansion estimate of their utility, is also not guaranteed to exist as illustrated in Example 1. What’s more, it is NP-hard to check whether a given strategy profile is a strict local Nash equilibrium or a second-order

local Nash equilibrium, as implied by the result of Murty and Kabadi [1987].⁶ Finally, one can consider *local Nash equilibrium*, a first order stationary solution, which is guaranteed to exist [Daskalakis et al., 2021b]. Unlike non-convex optimization, where targeting first-order local optima sidesteps the intractability of global optima, this first-order local Nash equilibrium has been recently shown to be intractable, even in two-player zero-sum non-concave games with joint feasibility constraints [Daskalakis et al., 2021b].⁷ See Table 1 for a summary of solution concepts in non-concave games.

Online Learning with Non-Convex Losses. A line of work has studied online learning against non-convex losses. To circumvent the computational intractability of this problem, various approaches have been pursued: some works assume a restricted set of non-convex loss functions [Gao et al., 2018], while others assume access to a sampling oracle [Maillard and Munos, 2010, Krichene et al., 2015] or access to an offline optimization oracle [Agarwal et al., 2019, Suggala and Netrapalli, 2020, Héliou et al., 2020] or a weaker notion of regret [Hazan et al., 2017, Aydore et al., 2019, Hallak et al., 2021, Guan et al., 2023]. The work most closely related to ours is [Hazan et al., 2017]. The authors propose a notion of *w-smoothed local regret* against non-convex losses, and they also define a local equilibrium concept for non-concave games. They use the idea of *smoothing* to average the loss functions in the previous w iterations and design algorithms with optimal w -smoothed local regret. The concept of regret they introduce suggests a local equilibrium concept. However, their local equilibrium concept is non-standard in that its local stability is not with respect to a distribution over strategy profiles sampled by this equilibrium concept. Moreover, the path to attaining this local equilibrium through decentralized learning dynamics remains unclear. The algorithms provided in [Hazan et al., 2017, Guan et al., 2023] require that every agent i experiences (over several rounds) the average utility function of the previous w iterates, denoted as $F_{i,w}^t := \frac{1}{w} \sum_{\ell=0}^{w-1} u_i^{t-\ell}(\cdot, x_{-i}^{t-\ell})$. Implementing this imposes a significant coordination burden on the agents. In contrast, we focus on a natural concept of $\Phi(\delta)$ -equilibrium, which is incomparable to that of Hazan et al. [2017], and we also show that efficient convergence to this concept is achieved via decentralized gradient-based learning dynamics.

Φ -regret and Φ -equilibrium. The concept of Φ -regret and the associated Φ -equilibrium is introduced by Greenwald and Jafari [2003] and has been broadly investigated in the context of concave games [Greenwald and Jafari, 2003, Stoltz and Lugosi, 2007, Gordon et al., 2008, Rakhlin et al., 2011, Piliouras et al., 2022, Bernasconi et al., 2023] and extensive-form games [Von Stengel and Forges, 2008, Morrill et al., 2021a,b, Farina et al., 2022b, Bai et al., 2022, Song et al., 2022, Anagnostides et al., 2023, Zhang et al., 2024]. The work of Sharma [2024] studies internal regret minimization against non-convex losses. To our knowledge, no efficient algorithm exists for the classes of Φ -equilibria we consider for non-concave games.

B Additional Preliminaries: Solution Concepts in Non-Concave Games

We present definitions of several solution concepts in the literature as well as the existence and computational complexity of each solution concept.

⁶Murty and Kabadi [1987] shows that checking whether a point is a local maximum of a multi-variate quadratic function is NP-hard.

⁷In general sum games, it is not hard to see that the intractability results [Daskalakis et al., 2009, Chen et al., 2009] for computing *global* Nash equilibria in bimatrix games imply intractability for computing *local* Nash equilibria.

Definition 4 (Nash Equilibrium). *In a continuous game, a strategy profile $x \in \prod_{j=1}^n \mathcal{X}_j$ is a Nash equilibrium (NE) if and only if for every player $i \in [n]$,*

$$u_i(x'_i, x_{-i}) \leq u_i(x), \forall x'_i \in \mathcal{X}_i$$

Definition 5 (Mixed Nash Equilibrium). *In a continuous game, a mixed strategy profile $p \in \prod_{j=1}^n \Delta(\mathcal{X}_j)$ (here we denote $\Delta(\mathcal{X}_i)$ as the set of probability measures over \mathcal{X}_i) is a mixed Nash equilibrium (MNE) if and only if for every player $i \in [n]$,*

$$u_i(p'_i, p_{-i}) \leq u_i(p), \forall p'_i \in \Delta(\mathcal{X}_i)$$

Definition 6 ((Coarse) Correlated Equilibrium). *In a continuous game, a distribution σ over joint strategy profiles $\prod_{i=1}^n \mathcal{X}_i$ is a correlated equilibrium (CE) if and only if for all player $i \in [n]$,*

$$\max_{\phi_i: \mathcal{X}_i \rightarrow \mathcal{X}_i} \mathbb{E}_{x \sim \sigma} [u_i(\phi_i(x_i), x_{-i})] \leq \mathbb{E}_{x \sim \sigma} [u_i(x)].$$

Similarly, a distribution σ over joint strategy profiles $\prod_{i=1}^n \mathcal{X}_i$ is a coarse correlated equilibrium (CCE) if and only if for all player $i \in [n]$,

$$\max_{x'_i \in \mathcal{X}_i} \mathbb{E}_{x \sim \sigma} [u_i(x'_i, x_{-i})] \leq \mathbb{E}_{x \sim \sigma} [u_i(x)].$$

Definition 7 (Strict Local Nash Equilibrium). *In a continuous game, a strategy profile $x \in \prod_{j=1}^n \mathcal{X}_j$ is a strict local Nash equilibrium if and only if for every player $i \in [n]$, there exists $\delta > 0$ such that*

$$u_i(x'_i, x_{-i}) \leq u_i(x), \forall x'_i \in B_{d_i}(x_i, \delta) \cap \mathcal{X}_i.$$

Definition 8 (Second-order Local Nash Equilibrium). *Consider a continuous game where each utility function $u_i(x_i, x_{-i})$ is twice-differentiable with respect to x_i for any fixed x_{-i} . A strategy profile $x \in \prod_{j=1}^n \mathcal{X}_j$ is a second-order local Nash equilibrium if and only if for every player $i \in [n]$, x_i maximizes the second-order Taylor expansion of its utility functions at x_i , or formally,*

$$\langle \nabla_{x_i} u_i(x), x'_i - x_i \rangle + (x'_i - x_i)^\top \nabla_{x_i}^2 u_i(x) (x'_i - x_i) \leq 0, \forall x'_i \in \mathcal{X}_i.$$

Existence Mixed Nash equilibria exist in continuous games, thus smooth games [Debreu, 1952, Glicksberg, 1952, Fan, 1953]. By definition, an MNE is also a CE and a CCE. This also proves the existence of CE and CCE. In contrast, strict local Nash equilibria, second-order Nash equilibria, or (pure) Nash equilibria may not exist in a smooth non-concave game, as we show in the following example.

Example 1. *Consider a two-player zero-sum non-concave game: the action sets are $\mathcal{X}_1 = \mathcal{X}_2 = [-1, 1]$ and the utility functions are $u_1(x_1, x_2) = -u_2(x_1, x_2) = (x_1 - x_2)^2$. Let $x = (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ be any strategy profile: if $x_1 = x_2$, then player 1 is not at a local maximum; if $x_1 \neq x_2$, then player 2 is not at a local maximum. Thus x is not a strict local Nash equilibrium. Since the utility function is quadratic, we conclude that the game also has no second-order local Nash equilibrium.*

Computational Complexity Consider a single-player smooth non-concave game with a quadratic utility function $f : \mathcal{X} \rightarrow \mathbb{R}$. The problem of finding a local maximum of f can be reduced to the problem of computing a NE, a MNE, a CE, a CCE, a strict local Nash equilibrium, or a second-order local Nash equilibrium. Since computing a local maximum or checking if a given point is a local maximum is NP-hard [Murty and Kabadı, 1987], we know that the computational complexities of NE, MNE, CE, CCE, strict local Nash equilibria, and second-order local Nash equilibria are all NP-hard.

Representation Complexity Karlin [1959] present a two-player zero-sum non-concave game whose unique MNE has infinite support. Since in a two-player zero-sum game, the marginal distribution of a CE or a CCE is an MNE, it also implies that the representation complexity of any CE or CCE is infinite. We present the example in Karlin [1959] here for completeness and also prove that the game is Lipschitz and smooth.

Example 2 ([Karlin, 1959, Chapter 7.1, Example 3]). *We consider a two-player zero-sum game with action sets $\mathcal{X}_1 = \mathcal{X}_2 = [0, 1]$. Let p and q be two distributions over $[0, 1]$. The only requirement for p and q is that their cumulative distribution functions are not finite-step functions. For example, we can take $p = q$ to be the uniform distribution.*

Let μ_n and ν_n denote the n -th moments of p and q , respectively. Define the utility function

$$u(x, y) = u_1(x, y) = -u_2(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^n} (x^n - \mu_n)(y^n - \nu_n), \quad 0 \leq x, y \leq 1.$$

Claim 2. *The game in Example 2 is 2-Lipschitz and 6-smooth, and (p, q) is its unique (mixed) Nash equilibrium.*

Proof. Fix any $y \in [0, 1]$, since $|\frac{1}{2^n}(y^n - \nu_n)nx^{n-1}| \leq \frac{n}{2^n}$, the series of $\nabla_x u(x, y)$ is uniformly convergent. We have $|\nabla_x u(x, y)| \leq \sum_{n=0}^{\infty} \frac{n}{2^n} \leq 2$, $y \in [0, 1]$. Similarly, we have $|\nabla_x^2 u(x, y)| \leq \sum_{n=0}^{\infty} \frac{n^2}{2^n} \leq 6$ for all $y \in [0, 1]$. By symmetry, we also have $|\nabla_y u(x, y)| \leq 2$ and $|\nabla_y^2 u(x, y)| \leq 6$ for all $x, y \in [0, 1]$. Thus, the game is 2-Lipschitz and 6-smooth.

Since $|\frac{1}{2^n}(x^n - \mu_n)(y^n - \nu_n)| \leq \frac{1}{2^n}$, the series of $u(x, y)$ is absolutely and uniformly convergent. We have

$$\begin{aligned} \int_0^1 u(x, y) dF_p(x) &= \sum_{n=0}^{\infty} \frac{1}{2^n} (y^n - \nu_n) \int_0^1 (x^n - \mu_n) dF_p(x) \equiv 0, \\ \int_0^1 u(x, y) F_q(y) &= \sum_{n=0}^{\infty} \frac{1}{2^n} (x^n - \mu_n) \int_0^1 (y^n - \nu_n) dF_q(y) \equiv 0. \end{aligned}$$

In particular, (p, q) is a mixed Nash equilibrium, and the value of the game is 0. Suppose (p', q') is also a mixed Nash equilibrium. Then (p, q') is a mixed Nash equilibrium. Note that p supports on every point in $[0, 1]$. As a consequence, we have

$$0 \equiv \int_0^1 u(x, y) dF_{q'}(y) = \sum_{n=0}^{\infty} \frac{1}{2^n} (x^n - \mu_n)(\nu'_n - \nu_n)$$

for all $x \in [0, 1]$, where ν'_n is the n -th moment of q' . Since the series vanished identically, the coefficients of each power of x must vanish. Thus $\nu'_n = \nu_n$ and $q' = q$. Similarly, we have $p' = p$, and the mixed Nash equilibrium is unique. \square

C Proof of Lemma 1

Let $\{f^t\}_{t \in [T]}$ be a sequence of non-convex L -smooth loss functions satisfying Assumption 1. Let $\{x^t\}_{t \in [T]}$ be the iterates produced by \mathcal{A} against $\{f^t\}_{t \in [T]}$. Then $\{x^t\}_{t \in [T]}$ is also the iterates produced by \mathcal{A} against a sequence of linear loss functions $\{\langle \nabla f^t(x^t), \cdot \rangle\}$. For the latter, we know

$$\max_{\phi \in \Phi^{\mathcal{X}}(\delta)} \sum_{t=1}^T \langle \nabla f^t(x^t), x^t - \phi(x^t) \rangle \leq \text{Reg}_{\Phi^{\mathcal{X}}(\delta)}^T.$$

Then using L -smoothness of $\{f^t\}$ and the fact that $\|\phi(x) - x\| \leq \delta$ for all $\phi \in \Phi(\delta)$, we get

$$\begin{aligned} \max_{\phi \in \Phi^{\mathcal{X}}(\delta)} \sum_{t=1}^T f^t(x^t) - f^t(\phi(x^t)) &\leq \max_{\phi \in \Phi^{\mathcal{X}}(\delta)} \sum_{t=1}^T \left(\langle \nabla f^t(x^t), x^t - \phi(x^t) \rangle + \frac{L}{2} \|x^t - \phi(x^t)\|^2 \right) \\ &\leq \text{Reg}_{\Phi^{\mathcal{X}}(\delta)}^T + \frac{\delta^2 LT}{2}. \end{aligned}$$

This completes the proof of the first part.

