
On the Interplay between Social Welfare and Tractability of Equilibria

Ioannis Anagnostides
Carnegie Mellon University
ianagnos@cs.cmu.edu

Tuomas Sandholm
Carnegie Mellon University
Strategic Machine, Inc.
Strategy Robot, Inc.
Optimized Markets, Inc.
sandholm@cs.cmu.edu

Abstract

Computational tractability and social welfare (aka. efficiency) of equilibria are two fundamental but in general orthogonal considerations in algorithmic game theory. Nevertheless, we show that when (approximate) *full efficiency* can be guaranteed via a *smoothness* argument à la Roughgarden, Nash equilibria are approachable under a family of no-regret learning algorithms, thereby enabling fast and decentralized computation. We leverage this connection to obtain new convergence results in *large games*—wherein the number of players $n \gg 1$ —under the well-documented property of full efficiency via smoothness in the limit. Surprisingly, our framework unifies equilibrium computation in disparate classes of problems including games with vanishing *strategic sensitivity* and two-player zero-sum games, illuminating en route an immediate but overlooked equivalence between smoothness and a well-studied condition in the optimization literature known as the *Minty property*. Finally, we establish that a family of no-regret dynamics attains a welfare bound that improves over the smoothness framework while at the same time guaranteeing convergence to the set of coarse correlated equilibria. We show this by employing the *clairvoyant* mirror descent algorithm recently introduced by Piliouras *et al.*

1 Introduction

The *Nash equilibrium (NE)* [76] formalizes the notion of a *stable* outcome in a multiagent strategic interaction, and has arguably served as the most influential solution concept in the development of game theory. Indeed, algorithms designed to approximate Nash equilibria in two-player zero-sum games have recently resolved major challenges in AI [7, 10, 81]. Its prescriptive power, however, has been severely undermined in multi-player general-sum games by intrinsic computational barriers [30, 20, 94, 35, 28, 44], limitations which also manifest in the inability of natural learning algorithms to converge [73, 105, 60, 70]. Another well-established but orthogonal critique is the *equilibrium selection* problem [52]: a general-sum game may have multiple Nash equilibria with widely different welfare. As a result, a modern research agenda in computational game theory has been to identify and characterize natural classes of games that circumvent those fundamental limitations.

In this paper, we uncover a new natural class of games for which the aforementioned caveats of Nash equilibria can be effectively addressed. In particular, our investigation originates from a natural question: when are *efficient*—in terms of social welfare—Nash equilibria easy to compute? The answer to this question is, at first glance, unsatisfactory: even if Nash equilibria are fully efficient, computational hardness still persists given that constant-sum multi-player games are hard. Indeed, efficiency and computational tractability are, in general, two orthogonal considerations. We show,

however, an interesting twist, an unexpected interplay between efficiency—when viewed from a specific lens—and the behavior of a family of *no-regret* learning algorithms.

1.1 Our results

To elucidate the alluded connection that drives much of our results, we first have to recall that the canonical paradigm for establishing the efficiency of equilibria is Roughgarden’s celebrated *smoothness* framework [90] (exposed thoroughly in Section 2). In this context, we observe that *if full efficiency of equilibria can be guaranteed via a smoothness argument (with bounded parameters), then Nash equilibria are approachable under a family of no-regret learning algorithms*. In other words, full efficiency *via smoothness* implies computational tractability of NE. This is surprising in that Roughgarden’s smoothness framework was developed primarily in order to automatically extend *price of anarchy (PoA)* bounds to more permissive equilibrium concepts, such as *coarse correlated equilibria (CCE)*; tractability of NE appears at first glance entirely unrelated. In fact, while applying the smoothness framework to (multi-player) constant-sum games appears to make little sense, given that PoA considerations are trivial in such games (all outcomes attain the same social welfare), it turns out that a certain regime of smoothness in (multi-player) constant-sum games is equivalent to a well-known condition in the optimization literature called the *Minty property* [36] (Observation 3.4).

The condition described above already captures—somewhat unexpectedly—well-studied settings, such as games that admit a minimax theorem [13], but we have found that the most fertile and novel ground to apply this theory revolves around *large games*, that is, games with a large number of players $n \gg 1$.¹ The reason we focus on large games is a well-documented economic phenomenon: equilibria in large games approach—under natural conditions—full efficiency as $n \rightarrow +\infty$, a property often established via smoothness [39, 23]—a crucial ingredient in our framework. (One rough intuition for this is that in large games each player’s influence on the outcome becomes negligible, making optimal behavior easier to characterize [39].) To state our first main result, we denote by (λ, μ) a pair of bounded *smoothness* parameters of a game \mathcal{G} ; the ratio $\rho(\lambda, \mu) := \frac{\lambda}{1+\mu}$ circumscribes the (in)efficiency of equilibria of \mathcal{G} , in that all equilibria will attain at least a ρ fraction of the optimal welfare.

Theorem 1.1 (Informal). *Consider a sequence of n -player (λ_n, μ_n) -smooth games $(\mathcal{G}_n)_{n \geq 1}$ such that $\rho_n := \frac{\lambda_n}{1+\mu_n} \rightarrow 1$ with a sufficiently fast rate. Then, there are decentralized and computationally efficient no-regret dynamics approaching an $(o_n(1), o_n(1))$ -weak Nash equilibrium.*

A few remarks are in order. First, in Theorems 3.1 and A.2 we give a precise non-asymptotic characterization that quantifies the number of iterations as well as the approximation error. We also recall that in an $(o_n(1), o_n(1))$ -weak Nash equilibrium *almost all* players are *almost best responding* (Definition 2.2); this is a well-studied relaxation for which hardness results carry over in general [4, 2, 5, 94]. A limitation of a weak Nash equilibrium is that it can prescribe a strategy profile in which a large number of players—albeit of vanishing fraction—can significantly benefit from deviating; this limitation is inherent in a certain regime of Theorem 1.1. Depending on the rate with which ρ_n approaches full efficiency, Theorem 1.1 can also imply convergence to the usual notion of Nash equilibrium—wherein *all* players are almost best responding (Corollary A.5). More generally, for a broad class of games that includes *graphical* games with bounded degree, *polymatrix* games, and games exhibiting small *strategic sensitivity*, the conclusion of Theorem 1.1 applies without imposing any restrictions on the rate of convergence of ρ_n ; those are the most commonly studied classes of games when $n \gg 1$.

There are ample compelling aspects in connecting the convergence of no-regret learning algorithms with Roughgarden’s smoothness framework. First, there has been a considerable interest in understanding smoothness, insights that can now be inherited to a seemingly entirely different but equally fundamental problem. For example, smoothness naturally extends to *Bayesian games* [101, 91, 53] (see also [97] for the related class of contextual games), a property that we leverage in Theorem 3.8 to expand our scope to mechanisms of incomplete information. Additional extensions based on *local smoothness* [92, 78] and the refined primal-dual framework of Nadav and Roughgarden [75] are also described in Appendices A.3 and A.4. Finally, our criterion has a clear and natural economic interpretation, and is part of an ongoing research effort to identify tractable classes of variational inequalities (VIs) beyond the Minty property [33].

¹We use the terminology large games here—in accordance with much of the economics literature—to refer to games with a large number of players; it should not be confused with another common usage referring to games with a “large” action space.

As a concrete example, our framework subsumes games wherein each player’s effect on the outcome vanishes as $n \rightarrow \infty$; this captures, for example, simple voting settings [57], as well as general auction design problems where establishing that property turns out to be highly non-trivial and quite delicate [39]. More concretely, for games with vanishing *sensitivity* $\epsilon_n \rightarrow 0$, in that unilateral deviations can only affect a player’s utility by an additive ϵ_n (Definition 3.5), we show that the conclusion of Theorem 1.1 applies as long as $\epsilon_n \leq o_n(1/\sqrt{n})$ (Theorem 3.6). At the same time, the condition that $\rho_n \rightarrow 1$ goes much deeper than games with vanishing sensitivity, and, as we explained earlier, surprisingly applies to two-player zero-sum games as well. We find it conceptually appealing that a unifying framework can establish tractability of Nash equilibria in two seemingly disparate classes of problems such as two-player zero-sum games and games with vanishing strategic sensitivity.

Remaining on smooth games, but relaxing the assumption $\rho \approx 1$, we next study conditions under which the efficiency guaranteed by the smoothness framework can be improved, while at the same time ensuring the no-regret property, thereby implying convergence to the set of CCE. The smoothness bound is known to be applicable to *any* outcome of no-regret dynamics, but here we are instead interested in more refined guarantees when specific learning algorithms are in place. Building on a recent result [1], we show in Theorem 4.1 that the *clairvoyant* variant of gradient descent, introduced by Piliouras et al. [85], enjoys an improved welfare bound *and* ensures fast convergence to the set of CCE for the average correlated distribution of play. Crucially, compared to an earlier result [1], the clairvoyant algorithm manages to satisfy an appealing notion of per-player incentive compatibility, in the form of convergence to CCE. In other words, improving the welfare predicted by smoothness is not at odds with incentive compatibility (Corollary 4.3).

2 Background

Notation We let $\mathbb{N} = \{1, 2, \dots\}$ be the set of natural numbers. For $n \in \mathbb{N}$, we denote by $\llbracket n \rrbracket := \{1, 2, \dots, n\}$. For a vector $\mathbf{x} \in \mathbb{R}^d$, we denote by $\|\mathbf{x}\|_2$ its (Euclidean) ℓ_2 norm. For a convex, compact and nonempty set \mathcal{X} , we let $\mathcal{P}_{\mathcal{X}}(\cdot)$ represent the Euclidean projection operator with respect to \mathcal{X} . We let $D_{\mathcal{X}}$ be the ℓ_2 -diameter of \mathcal{X} . To simplify the exposition, we often use the $O(\cdot)$ notation in the main body to suppress the dependence on certain parameters; we also write $O_n(\cdot)$ to indicate the asymptotic growth solely as a function of n , so as to lighten the exposition.

Multilinear games We consider n -player *multilinear* games. In a multilinear game \mathcal{G} each player $i \in \llbracket n \rrbracket$ has a convex, compact and nonempty set of feasible strategies $\mathcal{X}_i \subseteq \mathbb{R}^{d_i}$, for some dimension $d_i \in \mathbb{N}$. Under a joint strategy profile $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \prod_{i=1}^n \mathcal{X}_i =: \mathcal{X}$, there is a continuous utility function $u_i : (\mathbf{x}_1, \dots, \mathbf{x}_n) \mapsto \mathbb{R}$ such that $u_i(\mathbf{x}) = \langle \mathbf{x}_i, \mathbf{u}_i(\mathbf{x}_{-i}) \rangle$, for some function $\mathbf{u}_i : \mathbf{x}_{-i} \mapsto \mathbb{R}^{d_i}$; here, we used for convenience the standard notation $\mathbf{x}_{-i} := (\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n)$. This setup readily captures *normal-* as well as *extensive-form* games.

Specifically, in a normal-form game every player $i \in \llbracket n \rrbracket$ selects as strategy a probability distribution $\mathbf{x}_i \in \Delta(\mathcal{A}_i)$ over a finite set of available actions \mathcal{A}_i . There is an arbitrary utility function $u_i : \mathcal{A} := \prod_{i=1}^n \mathcal{A}_i \rightarrow [-1, 1]$ that maps a joint action profile to a utility for Player i ; we will be making the standard assumption that the range of each player’s utilities is bounded by an absolute constant, which is in particular independent of the number of players n . In this setting, the mixed extension of the utility function indeed satisfies the multilinearity condition imposed above: for $\mathbf{x} \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$, $u_i(\mathbf{x}) := \mathbb{E}_{\mathbf{a} \sim \mathbf{x}}[u_i(\mathbf{a})] = \langle \mathbf{x}_i, \nabla_{\mathbf{x}_i} u_i(\mathbf{x}) \rangle$, where we overloaded the notation $u_i(\cdot)$.

Returning to the general setting of multilinear games, we let $F_{\mathcal{G}} : \prod_{i=1}^n \mathcal{X}_i \rightarrow \prod_{i=1}^n \mathcal{X}_i$ denote the underlying *operator* of the game \mathcal{G} , defined as $F_{\mathcal{G}} : (\mathbf{x}_1, \dots, \mathbf{x}_n) \mapsto (\mathbf{u}_1(\mathbf{x}_{-1}), \dots, \mathbf{u}_n(\mathbf{x}_{-n}))$, where function \mathbf{u}_i was introduced earlier. For notational simplicity, we will often omit the subscript \mathcal{G} when it is clear from the context. We will say that F is *L-Lipschitz continuous* (w.r.t. the ℓ_2 norm) if for any $\mathbf{x}, \mathbf{x}' \in \prod_{i=1}^n \mathcal{X}_i$ it holds that $\|F(\mathbf{x}) - F(\mathbf{x}')\|_2 \leq L\|\mathbf{x} - \mathbf{x}'\|_2$.

Welfare and the price of anarchy For a joint strategy profile $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$, we define the *social welfare* attained under \mathbf{x} as $\text{SW}(\mathbf{x}) := \sum_{i=1}^n u_i(\mathbf{x})$. The maximum possible social welfare attainable in a game \mathcal{G} will be denoted by $\text{OPT}_{\mathcal{G}}$. Without any essential loss of generality, it will be assumed that $\text{OPT}_{\mathcal{G}} > 0$. We say that the game is *constant-sum* if $\text{SW}(\mathbf{x}) = V \in \mathbb{R}_{>0}$ for any $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$. The *price of anarchy (PoA)* of a game \mathcal{G} quantifies the loss in efficiency incurred on

account of strategic players [61]. Formally, if $\text{NE}_{\mathcal{G}}$ is the nonempty set of (mixed) Nash equilibria of \mathcal{G} (Definition 2.2) [88], we define $\text{PoA}_{\mathcal{G}} := \sup_{\mathbf{x} \in \text{NE}_{\mathcal{G}}} \left\{ \frac{\text{SW}(\mathbf{x})}{\text{OPT}_{\mathcal{G}}} \right\}$.

Smooth games We are now ready to recall the seminal notion of a *smooth game*,² conceived in the pioneering work of Roughgarden [90] as a technique to (lower) bound the price of anarchy.

Definition 2.1 (Smooth game [90]). An n -player game \mathcal{G} is called (λ, μ) -smooth, where $\lambda > 0$ and $\mu > -1$, if there exists $\mathbf{x}^* \in \prod_{i=1}^n \mathcal{X}_i$ with $\text{SW}(\mathbf{x}^*) = \text{OPT}_{\mathcal{G}}$ such that for every $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$,

$$\sum_{i=1}^n u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}) \geq \lambda \text{OPT}_{\mathcal{G}} - \mu \text{SW}(\mathbf{x}). \quad (1)$$

(In the definition above, and throughout this paper, we slightly abuse notation by parsing $u_i(\mathbf{x}_i^*, \mathbf{x}_{-i})$ as $u_i(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i^*, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n)$.) Roughgarden [90] observed that in a (λ, μ) -smooth game, in the sense of Definition 2.1, every Nash equilibrium attains at least a $\rho(\lambda, \mu) := \frac{\lambda}{1+\mu}$ fraction of the optimal social welfare $\text{OPT}_{\mathcal{G}}$. Importantly, this efficiency guarantee immediately carries over to outcomes of no-regret learning algorithms as well. The *robust price of anarchy* ($\text{rPoA}_{\mathcal{G}}$) is the best (i.e., largest) price of anarchy bound provable via a smoothness argument, and can be defined as the solution to the linear program induced by the smoothness constraints given in (1) (see (24) in Appendix A.5). One delicate point here is that the value of $\text{rPoA}_{\mathcal{G}}$ could be associated with unbounded smoothness parameters, which is a pathological and rather trivial manifestation of smoothness (Remark A.10 elaborates on this point). To be clear, when we say a game \mathcal{G} attains a certain value $\rho_{\mathcal{G}}$, we mean that there exists a finite pair of legitimate smoothness parameters (λ, μ) such that $\rho_{\mathcal{G}} = \frac{\lambda}{1+\mu}$. With this convention, it might be the case that $\rho_{\mathcal{G}}(\lambda, \mu) \neq \text{rPoA}_{\mathcal{G}}$ under any finite pair of smoothness parameters (λ, μ) (Remark A.10). It is also easy to see that $\text{PoA}_{\mathcal{G}} \geq \text{rPoA}_{\mathcal{G}}$.

Nash equilibrium We next recall the concept of a *weak Nash equilibrium*, a natural generalization of the standard notion which is meaningful in multi-player games.

Definition 2.2 (Weak Nash equilibrium [5]). Let $\delta \in [0, 1]$ and $\epsilon \in \mathbb{R}_{\geq 0}$. A joint strategy profile $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$ is an (ϵ, δ) -weak Nash equilibrium if at least a $1 - \delta$ fraction of the players are ϵ -best responding. An $(\epsilon, 0)$ -weak Nash equilibrium will be simply referred to as ϵ -Nash equilibrium.

In the definition above, we clarify that a player $i \in \llbracket n \rrbracket$ is said to be ϵ -best responding if $\text{BRGAP}_i(\mathbf{x}_1, \dots, \mathbf{x}_n) := \max_{\mathbf{x}'_i \in \mathcal{X}_i} \langle \mathbf{x}'_i, \mathbf{u}_i(\mathbf{x}_{-i}) \rangle - \langle \mathbf{x}_i, \mathbf{u}_i(\mathbf{x}_{-i}) \rangle \leq \epsilon$. We also define the *Nash equilibrium gap* as $\text{NEGAP}(\mathbf{x}) := \max_{1 \leq i \leq n} \text{BRGAP}_i(\mathbf{x})$. It has been shown that hardness results in multi-player general-sum games persist under the weak Nash equilibrium concept introduced above, even when ϵ and δ are absolute constants bounded away from 0 [4, 94].