Let each player $i \in [n]$ employ algorithm \mathcal{A} in a smooth game independently and produce iterates $\{x^t\}$. The averaged joint strategy profile σ^T that chooses x^t uniformly at random from $t \in [T]$ satisfies for any player $i \in [n]$,

$$\begin{aligned} &\max_{\phi \in \Phi^{\mathcal{X}_i}(\delta)} \mathbb{E}_{x \sim \sigma} [u_i(\phi(x_i), x_{-i})] - \mathbb{E}_{x \sim \sigma} [u_i(x)] \\ &= \max_{\phi \in \Phi^{\mathcal{X}_i}(\delta)} \frac{1}{T} \sum_{t=1}^T (u_i(\phi(x_i^t), x_{-i}^t) - u_i(x^t)) \\ &\leq \frac{\text{Reg}_{\Phi^{\mathcal{X}_i}(\delta)}^T}{T} + \frac{\delta^2 L}{2}. \end{aligned}$$

Thus σ^T is a $(\max_{i \in [n]} \{\text{Reg}_{\Phi^{\mathcal{X}_i}(\delta)}^T\} \cdot T^{-1} + \frac{\delta^2 L}{2})$ -approximate $\Phi(\delta)$ -equilibrium. This completes the proof of the second part.

D Missing Details in Section 4.1

D.1 Differences between External Regret and $\Phi_{\text{Proj}}^{\mathcal{X}}$ -regret

The following two examples show that $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is incomparable with external regret for convex loss functions. A sequence of actions may suffer high Reg^T but low $\text{Reg}_{\text{Proj}, \delta}^T$ (Example 3), and vice versa (Example 4).

Example 3. Let $f^1(x) = f^2(x) = |x|$ for $x \in \mathcal{X} = [-1, 1]$. Then the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of the sequence $\{x^1 = \frac{1}{2}, x^2 = -\frac{1}{2}\}$ for any $\delta \in (0, \frac{1}{2})$ is 0. However, the external regret of the same sequence is 1. By repeating the construction for $\frac{T}{2}$ times, we conclude that there exists a sequence of actions with $\text{Reg}_{\text{Proj}, \delta}^T = 0$ and $\text{Reg}^T = \frac{T}{2}$ for all $T \geq 2$.

Example 4. Let $f^1(x) = -2x$ and $f^2(x) = x$ for $x \in \mathcal{X} = [-1, 1]$. Then the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of the sequence $\{x^1 = \frac{1}{2}, x^2 = 0\}$ for any $\delta \in (0, \frac{1}{2})$ is δ . However, the external regret of the same sequence is 0. By repeating the construction for $\frac{T}{2}$ times, we conclude that there exists a sequence of actions with $\text{Reg}_{\text{Proj}, \delta}^T = \frac{\delta T}{2}$ and $\text{Reg}^T = 0$ for all $T \geq 2$.

At a high level, the external regret competes against a fixed action, whereas $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is more akin to the notion of *dynamic regret*, competing with a sequence of varying actions. When the environment is stationary, i.e., $f^t = f$ (Example 3), a sequence of actions that are far away from the global minimum must suffer high regret, but may produce low $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret since the change to the cumulative loss caused by a fixed-direction deviation could be neutralized across different actions in the sequence. In contrast, in

a non-stationary (dynamic) environment (Example 4), every fixed action performs poorly, and a sequence of actions could suffer low regret against a fixed action but the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret that competes with a fixed-direction deviation could be large. Nevertheless, despite these differences between the two notions of regret as shown above, they are *compatible* for convex loss functions: our main results in this section provide algorithms that minimize external regret and $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret simultaneously.

D.2 Lower bounds for $\Phi_{\text{Proj}}^{\mathcal{X}}$ -Regret

Theorem 8 (Lower bound for $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret against convex losses). *For any $T \geq 1$, $D_{\mathcal{X}} > 0$, $0 < \delta \leq D_{\mathcal{X}}$, and $G \geq 0$, there exists a distribution \mathcal{D} on G -Lipschitz linear loss functions f^1, \dots, f^T over $\mathcal{X} = [-D_{\mathcal{X}}, D_{\mathcal{X}}]$ such that for any online algorithm, its $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret on the loss sequence satisfies $\mathbb{E}_{\mathcal{D}}[\text{Reg}_{\text{Proj}, \delta}^T] = \Omega(\delta G \sqrt{T})$.*

Remark 2. *A keen reader may notice that the $\Omega(G\delta\sqrt{T})$ lower bound in Theorem 8 does not match the $O(G\sqrt{\delta D_{\mathcal{X}} T})$ upper bound in Theorem 3, especially when $D_{\mathcal{X}} \gg \delta$. A natural question is: which of them is tight? We conjecture that the lower bound is tight. In fact, for the special case where the feasible set \mathcal{X} is a box, we obtain a $D_{\mathcal{X}}$ -independent bound $O(d^{\frac{1}{4}} G \delta \sqrt{T})$ using a modified version of **GD**, which is tight when $d = 1$. See [Appendix J](#) for a detailed discussion.*

This lower bound suggests that **GD** achieves near-optimal $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret for convex losses. For L -smooth *non-convex* loss functions, we provide another $\Omega(\delta^2 L T)$ lower bound for algorithms that satisfy the linear span assumption. The *linear span* assumption states that the algorithm produces $x^{t+1} \in \{\Pi_{\mathcal{X}}[\sum_{i \in [t]} a_i \cdot x^i + b_i \cdot \nabla f^i(x^i)] : a_i, b_i \in \mathbb{R}, \forall i \in [t]\}$ as essentially the linear combination of the previous iterates and their gradients. Many online algorithms, such as online gradient descent and optimistic gradient, satisfy the linear span assumption. Combining with [Lemma 1](#), this lower bound suggests that **GD** attains nearly optimal $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret, even in the non-convex setting, among a natural family of gradient-based algorithms.

Proposition 1 (Lower bound for $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret against non-convex losses). *For any $T \geq 1$, $\delta \in (0, 1)$, and $L \geq 0$, there exists a sequence of L -Lipschitz and L -smooth non-convex loss functions f^1, \dots, f^T on $\mathcal{X} = [-1, 1]$ such that for any algorithm that satisfies the linear span assumption, its $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret on the loss sequence is $\text{Reg}_{\text{Proj}, \delta}^T \geq \frac{\delta^2 L T}{2}$.*

D.2.1 Proof of [Theorem 8](#)

Our proof technique comes from the standard one used in multi-armed bandits [[Auer et al., 2002](#), Theorem 5.1]. Suppose that $f^t(x) = g^t x$. We construct two possible environments. In the first environment, $g^t = G$ with probability $\frac{1+\varepsilon}{2}$ and $g^t = -G$ with probability $\frac{1-\varepsilon}{2}$; in the second environment, $g^t = G$ with probability $\frac{1-\varepsilon}{2}$ and $g^t = -G$ with probability $\frac{1+\varepsilon}{2}$. We use \mathbb{E}_i and \mathbb{P}_i to denote the expectation and probability measure under environment i , respectively, for $i = 1, 2$. Suppose that the true environment is uniformly chosen from one of these two environments. Below, we show that the expected regret of the learner is at least $\Omega(\delta G \sqrt{T})$.

Define $N_+ = \sum_{t=1}^T \mathbb{I}\{x^t \geq 0\}$ be the number of times x^t is non-negative, and define $f^{1:T} = (f^1, \dots, f^T)$.

Then we have

$$\begin{aligned}
|\mathbb{E}_1[N_+] - \mathbb{E}_2[N_+]| &= \left| \sum_{f^{1:T}} \left(\mathbb{P}_1(f^{1:T}) \mathbb{E}[N_+ | f^{1:T}] - \mathbb{P}_2(f^{1:T}) \mathbb{E}[N_+ | f^{1:T}] \right) \right| \\
&\quad \text{(enumerate all possible sequences of } f^{1:T}\text{)} \\
&\leq T \sum_{f^{1:T}} |\mathbb{P}_1(f^{1:T}) - \mathbb{P}_2(f^{1:T})| \\
&= T \|\mathbb{P}_1 - \mathbb{P}_2\|_{\text{TV}} \\
&\leq T \sqrt{(2 \ln 2) \text{KL}(\mathbb{P}_1, \mathbb{P}_2)} \quad \text{(Pinsker's inequality)} \\
&= T \sqrt{(2 \ln 2) T \cdot \text{KL} \left(\text{Bernoulli} \left(\frac{1+\varepsilon}{2} \right), \text{Bernoulli} \left(\frac{1-\varepsilon}{2} \right) \right)} \\
&= T \sqrt{(2 \ln 2) T \varepsilon \ln \frac{1+\varepsilon}{1-\varepsilon}} \leq T \sqrt{(4 \ln 2) T \varepsilon^2}. \tag{3}
\end{aligned}$$

In the first environment, we consider the regret with respect to $v = \delta$. Then we have

$$\begin{aligned}
\mathbb{E}_1 [\text{Reg}_{\text{Proj}, \delta}^T] &\geq \mathbb{E}_1 \left[\sum_{t=1}^T f^t(x^t) - f^t(\Pi_{\mathcal{X}}[x^t - \delta]) \right] = \mathbb{E}_1 \left[\sum_{t=1}^T g^t(x^t - \Pi_{\mathcal{X}}[x^t - \delta]) \right] \\
&= \mathbb{E}_1 \left[\sum_{t=1}^T \varepsilon G(x^t - \Pi_{\mathcal{X}}[x^t - \delta]) \right] \geq \varepsilon \delta G \mathbb{E}_1 \left[\sum_{t=1}^T \mathbb{I}\{x^t \geq 0\} \right] = \varepsilon \delta G \mathbb{E}_1 [N_+],
\end{aligned}$$

where in the last inequality, we use the fact that if $x^t \geq 0$, then $x^t - \Pi_{\mathcal{X}}[x^t - \delta] = x^t - (x^t - \delta) = \delta$ because $D \geq \delta$. In the second environment, we consider the regret with respect to $v = -\delta$. Then similarly, we have

$$\begin{aligned}
\mathbb{E}_2 [\text{Reg}_{\text{Proj}, \delta}^T] &\geq \mathbb{E}_2 \left[\sum_{t=1}^T f^t(x^t) - f^t(\Pi_{\mathcal{X}}[x^t + \delta]) \right] = \mathbb{E}_2 \left[\sum_{t=1}^T g^t(x^t - \Pi_{\mathcal{X}}[x^t + \delta]) \right] \\
&= \mathbb{E}_2 \left[\sum_{t=1}^T -\varepsilon G(x^t - \Pi_{\mathcal{X}}[x^t + \delta]) \right] \geq \varepsilon \delta G \mathbb{E}_2 \left[\sum_{t=1}^T \mathbb{I}\{x^t < 0\} \right] = \varepsilon \delta G (T - \mathbb{E}_2 [N_+]).
\end{aligned}$$

Summing up the two inequalities, we get

$$\begin{aligned}
\frac{1}{2} (\mathbb{E}_1 [\text{Reg}_{\text{Proj}, \delta}^T] + \mathbb{E}_2 [\text{Reg}_{\text{Proj}, \delta}^T]) &\geq \frac{1}{2} (\varepsilon \delta G T + \varepsilon \delta G (\mathbb{E}_1 [N_+] - \mathbb{E}_2 [N_+])) \\
&\geq \frac{1}{2} (\varepsilon \delta G T - \varepsilon \delta G T \varepsilon \sqrt{(4 \ln 2) T}). \quad \text{(by (3))}
\end{aligned}$$

Choosing $\varepsilon = \frac{1}{\sqrt{(16 \ln 2) T}}$, we can lower bound the last expression by $\Omega(\delta G \sqrt{T})$. The theorem is proven by noticing that $\frac{1}{2} (\mathbb{E}_1 [\text{Reg}_{\text{Proj}, \delta}^T] + \mathbb{E}_2 [\text{Reg}_{\text{Proj}, \delta}^T])$ is the expected regret of the learner.

D.2.2 Proof of Proposition 1

Proof. Consider $f : [-1, 1] \rightarrow \mathbb{R}$ such that $f(x) = -\frac{L}{2}x^2$ and let $f^t = f$ for all $t \in [T]$. Then any first-order methods that satisfy the linear span assumption with initial point $x^1 = 0$ will produce $x^t = 0$ for all $t \in [T]$. The $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret is thus $\sum_{t=1}^T (f(0) - f(\delta)) = \frac{\delta^2 L T}{2}$. \square

D.3 Improved $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -Regret in the Game Setting and Proof of Theorem 4

We first prove an $O(\sqrt{T})$ $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret upper bound for **OG** in the adversarial setting.

Theorem 9 (Adversarial Regret Bound for **OG**). *Let $\delta > 0$ and $T \in \mathbb{N}$. For convex and G -Lipschitz loss functions $\{f^t : \mathcal{X} \rightarrow \mathbb{R}\}_{t \in [T]}$, the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of (**OG**) with step size $\eta > 0$ is*

$$\text{Reg}_{\text{Proj}, \delta}^T \leq \frac{\delta D_{\mathcal{X}}}{\eta} + \eta \sum_{t=1}^T \|g^t - g^{t-1}\|^2.$$

Choosing step size $\eta = \frac{\sqrt{\delta D_{\mathcal{X}}}}{2G\sqrt{T}}$, we have $\text{Reg}_{\text{Proj}, \delta}^T \leq 4G\sqrt{\delta D_{\mathcal{X}}T}$.

In the analysis of **Theorem 9** for the adversarial setting, the term $\|g^t - g^{t-1}\|^2$ can be as large as $4G^2$. In the game setting where every player i employs **OG**, g_i^t , i.e., $-\nabla_{x_i} u_i(x)$, depends on other players' action x_{-i}^t . Note that the change of the players' actions $\|x^t - x^{t-1}\|^2$ is only $O(\eta^2)$. Such stability of the updates leads to an improved upper bound on $\|g_i^t - g_i^{t-1}\|^2$ and hence also an improved $O(T^{\frac{1}{4}})$ $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret for the player.

Proof. Fix any deviation v that is bounded by δ . Let us define $p^0 = w^0$ and $p^t = \Pi_{\mathcal{X}}[x^t - v]$. Following standard analysis of **OG** [Rakhlin and Sridharan, 2013], we have

$$\begin{aligned} \sum_{t=1}^T f^t(x^t) - f^t(p^t) &\leq \sum_{t=1}^T \langle \nabla f^t(x^t), x^t - p^t \rangle \\ &\leq \sum_{t=1}^T \frac{1}{2\eta} \left(\|w^{t-1} - p^t\|^2 - \|w^t - p^t\|^2 \right) + \eta \|g^t - g^{t-1}\|^2 - \frac{1}{2\eta} \left(\|x^t - w^t\|^2 + \|x^t - w^{t-1}\|^2 \right) \\ &\leq \sum_{t=1}^T \left(\frac{1}{2\eta} \|w^{t-1} - p^t\|^2 - \frac{1}{2\eta} \|w^{t-1} - p^{t-1}\|^2 + \eta \|g^t - g^{t-1}\|^2 - \frac{1}{2\eta} \|x^t - w^{t-1}\|^2 \right) \end{aligned} \quad (4)$$

Now we apply a similar analysis from **Theorem 3** to upper bound the term $\|w^{t-1} - p^t\|^2 - \|w^{t-1} - p^{t-1}\|^2$:

$$\begin{aligned} &\|w^{t-1} - p^t\|^2 - \|w^{t-1} - p^{t-1}\|^2 \\ &= \langle p^{t-1} - p^t, 2w^{t-1} - p^{t-1} - p^t \rangle \\ &= \langle p^{t-1} - p^t, 2w^{t-1} - 2p^t \rangle - \|p^t - p^{t-1}\|^2 \\ &= 2\langle p^{t-1} - p^t, v \rangle + 2\langle p^{t-1} - p^t, w^{t-1} - v - p^t \rangle - \|p^t - p^{t-1}\|^2 \\ &= 2\langle p^{t-1} - p^t, v \rangle + 2\langle p^{t-1} - p^t, x^t - v - p^t \rangle + 2\langle p^{t-1} - p^t, w^{t-1} - x^t \rangle - \|p^t - p^{t-1}\|^2 \\ &\leq 2\langle p^{t-1} - p^t, v \rangle + \|x^t - w^{t-1}\|^2, \end{aligned}$$

where in the last-inequality we use $\langle p^{t-1} - p^t, x^t - v - p^t \rangle \leq 0$ since $p^t = \Pi_{\mathcal{X}}[x^t - v]$ and \mathcal{X} is a compact convex set; we also use $2\langle a, b \rangle - b^2 \leq a^2$. In the analysis above, unlike the analysis of **GD** where we drop the negative term $-\|p^t - p^{t-1}\|^2$, we use $-\|p^t - p^{t-1}\|^2$ to get a term $\|x^t - w^{t-1}\|^2$ which can be canceled by the last term in (4).

Now we combine the above two inequalities. Since the term $\|x^t - w^{t-1}\|^2$ cancels out and $2\langle p^{t-1} - p^t, v \rangle$ telescopes, we get

$$\sum_{t=1}^T f^t(x^t) - f^t(p^t) \leq \frac{\langle p^0 - p^T, u \rangle}{\eta} + \sum_{t=1}^T \eta \|g^t - g^{t-1}\|^2 \leq \frac{\delta D \mathcal{X}}{\eta} + \eta \sum_{t=1}^T \|g^t - g^{t-1}\|^2.$$

□

D.3.1 Proof of Theorem 4

Proof. Let us fix any player $i \in [n]$ in the smooth game. In every step t , player i 's loss function $f^t : \mathcal{X}_i \rightarrow \mathbb{R}$ is $\langle -\nabla_{x_i} u_i(x^t), \cdot \rangle$ determined by their utility function u_i and all players' actions x^t . Therefore, their gradient feedback is $g^t = -\nabla_{x_i} u_i(x^t)$. For all $t \geq 2$, we have

$$\begin{aligned} \|g^t - g^{t-1}\|^2 &= \|\nabla u_i(x^t) - \nabla u_i(x^{t-1})\|^2 \\ &\leq L^2 \|x^t - x^{t-1}\|^2 \\ &= L^2 \sum_{i=1}^n \|x_i^t - x_i^{t-1}\|^2 \\ &\leq 3L^2 \sum_{i=1}^n \left(\|x_i^t - w_i^t\|^2 + \|w_i^t - w_i^{t-1}\|^2 + \|w_i^{t-1} - x_i^{t-1}\|^2 \right) \\ &\leq 3nL^2 \eta^2 G^2, \end{aligned}$$

where we use L -smoothness of the utility function u_i in the first inequality; we use the update rule of **OG** and the fact that gradients are bounded by G in the last inequality.

Applying the above inequality to the regret bound obtained in **Theorem 9**, the individual $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of player i is upper bounded by

$$\text{Reg}_{\text{Proj}, \delta}^{T, i} \leq \frac{\delta D}{\eta} + \eta G^2 + 3nL^2 G^2 \eta^3 T.$$

Choosing $\eta = \min\{(\delta D / (nL^2 G^2 T))^{\frac{1}{4}}, (\delta D)^{\frac{1}{2}} / G\}$, we have $\text{Reg}_{\text{Proj}, \delta}^{T, i} \leq 4(\delta D)^{\frac{3}{4}} (nL^2 G^2 T)^{\frac{1}{4}} + 2\sqrt{\delta D} G$. Using **Lemma 1**, we have the empirical distribution of played strategy profiles converge to an $(\varepsilon + \frac{\delta^2 L}{2})$ -approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium in $O(1/\varepsilon^{\frac{4}{3}})$ iterations. □

E Missing Details in Section 4.2

E.1 Proof of Theorem 5

Proof. For a sequence of strategies $\{x^t\}_{t \in [T]}$, its $\text{Conv}(\Phi)$ -regret is

$$\begin{aligned} \text{Reg}_{\text{Conv}(\Phi)}^T &= \max_{\phi \in \text{Conv}(\Phi)} \left\{ \sum_{t=1}^T (u^t(\phi(x^t)) - u^t(x^t)) \right\} \\ &= \underbrace{\max_{p \in \Delta(\Phi)} \left\{ \sum_{t=1}^T u^t(\phi_p(x^t)) - u^t(\phi_{p^t}(x^t)) \right\}}_{\text{I: external regret over } \Delta(\Phi)} + \underbrace{\sum_{t=1}^T u^t(\phi_{p^t}(x^t)) - u^t(x^t)}_{\text{II: approximation error of fixed point}}. \end{aligned}$$

Algorithm 3: $\text{Conv}(\Phi)$ -regret minimization for Lipschitz smooth non-concave rewards

Input: $x_1 \in \mathcal{X}$, $K \geq 2$, a no-external-regret algorithm \mathfrak{R}_Φ against linear reward over $\Delta(\Phi)$

Output: A $\text{Conv}(\Phi)$ -regret minimization algorithm over \mathcal{X}

1 **function** NEXTSTRATEGY()

2 $p^t \leftarrow \mathfrak{R}_\Phi.\text{NEXTSTRATEGY}()$. Note that p_t is a distribution over Φ .

3 $x_k \leftarrow \phi_{p^t}(x_{k-1})$, for all $2 \leq k \leq K$

4 **return** $x^t \leftarrow$ uniformly at random from $\{x_1, \dots, x_K\}$.

5 **function** OBSERVEREWARD($\nabla_x u^t(x^t)$)

6 $u_\Phi^t(\cdot) \leftarrow$ a linear reward over $\Delta(\Phi)$ with $u_\Phi^t(\phi) = \langle \nabla_x u^t(x^t), \phi(x^t) - x^t \rangle$ for all $\phi \in \Phi$.

7 $\mathfrak{R}_\Phi.\text{OBSERVEREWARD}(u_\Phi^t(\cdot))$.