Regret We are operating in the usual online learning setting. At every time $t \in \mathbb{N}$ a player $i \in \llbracket n \rrbracket$ selects a strategy $\mathbf{x}_i^{(t)} \in \mathcal{X}_i$, and then receives as feedback the linear utility function $\mathbf{x}_i \mapsto \langle \mathbf{x}_i, \mathbf{u}_i^{(t)} \rangle$, where we overload notation so that $\mathbf{u}_i^{(t)} := \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})$. The *regret* of player $i \in \llbracket n \rrbracket$ under a time horizon $T \in \mathbb{N}$ is defined as

$$\text{Reg}_i^{(T)} := \max_{\mathbf{x}_i^* \in \mathcal{X}_i} \left\{ \sum_{t=1}^T u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}^{(t)}) \right\} - \sum_{t=1}^T u_i(\mathbf{x}_i^{(t)}).$$

Optimistic gradient descent By now, there are many online algorithms known to guarantee *sublinear* regret, $\text{Reg}_i^{(T)} = o(T)$, even if the observed utilities are selected adversarially [18, 98]. The main no-regret learning algorithm we consider in this paper is *optimistic gradient descent* (henceforth OGD) [86, 22], which is known to yield improved regret guarantees in the setting of learning in games [103]. For each player $i \in \llbracket n \rrbracket$, OGD is defined through the following update rule for $t \in \mathbb{N}$.

$$\begin{aligned} \mathbf{x}_i^{(t)} &:= \mathcal{P}_{\mathcal{X}_i} \left(\hat{\mathbf{x}}_i^{(t)} + \eta \mathbf{m}_i^{(t)} \right), \\ \hat{\mathbf{x}}_i^{(t+1)} &:= \mathcal{P}_{\mathcal{X}_i} \left(\hat{\mathbf{x}}_i^{(t)} + \eta \mathbf{u}_i^{(t)} \right). \end{aligned} \quad (\text{OGD})$$

²Smoothness in the sense of Definition 2.1 should not be confused with the unrelated notion of smoothness in the optimization nomenclature.

Here, $\eta > 0$ is the *learning rate*; $\mathbf{m}_i^{(t)} \in \mathbb{R}^{d_i}$ is the *prediction vector*; and $\hat{\mathbf{x}}_i^{(1)} \in \mathcal{X}_i$ is the *initialization*. It is assumed that the strategy set \mathcal{X}_i is such that the Euclidean projection $\mathcal{P}_{\mathcal{X}_i}(\cdot)$ can be implemented efficiently. We will let $\mathbf{m}_i^{(t)} := \mathbf{u}_i^{(t-1)}$ for any $t \in \mathbb{N}$, where $\mathbf{u}_i^{(0)} := \mathbf{u}_i(\hat{\mathbf{x}}_i^{(1)})$.

3 Convergence to Nash equilibria via smoothness

In this section, we study the convergence of optimistic gradient descent (OGD) in large games, that is to say, in the regime $n \gg 1$. In our first main result, stated below as Theorem 3.1, we show that when $\lim_{n \rightarrow +\infty} \rho_n = 1$ with a sufficiently fast rate, not only are Nash equilibria approaching the optimal social welfare, but they can also be computed efficiently in a decentralized fashion via OGD.

In the sequel, when considering a sequence of games $(\mathcal{G}_n)_{n \geq 1}$, with each game \mathcal{G}_n being parameterized by the number of players n , we will use a subscript with variable n to index the n th game in the sequence; that notation will also be used to refer to the other underlying parameters of the game that depend on the number of players.

Theorem 3.1. *Consider an n -player (λ, μ) -smooth game \mathcal{G}_n such that the game operator F_n is L_n -Lipschitz continuous and $\lambda \geq (1 - \epsilon_n)(1 + \mu)$, with $\epsilon_n \neq 0$. Suppose further that all players follow OGD with learning rate $\eta_n = 1/(4L_n)$. Then, for $\delta \in (0, 1)$ and a sufficiently large number of iterations $T = O_n(nL_n^2/\gamma_n)$, where $\gamma_n := L_n \text{OPT}_{\mathcal{G}_n} \epsilon_n$, there is time $t^* \in \llbracket T \rrbracket$ such that $\mathbf{x}^{(t^*)} := (\mathbf{x}_1^{(t^*)}, \dots, \mathbf{x}_n^{(t^*)})$ is a*

$$\left(\frac{1}{\sqrt{\delta}} O_n \left(\sqrt{\frac{\gamma_n}{n}} \right), \delta \right) \text{--weak Nash equilibrium.} \quad (2)$$

In particular, if $\gamma_n = O_n(n^{1-\alpha})$, for $\alpha \in (0, 1)$, OGD yields an $(O_n(n^{-\frac{\alpha}{3}}), O_n(n^{-\frac{\alpha}{3}}))$ -weak NE.

We remark that if it further holds that $\gamma_n = o_n(1)$, the above theorem establishes convergence to the standard notion of Nash equilibrium—wherein *all* players are (almost) best responding (Corollary A.5). Parameter γ_n , which is proportional to the error term ϵ_n , controls both the number of iterations and the approximation guarantee in (2); we refer to Theorem A.2 (in Appendix A) for a more precise non-asymptotic characterization, which bounds the players' cumulative best response gap as a function of the growth of γ_n . We also note that Theorem 3.1 can be strengthened so that (2) holds for *most* iterates of OGD (say 99%), not just a single one (Remark A.4).

Now, to be more concrete regarding the preconditions of Theorem 3.1, we first observe that the growth of the Lipschitz constant L_n depends on the normalization assumptions as well as the structure of the underlying game \mathcal{G}_n . In particular, let us assume—as is standard—that the range of the utility functions is independent of n , in which case $\text{OPT}_{\mathcal{G}_n} = O_n(n)$. We then show that $L_n = O_n(1)$ in each of the following cases: graphical games with bounded degree (Lemma A.7), games with $O_n(1/n)$ strategic sensitivity per Definition 3.5 (Lemma A.8), and polymatrix (general-sum) games even with unbounded neighborhoods (Lemma A.9); the first two of the aforementioned classes are the most commonly studied classes in the literature under the regime $n \gg 1$. For all of those classes, applying Theorem 3.1 yields an $(o_n(1), o_n(1))$ -weak Nash equilibrium for any $\epsilon_n \leq o_n(1)$. More generally, the Lipschitz constant L_n can grow with n (Lemma A.6), in which case ϵ_n must vanish with a faster rate for the conclusion of Theorem 3.1 to kick in. One important limitation of Theorem 3.1 is that, at least in a certain regime, it prescribes a strategy in which many players—albeit a vanishing fraction—may have a profitable deviation (in accordance with Definition 2.2); this is admittedly an inherent feature of our framework.

Remark 3.2. In a decentralized environment, one question that arises from Theorem 3.1 concerns the identification of a time index $t^* \in \llbracket T \rrbracket$ that satisfies (2), which can be viewed as the stopping condition of the algorithm. We suggest two possible approaches. First, if we accept that players have access to a common source of randomness, then all players can sample the same index t^* uniformly at random from the set $\llbracket T \rrbracket$. As we point out in Remark A.4, this suffices to provide a guarantee with high probability with only a small degradation in the solution quality. The second approach, which does not rest any having a common source of randomness, involves a coordinator who can communicate with the players but possesses no information whatsoever about the underlying game. The coordinator sets a target solution quality parameterized by (ϵ, δ) (per Definition 2.2), and after each iteration t elicits from each player $i \in \llbracket n \rrbracket$ a single bit, encoding whether $\mathbb{1}\{\text{BRGAP}_i(\mathbf{x}^{(t)}) > \epsilon\}$. (We note that each player can indeed determine its best response gap with only its local information—namely,

the utility feedback.) The coordinator can then evaluate whether the fraction of the players with at most an ϵ best response gap matches the desired accuracy. While the second approach makes for a less decentralized protocol, the communication overhead described above is arguably very limited.

The key precondition of Theorem 3.1 pertains the behavior of the smoothness parameters, to be discussed next after we first sketch the proof of Theorem 3.1, which extends a recent technique [1].

Proof sketch of Theorem 3.1 The proof is based on the fact that the sum of the players' regrets $\sum_{i=1}^n \text{Reg}_i^{(T)}$ cannot be too negative, which in turn follows from the assumption that $\rho_n \geq 1 - \epsilon_n$. The argument then proceeds by bounding the players' cumulative best response gap across the T iterations $\sum_{t=1}^T \sum_{i=1}^n (\text{BRGAP}_i(\mathbf{x}^{(t)}))^2$ as a function of γ_n and the time horizon $T \in \mathbb{N}$, ultimately leading to the conclusion of Theorem 3.1.

Connection with the Minty property While we have stated Theorem 3.1 in the regime $n \gg 1$, under the premise that ρ_n is sufficiently close to 1 (as a function of n), its conclusion is in fact interesting beyond that regime. Indeed, an important observation is that any two-player constant-sum game $\mathcal{G} := (\mathbf{A}, \mathbf{B})$, with $\langle \mathbf{x}_1, \mathbf{A}\mathbf{x}_2 \rangle + \langle \mathbf{x}_1, \mathbf{B}\mathbf{x}_2 \rangle = V$ for any $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, satisfies $\rho_{\mathcal{G}} = 1$. This is indeed a consequence of Von Neumann's minimax theorem: $\exists(\mathbf{x}_1^*, \mathbf{x}_2^*) \in \mathcal{X}_1 \times \mathcal{X}_2$ such that $u_1(\mathbf{x}_1^*, \mathbf{x}_2) + u_2(\mathbf{x}_2^*, \mathbf{x}_1) - V = \langle \mathbf{x}_1^*, \mathbf{A}\mathbf{x}_2 \rangle - \langle \mathbf{x}_1, \mathbf{A}\mathbf{x}_2^* \rangle \geq 0$ for any $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, in turn implying that $u_1(\mathbf{x}_1^*, \mathbf{x}_2) + u_2(\mathbf{x}_2^*, \mathbf{x}_1) \geq (1 + \mu)\text{OPT}_{\mathcal{G}} - \mu\text{SW}(\mathbf{x}) = V$; that is, any two-player constant-sum game is $(1 + \mu, \mu)$ -smooth, thereby making the conclusion of Theorem 3.1 readily applicable (by taking $\rho \geq 1 - \epsilon^2$ for any $\epsilon > 0$; see the explicit statement of Theorem A.2). This captures and unifies earlier iteration complexity bounds under OGD and the extra-gradient method [15, 48, 46].

Proposition 3.3. Any two-player constant-sum game \mathcal{G} is $(1 + \mu, \mu)$ -smooth for $\mu > -1$, implying that $\rho_{\mathcal{G}} = 1$. Thus, $\tilde{O}(1/\epsilon^2)$ iterations of OGD suffice to obtain an ϵ -Nash equilibrium, for any $\epsilon > 0$.

More broadly, there is a surprisingly overlooked but immediate connection between smooth games (per Definition 2.1) and the *Minty property*, a well-known condition in the literature on variational inequalities (VIs) [36, 66, 71]. More precisely, the Minty property postulates the existence of a strategy profile $\mathbf{x}^* \in \prod_{i=1}^n \mathcal{X}_i$ such that $\langle \mathbf{x}^* - \mathbf{x}, F(\mathbf{x}) \rangle \geq 0$ for any $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$, where we recall that F is the operator of the game. (We caution that the last inequality is typically stated with the opposite sign since the operator F is defined oppositely.) The following connection is thus immediate from the fact that $\langle \mathbf{x}^*, F(\mathbf{x}) \rangle = \sum_{i=1}^n \langle \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{x}_{-i}) \rangle = \sum_{i=1}^n u_i(\mathbf{x}_i^*, \mathbf{x}_{-i})$ and $\langle \mathbf{x}, F(\mathbf{x}) \rangle = \sum_{i=1}^n \langle \mathbf{x}_i, \mathbf{u}_i(\mathbf{x}_{-i}) \rangle = \text{SW}(\mathbf{x})$ (by multilinearity).

Observation 3.4. For any (multi-player) constant-sum game \mathcal{G} , the Minty property is equivalent to \mathcal{G} being $(1 + \mu, \mu)$ -smooth for some $\mu > -1$.

Indeed, the Minty property implies that $\sum_{i=1}^n u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}) \geq V = (1 + \mu)\text{OPT}_{\mathcal{G}} - \mu\text{SW}(\mathbf{x})$, and the converse direction is also immediate. In fact, for (multi-player) zero-sum games, the Minty property is equivalent to \mathcal{G} satisfying (1) under some pair $(\lambda, \mu) \in \mathbb{R}^2$. We stress that even if $\rho_{\mathcal{G}} = 1$, traditional no-regret learning algorithms such as online mirror descent do not generally enjoy iterate convergence to Nash equilibria [70], which stands in stark contrast to the behavior of OGD (Proposition 3.3). In light of Observation 3.4, Theorem 3.1 should also be viewed as part of an ongoing effort to establish sufficient conditions of tractability that are more permissive than the Minty property (e.g., [33, 12, 14, 84, 11]). The criterion we furnish herein, based on the smoothness framework, has the important benefit of enjoying a natural economic interpretation, as well as having being extensively studied in the literature. Indeed, we will leverage insights from prior work to obtain several interesting extensions in the remainder of this section.

Games with vanishing sensitivity Returning to the regime $n \gg 1$, why should we expect $\rho_n \rightarrow 1$? Indeed, if anything large games are more general than games with a small number of players since one can always incorporate “dummy” players into the game. Yet, the point is that large games oftentimes exhibit a structure that leads to more efficient outcomes. For example, one immediate implication of our framework relates to games with a *vanishing (strategic) sensitivity* (see, e.g., [57, 3, 72, 31, 45]). There are various ways of defining sensitivity; here, we adopt the following standard definition.

Definition 3.5. The strategic *sensitivity* $\epsilon \in \mathbb{R}_{>0}$ of an n -player game in normal form is defined as

$$\epsilon := \max_{1 \leq i \leq n} \max_{\mathbf{a} \in \mathcal{A}} \max_{1 \leq i' \leq n} \max_{\mathbf{a}'_{i'} \in \mathcal{A}_{i'}} |u_i(\mathbf{a}'_{i'}, \mathbf{a}_{-i}) - u_i(\mathbf{a})|.$$

In words, a unilateral deviation can only impact a player's utility by an additive ϵ . Now, as long as the sensitivity decays fast enough, a proof analogous to that of Theorem 3.1 implies the following.

Theorem 3.6. *Consider an n -player game \mathcal{G}_n with sensitivity $\epsilon_n \in \mathbb{R}_{>0}$. Then, $T = O_n(n)$ iterations of OGD suffice to obtain a $\left(\frac{1}{\sqrt{\delta}} O_n(\epsilon_n \sqrt{n}), \delta\right)$ -weak Nash equilibrium, for $\delta \in (0, 1)$.*

In particular, Theorem 3.6 yields an $(o_n(1), o_n(1))$ -weak Nash equilibrium as long as $\epsilon_n = o_n(\sqrt{n})$. Further, in the canonical regime where $\epsilon_n = O_n(1/n)$, Theorem 3.6 circumscribes the best response gap for all but a constant number of players (by taking $\delta = O_n(1/n)$). There are many natural settings where we should expect results such as Theorem 3.6 to be applicable [57, 58, 69]. We find it conceptually compelling that our framework can provide in a unifying way equilibrium guarantees for two seemingly disparate classes of games, namely two-player zero-sum games and games with vanishing strategic sensitivity.

In a similar vein, Feldman et al. [39] showed that $\rho_n \rightarrow 1$ with a rate of $1/\sqrt{n}$ in a general auction design problem (see also [23]) under the relatively mild assumption that each player participates in the market with some constant probability (aka. probabilistic demand), thereby bypassing known barriers regarding the inefficiency of equilibria in general combinatorial domains. Their proof is based on the fact that each bidder's impact on the prices—under a simultaneous uniform-price auction format—becomes asymptotically negligible, in the spirit of Definition 3.5 introduced above. The difficulty that arises in that setting is that the natural representation of the utility functions violates our multilinearity assumption (postulated in Section 2). Instead, one would have to resort to some form of discretization before attempting to apply Theorem 3.1, which could be computationally prohibitive; the other prerequisite is that $L_n = o_n(\sqrt{n})$, for which our approach in Lemma A.8 in conjunction with the insights of Feldman et al. [39] could be useful. Understanding whether our techniques can be applied in the combinatorial auction setting of Feldman et al. [39] is left as a challenging open question.

Efficiency of equilibria does not suffice for tractability A natural question arising from Theorem 3.1 is whether a similar statement applies under the assumption that $\text{PoA} \rightarrow 1$, that is, assuming that all Nash equilibria of \mathcal{G} are (approximately) fully efficient. This is clearly a weaker assumption, but it is unfortunately not sufficient to yield any non-trivial guarantees even in normal-form games:

Proposition 3.7. *Even under the promise that $\text{PoA}_{\mathcal{G}} = 1$, computing a $(1/\text{poly}(\mathcal{G}))$ -Nash equilibrium in normal-form games in polynomial time is impossible when $n \geq 3$, unless $\text{PPAD} \subseteq \text{P}$.*

This stands in contrast to the class of $(1 + \mu, \mu)$ -smooth games, where a fully polynomial-time approximation scheme (FPTAS) is implied by Theorem 3.1—assuming access to a utility and a projection oracle, both of which are available in, for example, most succinct normal-form games. Proposition 3.7 is a straightforward consequence of the fact that Nash equilibria are hard to compute even in constant-sum 3-player games [20]. Furthermore, in Example A.11 we identify a specific 3-player game in which OGD fails (unconditionally) to converge to ϵ -Nash equilibria for a constant value of $\epsilon > 0$, even though $\text{PoA} = 1$. Our example is based on a variant of *Shapley's game* [99].

Another notable advantage of smoothness as a criterion of convergence is that, at least when the game is represented explicitly, it is easy to compute; this is in contrast to PoA , whose identification even in two-player games is NP-hard (Proposition A.12).