Bounding External Regret over $\Delta(\Phi)$ We can define a new reward function $f^t(p) := u^t(\phi_p(x^t))$ over $p \in \Delta(\Phi)$. Since u^t is non-concave, the reward f^t is also non-concave, and it is computationally intractable to minimize external regret. We use locality to avoid computational barriers. Here we use the fact that $\Phi = \Phi(\delta)$ contains only δ -local strategy modifications. Then by L -smoothness of u^t , we know for any $p \in \Delta(\Phi)$

$$|u^t(\phi_p(x^t)) - u^t(x^t) - \langle \nabla u^t(x^t), \phi_p(x^t) - x^t \rangle| \leq \frac{L}{2} \|\phi_p(x^t) - x^t\|^2 \leq \frac{\delta^2 L}{2}.$$

Thus we can approximate the non-concave optimization problem by a linear optimization problem over $\Delta(\Phi)$ with only second-order error $\frac{\delta^2 L}{2}$. Here we use the notation $a = b \pm c$ to mean $b - c \leq a \leq b + c$.

$$\begin{aligned} u^t(\phi_p(x^t)) - u^t(x^t) &= \langle \nabla u^t(x^t), \phi_p(x^t) - x^t \rangle \pm \frac{\delta^2 L}{2} \\ &= \left\langle \nabla u^t(x^t), \sum_{\phi \in \Phi} p(\phi) \phi(x^t) - x^t \right\rangle \pm \frac{\delta^2 L}{2} \\ &= \sum_{\phi \in \Phi} p(\phi) \langle \nabla u^t(x^t), \phi(x^t) - x^t \rangle \pm \frac{\delta^2 L}{2}. \end{aligned}$$

We can then instantiate the external regret \mathfrak{R}_Φ as the Hedge algorithm over reward

$$f^t(p) = \sum_{\phi \in \Phi} p(\phi) \langle \nabla u^t(x^t), \phi(x^t) - x^t \rangle$$

and get

$$\begin{aligned} &\max_{p \in \Delta(\Phi)} \left\{ \sum_{t=1}^T u^t(\phi_p(x^t)) - u^t(\phi_{p^t}(x^t)) \right\} \\ &\leq \max_{p \in \Delta(\Phi)} \left\{ \sum_{t=1}^T \sum_{\phi \in \Phi} (p(\phi) - p^t(\phi)) \langle \nabla u^t(x^t), \phi(x^t) - x^t \rangle \right\} + \delta^2 L T \\ &\leq 2G\delta \sqrt{T \log |\Phi|} + \delta^2 L T, \end{aligned}$$

where we use the fact that $\langle \nabla u^t(x^t), \phi(x^t) - x^t \rangle \leq \|\nabla u^t(x^t)\| \cdot \|\phi(x^t) - x^t\| \leq G\delta$.

Bounding error due to sampling from a fixed point in expectation We choose x_1 as an arbitrary point in \mathcal{X} . Then we recursively apply ϕ_{p^t} to get

$$x_k = \phi_{p^t}(x_{k-1}) = \sum_{\phi \in \Phi} p^t(\phi) \phi(x_{k-1}), \forall 2 \leq k \leq K.$$

We denote $\mu^t = \text{Uniform}\{x_k : 1 \leq k \leq K\}$. Then the strategy $x^t \sim \mu^t$ is sampled from μ^t . We have that μ^t is an approximate fixed-point in expectation / stationary distribution in the sense that

$$\begin{aligned} \mathbb{E}_{\mu^t} [u^t(\phi_{p^t}(x^t)) - u^t(x^t)] &= \frac{1}{K} \sum_{k=1}^K u^t(\phi_{p^t}(x_k)) - u^t(x_k) \\ &= \frac{1}{K} (u^t(\phi_{p^t}(x_K)) - u^t(x_1)) \\ &\leq \frac{1}{K}. \end{aligned}$$

Thanks to the boundedness of u^t , we can use Hoeffding-Azuma's inequality to conclude that

$$\Pr \left[\sum_{t=1}^T \left(u^t(\phi_{p^t}(x^t)) - u^t(x^t) - \frac{1}{K} \right) \geq \varepsilon \right] \leq \exp \left(-\frac{\varepsilon^2}{8T} \right). \quad (5)$$

for any $\varepsilon > 0$. Combining the above with $\varepsilon = \sqrt{8T \log(1/\beta)}$ and $K = \sqrt{T}$, we get with probability at least $1 - \beta$,

$$\begin{aligned} \text{Reg}_{\text{Conv}(\Phi)}^T &\leq 2G\delta\sqrt{T \log |\Phi|} + \delta^2 LT + \sqrt{T} + \sqrt{8T \log(1/\beta)} \\ &\leq 8\sqrt{T} \left(G\delta\sqrt{\log |\Phi|} + \sqrt{\log(1/\beta)} \right) + \delta^2 LT. \end{aligned}$$

Convergence to Φ -equilibrium If all players in a non-concave continuous game employ [Algorithm 1](#), then we know for each player i , with probability $1 - \frac{\beta}{n}$, its $\Phi^{\mathcal{X}_i}$ -regret is upper bounded by

$$8\sqrt{T} \left(G\delta\sqrt{\log |\Phi^{\mathcal{X}_i}|} + \sqrt{\log(n/\beta)} \right) + \delta^2 LT.$$

By a union bound over all n players, we get with probability $1 - \beta$, every player i 's $\Phi^{\mathcal{X}_i}$ -regret is upper bounded by $8\sqrt{T}(G\delta\sqrt{\log |\Phi^{\mathcal{X}_i}|} + \sqrt{\log(n/\beta)}) + \delta^2 LT$. Now by [Theorem 1](#), we know the empirical distribution of strategy profiles played forms an $(\varepsilon + \delta^2 L)$ -approximate $\Phi = \prod_{i=1}^n \Phi^{\mathcal{X}_i}$ -equilibrium, as long as $T \geq \frac{128(G^2\delta^2 \log |\Phi^{\mathcal{X}_i}| + \log(n/\beta))}{\varepsilon^2}$ iterations. \square

F Missing details in [Section 4.3](#)

We introduce a natural set of local strategy modifications and the corresponding local equilibrium notion. Given any set of (possibly non-local) strategy modifications $\Psi = \{\psi : \mathcal{X} \rightarrow \mathcal{X}\}$, we define a set of *local* strategy modifications as follows: for $\delta \leq D_{\mathcal{X}}$ and $\lambda \in [0, 1]$, each strategy modification $\phi_{\lambda, \psi}$ interpolates the input strategy x with the modified strategy $\psi(x)$: formally,

$$\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta) := \{\phi_{\lambda, \psi}(x) := (1 - \lambda)x + \lambda\psi(x) : \psi \in \Psi, \lambda \leq \delta/D_{\mathcal{X}}\}.$$

Note that for any $\psi \in \Psi$ and $\lambda \leq \frac{\delta}{D_{\mathcal{X}}}$, we have $\|\phi_{\lambda, \psi}(x) - x\| = \lambda\|x - \psi(x)\| \leq \delta$, respecting the locality constraint. The induced $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret can be written as

$$\text{Reg}_{\text{Int}, \Psi, \delta}^T := \max_{\psi \in \Psi, \lambda \leq \frac{\delta}{D_{\mathcal{X}}}} \sum_{t=1}^T (f^t(x^t) - f^t((1-\lambda)x^t + \lambda\psi(x^t))).$$

We now define the corresponding $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium.

Definition 9. Define $\Phi_{\text{Int}, \Psi}(\delta) = \prod_{j=1}^n \Phi_{\text{Int}, \Psi_j}^{\mathcal{X}_j}(\delta)$. In a continuous game, a distribution σ over strategy profiles is an ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium if and only if for all player $i \in [n]$,

$$\max_{\psi \in \Psi_i, \lambda \leq \delta/D_{\mathcal{X}_i}} \mathbb{E}_{x \sim \sigma} [u_i((1-\lambda)x_i + \lambda\psi(x_i), x_{-i})] \leq \mathbb{E}_{x \sim \sigma} [u_i(x)] + \varepsilon.$$

Intuitively speaking, when a correlation device recommends strategies to players according to an ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium, no player can increase their utility by more than ε through a local deviation by interpolating with a (possibly global) strategy modification $\psi \in \Psi$. The richness of Ψ determines the incentive guarantee provided by an ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium and its computational complexity. When we choose Ψ to be the set of all possible strategy modifications, the corresponding notion of local equilibrium—limiting a player’s gain by interpolating with any strategy—resembles that of a *correlated equilibrium*.

Computation of ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -Equilibrium. By [Lemma 1](#), we know computing an ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium reduces to minimizing $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret against convex loss functions. We show that minimizing $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret against convex loss functions further reduces to Ψ -regret minimization against linear loss functions.

Theorem 10. Let \mathcal{A} be an algorithm with Ψ -regret $\text{Reg}_{\Psi}^T(G, D_{\mathcal{X}})$ for linear and G -Lipschitz loss functions over \mathcal{X} . Then, for any $\delta > 0$, the $\Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)$ -regret of \mathcal{A} for convex and G -Lipschitz loss functions over \mathcal{X} is at most $\frac{\delta}{D_{\mathcal{X}}} \cdot [\text{Reg}_{\Psi}^T(G, D_{\mathcal{X}})]^+$.

Proof. By definition and convexity of f^t , we get

$$\begin{aligned} \max_{\phi \in \Phi_{\text{Int}, \Psi}^{\mathcal{X}}(\delta)} \sum_{t=1}^T f^t(x^t) - f^t(\phi(x^t)) &= \max_{\psi \in \Psi, \lambda \leq \frac{\delta}{D_{\mathcal{X}}}} \sum_{t=1}^T f^t(x^t) - f^t((1-\lambda)x^t + \lambda\psi(x^t)) \\ &\leq \frac{\delta}{D_{\mathcal{X}}} \left[\max_{\psi \in \Psi} \sum_{t=1}^T \langle \nabla f^t(x^t), x^t - \psi(x^t) \rangle \right]^+. \end{aligned}$$

□

Note that when f^t is linear, the reduction is without loss. Thus, any worst-case $\Omega(r(T))$ -lower bound for Ψ -regret implies a $\Omega(\frac{\delta}{D_{\mathcal{X}}} \cdot r(T))$ lower bound for $\Phi_{\text{Int}, \Psi}(\delta)$ -regret. Moreover, for any set Ψ that admits efficient Ψ -regret minimization algorithms such as swap transformations over the simplex and more generally any set such that (i) all modifications in the set can be represented as linear transformations in some finite-dimensional space and (ii) fixed point computation can be carried out efficiently for any linear transformations [[Gordon et al., 2008](#)], we also get an efficient algorithm for computing an ε -approximate $\Phi_{\text{Int}, \Psi}(\delta)$ -equilibrium in the first-order stationary regime.

CCE-like Instantiation In the special case where Ψ contains only *constant* strategy modifications (i.e., $\psi(x) = x^*$ for all x), we get a coarse correlated equilibrium (CCE)-like instantiation of local equilibrium, which limits the gain by interpolating with any *fixed* strategy. We denote the resulting set of local strategy modification simply as $\Phi_{\text{Int}}^{\mathcal{X}}$. We can apply any no-external regret algorithm for efficient $\Phi_{\text{Int}}^{\mathcal{X}}$ -regret minimization and computation of ε -approximate $\Phi_{\text{Int}}(\delta)$ -equilibrium in the first-order stationary regime as summarized in [Theorem 6](#).

The above $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -regret bound of $O(\sqrt{T})$ is derived for the adversarial setting. In the game setting, where each player employs the same algorithm, players may have substantially lower external regret [[Syrgkanis et al., 2015](#), [Chen and Peng, 2020](#), [Daskalakis et al., 2021a](#), [Anagnostides et al., 2022a,b](#), [Farina et al., 2022a](#)] but we need a slightly stronger smoothness assumption than [Assumption 1](#). This assumption is naturally satisfied by finite normal-form games and is also made for results about concave games [[Farina et al., 2022a](#)]. Using [Assumption 2](#) and [Lemma 1](#), the no-regret learning dynamics of [[Farina et al., 2022a](#)] that guarantees $O(\log T)$ individual external regret in concave games can be applied to smooth non-concave games so that the individual $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -regret of each player is at most $O(\log T) + \frac{\delta^2 LT}{2}$. This gives an algorithm with faster $\tilde{O}(1/\varepsilon)$ convergence to an $(\varepsilon + \frac{\delta^2 L}{2})$ -approximate $\Phi_{\text{Int}}(\delta)$ -equilibrium than [GD](#).

G Proof of Theorem 7

Let $\mathcal{X} \subset \mathbb{R}^2$ be a triangle region with vertices $A = (0, 0)$, $B = (1, 1)$, $C = (\delta, 0)$. Consider $v = (-\delta, 0)$. The initial point is $x_1 = (0, 0)$.

The adversary will choose ℓ_t adaptively so that x_t remains on the boundary of \mathcal{X} and cycles clockwise (i.e., $A \rightarrow \dots \rightarrow B \rightarrow \dots \rightarrow C \rightarrow \dots \rightarrow A \rightarrow \dots$). To achieve this, the adversary will repeat the following three phases:

1. Keep choosing $\ell_t = u_{\overrightarrow{B\bar{A}}}$ ($u_{\overrightarrow{B\bar{A}}}$ denotes the unit vector in the direction of $\overrightarrow{B\bar{A}}$) until x_{t+1} reaches B .
2. Keep choosing $\ell_t = u_{\overrightarrow{CB}}$ until x_{t+1} reaches C .
3. Keep choosing $\ell_t = u_{\overrightarrow{AC}}$ until x_{t+1} reaches A .

In Phase 1, $x_t \in \overline{AB}$. By the choice of $v = (-\delta, 0)$, we have $x_t - \phi_v(x_t) = (-\delta(1 - x_{t,1}), 0)$, and the instantaneous regret is $\frac{\delta(1-x_{t,1})}{\sqrt{2}} \geq 0$.

In Phase 2, $x_t \in \overline{BC}$. By the choice of $v = (-\delta, 0)$, we have $x_t - \phi_v(x_t) = (0, 0)$, and the instantaneous regret is 0.

In Phase 3, $x_t \in \overline{CA}$. By the choice of $v = (-\delta, 0)$, we have $x_t - \phi_v(x_t) = (-\delta + x_{t,1}, 0)$, and the instantaneous regret is $-\delta + x_{t,1} \leq 0$.

In each cycle, the number of rounds in Phase 1 is of order $\Theta(\frac{\sqrt{2}}{\eta})$, the number of rounds in Phase 2 is between $O(\frac{1}{\eta})$ and $O(\frac{\sqrt{2}}{\eta})$, the number of rounds in Phase 3 is of order $\Theta(\frac{\delta}{\eta})$.

Therefore, the cumulative regret in each cycle is roughly

$$\frac{\sqrt{2}}{\eta} \times \frac{0.5\delta}{\sqrt{2}} + 0 + \frac{\delta}{\eta} (-0.5\delta) = \frac{0.5\delta - 0.5\delta^2}{\eta}.$$

On the other hand, the number of cycles is no less than $\frac{T}{\frac{\sqrt{2}}{\eta} + \frac{\sqrt{2}}{\eta} + \frac{\delta}{\eta}} = \Theta(\eta T)$. Overall, the cumulative regret is at least $\frac{0.5\delta - 0.5\delta^2}{\eta} \times \Theta(\eta T) = \Theta(\delta T)$ as long as $\delta < 0.5$.

H First-Order Stationary Regime

For an infinite set of δ -local strategy modifications $\Phi(\delta)$, we focus on computing ε -approximate $\Phi(\delta)$ -equilibrium when $\varepsilon = \Omega(\delta^2)$. We call $\varepsilon = \Omega(\delta^2)$ the *first-order stationary regime* since a deviation bounded by δ only gives δ^2 gain in utility. We show that by considering $\Phi(\delta)$ -equilibria that correlate players' strategy, efficient algorithms exist for $\varepsilon = O(L\delta^2)$ for certain classes of $\Phi(\delta)$ (Sections 4.1 to 4.3). A natural question on the complexity of computing an ε -approximate $\Phi(\delta)$ -equilibrium is

Is first-order stationary regime the best we can hope for? Can we get $\varepsilon = o(\delta^2)$ efficiently? (6)

We show that it is NP-hard to compute an ε -approximate $\Phi(\delta)$ -equilibrium for $\varepsilon = o(\delta^2)$. This hardness result complements our positive results in the first-order stationary regime. Together, they present a more complete picture of the complexity landscape of approximate $\Phi(\delta)$ -equilibrium.

H.1 Hardness when $\Phi(\delta)$ contains all local strategy modifications

We start with the case when $\Phi(\delta)$ contains all δ -local strategy modifications. We consider a single-player game with an action set $\mathcal{X} = [0, 1]^d$. As a byproduct of our analysis, we also prove the NP-hardness of deciding/computing a (ε, δ) -local maximum of a smooth and Lipschitz function.

We will construct a reduction from the maximum clique problem denoted as CLIQUE. We remark that MAX CLIQUE is NP-hard. Moreover, even approximating MAX CLIQUE with a factor of $O(d^{1-o(1)})$ is NP-hard [Zuckerman, 2006], where d is the number of nodes in the graph.

Definition 10 (MAX CLIQUE). *Given a graph $G = (V, E)$ with vertices V and edges E . The MAX CLIQUE problem asks to compute the clique number $\omega(G)$ defined as the size of the maximum clique in the graph G .*

Fix any graph $G = (V, E)$ with $|V| = d \geq 3$. We denote $A \in \{0, 1\}^{d \times d}$ as G 's adjacent matrix. Then by then Motzkin-Straus Theorem [Motzkin and Straus, 1965], we have

$$\max_{x \in \Delta^d} x^\top A x = 1 - \frac{1}{\omega(G)}.$$

We denote $x^* \in \Delta^d$ as an optimal solution to the above optimization problem. We also define d functions $f_k : \mathcal{X} := [0, 1]^d \rightarrow \mathbb{R}$ for $1 \leq k \leq d$.

$$f_k(x) := \left(x^\top A x - \left(1 - \frac{1}{k} \right) \cdot \|x\|_1^2 \right), \quad 1 \leq k \leq d. \quad (7)$$

We can verify that each f_k is G -Lipschitz and L -smooth with $G, L = O(\text{poly}(d))$.

Lemma 2. *The following holds for any f_k :*

1. $f_k(x) \in [-d^2, d^2]$ for all $x \in [0, 1]^d$. If $k > \omega(G)$, $f_k(x) < 0$ for all $x \in [0, 1]^d$.
2. **(Local) Lipschitzness:** For any $x \in [0, 1]^d$, $\|\nabla_x f_k(x)\|_2 \leq 3\sqrt{d}\|x\|_1$. Therefore, f_k is G -Lipschitz with $G = 3d^{\frac{3}{2}}$.
3. **Smoothness:** f_k is L -smooth with $L = 2d$.

Proof. For any $x \in [0, 1]^d$ such that $x \neq 0$, we have

$$\begin{aligned}
f_k(x) &= x^\top Ax - \left(1 - \frac{1}{k}\right) \cdot \|x\|_1^2 \\
&= \|x\|_1^2 \cdot \left(\frac{x}{\|x\|_1}^\top A \frac{x}{\|x\|_1} - \left(1 - \frac{1}{k}\right)\right) \\
&\leq \|x\|_1^2 \left(1 - \frac{1}{\omega(G)} - \left(1 - \frac{1}{k}\right)\right) \\
&= \|x\|_1^2 \left(\frac{1}{k} - \frac{1}{\omega(G)}\right). \tag{8}
\end{aligned}$$

We note that the inequality becomes equality when $\frac{x}{\|x\|_1} = x^*$ the optimal solution for $\max_{x \in \Delta^d} x^\top Ax$. Thus $f_k(x) \leq \|x\|_1^2 \leq d^2$. Moreover, we have $f_k(x) \geq -(1 - \frac{1}{k})\|x\|_1^2 \geq -\|x\|_1^2 \geq -d^2$. Thus $f_k(x) \in [-d^2, d^2]$. If $k > \omega(G)$, $f_k(x) \leq \|x\|_1^2(\frac{1}{k} - \frac{1}{\omega(G)}) < 0$.

For the second part, we first compute the gradient of f_k ,

$$\nabla_x f_k(x) = Ax - 2\left(1 - \frac{1}{k}\right)\|x\|_1 \cdot 1_d.$$

Then

$$\begin{aligned}
\|\nabla_x f_k(x)\|_2 &= \left\| Ax - 2\left(1 - \frac{1}{k}\right)\|x\|_1 \cdot 1_d \right\|_2 \\
&\leq \|Ax\|_2 + 2\|\|x\|_1 \cdot 1_d\|_2 \\
&\leq \sqrt{d}\|x\|_1 + 2\sqrt{d}\|x\|_1 \\
&= 3\sqrt{d}\|x\|_1.
\end{aligned}$$

where in the first inequality, we apply the triangle inequality; in the second inequality, we use the fact that $A \in [0, 1]^d$.

The Hessian of f_k is

$$\nabla_x^2 f(x) = A - 2\left(1 - \frac{1}{k}\right) \cdot 1_{d \times d},$$

where we use $1_{d \times d}$ to denote the all-one matrix. Since $A \in \{0, 1\}^{d \times d}$, we know the absolute eigenvalues of $\nabla_x^2 f(x)$ is bounded by $2d$. Thus f_k is $2d$ -smooth. \square

The following technical lemma relates the local maximum of f_k for $1 \leq k \leq d$ and the clique number $\omega(G)$. Specifically, [Lemma 3](#) shows that when $k \geq \omega(G) + 1$, all $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum must have small ℓ_2 -norms, i.e., $\|x\| < \frac{\delta}{2}$; when $k < \omega(G)$, all $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum must have large ℓ_2 -norm, i.e., $\|x\| > \frac{\delta}{2}$.

Lemma 3. *The following holds for all $\delta \in (0, 1]$:*

1. *If $k \geq \omega(G) + 1$, then any x with $\|x\|_2 \geq \frac{\delta}{2}$ is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum.*
2. *If $k < \omega(G)$, then any x with $\|x\|_2 \leq \frac{\delta}{2}$ is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum.*

Proof. We first prove the case when $k \geq \omega(G) + 1$.

When $k \geq \omega(G) + 1$: Let $x \in [0, 1]^d$ be any point such that $\|x\|_2 \geq \frac{\delta}{2}$. Define $x' := (1 - \frac{\delta}{2\|x\|_2})x$ the point by moving x towards the origin by $\frac{\delta}{2}$. Then we know the distance between x and x' is bounded by

$$\|x - x'\|_2 = \frac{\delta}{2\|x\|_2} \|x\|_2 = \frac{\delta}{2} \leq \delta.$$

Moreover, we have $\|x'\|_1 = (1 - \frac{\delta}{2\|x\|_2}) \cdot \|x\|_1$.

$$\begin{aligned} f_k(x') - f_k(x) &= \left(\left(1 - \frac{\delta}{2\|x\|_2}\right)^2 - 1 \right) \cdot f_k(x) \\ &= \left(\frac{\delta^2}{4\|x\|_2^2} - 2 \cdot \frac{\delta}{2\|x\|_2} \right) \cdot f_k(x) \\ &\geq \frac{\delta}{2\|x\|_2} \cdot (-f_k(x)) && (0 < \frac{\delta}{2\|x\|_2} \leq 1 \text{ and } f_k(x) < 0) \\ &\geq \frac{\delta}{2\|x\|_2} \cdot \|x\|_1^2 \cdot \left(-\frac{1}{k} + \frac{1}{\omega(G)} \right) && (\text{by (8)}) \\ &\geq \frac{\delta^2}{4} \cdot \left(-\frac{1}{k} + \frac{1}{\omega(G)} \right) && (\|x\|_1 \geq \|x\|_2 \geq \frac{\delta}{2}) \\ &> \frac{\delta^2}{4d^2}. && (d \geq k \geq \omega(G) + 1) \end{aligned}$$

Thus x is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum.

When $k < \omega(G)$: Let $x \in [0, 1]^d$ be any point such that $\|x\|_2 \leq \frac{\delta}{2}$. We further define a threshold v_x for x :

$$v_x := \frac{f_k(x)}{\|x\|_2^2}.$$

and proceed by two cases depending on v_x .