Extensions It turns out that smoothness per Definition 2.1 can be further sharpened using a primal-dual framework [75]. Such refined guarantees can also be translated into our setting (in the context of Theorem 3.1), as we elaborate on in Appendix A.3. The upshot is that the modification of Nadav and Roughgarden [75] necessitates analyzing the *weighted* sum of the players' regrets $\sum_{i=1}^n z_i \text{Reg}_i^{(T)}$ under a dual set of variables $\{z_i\}_{i=1}^n$. Another interesting extension worth noting relies on the *local* smoothness framework [92, 78], as we explain in Appendix A.4; the key observation is that local smoothness per Nguyen [78] can be associated with a *linearized* notion of regret, at which point the analysis of Theorem 3.1 readily carries over. Finally, we also expand our scope to Bayesian mechanisms, as we expound in the upcoming subsection.

Before we proceed, it is important to point out that, unsurprisingly, smoothness is not merely enough to guarantee convergence of OGD; see Example A.13. It is instead crucial to additionally ensure that $\rho \approx 1$ in order to obtain interesting guarantees for the behavior of algorithms such as OGD.

3.1 Bayesian mechanisms

Next, we leverage the connection between smoothness and convergence to Nash equilibria to extend the scope of our framework to Bayesian mechanisms. In particular, analogously to Definition 2.1, Syrgkanis and Tardos [102] have introduced the notion of a *smooth mechanism* (Definition A.14); detailed background on Bayesian mechanisms and smoothness in that realm is provided in Appendix A.7. In this context, we leverage a reduction of Hartline et al. [53] from an incomplete-information to a complete-information mechanism to arrive at the following theorem. (The standard notion of a *Bayes-Nash equilibrium* is analogous to Definition 2.2, and is recalled in Definition A.15 of Appendix A.7.)

Theorem 3.8. *Consider a Bayesian mechanism \mathcal{M} such that $\rho_{\mathcal{M}} = 1$. Then, for any $\epsilon > 0$, $T = O(1/\epsilon^2)$ iterations of OGD suffice to obtain an ϵ -Bayes-Nash equilibrium of \mathcal{M} .*

In particular, OGD above is executed on the so-called *agent-form* representation of \mathcal{M} (Appendix A.7). Analogously to Theorem 3.1, the above theorem can also be extended in the large $n \gg 1$ under the assumption that $\rho_n \rightarrow 1$. It is worth noting that Theorem 3.1 already can be applied to certain games of incomplete information (such as imperfect-information extensive-form games), but Theorem 3.8 additionally makes a connection with the literature on smoothness in mechanism design, which facilitates characterizing the smoothness parameters.

4 Improved welfare for no-regret dynamics

Roughgarden’s seminal work [90] established that no-regret learning algorithms always attain asymptotically at least rPoA fraction of the optimal social welfare (on average). This guarantee is satisfactory for many classes of games where rPoA is close to 1 (emphatically those studied earlier in Section 3), but smoothness is certainly not a universal phenomenon: there are simple games in which the smoothness framework only provides vacuous guarantees; one such example is Shapley’s game, discussed in Appendix A.10. As a result, one important question arising is whether it is possible to improve the efficiency bound predicted by smoothness when specific learning algorithms are in place, while at the same time still guaranteeing convergence to the set of *coarse correlated equilibria* (CCE). We stress that optimizing over the set of CCE is typically NP-hard in succinct games [79, 6], making this question interesting also from a complexity-theoretic standpoint.

In this section, we show that it is indeed possible to obtain improved efficiency bounds under a generic condition, while at the same time guaranteeing the no-regret property for each player. A key ingredient in our improvement is the use of *clairvoyant* mirror descent, an algorithm recently introduced by Piliouras et al. [85]. More precisely, we will instantiate that algorithm with (squared) Euclidean regularization, which can be defined as follows. Let $\Pi_{\mathbf{x}_i}(\mathbf{u}_i) := \arg \max_{\mathbf{x}'_i \in \mathcal{X}_i} \left\{ \langle \mathbf{x}'_i, \mathbf{u}_i \rangle - \frac{1}{2} \|\mathbf{x}_i - \mathbf{x}'_i\|_2^2 \right\}$ be the induced *prox operator*, where $\mathbf{x}_i \in \mathcal{X}_i$ and $\mathbf{u}_i \in \mathbb{R}^{d_i}$. *Clairvoyant gradient descent* (henceforth CGD) at time $t \in \mathbb{N}$ outputs $\mathbf{x}^{(t)} = \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}^{(t)})) := (\Pi_{\mathbf{x}_1^{(t-1)}}(\eta \mathbf{u}_1(\mathbf{w}_1^{(t)})), \dots, \Pi_{\mathbf{x}_n^{(t-1)}}(\eta \mathbf{u}_n(\mathbf{w}_n^{(t)})))$, where $\mathbf{w}^{(t)}$ is any $\epsilon^{(t)}$ -approximate fixed point of the map $\prod_{i=1}^n \mathcal{X}_i \ni \mathbf{w} \mapsto \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}))$, and $\mathbf{x}^{(0)} \in \mathcal{X}$ is an arbitrary initialization. It turns out that for $\eta < 1/L$, this map is a *contraction* [37, 85, 19], thereby making approximate fixed points easy to compute. Furthermore, there is also an uncoupled implementation of CGD [85], making the algorithm compelling from a decentralized standpoint as well, but we will not dwell on this issue here. We are now ready to state the main result of this section.

Theorem 4.1. *Suppose that all players are updating their strategies using CGD with $\epsilon^{(t)} \leq \frac{\min_i D_{\mathcal{X}_i}}{t^2}$ and learning rate $\eta = \frac{1}{2L}$ in a (λ, μ) -smooth game \mathcal{G} , where L is the Lipschitz-continuity parameter of F . Then, for any $\epsilon_0 > 0$ and $T \geq \frac{64L^2 D_{\mathcal{X}}^4}{\epsilon_0^2}$ iterations,*

1. *the average correlated distribution of play is a $\frac{4LD_{\mathcal{X}}^2}{T} - \text{CCE}$;*
2. *there is a time $t^* \in \llbracket T \rrbracket$ such that*

$$\text{SW}(\mathbf{x}^{(t^*)}) \geq \sup_{\epsilon \geq \epsilon_0} \min \left\{ \rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} + \frac{\epsilon^2}{16(\mu + 1)LD_{\mathcal{X}}^2}, \text{PoA}_{\mathcal{G}}^{\epsilon} \cdot \text{OPT}_{\mathcal{G}} \right\}. \quad (3)$$

This is the first result that establishes simultaneously these properties under a computationally efficient algorithm, improving a recent work [1] (see also [42, 67] for related results) that failed to

guarantee convergence to CCE. In particular, that earlier work was analyzing OGD, and as it turns out, under a time-invariant learning rate η it is not even known whether OGD ensures *sublinear* per-player regret, let alone constant (as in Corollary 4.3). The basic ingredient to this improvement is a new property of CGD, which we explain below. Before we sketch the proof, we note that Item 2 above can be readily strengthened so that the improvement holds for the average welfare of most of the strategies, not just a single one (Remark A.22).

Proof sketch of Theorem 4.1 The key step in the proof is showing (in Corollary A.21) that CGD satisfies the remarkable per-player regret bound $\text{Reg}_i^{(T)} \leq \alpha - \gamma \sum_{t=1}^T (\text{BRGAP}_i(\mathbf{x}^{(t)}))^2$, where $\alpha > 0$ depends on the approximation error of the fixed points of CGD—and can be made time-invariant with only an $O(\log T)$ per-iteration overhead—and $\gamma > 0$. To do this, we crucially rely on a certain property of the Euclidean regularizer (Lemma A.20), which we use in conjunction with the analysis of Farina et al. [37] who extended the original argument of Piliouras et al. [85] beyond entropic regularization.

It is worth noting that the above per-player regret bound (Corollary A.21) implies that a player with nonnegative regret will be almost always approximately best responding, a rather singular occurrence in the context of learning in games; this has interesting implications and goes well-beyond what is currently known for OGD. In particular, it is an open question whether Theorem 4.1 holds under OGD.

Next, we shall describe a concrete implication of Theorem 4.1 under a generic condition. To do so, let us denote by PoA_G^ϵ the price of anarchy in \mathcal{G} with respect to the worst-case ϵ -Nash equilibrium (so that $\text{PoA}_G^0 \equiv \text{PoA}_G$).

Condition 4.2. Consider a game \mathcal{G} and some game-dependent parameter $C = C(\mathcal{G}) > 0$. There exists an $\epsilon_0 > 0$ such that $\text{PoA}_G^{\epsilon_0} > \text{rPoA}_G + \epsilon_0^2 C$.

Naturally, it is always the case that $\text{PoA}_G \geq \text{rPoA}_G$. Further, rPoA_G is in general strictly smaller since it measures the worst-case welfare over a larger set than PoA_G (even broader than outcomes of no-regret learning); Figure 1 in Appendix A.8 further corroborates this premise in a sequence of random normal-form games. Now assuming that $\text{PoA}_G > \text{rPoA}_G$, Condition 4.2 is met if $\lim_{\epsilon \rightarrow 0} \text{PoA}_G^\epsilon = \text{PoA}_G$, a mild continuity condition (see, for example, the discussion by Roughgarden [89]).

Corollary 4.3. Consider a (λ, μ) -smooth game \mathcal{G} that satisfies Condition 4.2 under some $\epsilon_0 > 0$. Then, CGD after $T \geq \frac{64L^2 D^4 x}{\epsilon_0^2}$ iterations and $\eta = \frac{1}{2L}$ satisfies the following:

1. the average correlated distribution of play is an $O\left(\frac{1}{T}\right)$ -CCE;
2. there is a time $t^* \in \llbracket T \rrbracket$ and $C'(\mathcal{G}) > 0$ such that $\text{SW}(\mathbf{x}^{(t^*)}) \geq (\rho_G(\lambda, \mu) + \epsilon_0^2 C'(\mathcal{G})) \text{OPT}_G$.

A fundamental question that arises from Theorem 4.1 is whether there exists a computationally efficient algorithm that determines a CCE with social welfare at least a PoA fraction of the optimal welfare.³ In games where $\text{PoA} = \text{rPoA}$ this is clearly possible; in contrast, while Theorem 4.1 improves over the smoothness bound, it does not always guarantee welfare up to PoA. This is a central question in light of the intractability of Nash equilibria [30, 20], which has indeed served as a primary critique to the literature quantifying the price of anarchy of Nash equilibria [90].

Another promising avenue to improving the welfare predicted by the smoothness framework revolves around eliminating certain strategy profiles by arguing that they are reached with negligible probability. For example, in Appendix A.10 we identify an example where iteratively eliminating strictly dominated actions can improve the predictive power of the smoothness framework.

5 Further related work

Large games The study of non-cooperative games with many players (*i.e.*, large games) has been a classical topic in economic theory [96, 41, 68, 95, 83, 38, 50], most recently revived in the context of *mean-field games* (*e.g.*, [64, 47, 74, 51, 16, 82, 100, 80, 29, 65]). Indeed, many traditional motivating scenarios in algorithmic game theory, including markets and Internet routing, often feature a large number of players in practice. In particular, it has emerged that, under certain conditions, equilibria

³We clarify that Theorem 4.1 could have also been stated as follows: CGD outputs an approximate CCE with social welfare attaining the right-hand side of (3); this is evident from the proof in Appendix A.9.

in large games exhibit certain remarkable robustness and stability properties; see, for example, the recent survey of Gradwohl and Kalai [49], as well as the older treatment of Kalai [56] on the subject. Furthermore, mechanism design in large games, along with privacy guarantees, is explored in the work of Kearns et al. [58] (see also [57, 59]).

Efficiency in large games Of particular importance to our work, and specifically the precondition of Theorem 3.1, is the line of work uncovering the by now well-documented phenomenon in economics that large games exhibit, under certain relatively mild assumptions, fully efficient equilibria. Our framework additionally requires that the efficiency of equilibria can be derived via a smoothness argument, in the sense of Roughgarden [90]; we stress again that efficiency alone is of little use when it comes to equilibrium computation (Proposition 3.7). Fortunately, smoothness has emerged as the canonical paradigm for bounding the price of anarchy (e.g., see the survey of Roughgarden et al. [93]), albeit with some notable exceptions [40, 55]. In particular, Feldman et al. [39] quantify the price of anarchy in large games via the smoothness framework. They show that in a general combinatorial domain with simultaneous uniform-price auctions, it holds that $\rho_n \rightarrow 1$ with a rate of $1/\sqrt{n}$ as long as there is *probabilistic demand*, meaning that every buyer abstains from the auction with a constant probability. Several other papers have studied the price of anarchy in large games [63, 23, 24, 26, 25, 17]. In particular, we highlight the work of Cole and Tao [23] which, as Feldman et al. [39], relies on a smoothness argument to establish full efficiency in the limit with a rate of $1/\sqrt{n}$ in a Walrasian auction, while asymptotic full efficiency is also shown for Fisher markets under the gross substitutes condition. Further, Carmona et al. [17] provide sufficient conditions under which equilibria are fully efficient in a class of mean-field games; understanding thus whether our framework has new implications in such games is an interesting direction for the future. We finally point out that many other papers have focused on learning in auctions and markets; see [21, 43, 104, 8, 9, 32], and references therein.

6 Conclusions and future work

In conclusion, we have furnished a new sufficient condition under which a family of no-regret learning algorithms, including optimistic gradient descent (OGD), approaches (weak) Nash equilibria. Our criterion has a natural economic interpretation, being intricately connected with Roughgarden’s smoothness framework, and captures other well-studied conditions such as the Minty property. We have also shown that *clairvoyant* gradient descent attains an improved welfare bound compared to that predicted by the smoothness framework, while ensuring at the same time fast convergence to the set of CCE.

There are many promising directions for future work. First, we have seen that under the condition $\rho = 1$ there exists an algorithm that computes an ϵ -NE in time $\text{poly}(1/\epsilon)$, leading to a *pseudo* polynomial-time algorithm (under natural game representations); is there an algorithm that instead runs in time $\text{poly}(\log(1/\epsilon))$?

Convergence to Nash equilibria via computational hardness? Another promising approach for showing convergence to Nash equilibria is by harnessing computational hardness results for the underlying welfare maximization problem. To be specific, we consider the following condition.

Condition 6.1. Consider a multi-player (λ, μ) -smooth game \mathcal{G} with $\rho_{\mathcal{C}} := \frac{\lambda}{1+\mu}$ from a class of games \mathcal{C} with the polynomial expectation property [79]. For any $\mathcal{G} \in \mathcal{C}$, computing a joint strategy profile $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$ such that $\text{SW}(\mathbf{x}) \geq \rho_{\mathcal{C}} \cdot \text{OPT}_{\mathcal{G}} + 1/\text{poly}(\mathcal{G})$ is NP-hard, for any $\text{poly}(\mathcal{G})$.

Indeed, smoothness often circumscribes the welfare of polynomial algorithms, such as combinatorial auctions under XOS valuations—in fact, unconditionally under polynomial communication; see [34, Theorem 1.4] and [102, Appendix A.7]. Now, the role of Condition 6.1 is that (unless $\text{P} = \text{NP}$) a polynomially-bounded algorithm such as OGD—which is efficiently implementable (for games with a polynomial number of actions) under the polynomial expectation property—will have the property that $\frac{1}{T} \sum_{i=1}^n \text{Reg}_i^{(T)} \geq -1/\text{poly}(\mathcal{G})$, for any $\mathcal{G} \in \mathcal{C}$ and $\text{poly}(\mathcal{G})$, which in turn leads to the following.

Theorem 6.2. Consider a class \mathcal{C} satisfying Condition 6.1. For any $\mathcal{G} \in \mathcal{C}$ and $\epsilon = 1/\text{poly}(\mathcal{G})$, there is a polynomial-time algorithm for computing an ϵ -Nash equilibrium, unless $\text{P} = \text{NP}$.

By virtue of Corollary 4.3, the same conclusion applies even under the weaker condition that computing a CCE with welfare improving over the smoothness bound is hard; this is related to the hardness result of Barman and Ligett [6], discussed in the full version of this paper.

Acknowledgments and Disclosure of Funding

We are grateful to anonymous reviewers and the area chair at NeurIPS for valuable feedback. We also thank Brendan Lucier for helpful pointers to the literature. This material is based on work supported by the Vannevar Bush Faculty Fellowship ONR N00014-23-1-2876, National Science Foundation grants RI-2312342 and RI-1901403, ARO award W911NF2210266, and NIH award A240108S001.

References

- [1] I. Anagnostides, I. Panageas, G. Farina, and T. Sandholm. On last-iterate convergence beyond zero-sum games. In *International Conference on Machine Learning (ICML)*, volume 162 of *Proceedings of Machine Learning Research*, pages 536–581. PMLR, 2022.
- [2] I. Arieli and Y. Babichenko. Simple approximate equilibria in games with many players. In *ACM Conference on Economics and Computation (EC)*, pages 681–691. ACM, 2017.
- [3] Y. Azrieli and E. Shmaya. Lipschitz games. *Mathematics of Operation Research*, 38(2): 350–357, 2013.
- [4] Y. Babichenko and A. Rubinstein. Communication complexity of approximate nash equilibria. *Games and Economic Behavior*, 134:376–398, 2022.
- [5] Y. Babichenko, C. H. Papadimitriou, and A. Rubinstein. Can almost everybody be almost happy? In *Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 1–9. ACM, 2016.
- [6] S. Barman and K. Ligett. Finding any nontrivial coarse correlated equilibrium is hard. In *ACM Conference on Economics and Computation (EC)*, pages 815–816. ACM, 2015.
- [7] M. Bowling, N. Burch, M. Johanson, and O. Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218), Jan. 2015.
- [8] S. Brânzei. Exchange markets: proportional response dynamics and beyond. *SIGecom Exch.*, 19(2):37–45, 2021.
- [9] S. Brânzei, N. R. Devanur, and Y. Rabani. Proportional dynamics in exchange economies. In *ACM Conference on Economics and Computation (EC)*, pages 180–201. ACM, 2021.
- [10] N. Brown and T. Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.
- [11] A. Böhm. Solving nonconvex-nonconcave min-max problems exhibiting weak minty solutions, 2022.
- [12] Y. Cai and W. Zheng. Accelerated single-call methods for constrained min-max optimization. *CoRR*, abs/2210.03096, 2022.
- [13] Y. Cai, O. Candogan, C. Daskalakis, and C. Papadimitriou. Zero-sum polymatrix games: A generalization of minmax. *Mathematics of Operations Research*, 41(2):648–655, 2016.
- [14] Y. Cai, A. Oikonomou, and W. Zheng. Accelerated algorithms for monotone inclusions and constrained nonconvex-nonconcave min-max optimization. *CoRR*, abs/2206.05248, 2022.
- [15] Y. Cai, A. Oikonomou, and W. Zheng. Finite-time last-iterate convergence for learning in multi-player games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- [16] L. Campi and M. Fischer. Correlated equilibria and mean field games: A simple model. *Mathematics of Operations Research*, 47(3):2240–2259, 2022.
- [17] R. Carmona, C. V. Graves, and Z. Tan. Price of anarchy for mean field games. *ESAIM: Proceedings and Surveys*, 65:349–383, 2019.

- [18] N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- [19] V. Cevher, G. Piliouras, R. Sim, and S. Skoulakis. Min-max optimization made simple: Approximating the proximal point method via contraction maps. In *Symposium on Simplicity in Algorithms (SOSA)*, pages 192–206. SIAM, 2023.
- [20] X. Chen, X. Deng, and S.-H. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 2009.
- [21] Y. K. Cheung, S. Leonardos, and G. Piliouras. Learning in markets: Greed leads to chaos but following the price is right. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 111–117. ijcai.org, 2021.
- [22] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In *Conference on Learning Theory (COLT)*, pages 6–1, 2012.
- [23] R. Cole and Y. Tao. Large market games with near optimal efficiency. In *ACM Conference on Economics and Computation (EC)*, pages 791–808. ACM, 2016.
- [24] R. Colini-Baldeschi, R. Cominetti, P. Mertikopoulos, and M. Scarsini. The asymptotic behavior of the price of anarchy. In *International Workshop On Internet And Network Economics (WINE)*, pages 133–145. Springer, 2017.
- [25] R. Colini-Baldeschi, R. Cominetti, and M. Scarsini. Price of anarchy for highly congested routing games in parallel networks. *Theory of Computing Systems*, 63:90–113, 2019.
- [26] R. Colini-Baldeschi, R. Cominetti, P. Mertikopoulos, and M. Scarsini. When is selfish routing bad? the price of anarchy in light and heavy traffic. *Operations Research*, 68(2):411–434, 2020.
- [27] V. Conitzer and T. Sandholm. Complexity of constructing solutions in the core based on synergies among coalitions. *Artificial Intelligence*, 170(6-7):607–619, 2006. Earlier version in IJCAI-03.
- [28] V. Conitzer and T. Sandholm. New complexity results about Nash equilibria. *Games and Economic Behavior*, 63(2):621–641, 2008. Early version in IJCAI-03.
- [29] K. Cui and H. Koepl. Learning graphon mean field games and approximate nash equilibria. In *International Conference on Learning Representations (ICLR)*. OpenReview.net, 2022.
- [30] C. Daskalakis, P. Goldberg, and C. Papadimitriou. The complexity of computing a Nash equilibrium. In *Symposium on Theory of Computing (STOC)*, 2006.
- [31] J. Deb and E. Kalai. Stability in large bayesian games with heterogeneous players. *Journal of Economic Theory*, 157:1041–1055, 2015.
- [32] X. Deng, X. Hu, T. Lin, and W. Zheng. Nash convergence of mean-based learning algorithms in first price auctions. In *WWW '22: The ACM Web Conference*, pages 141–150. ACM, 2022.
- [33] J. Diakonikolas, C. Daskalakis, and M. I. Jordan. Efficient methods for structured nonconvex-nonconcave min-max optimization. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 130 of *Proceedings of Machine Learning Research*, pages 2746–2754. PMLR, 2021.
- [34] S. Dobzinski, N. Nisan, and M. Schapira. Approximation algorithms for combinatorial auctions with complement-free bidders. *Mathematics of Operation Research*, 35(1):1–13, 2010.
- [35] K. Etessami and M. Yannakakis. On the complexity of Nash equilibria and other fixed points (extended abstract). In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 113–123, 2007.
- [36] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer, 2003.

- [37] G. Farina, C. Kroer, C. Lee, and H. Luo. Clairvoyant regret minimization: Equivalence with nemirovski’s conceptual prox method and extension to general convex games. *CoRR*, abs/2208.14891, 2022.
- [38] T. Feddersen and W. Pesendorfer. Voting behavior and information aggregation in elections with private information. *Econometrica*, 65(5):1029–1058, 1997.
- [39] M. Feldman, N. Immorlica, B. Lucier, T. Roughgarden, and V. Syrgkanis. The price of anarchy in large games. In *Symposium on Theory of Computing (STOC)*, pages 963–976. ACM, 2016.
- [40] M. Feldman, H. Fu, N. Gravin, and B. Lucier. Simultaneous auctions without complements are (almost) efficient. *Games and Economic Behavior*, 123:327–341, 2020.
- [41] D. Fudenberg and D. K. Levine. Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory*, 44(1):1–18, 1988.
- [42] J. Gaitonde, Y. Li, B. Light, B. Lucier, and A. Slivkins. Budget pacing in repeated auctions: Regret and efficiency without convergence. In *Innovations in Theoretical Computer Science Conference (ITCS)*, volume 251 of *LIPICs*, pages 52:1–52:1. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.
- [43] Y. Gao. *New Optimization Models and Methods for Classical, Infinite-Dimensional, and Online Fisher Markets*. Columbia University, 2022.
- [44] I. Gilboa and E. Zemel. Nash and correlated equilibria: Some complexity considerations. *Games and Economic Behavior*, 1:80–93, 1989.
- [45] P. W. Goldberg and M. J. Katzman. Lower bounds for the query complexity of equilibria in lipschitz games. *Theoretical Comput. Sci.*, 962, 2023.
- [46] N. Golowich, S. Pattathil, and C. Daskalakis. Tight last-iterate convergence rates for no-regret learning in multi-player games. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
- [47] D. A. Gomes and J. Saúde. Mean field games models—a brief survey. *Dynamic Games and Applications*, 4:110–154, 2014.
- [48] E. Gorbunov, N. Loizou, and G. Gidel. Extragradient method: $O(1/K)$ last-iterate convergence for monotone variational inequalities and connections with cocoercivity. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 151 of *Proceedings of Machine Learning Research*, pages 366–402. PMLR, 2022.
- [49] R. Gradwohl and E. Kalai. Large games: Robustness and stability. *Annual Review of Economics*, 13:39–56, 2021.
- [50] R. Gradwohl and O. Reingold. Fault tolerance in large games. *Games and Economic Behavior*, 86:438–457, 2014.
- [51] Y. Guan, M. Afshari, and P. Tsiotras. Zero-sum games between mean-field teams: A common information and reachability based analysis. *CoRR*, abs/2303.12243, 2023.
- [52] J. C. Harsanyi, R. Selten, et al. A general theory of equilibrium selection in games. *MIT Press Books*, 1, 1988.
- [53] J. D. Hartline, V. Syrgkanis, and É. Tardos. No-regret learning in bayesian games. In *Conference on Neural Information Processing Systems (NeurIPS)*, pages 3061–3069, 2015.
- [54] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.
- [55] Y. Jin and P. Lu. First price auction is $1 - 1/e^2$ efficient. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 179–187. IEEE, 2022.
- [56] E. Kalai. Large robust games. *Econometrica*, 72(6):1631–1665, 2004.

- [57] M. J. Kearns and Y. Mansour. Efficient nash computation in large population games with bounded influence. In *Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 259–266. Morgan Kaufmann, 2002.
- [58] M. J. Kearns, M. M. Pai, A. Roth, and J. R. Ullman. Mechanism design in large games: incentives and privacy. In *Innovations in Theoretical Computer Science (ITCS)*, pages 403–410. ACM, 2014.
- [59] M. J. Kearns, M. M. Pai, R. M. Rogers, A. Roth, and J. R. Ullman. Robust mediators in large games. *CoRR*, abs/1512.02698, 2015.
- [60] D. Kim, M. Riemer, M. Liu, J. N. Foerster, M. Everett, C. Sun, G. Tesauro, and J. P. How. Influencing long-term behavior in multiagent reinforcement learning. *CoRR*, abs/2203.03535, 2022.
- [61] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Symposium on Theoretical Aspects in Computer Science (STACS)*, 1999.
- [62] J. Kulkarni and V. S. Mirrokni. Robust price of anarchy bounds via LP and fenchel duality. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1030–1049. SIAM, 2015.
- [63] D. Lacker and K. Ramanan. Rare nash equilibria and the price of anarchy in large static games. *Mathematics of Operations Research*, 44(2):400–422, 2019.
- [64] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese journal of mathematics*, 2(1): 229–260, 2007.
- [65] M. Laurière, S. Perrin, S. Girgin, P. Muller, A. Jain, T. Cabannes, G. Piliouras, J. Pérolat, R. Elie, O. Pietquin, and M. Geist. Scalable deep reinforcement learning algorithms for mean field games. In *International Conference on Machine Learning (ICML)*, volume 162 of *Proceedings of Machine Learning Research*, pages 12078–12095. PMLR, 2022.
- [66] T. Lin, Z. Zhou, P. Mertikopoulos, and M. I. Jordan. Finite-time last-iterate convergence for multi-agent learning in games. In *International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pages 6161–6171. PMLR, 2020.
- [67] B. Lucier, S. Pattathil, A. Slivkins, and M. Zhang. Autobidders with budget and ROI constraints: Efficiency, regret, and pacing dynamics. *CoRR*, abs/2301.13306, 2023.
- [68] G. J. Mailath and A. Postlewaite. Asymmetric information bargaining problems with many agents. *The Review of Economic Studies*, 57(3):351–367, 1990.
- [69] F. McSherry and K. Talwar. Mechanism design via differential privacy. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 94–103. IEEE Computer Society, 2007.
- [70] P. Mertikopoulos, C. H. Papadimitriou, and G. Piliouras. Cycles in adversarial regularized learning. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2703–2717. SIAM, 2018.
- [71] P. Mertikopoulos, B. Lecouat, H. Zenati, C. Foo, V. Chandrasekhar, and G. Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *International Conference on Learning Representations (ICLR)*. OpenReview.net, 2019.
- [72] P. Milgrom and R. Weber. Distributional strategies for games with incomplete information. *Mathematics of Operations Research*, 10:619–632, 1985.
- [73] J. Milionis, C. H. Papadimitriou, G. Piliouras, and K. Spendlove. Nash, conley, and computation: Impossibility and incompleteness in game dynamics. *CoRR*, abs/2203.14129, 2022.
- [74] P. Muller, M. Rowland, R. Elie, G. Piliouras, J. Pérolat, M. Laurière, R. Marinier, O. Pietquin, and K. Tuyls. Learning equilibria in mean-field games: Introducing mean-field PSRO. In *Autonomous Agents and Multi-Agent Systems*, pages 926–934. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2022.

- [75] U. Nadav and T. Roughgarden. The limits of smoothness: A primal-dual framework for price of anarchy bounds. In *International Workshop On Internet And Network Economics (WINE)*, volume 6484 of *Lecture Notes in Computer Science*, pages 319–326. Springer, 2010.
- [76] J. Nash. Equilibrium points in n-person games. *National Academy of Sciences*, 36:48–49, 1950.
- [77] A. Nemirovski. Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1), 2004.
- [78] K. T. Nguyen. Game efficiency through linear programming duality. In A. Blum, editor, *Innovations in Theoretical Computer Science Conference (ITCS)*, volume 124 of *LIPICs*, pages 66:1–66:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
- [79] C. H. Papadimitriou and T. Roughgarden. Computing correlated equilibria in multi-player games. *Journal of the ACM*, 55(3):14:1–14:29, 2008.
- [80] J. Pérolat, S. Perrin, R. Elie, M. Laurière, G. Piliouras, M. Geist, K. Tuyls, and O. Pietquin. Scaling mean field games by online mirror descent. In *Autonomous Agents and Multi-Agent Systems*, pages 1028–1037. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2022.
- [81] J. Perolat, B. D. Vyllder, D. Hennes, E. Tarassov, F. Strub, V. de Boer, P. Muller, J. T. Connor, N. Burch, T. Anthony, S. McAleer, R. Elie, S. H. Cen, Z. Wang, A. Gruslys, A. Malysheva, M. Khan, S. Ozair, F. Timbers, T. Pohlen, T. Eccles, M. Rowland, M. Lanctot, J.-B. Lespiau, B. Piot, S. Omidshafiei, E. Lockhart, L. Sifre, N. Beauguerlange, R. Munos, D. Silver, S. Singh, D. Hassabis, and K. Tuyls. Mastering the game of stratego with model-free multiagent reinforcement learning. *Science*, 378(6623):990–996, 2022.
- [82] S. Perrin, M. Laurière, J. Pérolat, R. Élie, M. Geist, and O. Pietquin. Generalization in mean field games by learning master policies. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 9413–9421. AAAI Press, 2022.
- [83] W. Pesendorfer and J. M. Swinkels. The loser’s curse and information aggregation in common value auctions. *Econometrica*, 65(6):1247–1281, 1997.
- [84] T. Pethick, P. Latafat, P. Patrinos, O. Fercoq, and V. Cevher. Escaping limit cycles: Global convergence for constrained nonconvex-nonconcave minimax problems. *CoRR*, abs/2302.09831, 2023.
- [85] G. Piliouras, R. Sim, and S. Skoulakis. Beyond time-average convergence: Near-optimal uncoupled online learning via clairvoyant multiplicative weights update. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- [86] A. Rakhlin and K. Sridharan. Online learning with predictable sequences. In *Conference on Learning Theory*, pages 993–1019, 2013.
- [87] S. Rakhlin and K. Sridharan. Optimization, learning, and games with predictable sequences. In *Advances in Neural Information Processing Systems*, pages 3066–3074, 2013.
- [88] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.
- [89] T. Roughgarden. Barriers to near-optimal equilibria. In *Annual Symposium on Foundations of Computer Science (FOCS)*, pages 71–80. IEEE Computer Society, 2014.
- [90] T. Roughgarden. Intrinsic robustness of the price of anarchy. *Journal of the ACM*, 62(5):32:1–32:42, 2015.
- [91] T. Roughgarden. The price of anarchy in games of incomplete information. *ACM Trans. Economics and Comput.*, 3(1):6:1–6:20, 2015.

- [92] T. Roughgarden and F. Schoppmann. Local smoothness and the price of anarchy in splittable congestion games. *Journal of Economic Theory*, 156:317–342, 2015.
- [93] T. Roughgarden, V. Syrgkanis, and É. Tardos. The price of anarchy in auctions. *Journal of Artificial Intelligence Research*, 59:59–101, 2017.
- [94] A. Rubinstein. Inapproximability of nash equilibrium. *SIAM Journal on Computing*, 47(3): 917–959, 2018.
- [95] A. Rustichini, M. Satterthwaite, and S. Williams. Convergence to efficiency in a simple market with incomplete information. *Econometrica*, 62:1041–1063, 1994.
- [96] D. Schmeidler. Equilibrium points of nonatomic games. *Journal of statistical Physics*, 7: 295–300, 1973.
- [97] P. G. Sessa, I. Bogunovic, A. Krause, and M. Kamgarpour. Contextual games: Multi-agent learning with side information. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
- [98] S. Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2), 2012. ISSN 1935-8237.
- [99] L. S. Shapley. Some topics in two-person games. In M. Drescher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*. Princeton University Press, 1964.
- [100] S. G. Subramanian, M. E. Taylor, M. Crowley, and P. Poupart. Decentralized mean field games. In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 9439–9447. AAAI Press, 2022.
- [101] V. Syrgkanis. Bayesian games and the smoothness framework. *CoRR*, abs/1203.5155, 2012.
- [102] V. Syrgkanis and É. Tardos. Composable and efficient mechanisms. In *Symposium on Theory of Computing (STOC)*, pages 211–220. ACM, 2013.
- [103] V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems*, pages 2989–2997, 2015.
- [104] Y. Tao. *Market Efficiency and Dynamics*. PhD thesis, New York University, 2020.
- [105] E. Vlatakis-Gkaragkounis, L. Flokas, T. Lianas, P. Mertikopoulos, and G. Piliouras. No-regret learning and mixed nash equilibria: They do not mix. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020*, 2020.

A Omitted proofs

In this section, we provide the proofs and a number of results omitted from the main body.

A.1 Proof of Theorem 3.1

We commence with the proof of Theorem 3.1. Below, we give a more detailed version of the statement provided earlier in the main body. We first state an auxiliary lemma which will be useful for the proof, and can be extracted from earlier work [1]

Lemma A.1. *Suppose that player $i \in \llbracket n \rrbracket$ updates its strategy using OGD with learning rate $\eta > 0$. Then, for any time $t \in \mathbb{N}$,*

$$\text{BRGAP}_i(\mathbf{x}^{(t)}) \leq \left(\frac{D_{\mathcal{X}_i}}{\eta} + \|\mathbf{u}_i^{(t)}\|_2 \right) \left(\|\mathbf{x}_i^{(t)} - \hat{\mathbf{x}}_i^{(t)}\|_2 + \|\mathbf{x}_i^{(t)} - \hat{\mathbf{x}}_i^{(t+1)}\|_2 \right).$$

In the sequel, we denote by $B_i \in \mathbb{R}_{>0}$ any number such that $\|\mathbf{u}_i(\mathbf{x}_{-i})\|_2 \leq B_i$, for any $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$. In the asymptotic notation below, we make the standard assumption that the parameters B_i and $D_{\mathcal{X}_i}$ do not depend on the number of players n .