Case 1: $v_x \geq \frac{1}{2d^2}$. In this case, we consider a deviation $x' := x + \frac{\delta}{2\|x\|_2}x$. We note that x' still lies in $[0, 1]^d$ since $\|x'\|_2 \leq \delta \leq 1$. The magnitude of deviation is clearly bounded by $\|x' - x\|_2 = \frac{\delta}{2}$. Moreover,

$$\begin{aligned} f_k(x') - f_k(x) &= \left(\left(1 + \frac{\delta}{2\|x\|_2}\right)^2 - 1 \right) \cdot f_k(x) \\ &> \frac{\delta^2}{4\|x\|_2^2} \cdot f_k(x) \\ &\geq \frac{\delta^2}{8d^2}. && (v_x = \frac{f_k(x)}{\|x\|_2^2} \geq \frac{1}{2d^2}) \end{aligned}$$

Thus x is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum.

Case 2: $v_x < \frac{1}{2d^2}$. We consider deviation towards the optimal solution $x^* \in \Delta^d$ with function value $f_k(x^*) = \frac{1}{k} - \frac{1}{\omega(G)} \geq \frac{1}{d^2}$. Specifically, we let $x' = \frac{\delta}{2}x^* \in \mathcal{B}_d(\frac{\delta}{2})$. Since both x and x' lie in $\mathcal{B}_d(\frac{\delta}{2})$, we know the deviation is bounded by $\|x' - x\| \leq \delta$. Moreover, deviation to x' increases the function value by

$$\begin{aligned} f_k(x') - f_k(x) &= \frac{\delta^2}{4} \cdot f_k(x^*) - f_k(x) \\ &> \frac{\delta^2}{4d^2} - f_k(x). \end{aligned}$$

If $f_k(x) \leq 0$, then we know $f_k(x') - f_k(x) \geq \frac{\delta^2}{4d^2}$. Otherwise, we can further lower bound the increase in function value by

$$\begin{aligned} f_k(x') - f_k(x) &\geq \frac{\delta^2}{4d^2} - f_k(x) \\ &\geq \frac{\delta^2}{4d^2} - \frac{\delta^2}{4} \cdot \frac{1}{\|x\|_2^2} \cdot f_k(x) && (\|x\|_2^2 \leq \frac{\delta^2}{4}) \\ &> \frac{\delta^2}{8d^2}. && (v_x = \frac{f_k(x)}{\|x\|_2^2} < \frac{1}{2d^2}) \end{aligned}$$

Combining the above, we know x is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum. \square

Denote $\Phi_{\text{All}}(\delta)$ the set of all δ -local strategy modifications. Using [Lemma 3](#), we can prove NP-hardness of ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium for two cases: (1) $\varepsilon \leq \frac{\delta^2}{16d}$; (2) $\varepsilon \leq \text{poly}(G, L, d) \cdot \delta^{2+c}$ for any $\text{poly}(G, L, d)$ and $c > 0$.

Theorem 11 (Hardness of Approximate $\Phi_{\text{All}}(\delta)$ -Equilibrium). *There exists a family of single-player games with utility functions $f : [0, 1]^d \rightarrow [-d^2, d^2]$, whose Lipschitzness G and smoothness L are both $\text{poly}(d)$, such that the followings are true.*

1. For any $0 < \delta \leq 1$, if there is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium for $\varepsilon \leq \frac{\delta^2}{16d^2}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.
2. For any $g = \text{poly}(G, L, d)$ and $c > 0$, if there is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium for all $\delta \in (0, 1]$ and $\varepsilon = g \cdot \delta^{2+c}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.

Proof. We consider utility functions $f_k : [0, 1]^d \rightarrow [-d^2, d^2]$ for $1 \leq k \leq d$ as defined in (7). We recall that f_k is G -Lipschitz and L -smooth with $G = 2d^{\frac{3}{2}}$ and $L = 2d$ ([Lemma 2](#)).

Part 1: We fix any $\delta \in (0, 1)$ such that $\delta = \frac{1}{\text{poly}(d)}$ and let $\varepsilon = \frac{\delta^2}{16d^2}$. Suppose \mathcal{A} is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium in time $\text{poly}(G, L, d, 1/\delta, 1/\varepsilon) = \text{poly}(d)$. Now consider the following algorithm for approximating MAX CLIQUE.

1. Run \mathcal{A} for the game f_k and denote the outcome as σ_k for $1 \leq k \leq d$.
2. Return the smallest $1 \leq k \leq d$ such that $\Pr_{x \sim \sigma_k}[\mathbb{I}[\|x\|_2 \leq \frac{\delta}{2}]] \geq \frac{1}{2}$.

We claim that the above algorithm outputs either $\omega(G)$ or $\omega(G) + 1$. We first prove that $\Pr_{x \sim \sigma_k} [\mathbb{I}[\|x\|_2 \leq \frac{\delta}{2}] < \frac{1}{2}] < \frac{1}{2}$ for all $k < \omega(G)$. Suppose the claim is false, then consider the following deviation for all x in σ_k 's support: (1) if $\|x\|_2 \leq \frac{\delta}{2}$, then deviate x to the point x' with $\|x' - x\| \leq \delta$ and $f_k(x') - f_k(x) > \frac{\delta^2}{8d^2}$. This deviation is possible since x is not a $(\frac{\delta^2}{8d^2}, \delta)$ -local maximum by [Lemma 3](#). (2) If $\|x\|_2 > \frac{\delta}{2}$, then keep x unchanged. We denote the resulting distribution after deviation as σ'_k . Then $f_k(\sigma'_k) - f_k(\sigma_k) > \frac{\delta^2}{8d^2} \cdot \frac{1}{2} = \frac{\delta^2}{16d^2}$, which contradicts the fact that σ_k is an $\frac{\delta^2}{16d^2}$ -approximate $\Phi(\delta)$ -equilibrium. Using a similar argument, we can prove that $\Pr_{x \sim \sigma_k} [\mathbb{I}[\|x\|_2 \leq \frac{\delta}{2}]] \geq \frac{1}{2}$ for all $k \geq \omega(G) + 1$. Combining the above, we know the algorithm either outputs $\omega(G)$ or $\omega(G) + 1$. Moreover, the algorithm runs in time $\text{poly}(d)$. But since it is NP-hard to approximate MAX CLIQUE even within $O(d^{1-o(1)})$ factor [[Zuckerman, 2006](#)], the above implies $\text{P} = \text{NP}$.

Part 2: We fix any $g = \text{poly}(G, L, d) = \text{poly}(d)$ and $c > 0$. We choose $\delta = (\frac{1}{16d^2 \cdot g})^{\frac{1}{c}} = \text{poly}(\frac{1}{d})$ so that $\varepsilon = g \cdot \delta^{2+c} \leq \frac{\delta^2}{16d^2}$. Then by Part 1, we know if there is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium in time $\text{poly}(G, L, d, 1/\delta, 1/\varepsilon) = \text{poly}(d)$ and, then $\text{P} = \text{NP}$. \square

As a byproduct of [Lemma 3](#), we can prove hardness results for (ε, δ) -local maximum similar to [Theorem 11](#). The proof is simpler than that of [Theorem 11](#), and we omitted it here.

Corollary 2 (Hardness of (ε, δ) -Local Maximum). *Consider a class of G -Lipschitz and L -smooth functions $f : [0, 1]^d \rightarrow [-d^2, d^2]$ with $G = \Theta(d^{\frac{3}{2}})$ and $L = \Theta(d)$. Then the following hold.*

1. For any $0 < \delta \leq 1$, if there is an algorithm that computes an (ε, δ) -local maximum for $\varepsilon \leq \frac{\delta^2}{16d^2}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.
2. For any $g = \text{poly}(G, L, d)$ and $c > 0$, if there is an algorithm that computes an (ε, δ) -local maximum for $\delta \in (0, 1]$ and $\varepsilon \leq g \cdot \delta^{2+c}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.

H.2 Restricted Deviations

In this section, we show the hardness of approximating a local maximum and approximating a $\Phi(\delta)$ -equilibrium when the class of strategy modifications $\Phi(\delta)$ is $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ plus one additional deviation. We let $\mathcal{X} = [0, 1]^d$ where $D_{\mathcal{X}} = \sqrt{d}$. Recall that each strategy modification $\phi_{x'}$ in $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ is of the form

$$\phi_{\lambda, x'}(x) = (1 - \lambda)x + \lambda x', \quad \forall \lambda \in [0, \frac{\delta}{\sqrt{d}}] \text{ and } x' \in [0, 1]^d$$

The additional strategy modification we consider is

$$\phi(x) = \begin{cases} (1 - \frac{\delta}{12d^3\|x\|_1})x, & \|x\|_1 \geq \frac{\delta}{12d^3}, \\ x, & \text{Otherwise.} \end{cases}$$

We can check that ϕ is a well-defined δ -local strategy modification for \mathcal{X} . We denote the union of $\{\phi\}$ and $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ as $\Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$.

The following technical lemma is similar to [Lemma 3](#). The main difference is that [Lemma 3](#) allows arbitrary δ -local strategy modifications, while in [Lemma 4](#), we only allow strategy modifications in $\Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$.

Lemma 4. *For all $\delta \in (0, 1]$, the following holds:*

1. If $k \geq \omega(G) + 1$, there exists a single $\phi \in \Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$ such that $f_k(\phi(x)) - f_k(x) > \frac{\delta^2}{144d^8}$ for all x with $\|x\|_1 \geq \frac{\delta}{12d^3}$.
2. If $k < \omega(G)$, there exists a single $\phi \in \Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$ such that $f_k(\phi(x)) - f_k(x) \geq \frac{\delta^2}{2d^3} > \frac{\delta^2}{144d^8}$ for all x with $\|x\|_1 \leq \frac{\delta}{12d^3}$.

Proof. We first prove the claim for $k \geq \omega(G) + 1$.

When $k \geq \omega(G) + 1$: Let $x \in [0, 1]^d$ be any point such that $\|x\|_1 \geq \frac{\delta}{12d^3}$. Define $x' = (1 - \frac{\delta}{12d^3\|x\|_1})x$. We note that $x' \in [0, 1]^d$ is well-defined and the distance between x and x' is bounded by

$$\|x - x'\|_2 = \frac{\delta}{12d^3\|x\|_1} \|x\|_2 \leq \frac{\delta}{12d^3} \leq \delta.$$

Moreover, we have $\|x'\|_1 = (1 - \frac{\delta}{12d^3\|x\|_1}) \cdot \|x\|_1$

$$\begin{aligned} f_k(x') - f_k(x) &= \left(\left(1 - \frac{\delta}{12d^3\|x\|_1}\right)^2 - 1 \right) \cdot f_k(x) \\ &= \left(\left(\frac{\delta}{12d^3\|x\|_1}\right)^2 - 2 \cdot \frac{\delta}{12d^3\|x\|_1} \right) \cdot f_k(x) \\ &\geq \frac{\delta}{12d^3\|x\|_1} \cdot (-f_k(x)) && \left(0 < \frac{\delta}{12d^3\|x\|_1} \leq 1 \text{ and } f_k(x) < 0\right) \\ &\geq \frac{\delta}{12d^3\|x\|_1} \cdot \|x\|_1^2 \cdot \left(-\frac{1}{k} + \frac{1}{\omega(G)}\right) && \text{(by (8))} \\ &\geq \frac{\delta^2}{144d^6} \cdot \left(-\frac{1}{k} + \frac{1}{\omega(G)}\right) && (\|x\|_1 \geq \frac{\delta}{12d^3}) \\ &> \frac{\delta^2}{144d^8}. && (d \geq k \geq \omega(G) + 1) \end{aligned}$$

Thus x is not a $(\frac{\delta^2}{144d^8}, \delta)$ -local maximum.

When $k < \omega(G)$: Let $x \in [0, 1]^d$ be any point such that $\|x\|_1 \leq \frac{\delta}{12d^3}$. We consider the deviation that interpolates with the optimal solution $x^* \in \Delta^d$, whose function value $f_k(x^*) = \frac{1}{k} - \frac{1}{\omega(G)} \geq \frac{1}{d^2}$. Specifically, we let $x' = (1 - \frac{\delta}{\sqrt{d}})x + \frac{\delta}{\sqrt{d}}x^*$. We know the deviation is bounded by $\|x' - x\|_2 \leq \delta$. Moreover, for any y lies in the line segment between x and x' , we know $\|y\|_1 = \|(1 - \frac{\delta}{\sqrt{d}})x\|_1 + \|\frac{\delta}{\sqrt{d}}x^*\|_1 \leq \frac{2\delta}{\sqrt{d}}$. Thus by the mean value theorem, there exists a y in the line segment between x and x' , such that

$$f_k(x') = f_k\left(\frac{\delta}{\sqrt{d}} \cdot x^*\right) + \left\langle \nabla f_k(y), x' - \frac{\delta}{\sqrt{d}}x^* \right\rangle \geq f_k\left(\frac{\delta}{\sqrt{d}} \cdot x^*\right) - \|\nabla f_k(y)\|_2 \cdot \left\|x' - \frac{\delta}{\sqrt{d}}x^*\right\|_2.$$

By the local Lipschitzness of f_k , the function value of $f_k(x')$ is at least

$$\begin{aligned}
f_k(x') &\geq f_k\left(\frac{\delta}{\sqrt{d}} \cdot x^*\right) - \|\nabla f_k(y)\|_2 \cdot \left\|x' - \frac{\delta}{\sqrt{d}}x^*\right\|_2 \\
&\geq \frac{\delta^2}{d} \cdot \frac{1}{d^2} - 3\sqrt{d} \cdot \frac{2\delta}{\sqrt{d}} \cdot \left\|\left(1 - \frac{\delta}{\sqrt{d}}\right)x\right\|_2 \\
&\geq \frac{\delta^2}{d^3} - 6\delta \cdot \|x\|_1 \\
&\geq \frac{\delta^2}{2d^3},
\end{aligned}$$

In the second inequality, we use $f_k(x^*) \geq \frac{1}{d^2}$ and $\|\nabla f_k(y)\|_2 \leq 3\sqrt{d}\|y\|_1 \leq 6\delta$ (Lemma 2); in the third inequality, we use the fact that the ℓ_1 -norm bounds ℓ_2 -norm.

On the other hand, we have

$$\begin{aligned}
f_k(x) &= \|x\|_1^2 \cdot f_k\left(\frac{x}{\|x\|_1}\right) \\
&\leq \frac{\delta^2}{144d^6} \cdot f_k(x^*) \\
&\leq \frac{\delta^2}{144d^6}.
\end{aligned}$$

Combining the above, we have $f_k(x') - f(x) \geq \frac{\delta^2}{3d^3}$ and we know x is not a $(\frac{\delta^2}{3d^3}, \delta)$ -local maximum. \square

Combining Lemma 4 and similar analysis as Theorem 11, we have the following hardness results for computing an ε -approximate $\Phi_{\text{Int}^+}(\delta)$ -equilibrium.

Theorem 12 (Hardness of Approximate $\Phi_{\text{Int}^+}(\delta)$ -Equilibrium). *There exists a family of single-player games, whose utility functions $f : [0, 1]^d \rightarrow [-d^2, d^2]$ with $G, L = \text{poly}(d)$, such that the following hold.*

1. For any $0 < \delta \leq 1$ such that $\delta = \frac{1}{\text{poly}(d)}$, if there is an algorithm that computes an ε -approximate $\Phi_{\text{Int}^+}(\delta)$ -equilibrium for $\varepsilon \leq \frac{\delta^2}{576d^8}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.
2. For any $g = \text{poly}(G, L, d)$ and $c > 0$, if there is an algorithm that computes an ε -approximate $\Phi_{\text{Int}^+}(\delta)$ -equilibrium for $\varepsilon \leq g \cdot \delta^{2+c}$ in time $\text{poly}(d)$, then $\text{P} = \text{NP}$.

Proof. We consider utility functions $f_k : [0, 1]^d \rightarrow [-d^2, d^2]$ for $1 \leq k \leq d$ as defined in (7). We recall that f_k is G -Lipschitz and L -smooth with $G = 2d^{\frac{3}{2}}$ and $L = 2d$ (Lemma 2).

Part I: We fix any $\delta \in (0, 1]$ such that $\delta = \frac{1}{\text{poly}(d)}$ and let $\varepsilon = \frac{\delta^2}{576d^8}$. Suppose \mathcal{A} is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium in time $\text{poly}(G, L, d, 1/\delta, 1/\varepsilon) = \text{poly}(d)$. Now consider the following algorithm for approximating MAX CLIQUE: define $p = \frac{\delta}{1152d^{10}} = \frac{1}{\text{poly}(d)}$,

1. Run \mathcal{A} for the game f_k and denote the outcome as σ_k for $1 \leq k \leq d$.
2. Return the smallest $1 \leq k \leq d$ such that $\Pr_{x \sim \sigma_k}[\mathbb{I}[\|x\|_1 \leq \frac{\delta}{12d^3}]] \geq 1 - p$.

We claim that the above algorithm outputs either $\omega(G)$ or $\omega(G) + 1$. We first prove that $\Pr_{x \sim \sigma_k}[\mathbb{I}[\|x\|_1 \leq \frac{\delta}{12d^3}]] < 1 - p$ for all $k < \omega(G)$. Suppose the claim is false. By Lemma 4, we know there is a single deviation $\phi \in \Phi_{\text{Int}^+}^{\mathcal{X}}(\delta)$ such that $f_k(\phi(x)) - f_k(x) > \frac{\delta^2}{144d^8}$ for all x with $\|x\|_1 \geq \frac{\delta}{12d^3}$. Moreover, since f_k is $2d^{\frac{3}{2}}$ -Lipschitz, we know for any $x \in [0, 1]^d$, $f_k(\phi(x)) - f_k(x) \geq -2d^{\frac{3}{2}}\delta$. Thus we have

$$\begin{aligned} f_k(\phi(\sigma_k)) - f_k(\sigma_k) &\geq \frac{\delta^2}{144d^8} \cdot (1 - p) - 2d^{\frac{3}{2}}\delta \cdot p \\ &> \frac{\delta^2}{288d^8} - \frac{\delta^2}{576d^8} = \varepsilon. \end{aligned}$$

This contradicts the fact that σ_k is an ε -approximate $\Phi(\delta)$ -equilibrium. Using a similar argument, we can prove that $\Pr_{x \sim \sigma_k}[\mathbb{I}[\|x\|_1 \leq \frac{\delta}{12d^3}]] \geq 1 - p$ for all $k \geq \omega(G) + 1$. Combining the above, we know the algorithm either outputs $\omega(G)$ or $\omega(G) + 1$. Moreover, the algorithm runs in time $\text{poly}(d)$. But since it is NP-hard to approximate MAX CLIQUE even within $O(d^{1-o(1)})$ factor [Zuckerman, 2006], the above implies $\text{P} = \text{NP}$.

Part 2: We fix any $g = \text{poly}(G, L, d) = \text{poly}(d)$ and $c > 0$. We choose $\delta = (\frac{1}{576d^8 \cdot g})^{\frac{1}{c}} = \frac{1}{\text{poly}(d)}$ so that $\varepsilon := g \cdot \delta^{2+c} \leq \frac{\delta^2}{576d^2}$. Then by Part 1, we know if there is an algorithm that computes an ε -approximate $\Phi_{\text{All}}(\delta)$ -equilibrium in time $\text{poly}(G, L, d, 1/\delta, 1/\varepsilon) = \text{poly}(d)$ and, then $\text{P} = \text{NP}$. \square

I Hardness for Approximate $\Phi_{\text{Proj}}(\delta)/\Phi_{\text{Int}}(\delta)$ -Equilibrium when $\delta = D$

In the first-order stationary regime $\delta \leq \sqrt{2\varepsilon/L}$, (ε, δ) -local Nash equilibrium is intractable, and we have shown polynomial-time algorithms for computing the weaker notions of ε -approximate $\Phi_{\text{Int}}(\delta)$ -equilibrium and ε -approximate $\Phi_{\text{Proj}}(\delta)$ -equilibrium. A natural question is whether correlation enables efficient computation of ε -approximate $\Phi(\delta)$ -equilibrium when δ is in the global regime, i.e., $\delta = \Omega(\sqrt{d})$. In this section, we prove both computational hardness and a query complexity lower bound for both notions in the global regime

To prove the lower bound results, we only require a single-player game. The problem of computing an ε -approximate $\Phi(\delta)$ -equilibrium becomes: given scalars $\varepsilon, \delta, G, L > 0$ and a polynomial-time Turing machine \mathcal{C}_f evaluating a G -Lipschitz and L -smooth function $f : [0, 1]^d \rightarrow [0, 1]$ and its gradient $\nabla f : [0, 1]^d \rightarrow \mathbb{R}^d$, we are asked to output a distribution σ that is an ε -approximate $\Phi(\delta)$ -equilibrium or \perp if such equilibrium does not exist.

Hardness of finding ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibria in the global regime When $\delta = \sqrt{d}$, which equals to the diameter D of $[0, 1]^d$, then the problem of finding an ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibrium is equivalent to finding a (ε, δ) -local minimum of f : assume σ is an ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibrium of f , then there exists $x \in [0, 1]^d$ in the support of σ such that

$$f(x) - \min_{x^* \in [0, 1]^d \cap B_d(x^*, \delta)} f(x^*) \leq \varepsilon.$$

Then hardness of finding an ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibrium follows from hardness of finding a (ε, δ) -local minimum of f [Daskalakis et al., 2021b]. The following Theorem is a corollary of Theorem 10.3 and 10.4 in [Daskalakis et al., 2021b].

Theorem 13 (Hardness of finding ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibria in the global regime). *In the worst case, the following two holds.*

- Computing an ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibrium for a game on $\mathcal{X} = [0, 1]^d$ with $G = \sqrt{d}$, $L = d$, $\varepsilon \leq \frac{1}{24}$, $\delta = \sqrt{d}$ is NP-hard.
- $\Omega(2^d/d)$ value/gradient queries are needed to determine an ε -approximate $\Phi_{\text{Int}}^{\mathcal{X}}(\delta)$ -equilibrium for a game on $\mathcal{X} = [0, 1]^d$ with $G = \Theta(d^{15})$, $L = \Theta(d^{22})$, $\varepsilon < 1$, $\delta = \sqrt{d}$.

Hardness of finding ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibria in the global regime

Theorem 14 (Hardness of finding ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibria in the global regime). *In the worst case, the following two holds.*

- Computing an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium for a game on $\mathcal{X} = [0, 1]^d$ with $G = \Theta(d^{15})$, $L = \Theta(d^{22})$, $\varepsilon < 1$, $\delta = \sqrt{d}$ is NP-hard.
- $\Omega(2^d/d)$ value/gradient queries are needed to determine an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium for a game on $\mathcal{X} = [0, 1]^d$ with $G = \Theta(d^{15})$, $L = \Theta(d^{22})$, $\varepsilon < 1$, $\delta = \sqrt{d}$.

The hardness of computing ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium also implies a lower bound on $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret in the global regime.

Corollary 3 (Lower bound of $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret against non-convex functions). *In the worst case, the $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -regret of any online algorithm is at least $\Omega(2^d/d, T)$ even for loss functions $f : [0, 1]^d \rightarrow [0, 1]$ with $G, L = \text{poly}(d)$ and $\delta = \sqrt{d}$.*

The proofs of [Theorem 14](#) and [Corollary 3](#) can be found in the next two sections.

I.1 Proof of [Theorem 14](#)

We will reduce the problem of finding an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium in smooth games to finding a satisfying assignment of a boolean function, which is NP-complete.