Theorem A.2 (Precise version of Theorem 3.1). *Consider an n -player (λ, μ) -smooth game \mathcal{G}_n such that the game operator F_n is L_n -Lipschitz continuous and $\lambda \geq (1 - \epsilon_n)(1 + \mu)$ ($\rho_n \geq 1 - \epsilon_n$). Suppose further that all players follow OGD with learning rate $\eta_n = \frac{1}{4L_n}$ and any initialization $(\hat{\mathbf{x}}_1^{(1)}, \dots, \hat{\mathbf{x}}_n^{(1)}) \in \prod_{i=1}^n \mathcal{X}_i$. If $\epsilon_n > 0$, then after $T := \frac{D_{\mathcal{X}}^2}{2\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n}}$ iterations there is a time $t^* \in \llbracket T \rrbracket$ such that*

$$\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \leq 32 \left(\frac{\max_{1 \leq i \leq n} D_{\mathcal{X}_i}^2}{(\eta_n)^2} + \max_{1 \leq i \leq n} B_i^2 \right) \eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n} \quad (4)$$

$$= O_n(L_n \text{OPT}_{\mathcal{G}_n} \epsilon_n). \quad (5)$$

In particular, for $\delta \in \{\frac{1}{n}, \frac{2}{n}, \dots, 1\}$, t^* constitutes a

$$\left(\frac{1}{\sqrt{\delta}} O_n \left(\sqrt{\frac{L_n \text{OPT}_{\mathcal{G}_n} \epsilon_n}{n}} \right), \delta \right) - \text{weak Nash equilibrium.}$$

On the other hand, if $\epsilon_n = 0$, then for any $T \in \mathbb{N}$ there is a time $t^* \in \llbracket T \rrbracket$ such that

$$\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \leq \frac{8}{T} \left(\frac{\max_{1 \leq i \leq n} D_{\mathcal{X}_i}^2}{(\eta_n)^2} + \max_{1 \leq i \leq n} B_i^2 \right) D_{\mathcal{X}}^2.$$

Before we proceed with the proof, we note that the underlying assumption $\mu = O_n(1)$ in the asymptotic notation above is consistent with known smoothness bounds in large games [39]; we also refer to Remark A.10 for an important point regarding the range of the smoothness parameters.

Proof of Theorem A.2. We will first translate the assumed property $\rho_n \geq 1 - \epsilon_n$ to a lower bound on the sum of the players' regrets. In particular, we have that the sum of the players' regrets $\sum_{i=1}^n \text{Reg}_i^{(T)}$ for any $T \in \mathbb{N}$ can be expressed as

$$\sum_{i=1}^n \max_{\mathbf{x}_i^* \in \mathcal{X}_i} \left\{ \sum_{t=1}^T u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}^{(t)}) \right\} - \sum_{t=1}^T \text{SW}(\mathbf{x}^{(t)}) \geq \lambda \text{OPT}_{\mathcal{G}_n} T - (1 + \mu) \sum_{t=1}^T \text{SW}(\mathbf{x}^{(t)}), \quad (6)$$

by Definition 2.1, where (λ, μ) are the assumed smoothness parameter of \mathcal{G}_n . Now, using the assumption that $\rho_n \geq 1 - \epsilon_n$, it follows from (6) that $\sum_{i=1}^n \text{Reg}_i^{(T)}$ is in turn lower bounded by

$$(1 - \epsilon_n)(1 + \mu) \text{OPT}_{\mathcal{G}_n} T - (1 + \mu) \sum_{t=1}^T \text{SW}(\mathbf{x}^{(t)}) \geq -\epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n} T, \quad (7)$$

where we used the fact that $\text{OPT}_{\mathcal{G}_n} \geq \text{SW}(\mathbf{x}^{(t)})$ (by definition of the optimal welfare). We next appropriately upper bound the sum of the players' regrets $\sum_{i=1}^n \text{Reg}_i^{(T)}$. Using a slight refinement of the RVU bound [103, 87], it follows that the regret $\text{Reg}_i^{(T)}$ can be upper bounded by

$$\frac{1}{2\eta_n} D_{\mathcal{X}_i}^2 + \eta_n \sum_{t=1}^T \|\mathbf{u}_i^{(t)} - \mathbf{m}_i^{(t)}\|_2^2 - \frac{1}{2\eta_n} \sum_{t=1}^T \left(\|\mathbf{x}_i^{(t)} - \hat{\mathbf{x}}_i^{(t)}\|_2^2 + \|\mathbf{x}_i^{(t)} - \hat{\mathbf{x}}_i^{(t+1)}\|_2^2 \right),$$

which in turn implies that the sum of the regrets $\sum_{i=1}^n \text{Reg}_i^{(T)}$ is upper bounded by

$$\frac{1}{2\eta_n} \sum_{i=1}^n D_{\mathcal{X}_i}^2 + \eta_n \sum_{t=1}^T \|\mathbf{u}^{(t)} - \mathbf{m}^{(t)}\|_2^2 - \frac{1}{2\eta_n} \sum_{t=1}^T \left(\|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|_2^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t+1)}\|_2^2 \right), \quad (8)$$

where we have defined $\mathbf{u}^{(t)} := (\mathbf{u}_1^{(t)}, \dots, \mathbf{u}_n^{(t)})$ and $\mathbf{m}^{(t)} := (\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_n^{(t)})$. Moreover, by the L_n -Lipschitz continuity of the game operator $F_{\mathcal{G}_n}$, it follows that $\|\mathbf{u}^{(t)} - \mathbf{m}^{(t)}\|_2 = \|F_{\mathcal{G}_n}(\mathbf{x}^{(t)}) - F_{\mathcal{G}_n}(\mathbf{x}^{(t-1)})\|_2 \leq L_n \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|_2$ for any $t \in \llbracket T \rrbracket$, where we also used the fact that $\mathbf{m}^{(t)} = \mathbf{u}^{(t-1)}$ and $\mathbf{x}^{(0)} := \hat{\mathbf{x}}^{(1)}$. Combining with (8), it follows that for any $\eta_n \leq \frac{1}{4L_n}$,

$$\sum_{i=1}^n \text{Reg}_i^T \leq \frac{1}{2\eta_n} \sum_{i=1}^n D_{\mathcal{X}_i}^2 - \frac{1}{4\eta_n} \sum_{t=1}^T \left(\|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|_2^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t+1)}\|_2^2 \right). \quad (9)$$

As a result, combining (7) and (9) we have that

$$\sum_{t=1}^T \left(\|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|_2^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t+1)}\|_2^2 \right) \leq 2 \sum_{i=1}^n D_{\mathcal{X}_i}^2 + 4\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n} T. \quad (10)$$

Next, applying Lemma A.1 yields that

$$\begin{aligned} \sum_{t=1}^T \sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 &\leq 4 \left(\frac{\max_i D_{\mathcal{X}_i}^2}{\eta_n^2} + \max_i B_i^2 \right) \sum_{t=1}^T \left(\|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|_2^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t+1)}\|_2^2 \right) \\ &\leq 4 \left(\frac{\max_i D_{\mathcal{X}_i}^2}{\eta_n^2} + \max_i B_i^2 \right) (2D_{\mathcal{X}}^2 + 4\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n} T), \end{aligned} \quad (11)$$

where the first bound uses the inequality $(\alpha + \beta)^2 \leq 2\alpha^2 + 2\beta^2$, while the second bound follows from (10) by noting that $D_{\mathcal{X}}^2 = \sum_{i=1}^n D_{\mathcal{X}_i}^2$. As a result, we conclude that there exists $t^* \in \llbracket T \rrbracket$, namely $t^* := \arg \min_{1 \leq t \leq T} \sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2$, such that

$$\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \leq 4 \left(\frac{\max_{1 \leq i \leq n} D_{\mathcal{X}_i}^2}{\eta_n^2} + \max_{1 \leq i \leq n} B_i^2 \right) \left(\frac{2}{T} D_{\mathcal{X}}^2 + 4\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n} \right). \quad (12)$$

Now, if $\epsilon_n \neq 0$, taking $T := \frac{D_{\mathcal{X}}^2}{2\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n}} = \frac{2L_n^2 D_{\mathcal{X}}^2}{(1 + \mu)\gamma_n}$ implies (4). The claimed approximation guarantee in terms of weak Nash equilibria (per Definition 2.2) in (2) can be derived from (4) as follows. We consider a parameter $\delta \in (0, 1)$ so that $\delta n = \lfloor \delta n \rfloor$. Let also ϵ be the minimum among the δn largest numbers from $\text{BRGAP}_1(\mathbf{x}^{(t^*)}), \dots, \text{BRGAP}_n(\mathbf{x}^{(t^*)})$. Then, $\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \geq \delta \epsilon^2 n$, and by (12) it follows that

$$\epsilon \leq \frac{1}{\sqrt{\delta}} \sqrt{32 \left(\frac{\max_{1 \leq i \leq n} D_{\mathcal{X}_i}^2}{\eta_n^2} + \max_{1 \leq i \leq n} B_i^2 \right) \frac{\eta_n \epsilon_n (1 + \mu) \text{OPT}_{\mathcal{G}_n}}{n}}.$$

This implies (2) since $\mathbf{x}^{(t^*)}$ is by definition an (ϵ, δ) -weak Nash equilibrium. Finally, if $\epsilon_n = 0$ the claimed bound follows directly from (12). \square

We note that the asymptotic notation in (2) and (5) applies in the regime $L_n \geq \Omega_n(1)$; this can be enforced by simply taking $L'_n := \max\{1, L_n\}$.

Remark A.3. One can also introduce a variant of weak Nash equilibria which is instead parameterized by the average over the players' best response gaps $\frac{1}{n} \sum_{i=1}^n \text{BRGAP}_i(\cdot)$. Inequality (4) implies that the average best response gap of $\mathbf{x}^{(t^*)}$ can be bounded by $O_n \left(\sqrt{\frac{L_n \text{OPT}_{\mathcal{G}_n} \epsilon_n}{n}} \right)$.

Remark A.4. While Theorem A.2 bounds the (weak) Nash equilibrium gap of a single iterate of the dynamics, (11) implies that at least a $1 - \gamma$ fraction of the iterates of the dynamics constitutes a $\left(\frac{1}{\sqrt{\delta\gamma}} O_n \left(\sqrt{\frac{L_n \text{OPT}_{\mathcal{G}_n} \epsilon_n}{n}} \right), \delta \right)$ -weak Nash equilibrium. So, by selecting a time index $t^* \in \llbracket T \rrbracket$ uniformly at random we obtain the desired guarantee with high probability, incurring only a small degradation in the solution quality.

In particular, if ϵ_n approaches to 0 with a sufficiently fast rate, Theorem A.2 also implies convergence to the more standard notion of Nash equilibrium (i.e., Definition 2.2 with $\delta := 0$), as we state below. In particular, the following corollary can be derived directly from (4).

Corollary A.5. *In the setting of Theorem A.2, if it additionally holds that $L_n \epsilon_n \text{OPT}_{\mathcal{G}_n} \leq o_n(1)$, OGD yields an $o_n(1)$ -Nash equilibrium after a sufficiently large number of iterations.*

On the Lipschitz constant Next, we make some remarks regarding the dependence of the Lipschitz constant L_n on the number of players in the context of general normal-form games. We first note that the Lipschitz constant L_n of the underlying game operator can always be bounded as $O_n(n)$.

Lemma A.6 (Lipschitz constant in normal-form games). *For any n -player normal-form game \mathcal{G} with utilities bounded in $[-1, 1]$, the Lipschitz constant L_n of the game operator satisfies $L_n \leq n \max_{1 \leq i \leq n} |\mathcal{A}_i|$.*

Proof. For any player $i \in \llbracket n \rrbracket$ and any joint strategies $\mathbf{x}, \mathbf{x}' \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$, we have

$$\|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_2 \leq \sqrt{|\mathcal{A}_i|} \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_\infty \quad (13)$$

$$\leq \sqrt{|\mathcal{A}_i|} \left\| \sum_{\mathbf{a}_{-i} \in \mathcal{A}_{-i}} u_i(\cdot, \mathbf{a}_{-i}) \left(\prod_{i' \neq i} \mathbf{x}_{i'}[a_{i'}] - \prod_{i' \neq i} \mathbf{x}'_{i'}[a_{i'}] \right) \right\|_1 \quad (14)$$

$$\leq \sqrt{|\mathcal{A}_i|} \left\| \sum_{\mathbf{a}_{-i} \in \mathcal{A}_{-i}} \left(\prod_{i' \neq i} \mathbf{x}_{i'}[a_{i'}] - \prod_{i' \neq i} \mathbf{x}'_{i'}[a_{i'}] \right) \right\|_1 \quad (15)$$

$$\leq \sqrt{|\mathcal{A}_i|} \sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_1 \leq \max_{1 \leq i \leq n} |\mathcal{A}_i| \sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2, \quad (16)$$

where (13) uses the equivalence between the ℓ_2 and the ℓ_∞ norm; (14) follows from the definition of the (expected) utility: $u_i(a_i, \mathbf{x}_{-i}) := \mathbb{E}_{\mathbf{a}_{-i} \sim \mathbf{x}_{-i}} [u_i(\mathbf{a})] = \sum_{\mathbf{a}_{-i} \in \mathcal{A}_{-i}} u_i(\mathbf{a}) \prod_{i' \neq i} \mathbf{x}_{i'}[a_{i'}]$, for any $a_i \in \mathcal{A}_i$; (15) uses the triangle inequality along with the assumption that $|u_i(\mathbf{a})| \leq 1$; and (16) follows from the well-known fact that the total variation distance between two product distributions can be upper bounded by the sum of the total variation distance of each individual component [54], as well as the equivalence between the ℓ_1 and the ℓ_2 norm. As a result, continuing from (16), we have

$$\begin{aligned} \|F(\mathbf{x}) - F(\mathbf{x}')\|_2^2 &= \sum_{i=1}^n \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_2^2 \leq \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \sum_{i=1}^n \left(\sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2 \right)^2 \\ &\leq n^2 \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \|\mathbf{x} - \mathbf{x}'\|_2^2, \end{aligned}$$

where the last inequality used that, by Jensen's inequality, $\left(\sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2 \right)^2 \leq (n - 1) \sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2^2 \leq n \|\mathbf{x} - \mathbf{x}'\|_2^2$. This concludes the proof. \square

Graphical games As a byproduct of the proof above, we next point out an important refinement of Lemma A.6 concerning *graphical games*. In particular, here we assume that the utility of each player $i \in \llbracket n \rrbracket$ only depends on the actions of players belonging to its *neighborhood* $\mathcal{N}_i \subseteq \llbracket n \rrbracket \setminus \{i\}$.

We further assume that $|\mathcal{N}_i| \leq \Delta$ for any player $i \in \llbracket n \rrbracket$, where $\Delta \in \mathbb{N}$ will be referred to as the *degree* of the graphical game. To conclude the definition, we also posit that each player $i \in \llbracket n \rrbracket$ can only affect the utilities of at most Δ other players: $|i' \in \llbracket n \rrbracket : i \in \mathcal{N}_{i'}| \leq \Delta$.

Lemma A.7 (Lipschitz constant in graphical games). *For any n -player graphical game with degree $\Delta \in \mathbb{N}$ and utilities bounded in $[-1, 1]$, the Lipschitz constant L_n of the game operator satisfies $L_n \leq \Delta \max_{1 \leq i \leq n} |\mathcal{A}_i|$.*

Proof. For any player $i \in \llbracket n \rrbracket$ and any joint strategies $\mathbf{x}, \mathbf{x}' \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$, we have

$$\begin{aligned} \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_2 &\leq \sqrt{|\mathcal{A}_i|} \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_\infty \\ &\leq \sqrt{|\mathcal{A}_i|} \left\| \sum_{\prod_{i' \in \mathcal{N}_i} \mathcal{A}_{i'}} \left(\prod_{i' \in \mathcal{N}_i} \mathbf{x}_{i'}[a_{i'}] - \prod_{i' \in \mathcal{N}_i} \mathbf{x}'_{i'}[a_{i'}] \right) \right\|_1 \\ &\leq \sqrt{|\mathcal{A}_i|} \sum_{i' \in \mathcal{N}_i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_1 \leq \max_{1 \leq i \leq n} |\mathcal{A}_i| \sum_{i' \in \mathcal{N}_i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2, \end{aligned}$$

where the derivation above is similar to that in Lemma A.6. As a result,

$$\begin{aligned} \|F(\mathbf{x}) - F(\mathbf{x}')\|_2^2 &= \sum_{i=1}^n \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_2^2 \\ &\leq \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \sum_{i=1}^n \left(\sum_{i' \in \mathcal{N}_i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2 \right)^2 \\ &\leq \Delta \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \sum_{i=1}^n \sum_{i' \in \mathcal{N}_i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2^2 \\ &= \Delta \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{x}'_i\|_2^2 |i' \in \llbracket n \rrbracket : i \in \mathcal{N}_{i'}| \\ &\leq \Delta^2 \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right)^2 \|\mathbf{x} - \mathbf{x}'\|_2^2. \end{aligned}$$

□

Games with vanishing sensitivity We next focus on a different subclass of normal-form games; namely, games with small sensitivity (per Definition 3.5). Taking a step back, in graphical games every player can only be impacted by (and have an impact to) a small number of other players. Instead, here we consider games where a player's utility can be impacted by all other players, but only by a small amount.

Lemma A.8 (Lipschitz constant in games with vanishing sensitivity). *For any n -player normal-form game with sensitivity $\epsilon_n \in \mathbb{R}_{>0}$, the Lipschitz constant L_n of the game operator satisfies $L_n \leq \epsilon_n n \max_{1 \leq i \leq n} |\mathcal{A}_i|$.*

Proof. Let $i \in \llbracket n \rrbracket$. For $\mathbf{x}_1, \mathbf{x}'_1 \in \Delta(\mathcal{A}_1)$ and $\mathbf{a}_{-1} \in \mathcal{A}_{-1}$ (restricting on Player 1 here is without any loss, and only made for the sake of simplicity in the notation), it follows that

$$\begin{aligned} u_i(\mathbf{x}_1, \mathbf{a}_{-1}) - u_i(\mathbf{x}'_1, \mathbf{a}_{-1}) &= \sum_{a_1 \in \mathcal{A}_1} \mathbf{x}_1[a_1] u_i(a_1, \mathbf{a}_{-1}) - \sum_{a_1 \in \mathcal{A}_1} \mathbf{x}'_1[a_1] u_i(a_1, \mathbf{a}_{-1}) \\ &= \sum_{a_1 \in \mathcal{A}_1 \setminus \{a'_1\}} (\mathbf{x}_1[a_1] - \mathbf{x}'_1[a_1]) u_i(a_1, \cdot) + (\mathbf{x}_1[a'_1] - \mathbf{x}'_1[a'_1]) u_i(a'_1, \cdot) \\ &= \sum_{a_1 \in \mathcal{A}_1 \setminus \{a'_1\}} (\mathbf{x}_1[a_1] - \mathbf{x}'_1[a_1]) (u_i(a_1, \mathbf{a}_{-1}) - u_i(a'_1, \mathbf{a}_{-1})), \quad (17) \end{aligned}$$

for some $a'_1 \in \mathcal{A}_1$, where we used that $(\mathbf{x}_1[a'_1] - \mathbf{x}'_1[a'_1]) = \sum_{a_1 \in \mathcal{A}_1 \setminus \{a'_1\}} (\mathbf{x}'_1[a_1] - \mathbf{x}_1[a_1])$ since $\mathbf{x}_1, \mathbf{x}'_1 \in \Delta(\mathcal{A}_1)$. Continuing from (17), we have

$$\begin{aligned} |u_i(\mathbf{x}_1, \mathbf{a}_{-1}) - u_i(\mathbf{x}'_1, \mathbf{a}_{-1})| &\leq \sum_{a_1 \in \mathcal{A}_1 \setminus \{a'_1\}} |\mathbf{x}_1[a_1] - \mathbf{x}'_1[a_1]| |u_i(a_1, \mathbf{a}_{-1}) - u_i(a'_1, \mathbf{a}_{-1})| \\ &\leq \epsilon_n \sum_{a_1 \in \mathcal{A}_1 \setminus \{a'_1\}} |\mathbf{x}_1[a_1] - \mathbf{x}'_1[a_1]| \leq \epsilon_n \|\mathbf{x}_1 - \mathbf{x}'_1\|_1, \end{aligned}$$

where ϵ_n is the sensitivity of the game. Similar reasoning yields that $|u_i(\mathbf{x}) - u_i(\mathbf{x}'_1, \mathbf{x}_{-1})| \leq \epsilon_n \|\mathbf{x}_1 - \mathbf{x}'_1\|_1$, for any $\mathbf{x}_{-1} \in \prod_{i=2}^n \Delta(\mathcal{A}_i)$. As a result, we have

$$\begin{aligned} \|\mathbf{u}_1(\mathbf{x}_{-1}) - \mathbf{u}_1(\mathbf{x}'_{-1})\|_\infty &\leq |u_1(\cdot, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n) - u_1(\cdot, \mathbf{x}'_2, \mathbf{x}_3, \dots, \mathbf{x}_n)| \\ &\quad + |u_1(\cdot, \mathbf{x}'_2, \mathbf{x}_3, \dots, \mathbf{x}_n) - u_1(\cdot, \mathbf{x}'_2, \mathbf{x}'_3, \dots, \mathbf{x}_n)| \\ &\quad + \dots \\ &\quad + |u_1(\cdot, \mathbf{x}'_2, \mathbf{x}'_3, \dots, \mathbf{x}'_{n-1}, \mathbf{x}_n) - u_1(\cdot, \mathbf{x}'_2, \mathbf{x}'_3, \dots, \mathbf{x}'_n)| \\ &\leq \epsilon_n \sum_{i \neq 1} \|\mathbf{x}_i - \mathbf{x}'_i\|_1. \end{aligned}$$

By symmetry, we have shown that $\|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_\infty \leq \epsilon_n \sum_{i' \neq i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_1$, and the rest of the argument is identical to that of Lemma A.6. \square

Polymatrix games A careful examination of the proof of Lemma A.8 reveals that its conclusion in fact applies under a more relaxed condition compared to what imposed by Definition 3.5; namely, we can define

$$\epsilon := \max_{1 \leq i \leq n} \max_{\mathbf{a} \in \mathcal{A}} \max_{i' \neq i} \max_{\mathbf{a}_{i'} \in \mathcal{A}_{i'}} |u_i(\mathbf{a}) - u_i(\mathbf{a}_{i'}, \mathbf{a}_{-i'})|. \quad (18)$$

In words, when considering the utility of a player $i \in \llbracket n \rrbracket$, we only bound deviations by players besides i . It is easy to see that Lemma A.8 in fact applies even if $\epsilon \in \mathbb{R}_{>0}$ is defined as in (18). This observation enables capturing other interesting classes of games under the premise that $L_n = O_n(1)$, such as *polymatrix games*. Specifically, a polymatrix game is defined with respect to an underlying directed graph $G = (\llbracket n \rrbracket, E)$, so that each node of G is uniquely associated with the corresponding player. For every edge $(i, i') \in E$ there is a matrix $\mathbf{A}_{i,i'} \in \mathbb{R}^{\mathcal{A}_i \times \mathcal{A}_{i'}}$ so that the utility of Player i is defined as

$$u_i(\mathbf{x}) := \frac{1}{n} \sum_{i' \in \mathcal{N}_i} \mathbf{x}_i^\top \mathbf{A}_{i,i'} \mathbf{x}_{i'}, \quad (19)$$

where $\mathcal{N}_i := \{i' \in \llbracket n \rrbracket : (i, i') \in E\}$. Unlike the class of graphical games we saw earlier, here we do not restrict the size of the neighborhoods. For this reason, we have normalized each player's utility by a $1/n$ factor in (19), for otherwise the utilities are not guaranteed to be bounded (independent of the number of players n). It is then easy to see that polymatrix games are subject to Lemma A.8 under ϵ defined in (18), which here satisfies $\epsilon = O(1/n)$. Below, we provide a simpler and sharper argument compared to Lemma A.8.

Lemma A.9. *For any n -player polymatrix game, the Lipschitz constant L_n of the game operator satisfies $L_n \leq \max_{(i,i') \in E} \|\mathbf{A}_{i,i'}\|_2$, where $\|\cdot\|_2$ here denotes the spectral norm.*

Proof. By definition of the utility functions in (19), we have that for any player $i \in \llbracket n \rrbracket$ and $\mathbf{x}, \mathbf{x}' \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$,

$$\begin{aligned} \|\mathbf{u}_i(\mathbf{x}_{-i}) - \mathbf{u}_i(\mathbf{x}'_{-i})\|_2 &\leq \frac{1}{n} \left\| \sum_{i' \in \mathcal{N}_i} \mathbf{A}_{i,i'}(\mathbf{x}_{i'} - \mathbf{x}'_{i'}) \right\|_2 \\ &\leq \frac{1}{n} \sum_{i' \in \mathcal{N}_i} \|\mathbf{A}_{i,i'}(\mathbf{x}_{i'} - \mathbf{x}'_{i'})\|_2 \\ &\leq \frac{1}{n} \sum_{i' \in \mathcal{N}_i} \|\mathbf{A}_{i,i'}\|_2 \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2 \\ &\leq \frac{\max_{(i,i') \in E} \|\mathbf{A}_{i,i'}\|_2}{n} \sum_{i' \in \mathcal{N}_i} \|\mathbf{x}_{i'} - \mathbf{x}'_{i'}\|_2, \end{aligned}$$

and the claim follows. \square

A.2 Games with vanishing sensitivity

We next establish an important implication of our framework concerning the class of games with vanishing strategic sensitivity (per Definition 3.5); the statement of the theorem is recalled below.

Theorem 3.6. Consider an n -player game \mathcal{G}_n with sensitivity $\epsilon_n \in \mathbb{R}_{>0}$. Then, $T = O_n(n)$ iterations of OGD suffice to obtain a $\left(\frac{1}{\sqrt{\delta}} O_n(\epsilon_n \sqrt{n}), \delta\right)$ -weak Nash equilibrium, for $\delta \in (0, 1)$.

Proof. Let $i \in \llbracket n \rrbracket$. Following the proof of Lemma A.8, the definition of sensitivity implies that $|u_i(\mathbf{x}'_i, \mathbf{x}_{-i}) - u_i(\mathbf{x}_i, \mathbf{x}_{-i})| \leq \epsilon_n \|\mathbf{x}'_i - \mathbf{x}_i\|_1 \leq 2\epsilon_n$, for any $\mathbf{x} \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$ and $\mathbf{x}'_i \in \Delta(\mathcal{A}_i)$. The proof now follows that of Theorem A.2. In particular, we can lower bound the sum of the players' regrets as follows.

$$\begin{aligned} \sum_{i=1}^n \text{Reg}_i^{(T)} &= \sum_{i=1}^n \left(\max_{\mathbf{x}'_i \in \Delta(\mathcal{A}_i)} \left\{ \sum_{t=1}^T u_i(\mathbf{x}'_i, \mathbf{x}_{-i}^{(t)}) \right\} - \sum_{t=1}^T u_i(\mathbf{x}^{(t)}) \right) \\ &\geq - \sum_{i=1}^n \sum_{t=1}^T |u_i(\mathbf{x}'_i, \mathbf{x}_{-i}^{(t)}) - u_i(\mathbf{x}^{(t)})| \geq -2Tn\epsilon_n. \end{aligned}$$

As a result, following the proof of Theorem A.2, we conclude that for learning rate $\eta_n := \frac{1}{4L_n}$ there exists a time $t^* \in \llbracket T \rrbracket$ such that

$$\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \leq 4 \left(\frac{\max_{1 \leq i \leq n} D_{\mathcal{X}_i}^2}{\eta_n^2} + \max_{1 \leq i \leq n} B_i^2 \right) \left(\frac{2D_{\mathcal{X}}^2}{T} + 4\eta_n n \epsilon_n \right),$$

Thus, setting

$$T := \frac{D_{\mathcal{X}}^2}{2\eta_n n \epsilon_n} = \frac{2D_{\mathcal{X}}^2 L_n}{n \epsilon_n} \leq 2D_{\mathcal{X}}^2 \left(\max_{1 \leq i \leq n} |\mathcal{A}_i| \right) = O_n(n),$$

by Lemma A.8, yields that

$$\sum_{i=1}^n \left(\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \right)^2 \leq O_n(L_n n \epsilon_n) = O_n(n^2 \epsilon_n^2),$$

where the asymptotic notation above applies in the regime $L_n \geq \Omega_n(1)$. The statement thus follows. \square

We note that the conclusion of Theorem 3.6 also applies under the weaker notion of sensitivity defined in (18), if we additionally assume that there exists $\mathbf{x}^* \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$ such that for any $\mathbf{x} \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$,

$$\sum_{i=1}^n u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}) - \sum_{i=1}^n u_i(\mathbf{x}) \geq -n\epsilon_n. \quad (20)$$

Condition (20) can be met even in games with large sensitivity (Observation 3.4), which is why we chose to state Theorem 3.6 under a stronger but more interpretable condition.

Next, we point out two interesting extensions of our approach based on developments following the original smoothness framework of Roughgarden [90].

A.3 Refined smoothness

Throughout this paper we have relied on the original smoothness framework [90] to derive much of our results. Nevertheless, it is worth noting that there are certain extensions documented in the literature that can make our predictions sharper. To be more precise, focusing on normal-form games, Nadav and Roughgarden [75] noted that the original smoothness framework is in fact able to provide efficiency bounds for a set of equilibria even more permissive than CCE, which they refer to as *average* CCE with respect to a welfare-maximizing strategy $\mathbf{x}^* \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$ (ACCE*). In particular, the latter relaxation only requires a guarantee for the average regret over the players, $\frac{1}{n} \sum_{i=1}^n \text{Reg}_i^{(T)}(\mathbf{x}_i^*)$, while CCE instead requires a guarantee for the maximum regret $\max_{1 \leq i \leq n} \text{Reg}_i^{(T)}$. Based on this observation, Nadav and Roughgarden [75] developed a primal-dual framework in order to provide refined guarantees (beyond the original smoothness framework) for CCE*, which boils down to solving the following linear program.

$$\begin{aligned} & \text{maximize} && \rho \\ & \text{subject to} && \sum_{i=1}^n z_i (u_i(\mathbf{a}_i^*, \mathbf{a}_{-i}) - u_i(\mathbf{a})) \geq \rho \text{SW}(\mathbf{a}^*) - \text{SW}(\mathbf{a}), \mathbf{a} \in \prod_{i=1}^n \mathcal{A}_i, \\ & && z_i \geq 0. \end{aligned} \quad (21)$$

Here, it is assumed that the underlying game is in normal form, with $\prod_{i=1}^n \mathcal{A}_i$ being the set of joint action profiles, and $\mathbf{a}^* \in \prod_{i=1}^n \mathcal{A}_i$ a welfare-maximizing joint action. (The formulation of Nadav and Roughgarden [75] is based on cost-minimization games, but of course their LP can be cast directly for payoff-maximization games in the form of (21).) The above LP is a simple transformation of the fractional-linear program corresponding to optimizing $\rho := \frac{\lambda}{1+\mu}$ subject to the smoothness constraints of Definition 2.1, but with the additional flexibility of optimizing over a vector $\mathbf{z} \in \mathbb{R}_{\geq 0}^n$. In particular, when restricting $z_1 = z_2 = \dots = z_n$, this exactly recovers the LP for computing the robust price of anarchy—given in (24). Nevertheless, Nadav and Roughgarden [75] pointed out that the additional flexibility of (21) can have an arbitrarily large impact on the predicted efficiency [75, Remark 1].

The sharper definition of smoothness introduced by Nadav and Roughgarden [75] can be leveraged in our framework as follows. We can define a generalized robust price of anarchy as the solution of the LP (21). If this quantity approaches to 1 with a sufficiently fast rate (in terms of the number of players), we can extend Theorem A.2 by analyzing the *weighted* sum of the players' regrets $\sum_{i=1}^n z_i \text{Reg}_i^{(T)}$. It is easy to see that the argument of Theorem 3.1 readily carries over under the condition that the ratio $z_i/z_{i'}$ is bounded for any $i, i' \in [n]$, as well as the assumption that each z_i is bounded away from 0. In contrast, if there exists a pair $i, i' \in [n]$ for which the ratio $z_i/z_{i'}$ is unbounded, our current techniques do not appear to be applicable; this is related to the well-known difficulty of deriving so-called RVU bounds for the maximum of the players' regret, instead of their sum [103]. In other words, our current techniques can sharpen Theorem 3.1 by considering solutions of (21) under the additional (linear) constraint that the variables $\{z_i\}_{i=1}^n$ have bounded pairwise ratio. Given that the more general framework of Nadav and Roughgarden [75] leads to improved bounds, we expect that this modification should have applications in our setting as well. In particular, we point out that the 2-player example of Nadav and Roughgarden [75] that separates ACCE* from CCE* indeed satisfies $z_1/z_2 \approx 2.3$ [75, Proposition 2], so the separation manifests itself even when the pairwise ratio is bounded by an absolute constant.

We finally refer to the works of Nguyen [78] and Kulkarni and Mirrokni [62] for a different primal-dual take on smoothness.

A.4 Local smoothness

Another interesting extension of our techniques can be obtained using the framework of *local smoothness*, first introduced by Roughgarden and Schoppmann [92] and recently refined by Nguyen [78] in the context of splittable congestion games. In what follows, we follow the treatment of Nguyen

[78] as it fits our framework. In this context, Nguyen [78] introduced the notion of a (λ, μ) -dual-smooth differentiable utility function $u : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{\geq 0}$ with the property that for every vectors $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{R}_{\geq 0}^n$ and $\mathbf{z}' = (z'_1, \dots, z'_n) \in \mathbb{R}_{\geq 0}^n$,

$$z'u(z) + \sum_{i=1}^n z_i(z'_i - z_i)u'(z) \geq \lambda z'u(z') - \mu zu(z), \quad (22)$$

where $z := \sum_{i=1}^n z_i$ and $z' := \sum_{i=1}^n z'_i$. We also note that u' above denotes the derivative of u . (Again, (22) has been translated to utility-maximization games compared to the original formulation of Nguyen [78].) Furthermore, a splittable congestion game is called (λ, μ) -dual-smooth [78] if for every resource $e \in E$ the utility function $u_e : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{\geq 0}$ is (λ, μ) -dual-smooth in the sense of (22). The importance of Nguyen's extension is that it can be applied to *coarse* correlated equilibria, as opposed to the original definition of Roughgarden and Schoppmann [92] that was applicable to correlated equilibria; this distinction is incidentally important for our framework.

We will connect this concept of local smoothness with the linearization of the players' regrets. In particular, the utility of a player $i \in [n]$ under a joint strategy \mathbf{x} is defined as $u_i(\mathbf{x}) := \sum_{e \in E} \mathbf{x}_i[e] u_e(\sum_{i=1}^n \mathbf{x}_i[e])$. Then,

$$\frac{\partial u_i(\mathbf{x})}{\partial \mathbf{x}_i[e]} = u_e \left(\sum_{i=1}^n \mathbf{x}_i[e] \right) + \mathbf{x}_i[e] u'_e \left(\sum_{i=1}^n \mathbf{x}_i[e] \right).$$

As a result, for any $t \in \mathbb{N}$,

$$\begin{aligned} \sum_{i=1}^n \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \nabla_{\mathbf{x}_i} u_i(\mathbf{x}^{(t)}) \rangle &= \sum_{e \in E} u_e(\mathbf{x}[e]) \mathbf{x}^*[e] - \sum_{e \in E} u_e(\mathbf{x}[e]) \mathbf{x}[e] \\ &\quad + \sum_{e \in E} \sum_{i=1}^n \mathbf{x}_i[e] (\mathbf{x}_i^*[e] - \mathbf{x}_i[e]) u'_e(\mathbf{x}[e]), \end{aligned}$$

with the understanding that $\mathbf{x}[e] := \sum_{i=1}^n \mathbf{x}_i[e]$ and $\mathbf{x}^*[e] := \sum_{i=1}^n \mathbf{x}_i^*[e]$ for any $e \in E$. Now let us define $\text{Reg}_{\mathcal{L},i}^{(T)}(\mathbf{x}_i^*) := \sum_{t=1}^T \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \nabla_{\mathbf{x}_i} u_i(\mathbf{x}^{(t)}) \rangle$. Combining the last displayed equality with local smoothness (22), we get that

$$\begin{aligned} \sum_{i=1}^n \text{Reg}_{\mathcal{L},i}^{(T)}(\mathbf{x}_i^*) &\geq \lambda \sum_{e \in E} \mathbf{x}^*[e] u_e(\mathbf{x}^*[e]) - (\mu + 1) \sum_{e \in E} \mathbf{x}[e] u_e(\mathbf{x}[e]) \\ &= \lambda \text{SW}(\mathbf{x}^*) - (\mu + 1) \text{SW}(\mathbf{x}). \end{aligned} \quad (23)$$

Consequently, if we define a local smoothness bound, $\rho_{\mathcal{L}} := \frac{\lambda}{1+\mu}$ with (λ, μ) being subject to (22), Theorem 3.1 can be readily extended in the regime $\rho_{\mathcal{L}} \rightarrow 1$ based on (23).