Fact 1. *Given only black-box access to a boolean formula $\phi : \{0, 1\}^d \rightarrow \{0, 1\}$, at least $\Omega(2^d)$ queries are needed in order to determine whether ϕ admits a satisfying assignment x^* such that $\phi(x^*) = 1$. The term black-box access refers to the fact that the clauses of the formula are not given, and the only way to determine whether a specific boolean assignment is satisfying is by querying the specific binary string. Moreover, the problem of finding a satisfying assignment of a general boolean function is NP-hard.*

We revisit the construction of the hard instance in the proof of [[Daskalakis et al., 2021b](#), Theorem 10.4] and use its specific structures. Given black-box access to a boolean formula ϕ as described in [Fact 1](#), following [[Daskalakis et al., 2021b](#)], we construct the function $f_\phi(x) : [0, 1]^d \rightarrow [0, 1]$ as follows:

1. for each corner $v \in V = \{0, 1\}^d$ of the $[0, 1]^d$ hypercube, we set $f_\phi(x) = 1 - \phi(x)$.
2. for the rest of the points $x \in [0, 1]^d/V$, we set $f_\phi(x) = \sum_{v \in V} P_v(x) \cdot f_\phi(v)$ where $P_v(x)$ are non-negative coefficients defined in [[Daskalakis et al., 2021b](#), Definition 8.9].

The function f_ϕ satisfies the following properties:

1. if ϕ is not satisfiable, then $f_\phi(x) = 1$ for all $x \in [0, 1]^d$ since $f_\phi(v) = 1$ for all $v \in V$; if ϕ has a satisfying assignment v^* , then $f_\phi(v^*) = 0$.
2. f_ϕ is $\Theta(d^{12})$ -Lipschitz and $\Theta(d^{25})$ -smooth.
3. for any point $x \in [0, 1]^d$, the set $V(x) := \{v \in V : P_v(x) \neq 0\}$ has cardinality at most $d + 1$ while $\sum_{v \in V} P_v(x) = 1$; any value / gradient query of f_ϕ can be simulated by $d + 1$ queries on ϕ .

In the case there exists a satisfying argument v^* , then $f_\phi(v^*) = 0$. Define the deviation e so that $e[i] = 1$ if $v^*[i] = 0$ and $e[i] = -1$ if $v^*[i] = 1$. It is clear that $\|e\| = \sqrt{d} = \delta$. By properties of projection on $[0, 1]^d$, for any $x \in [0, 1]^d$, we have $\Pi_{[0, 1]^n}[x - v] = v^*$. Then any ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium σ must include some $x^* \in \mathcal{X}$ with $f_\phi(x^*) < 1$ in the support, since $\varepsilon < 1$. In case there exists an algorithm \mathcal{A} that computes an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium, \mathcal{A} must have queried some x^* with $f_\phi(x^*) < 1$. Since $f_\phi(x^*) = \sum_{v \in V(x^*)} P_v(x^*) f_\phi(v) < 1$, there exists $\hat{v} \in V(x^*)$ such that $f_\phi(\hat{v}) = 0$. Since $|V(x^*)| \leq d + 1$, it takes addition $d + 1$ queries to find \hat{v} with $f_\phi(\hat{v}) = 0$. By [Fact 1](#) and the fact that we can simulate every value/gradient query of f_ϕ by $d + 1$ queries on ϕ , \mathcal{A} makes at least $\Omega(2^d/d)$ value/gradient queries.

Suppose there exists an algorithm \mathcal{B} that outputs an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium σ in time $T(\mathcal{B})$ for $\varepsilon < 1$ and $\delta = \sqrt{d}$. We construct another algorithm \mathcal{C} for SAT that terminates in time $T(\mathcal{B}) \cdot \text{poly}(d)$. \mathcal{C} : (1) given a boolean formula ϕ , construct f_ϕ as described above; (2) run \mathcal{B} and get output σ (3) check the support of σ to find $v \in \{0, 1\}^d$ such that $f_\phi(v) = 0$; (3) if finds $v \in \{0, 1\}^d$ such that $f_\phi(v) = 0$, then ϕ is satisfiable, otherwise ϕ is not satisfiable. Since we can evaluate f_ϕ and ∇f_ϕ in $\text{poly}(d)$ time and the support of σ is smaller than $T(\mathcal{B})$, the algorithm \mathcal{C} terminates in time $O(T(\mathcal{B}) \cdot \text{poly}(d))$. The above gives a polynomial time reduction from SAT to finding an ε -approximate $\Phi_{\text{Proj}}^{\mathcal{X}}(\delta)$ -equilibrium and proves the NP-hardness of the latter problem.

I.2 Proof of Corollary 3

Let $\phi : \{0, 1\}^d \rightarrow \{0, 1\}$ be a boolean formula and define $f_\phi : [0, 1]^d \rightarrow [0, 1]$ the same as that in [Theorem 14](#). We know f_ϕ is $\Theta(\text{poly}(d))$ -Lipschitz and $\Theta(\text{poly}(d))$ -smooth. Now we let the adversary pick f_ϕ each time. For any $T \leq O(2^d/d)$, in case there exists an online learning algorithm with $\text{Reg}_{\text{Proj}, \delta}^T < \frac{T}{2}$, then $\sigma := \frac{1}{T} \sum_{t=1}^T 1_{x^t}$ is an $(\frac{1}{2}, \delta)$ -equilibrium. Applying [Theorem 14](#) and the fact that in this case, $\text{Reg}_{\text{Proj}, \delta}^T$ is non-decreasing with respect to T concludes the proof.

J Removing the D dependence for $\Phi_{\text{Proj}}^{\mathcal{X}}$ -regret

For the regime $\delta \leq D_{\mathcal{X}}$ which we are more interested in, the lower bound in [Theorem 8](#) is $\Omega(G\delta\sqrt{T})$ while the upper bound in [Theorem 3](#) is $O(G\sqrt{\delta D_{\mathcal{X}} T})$. They are not tight, especially when $D_{\mathcal{X}} \gg \delta$. A natural question is: *which of them is the tight bound?* We conjecture that the lower bound is tight. In fact, for the special case where the feasible set \mathcal{X} is a *box*, we have a way to obtain a $D_{\mathcal{X}}$ -independent bound $O(d^{\frac{1}{4}} G\delta\sqrt{T})$, which is tight when $d = 1$. Below, we first describe the improved strategy in 1-dimension. Then we show how to extend it to the d -dimensional box setting.

J.1 One-Dimensional Case

In one-dimension, we assume that $\mathcal{X} = [a, b]$ for some $b - a \geq 2\delta$ (if $b - a \leq 2\delta$, then our original bound in [Theorem 3](#) is already of order $G\delta\sqrt{T}$). We first investigate the case where $f^t(x)$ is a linear function, i.e.,

$f^t(x) = g^t x$ for some $g^t \in [-G, G]$. The key idea is that we will only select x^t from the two intervals $[a, a + \delta]$ and $[b - \delta, b]$, and never play $x^t \in (a + \delta, b - \delta)$. To achieve this, we concatenate these two intervals and run an algorithm in this region whose diameter is only 2δ . The key property we would like to show is that the regret is preserved in this modified problem.

More precisely, given the original feasible set $\mathcal{X} = [a, b]$, we create a new feasible set $\mathcal{Y} = [-\delta, \delta]$ and apply our algorithm **GD** in this new feasible set. The loss function is kept as $f^t(x) = g^t x$. Whenever the algorithm for \mathcal{Y} outputs $y^t \in [-\delta, 0]$, we play $x^t = y^t + a + \delta$ in \mathcal{X} ; whenever it outputs $y^t \in (0, \delta]$, we play $x^t = y^t + b - \delta$. Below we show that the regret is the same in these two problems. Notice that when $y^t \leq 0$, we have for any $v \in [-\delta, \delta]$,

$$\begin{aligned}
x^t - \Pi_{\mathcal{X}}[x^t - v] &= x^t - \max(\min(x^t - v, b), a) \\
&= x^t - \max(x^t - v, a) \quad (x^t - v = y^t + a + \delta - v \leq a + 2\delta \leq b \text{ always holds}) \\
&= y^t + a + \delta - \max(y^t + a + \delta - v, a) \\
&= y^t - \max(y^t - v, -\delta) \\
&= y^t - \max(\min(y^t - v, \delta), -\delta) \quad (y^t - v \leq \delta \text{ always holds}) \\
&= y^t - \Pi_{\mathcal{Y}}[y^t - v]
\end{aligned}$$

Similarly, when $y^t > 0$, we can follow the same calculation and prove $x^t - \Pi_{\mathcal{X}}[x^t - v] = y^t - \Pi_{\mathcal{Y}}[y^t - v]$. Thus, the regret in the two problems:

$$g^t (x^t - \Pi_{\mathcal{X}}[x^t - v]) \quad \text{and} \quad g^t (y^t - \Pi_{\mathcal{Y}}[y^t - v])$$

are exactly the same for any v . Finally, observe that the diameter of \mathcal{Y} is only of order $O(\delta)$. Thus, the upper bound in Theorem 3 would give us an upper bound of $O(G\sqrt{\delta} \cdot \delta T) = O(G\delta\sqrt{T})$.

For convex f^t , we run the algorithm above with $g^t = \nabla f^t(x^t)$. Then by convexity, we have

$$f^t(x^t) - f^t(\Pi_{\mathcal{X}}[x^t - v]) \leq g^t(x^t - \Pi_{\mathcal{X}}[x^t - v]) = g^t(y^t - \Pi_{\mathcal{Y}}[y^t - v]),$$

so the regret in the modified problem (which is $O(G\delta\sqrt{T})$) still serves as a regret upper bound for the original problem.

J.2 d -Dimensional Box Case

A d -dimensional box is of the form $\mathcal{X} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$. The box case is easy to deal with because we can decompose the regret into individual components in each dimension. Namely, we have

$$\begin{aligned}
f^t(x^t) - f^t(\Pi_{\mathcal{X}}[x^t - v]) &\leq \nabla f^t(x^t)^\top (x^t - \Pi_{\mathcal{X}}[x^t - v]) \\
&= \sum_{i=1}^d g_i^t (x_i^t - \Pi_{\mathcal{X}_i}[x_i^t - v_i])
\end{aligned}$$

where we define $\mathcal{X}_i = [a_i, b_i]$, $g^t = \nabla f^t(x^t)$, and use subscript i to indicate the i -th component of a vector. The last equality above is guaranteed by the box structure. This decomposition allows us to view the problem as d independent 1-dimensional problems.

Now we follow the strategy described in Section ?? to deal with individual dimensions (if $b_i - a_i < 2\delta$, then we do not modify \mathcal{X}_i ; otherwise, we shrink \mathcal{X}_i to be of length 2δ). Applying the analysis of Theorem 3

to each dimension, we get

$$\begin{aligned}
& \sum_{i=1}^d g_i^t (x_i^t - \Pi_{\mathcal{X}_i}[x_i^t - v_i]) \\
& \leq \sum_{i=1}^d \left(\frac{v_i^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^T (g_i^t)^2 + \frac{|v_i| \times 2\delta}{\eta} \right) \quad (\text{the diameter in each dimension is now bounded by } 2\delta) \\
& \leq O \left(\frac{\delta \sum_{i=1}^d |v_i|}{\eta} + \eta G^2 T \right) \\
& \leq O \left(\frac{\delta^2 \sqrt{d}}{\eta} + \eta G^2 T \right). \quad (\text{by Cauchy-Schwarz, } \sum_i |v_i| \leq \sqrt{d} \sqrt{\sum_i |v_i|^2} \leq \delta \sqrt{d})
\end{aligned}$$

Choosing the optimal $\eta = \frac{d^{\frac{1}{4}} \delta}{G \sqrt{T}}$, we get the regret upper bound of order $O \left(d^{\frac{1}{4}} G \delta \sqrt{T} \right)$.