A.5 Considerations based on PoA

It is natural to ask if the conclusion of Theorem 3.1 can be relaxed to $\text{PoA} \rightarrow 1$. Here, we point out that such an assumption does not suffice to obtain interesting guarantees. In particular, we next note Proposition 3.7, stated earlier in the main body, which is an immediate byproduct of the hardness of computing Nash equilibria in constant-sum games.

Proposition 3.7. *Even under the promise that $\text{PoA}_{\mathcal{G}} = 1$, computing a $(1/\text{poly}(\mathcal{G}))$ -Nash equilibrium in normal-form games in polynomial time is impossible when $n \geq 3$, unless $\text{PPAD} \subseteq \text{P}$.*

Proof. Chen et al. [20] showed that computing a $(1/\text{poly}(\max_i |\mathcal{A}_i|))$ -Nash equilibrium in a general-sum two-player game in normal form is PPAD-hard. As a result, the same applies in constant-sum 3-player games by suitably incorporating an additional player who has no strategic impact on the game. Further, in any constant-sum game \mathcal{G} it clearly holds that $\text{PoA}_{\mathcal{G}} = 1$, concluding the proof. \square

Remark A.10. The positive result established in Theorem A.2 requires that $\rho \approx 1$ under a bounded pair of smoothness parameters (λ, μ) . One may wonder whether similar conclusions apply even

when the smoothness parameters are unbounded. In particular, the *robust price of anarchy* (rPoA) can be defined as the solution to the following linear program.

$$\begin{aligned} & \text{maximize} && \rho \\ & \text{subject to} && z \sum_{i=1}^n (u_i(a_i^*, \mathbf{a}_{-i}) - u_i(\mathbf{a})) \geq \rho \text{SW}(\mathbf{a}^*) - \text{SW}(\mathbf{a}), \mathbf{a} \in \prod_{i=1}^n \mathcal{A}_i, \\ & && z \geq 0. \end{aligned} \quad (24)$$

Above $\mathbf{a}^* = (a_1^*, \dots, a_n^*) \in \prod_{i=1}^n \mathcal{A}_i$ is a welfare-maximizing action profile, which can be assumed to be unique for the purposes of our discussion here. In this context, can we extend the conclusion of Theorem A.2 under the assumption that $\text{rPoA}_{\mathcal{G}_n} \rightarrow 1$? In general, that is not possible. Indeed, in any constant-sum game one can make the LP (24) feasible by taking $z = 0$ and $\rho = 1$; that is, $\text{rPoA}_{\mathcal{G}_n} = 1$ for any constant-sum game \mathcal{G} . Thus, assuming merely that $\text{rPoA}_{\mathcal{G}} \rightarrow 1$ is not enough to obtain interesting guarantees for computing Nash equilibria (in accordance with Proposition 3.7). In other words, our underlying assumption that the smoothness parameters are bounded is necessary. To further explain this discrepancy, we note that a constant-sum game is generally not $(1 + \mu, \mu)$ -smooth (for a finite $\mu > -1$), for otherwise any constant-sum game would satisfy the Minty property (Proposition 3.3), which would in turn contradict Proposition 3.7. The case where $z = 0$ in any optimal solution of (24) essentially corresponds to a pathological manifestation of smoothness in which the underlying parameters are unbounded. We stress again that throughout this paper, when we say that a game is (λ, μ) -smooth we, of course, posit that those smoothness parameters are finite. Further, when we say that $\rho_{\mathcal{G}} = 1$ in a game \mathcal{G} , we accept that there are finite smoothness parameters associated with $\rho_{\mathcal{G}}$. With this convention, we reiterate that there are games in which $\text{rPoA}_{\mathcal{G}} \neq \rho_{\mathcal{G}}$.

Next, we provide a concrete example based on Shapley's game [99] in which OGD (unconditionally) fails to converge to an ϵ -Nash equilibrium, for a constant $\epsilon > 0$.

Example A.11. We consider a 3-player game \mathcal{G} defined as follows. We let

$$\mathbf{A} := \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \mathbf{B} := \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}. \quad (25)$$

Then, we set $u_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) := \mathbf{x}_1^\top \mathbf{A} \mathbf{x}_2$, $u_2(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) := \mathbf{x}_1^\top \mathbf{B} \mathbf{x}_2$, and $u_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) := 3 - \mathbf{x}_1^\top \mathbf{A} \mathbf{x}_2 - \mathbf{x}_1^\top \mathbf{B} \mathbf{x}_2$. Thus, for any joint strategy $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ it holds that $\text{SW}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = 3$, implying that $\text{PoA}_{\mathcal{G}} = 1$. Further, \mathcal{G} is in normal form. Now, through a numerical simulation we draw the following conclusion: Although $\text{PoA}_{\mathcal{G}} = 1$, for $T \gg 1$ OGD with learning rate $\eta := 0.01$ and initialization $(\hat{\mathbf{x}}_1^{(1)}, \hat{\mathbf{x}}_2^{(1)}, \cdot) := ((0.5, 0.25, 0.25), (0.25, 0.5, 0.25), \cdot)$ satisfies $\text{NEGAP}(\mathbf{x}^{(t)}) \geq 0.1875$ for any $t \in \llbracket T \rrbracket$, where $(\mathbf{x}^{(t)})_{t \geq 1}$ is the sequence of iterates produced by OGD. We note that here we do not consider the initialization from the uniform distribution simply because that happens to be the unique Nash equilibrium in Shapley's game; the conclusion above readily applies for any initialization by suitably modifying the underlying game. We also note that the specific value for the learning rate specified above is used for concreteness, and the conclusion is not tied to that specific value.

We next note another separation between the bound predicted by smoothness and PoA, which gives an additional reason why the smoothness framework is more suited as a criterion for determining tractability. Below, we make for simplicity the assumption that the game has a unique welfare-maximizing action profile, which can always be enforced by incorporating an arbitrarily small noise in the players' utilities.

Proposition A.12. *Determining whether a game is (λ, μ) -smooth can be done in polynomial time in explicitly represented normal-form games. In contrast, even in two-player games, determining PoA is NP-hard.*

Proof. First of all, the welfare-maximizing action profile $\mathbf{a}^* \in \prod_{i=1}^n \mathcal{A}_i$ can be trivially computed in polynomial time (in the size of the input) since the game is explicitly represented. For a legitimate (λ, μ) pair, determining whether the game is (λ, μ) -smooth can be phrased as a feasibility linear program, with a number of constraints equal to the number of possible joint action profiles, each corresponding to a separate constraint in (1); namely, $\sum_{i=1}^n u_i(a_i^*, \mathbf{a}_{-i}) \geq \lambda \text{OPT} - \mu \sum_{i=1}^n u_i(\mathbf{a})$. As a result, the number of constraints is polynomial in the description of the game (since it is assumed that the game is explicitly represented). Furthermore, one can optimize over the smoothness parameters by considering the LP (of polynomial size) given in (24), which determines the robust price

of anarchy (rPoA). In accordance with Theorem A.2 (see Remark A.10), one can also incorporate the constraint $z \geq 1/\text{poly}(\mathcal{G})$. Regarding the second claim, hardness of determining PoA even in two-player games follows directly from the reduction of Conitzer and Sandholm [27, Theorem 1]. In particular, an algorithm computing PoA would enable determining the satisfiability of a SAT formula. \square

One important question is whether smoothness can be identified in polynomial time even in *succinctly* represented games [79]. Indeed, the obvious algorithm for identifying smoothness described above requires a number of constraints that scales exponentially with the number of players, which is especially problematic in the regime of large games we focus on in Section 3.

We need to clarify, however, that knowing the smoothness parameters is certainly not a prerequisite for applying our approach. In detail, one can first of all apply Theorem 3.1 in classes of games where there are available analytical bounds for ρ_n as a function of n , obviating the need to determine whether ρ_n is close to 1. Besides this point, and more importantly, one can always execute the algorithm prescribed by Theorem 3.1—which does not require any knowledge regarding the smoothness parameters—and then efficiently evaluate the solution quality (per Definition 2.2). If the desired accuracy has been reached, then this is precisely the initial goal; otherwise, we can safely assume that the preconditions of Theorem 3.1 are not met.

A.6 Smoothness does not suffice for convergence

Next, we provide for completeness an example of a smooth game where OGD fails to converge to Nash equilibria. This shows that, as expected, it is not merely enough to know that $\rho \neq 0$ to obtain interesting guarantees.

Example A.13. This example is based on a bimatrix game in normal form described with the payoff matrices

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.8 & 0.9 & 0.3 \\ 0.2 & 0.8 & 0.2 & 0.3 \\ 0.9 & 0.2 & 0.4 & 0.4 \\ 0.6 & 0.9 & 0.3 & 0.1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0.4 & 0.2 & 0 & 0.1 \\ 0.5 & 0 & 0.2 & 0.8 \\ 0.7 & 0.8 & 0 & 0.4 \\ 0 & 0 & 0.1 & 0.4 \end{bmatrix}. \quad (26)$$

We first claim that this bimatrix game \mathcal{G} satisfies $\rho_{\mathcal{G}} \geq 0.125$. Indeed, we first see that the welfare-maximizing profile of (26) reads $(\mathbf{x}_1^*, \mathbf{x}_2^*) = ((0, 0, 1, 0), (1, 0, 0, 0))$. We also claim that for any pair of actions $a_1 \in \mathcal{A}_1$ and $a_2 \in \mathcal{A}_2$ it holds that

$$\sum_{i=1}^2 u_i(\mathbf{x}_i^*, a_{-i}) \geq 0.125 \cdot \text{OPT}_{\mathcal{G}},$$

where $\text{OPT}_{\mathcal{G}} = 1.6$. As a result, it follows that for any $\mathbf{x}_1 \in \Delta(\mathcal{A}_1)$ and $\mathbf{x}_2 \in \Delta(\mathcal{A}_2)$ it holds that $\mathbb{E}_{a_1 \sim \mathbf{x}_1, a_2 \sim \mathbf{x}_2} [\sum_{i=1}^2 u_i(\mathbf{x}_i^*, a_{-i})] \geq 0.125 \cdot \text{OPT}_{\mathcal{G}}$, in turn implying that $\sum_{i=1}^2 u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}) \geq 0.125 \cdot \text{OPT}_{\mathcal{G}}$. This means that \mathcal{G} is $(0.125, 0)$ -smooth. Furthermore, through a numerical simulation we draw the following conclusion: for $T \gg 1$, OGD with learning rate $\eta := 0.01$ satisfies $\text{NEGAP}(\mathbf{x}_1^{(t)}, \mathbf{x}_2^{(t)}) \geq 0.046$ for any $t \in \llbracket T \rrbracket$, where $(\mathbf{x}_1^{(t)}, \mathbf{x}_2^{(t)})_{t \geq 1}$ is the sequence of iterates produced by OGD. Once again, it is worth noting that this conclusion is not tied to the specific choice of learning rate we chose above, which is only used for concreteness.

A.7 Bayesian mechanisms

In this subsection, we consider the standard *independent private value* model of Bayesian mechanisms. In particular, each player $i \in \llbracket n \rrbracket$ has a *type* v_i drawn from a distribution \mathcal{F}_i over a finite set of types \mathcal{V}_i ; without any loss, we may assume that \mathcal{F}_i is the uniform distribution over \mathcal{V}_i . It is further assumed that players' types are pairwise independent. After each player $i \in \llbracket n \rrbracket$ draws a type $v_i \sim \mathcal{F}_i$, i selects an action $a_i(v_i) \in \mathcal{A}_i$, which is a function of its type v_i . Now consider a fixed profile of types $\mathbf{v} \in \mathcal{V} := \mathcal{V}_1 \times \mathcal{V}_2 \dots \mathcal{V}_n$. The (expected) utility of Player i under a joint strategy profile $\mathbf{x}(\mathbf{v})$ is denoted by $u_i(\mathbf{x}; v_i) := \mathbb{E}_{\mathbf{a} \sim \mathbf{x}} [u_i(\mathbf{a}; v_i)]$. There is also a *principal* agent who does not take an action in the game, and whose utility under \mathbf{x} is given by $R(\mathbf{x}) := \mathbb{E}_{\mathbf{a} \sim \mathbf{x}} [R(\mathbf{a})]$. Accordingly, the social welfare is defined as $\text{SW}(\mathbf{x}, \mathbf{v}) := \sum_{i=1}^n u_i(\mathbf{x}; v_i) + R(\mathbf{x})$, while $\text{OPT}_{\mathcal{M}}(\mathbf{v})$ represents the optimal social welfare of mechanism \mathcal{M} as a function of the joint type $\mathbf{v} \in \mathcal{V}$. We will make the assumption that the utility functions assign solely nonnegative values.

Smooth mechanisms Analogously to the notion of a smooth game (Definition 2.1), Syrgkanis and Tardos [102] introduced the notion of a smooth mechanism, formally recalled below.

Definition A.14 (Smooth mechanism [102]). A mechanism \mathcal{M} is (λ, μ) -smooth, where $\lambda, \mu \geq 0$, if there exists an strategy profile $\mathbf{x}^*(\mathbf{v}) \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$, for every type profile $\mathbf{v} \in \mathcal{V}$, such that for any action profile $\mathbf{a} \in \mathcal{A}$ and type profile $\mathbf{v} = (v_1, \dots, v_n) \in \mathcal{V}$,

$$\sum_{i=1}^n u_i(\mathbf{x}_i^*(\mathbf{v}), \mathbf{a}_{-i}; v_i) \geq \lambda \text{OPT}_{\mathcal{M}}(\mathbf{v}) - \mu R(\mathbf{a}). \quad (27)$$

As it turns out, many important mechanisms satisfy the above definition under various parameters (λ, μ) [102]. For a (λ, μ) -smooth mechanism \mathcal{M} , we define $\rho_{\mathcal{M}} := \frac{\lambda}{\max\{1, \mu\}}$.

Definition A.15. Let $\mathbf{x}_i : \mathcal{V}_i \rightarrow \Delta(\mathcal{A}_i)$ be the strategy of each player $i \in [n]$. A joint strategy profile \mathbf{x} is a *Bayes-Nash equilibrium (BNE)* if for any player $i \in [n]$, any type $v_i \in \mathcal{V}_i$ and deviation $a'_i \in \mathcal{A}_i$,

$$\mathbb{E}_{\mathbf{v}_{-i} \sim \mathcal{F}_{-i}} [u_i(\mathbf{x}(\mathbf{v}); v_i)] \geq \mathbb{E}_{\mathbf{v}_{-i} \sim \mathcal{F}_{-i}} [u_i(a'_i, \mathbf{x}_{-i}(\mathbf{v}_{-i}); v_i)]. \quad (28)$$

This definition coincides with the standard notion of Nash equilibrium we saw in Definition 2.2 when each distribution \mathcal{F}_i is a point mass. We also note that an ϵ -BNE incorporates an $\epsilon \geq 0$ additive slackness in (28).

Population interpretation of Bayesian games In the *agent-form* representation of a Bayesian game [53], it is assumed that there are n finite subpopulations of players, each corresponding to a player $i \in [n]$. Each player belonging to population i corresponds to a type $v_i \in \mathcal{V}_i$, which is distinct in each population and across populations. In this induced population game, nature first draws one player from each population, and then each player v_i selects an action $a_i(v_i)$; the game is then played under those selected actions. In symbols, the utility of Player v_i from population i reads

$$u_{i, v_i}^{\text{AG}}(\mathbf{a}) := \mathbb{E}_{\mathbf{v} \sim \mathcal{F}} [u_i(\mathbf{a}(\mathbf{v}); v_i) \mathbb{1}\{v_i\}], \quad (29)$$

where by $\{v_i\}$ above we denote the event that type v_i is selected by nature among the population corresponding to Player i . The importance of the population interpretation of the Bayesian game is that it induces a mechanism of complete information, which will be denoted by $\mathcal{M}^{\text{AG}} = \mathcal{M}^{\text{AG}}(\mathcal{M})$. The following characterization highlights an important connection between \mathcal{M}^{AG} and \mathcal{M} .

Theorem A.16 ([53]). *If a mechanism \mathcal{M} is (λ, μ) -smooth, then the complete information mechanism $\mathcal{M}^{\text{AG}} = \mathcal{M}^{\text{AG}}(\mathcal{M})$ is also (λ, μ) -smooth.*

We now proceed with the proof of Theorem 3.8. To keep the exposition self-contained, we will not make explicit use of Theorem A.16.

Theorem 3.8. *Consider a Bayesian mechanism \mathcal{M} such that $\rho_{\mathcal{M}} = 1$. Then, for any $\epsilon > 0$, $T = O(1/\epsilon^2)$ iterations of OGD suffice to obtain an ϵ -Bayes-Nash equilibrium of \mathcal{M} .*

Proof. Under the assumption that $\rho_{\mathcal{M}} = 1$, it follows that there exists a pair $(\lambda, \mu) \in \mathbb{R}_{\geq 0}^2$ such that \mathcal{M} is (λ, μ) -smooth with $\lambda = \max\{1, \mu\}$. Further, by definition it holds that $\text{OPT}_{\mathcal{M}}(\mathbf{v}) \geq \sum_{i=1}^n u_i(\mathbf{a}; v_i) + R(\mathbf{a})$, for any action profile $\mathbf{a} \in \mathcal{A}$. Combining with (27), we have that there exists a strategy profile $\mathbf{x}^*(\mathbf{v}) \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$ such that for every type profile $\mathbf{v} \in \mathcal{V}$ and action profile $\mathbf{a} \in \mathcal{A}$,

$$\sum_{i=1}^n u_i(\mathbf{x}_i^*(\mathbf{v}), \mathbf{a}_{-i}; v_i) \geq \lambda \sum_{i=1}^n u_i(\mathbf{a}; v_i) + \lambda R(\mathbf{a}) - \mu R(\mathbf{a}) \geq \sum_{i=1}^n u_i(\mathbf{a}; v_i),$$

since $\lambda = \max\{1, \mu\}$ and utilities are nonnegative. As a result, for any $\mathbf{x}(\mathbf{v}) \in \prod_{i=1}^n \Delta(\mathcal{A}_i)$,

$$\sum_{i=1}^n \mathbb{E}_{\mathbf{v} \sim \mathcal{F}} [u_i(\mathbf{x}_i^*(\mathbf{v}), \mathbf{x}_{-i}(\mathbf{v}_{-i}); v_i)] \geq \sum_{i=1}^n \mathbb{E}_{\mathbf{v} \sim \mathcal{F}} [u_i(\mathbf{x}(\mathbf{v}); v_i)]. \quad (30)$$

Further, by definition of the agent-form utilities (29) and the law of total expectation, (30) can be equivalently cast as

$$\sum_{i=1}^n \sum_{v_i \in \mathcal{V}_i} u_{i, v_i}^{\text{AG}}(\mathbf{x}_i^*, \mathbf{x}_{-(i, v_i)}) \geq \sum_{i=1}^n \sum_{v_i \in \mathcal{V}_i} u_{i, v_i}^{\text{AG}}(\mathbf{x}),$$

where we used the notation $\mathbf{x}_{i,v_i} := \mathbf{x}_i(v_i)$. This implies that the sum of the players' regrets in the agent-form representation is nonnegative. The proof of Theorem 3.8 then follows from the correspondence between Nash equilibria in the agent-form representation of \mathcal{M} and Bayes-Nash equilibria in the original incomplete-information mechanism \mathcal{M} . \square

A.8 PoA vs robust PoA

In this subsection, we provide some additional justification for the condition $\text{rPoA} \neq \text{PoA}$ required in Corollary 4.3. In particular, we conduct experiments on a set of random normal-form games. Some illustrative results for 10 random games are demonstrated in Figure 1. Overall, we observe that not only $\text{rPoA} \neq \text{PoA}$, but in fact the gap between the two quantities is typically substantial. This discrepancy is expected since, as we have already stressed, rPoA quantifies the worst-case welfare over a (typically) much broader set of equilibria.

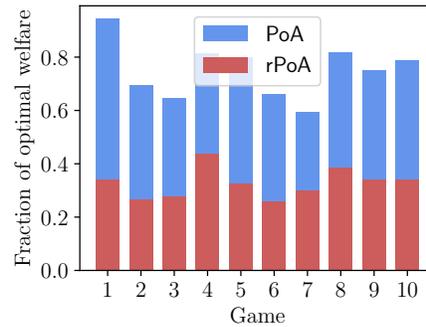


Figure 1: PoA versus rPoA in random normal-form games.

A.9 Proof of Theorem 4.1

In this subsection, we provide the proof of Theorem 4.1. For completeness, we also include the proofs of certain results known from prior work [37, 85, 77].

To this end, we first recall that the *prox* operator associated with the (squared) Euclidean regularizer $\frac{1}{2} \|\cdot\|_2^2$ is defined as

$$\Pi_{\mathcal{X}_i}(\mathbf{u}_i) := \arg \max_{\mathbf{x}'_i \in \mathcal{X}_i} \left\{ \langle \mathbf{x}'_i, \mathbf{u}_i \rangle - \frac{1}{2} \|\mathbf{x}_i - \mathbf{x}'_i\|_2^2 \right\}, \quad (31)$$

under some utility vector $\mathbf{u}_i \in \mathbb{R}^{d_i}$. Accordingly, we let $\Pi_{\mathbf{x}}(\mathbf{u}) := (\Pi_{\mathcal{X}_1}(\mathbf{u}_1), \dots, \Pi_{\mathcal{X}_n}(\mathbf{u}_n))$, where $\mathbf{u} := (\mathbf{u}_1, \dots, \mathbf{u}_n)$. With this definition in mind, we note the following property of the prox operator.

Lemma A.17 ([37, 77]). *The prox operator $\Pi_{\mathbf{x}}(\cdot)$ is 1-Lipschitz continuous with respect to $\|\cdot\|_2$ for any $\mathbf{x} \in \prod_{i=1}^n \mathcal{X}_i$.*

Proof. Let $i \in [n]$. For any $\mathbf{u}_i, \mathbf{u}'_i \in \mathbb{R}^{d_i}$ and $\mathbf{x}_i \in \mathcal{X}_i$,

$$\|\Pi_{\mathcal{X}_i}(\mathbf{u}_i) - \Pi_{\mathcal{X}_i}(\mathbf{u}'_i)\|_2 = \|\mathcal{P}_{\mathcal{X}_i}(\mathbf{x}_i + \mathbf{u}_i) - \mathcal{P}_{\mathcal{X}_i}(\mathbf{x}_i + \mathbf{u}'_i)\|_2 \leq \|\mathbf{u}_i - \mathbf{u}'_i\|_2,$$

where we used the fact that the projection operator is non-expansive with respect to $\|\cdot\|_2$, and that

$$\begin{aligned} \Pi_{\mathcal{X}_i}(\mathbf{u}_i) &= \arg \max_{\mathbf{x}'_i \in \mathcal{X}_i} \left\{ \langle \mathbf{x}'_i, \mathbf{u}_i \rangle - \frac{1}{2} \|\mathbf{x}_i - \mathbf{x}'_i\|_2^2 \right\} = \arg \max_{\mathbf{x}'_i \in \mathcal{X}_i} \left\{ -\frac{1}{2} \|\mathbf{x}'_i\|_2^2 + \langle \mathbf{x}'_i, \mathbf{u}_i + \mathbf{x}_i \rangle \right\} \\ &= \arg \min_{\mathbf{x}'_i \in \mathcal{X}_i} \left\{ \|\mathbf{x}'_i - (\mathbf{u}_i + \mathbf{x}_i)\|_2^2 \right\} \\ &= \mathcal{P}_{\mathcal{X}_i}(\mathbf{x}_i + \mathbf{u}_i). \end{aligned}$$

This implies that $\|\Pi_{\mathbf{x}}(\mathbf{u}) - \Pi_{\mathbf{x}}(\mathbf{u}')\|_2 \leq \|\mathbf{u} - \mathbf{u}'\|_2$, where $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ and $\mathbf{u}' = (\mathbf{u}'_1, \dots, \mathbf{u}'_n)$. \square

Using this lemma, a key observation for the analysis of *clairvoyant* mirror descent is the following contraction property [85, 37].

Proposition A.18 ([85, 37]). *Suppose that F is L -Lipschitz continuous. For any $\mathbf{x}' \in \prod_{i=1}^n \mathcal{X}_i$ the function*

$$\prod_{i=1}^n \mathcal{X}_i \ni \mathbf{w} \mapsto \Pi_{\mathbf{x}'}(\eta F(\mathbf{w})) \quad (32)$$

is (ηL) -Lipschitz continuous. As a result, function (32) is a contraction mapping as long as $\eta < \frac{1}{L}$.

This follows directly from Lemma A.17: $\|\Pi_{\mathbf{x}'}(\eta F(\mathbf{w})) - \Pi_{\mathbf{x}'}(\eta F(\mathbf{w}'))\|_2 \leq \eta \|F(\mathbf{w}) - F(\mathbf{w}')\|_2 \leq \eta L \|\mathbf{w} - \mathbf{w}'\|_2$.

Proposition A.18 reassures us that fixed points of the contraction mapping (32) not only exist, but ϵ -approximate fixed points can also be computed in a time proportional to $\log(1/\epsilon)$ [85]. In this context, if $\mathbf{x}^{(0)} \in \prod_{i=1}^n \mathcal{X}_i$ is an arbitrary point, we consider the update rule defined for $t \in \mathbb{N}$ via

$$\mathbf{x}^{(t)} = \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}^{(t)})), \quad (33)$$

where $\mathbf{w}^{(t)} \in \prod_{i=1}^n \mathcal{X}_i$ is any point such that $\|\mathbf{w}^{(t)} - \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}^{(t)}))\|_2 \leq \epsilon^{(t)}$. It is important to note that this sequence $(\mathbf{x}^{(t)})_{t \geq 1}$ is not uniquely defined, but with a slight abuse we refer to any sequence satisfying (33) as *clairvoyant gradient descent* (CGD). CGD satisfies the following remarkable regret bound.

Theorem A.19 ([85, 37]). *For any $T \in \mathbb{N}$, the regret of player $i \in \llbracket n \rrbracket$ under CGD satisfies*

$$\text{Reg}_i^{(T)} \leq \frac{D_{\mathcal{X}_i}^2}{2\eta} - \frac{1}{2\eta} \sum_{t=1}^T \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2^2 + D_{\mathcal{X}_i} L_i \sum_{t=1}^T \epsilon^{(t)}, \quad (34)$$

where $L_i \in \mathbb{R}_{>0}$ is the Lipschitz constant of \mathbf{u}_i with respect to $\|\cdot\|_2$.

Proof. Given that $\mathbf{x}_i^{(t)} := \Pi_{\mathbf{x}_i^{(t-1)}}(\eta \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}))$, the first-order optimality condition implies that for any $\mathbf{x}_i^* \in \mathcal{X}_i$ and $t \in \llbracket T \rrbracket$,

$$\langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \eta \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - (\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}) \rangle \geq 0. \quad (35)$$

We further have that

$$\langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)} \rangle = -\frac{1}{2} \|\mathbf{x}_i^* - \mathbf{x}_i^{(t)}\|_2^2 + \frac{1}{2} \|\mathbf{x}_i^* - \mathbf{x}_i^{(t-1)}\|_2^2 - \frac{1}{2} \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2^2,$$

thereby implying through a telescopic summation that

$$\sum_{t=1}^T \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)} \rangle \leq \frac{1}{2} D_{\mathcal{X}_i}^2 - \frac{1}{2} \sum_{t=1}^T \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2^2.$$

Combining with (35),

$$\sum_{t=1}^T \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) \rangle \leq \frac{1}{2\eta} D_{\mathcal{X}_i}^2 - \frac{1}{2\eta} \sum_{t=1}^T \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2^2.$$

The claim thus follows from the fact that

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle &\geq - \sum_{t=1}^T \|\mathbf{x}_i^* - \mathbf{x}_i^{(t)}\|_2 \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2 \\ &\geq -D_{\mathcal{X}_i} L_i \sum_{t=1}^T \epsilon^{(t)}. \end{aligned}$$

□

Unlike prior work, here we will make crucial use of the negative term in (34). The important lemma below will enable us to express the regret bound (34) in terms of i 's best response gap; it is analogous to Lemma A.1 we stated earlier for OGD.

Lemma A.20. Fix any $t \in \mathbb{N}$ and let $\mathbf{x}^{(t)} = \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}^{(t)}))$, where $\mathbf{w}^{(t)} \in \prod_{i=1}^n \mathcal{X}_i$ is any point such that $\|\mathbf{w}^{(t)} - \Pi_{\mathbf{x}^{(t-1)}}(\eta F(\mathbf{w}^{(t)}))\|_2 \leq \epsilon^{(t)}$. Then,

$$\text{BRGAP}_i(\mathbf{x}^{(t)}) \leq D_{\mathcal{X}_i} \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2 + \frac{D_{\mathcal{X}_i}}{\eta} \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2.$$

Proof. By the first-order optimality condition, it follows that for any $\mathbf{x}_i^* \in \mathcal{X}_i$,

$$\left\langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \eta \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - (\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}) \right\rangle \geq 0,$$

in turn implying that

$$\langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) \rangle \geq -\frac{1}{\eta} \langle \mathbf{x}_i^* - \mathbf{x}_i^{(t)}, \mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)} \rangle. \quad (36)$$

Furthermore,

$$\begin{aligned} \langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) \rangle &= \langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle + \langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle \\ &\leq \langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle + \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^*\|_2 \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2 \\ &\leq \langle \mathbf{x}_i^{(t)} - \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle + D_{\mathcal{X}_i} \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2. \end{aligned} \quad (37)$$

Combining (36) and (37) yields that

$$\langle \mathbf{x}_i^{(t)}, \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle - \max_{\mathbf{x}_i^* \in \mathcal{X}_i} \langle \mathbf{x}_i^*, \mathbf{u}_i(\mathbf{x}_{-i}^{(t)}) \rangle \geq -D_{\mathcal{X}_i} \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2 - \frac{D_{\mathcal{X}_i}}{\eta} \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2,$$

concluding the proof. \square

Corollary A.21. For any $T \in \mathbb{N}$, the regret of player $i \in [n]$ under CGD can be bounded as

$$\text{Reg}_i^{(T)} \leq \frac{D_{\mathcal{X}_i}^2}{2\eta} - \frac{\eta}{4D_{\mathcal{X}_i}^2} \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + L_i D_{\mathcal{X}_i} \sum_{t=1}^T \epsilon^{(t)} + \frac{1}{2} \eta \sum_{t=1}^T \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2^2.$$

In particular, if $\eta := \frac{1}{2L}$ and $\epsilon^{(t)} \leq \frac{D_{\mathcal{X}_i}}{t^2}$ for any $t \in [T]$,

$$\text{Reg}_i^{(T)} \leq 3LD_{\mathcal{X}_i}^2 - \frac{1}{8LD_{\mathcal{X}_i}^2} \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + \frac{1}{2} \eta \sum_{t=1}^T \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2^2.$$

Proof. By Lemma A.20, it follows that for any $t \in [T]$,

$$\frac{1}{2\eta} \|\mathbf{x}_i^{(t)} - \mathbf{x}_i^{(t-1)}\|_2^2 \geq \frac{\eta}{4D_{\mathcal{X}_i}^2} \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 - \frac{1}{2} \eta \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2^2.$$

Summing over all $t \in [T]$ and combining with Theorem A.19 implies the statement since $L_i D_{\mathcal{X}_i} \sum_{t=1}^T \epsilon^{(t)} \leq L_i D_{\mathcal{X}_i}^2 \sum_{t=1}^T \frac{1}{t^2} \leq 2LD_{\mathcal{X}_i}^2$. \square

We are now ready to prove Theorem 4.1, which is recalled below.

Theorem 4.1. Suppose that all players are updating their strategies using CGD with $\epsilon^{(t)} \leq \frac{\min_i D_{\mathcal{X}_i}}{t^2}$ and learning rate $\eta = \frac{1}{2L}$ in a (λ, μ) -smooth game \mathcal{G} , where L is the Lipschitz-continuity parameter of F . Then, for any $\epsilon_0 > 0$ and $T \geq \frac{64L^2 D_{\mathcal{X}}^4}{\epsilon_0^2}$ iterations,

1. the average correlated distribution of play is a $\frac{4LD_{\mathcal{X}}^2}{T}$ - CCE;
2. there is a time $t^* \in [T]$ such that

$$\text{SW}(\mathbf{x}^{(t^*)}) \geq \sup_{\epsilon \geq \epsilon_0} \min \left\{ \rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} + \frac{\epsilon^2}{16(\mu+1)LD_{\mathcal{X}}^2}, \text{PoA}_{\mathcal{G}}^{\epsilon} \cdot \text{OPT}_{\mathcal{G}} \right\}. \quad (3)$$

Proof. First, by Corollary A.21,

$$\begin{aligned}
 \sum_{i=1}^n \text{Reg}_i^{(T)} &\leq 3L \sum_{i=1}^n D_{\mathcal{X}_i}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \sum_{i=1}^n \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + \frac{1}{2}\eta \sum_{t=1}^T \sum_{i=1}^n \|\mathbf{u}_i(\mathbf{w}_{-i}^{(t)}) - \mathbf{u}_i(\mathbf{x}_{-i}^{(t)})\|_2^2 \\
 &= 3LD_{\mathcal{X}}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \sum_{i=1}^n \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + \frac{1}{2}\eta \sum_{t=1}^T \|F(\mathbf{w}^{(t)}) - F(\mathbf{x}^{(t)})\|_2^2 \\
 &\leq 3LD_{\mathcal{X}}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \sum_{i=1}^n \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + \frac{1}{2}\eta L^2 \sum_{t=1}^T \|\mathbf{w}^{(t)} - \mathbf{x}^{(t)}\|_2^2 \\
 &\leq 3LD_{\mathcal{X}}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \sum_{i=1}^n \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2 + \frac{1}{2}\eta L^2 \sum_{t=1}^T (\epsilon^{(t)})^2 \\
 &\leq 4LD_{\mathcal{X}}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \sum_{i=1}^n \sum_{t=1}^T \left(\text{BRGAP}_i(\mathbf{x}^{(t)}) \right)^2, \tag{38}
 \end{aligned}$$

where we used the fact that $\sum_{t=1}^T (\epsilon^{(t)})^2 \leq D_{\mathcal{X}}^2 \sum_{t=1}^T \frac{1}{t^4} \leq 2D_{\mathcal{X}}^2$ and $\eta = \frac{1}{2L}$. As a result, Item 1 follows directly from (38) by invoking the well-known fact that the CCE gap is bounded by $\max_{1 \leq i \leq n} \text{Reg}_i^{(T)}$.

For Item 2, let us fix any $\epsilon \geq \epsilon_0$. Suppose that at every time $t \in \llbracket T \rrbracket$ it holds that $\text{BRGAP}_i(\mathbf{x}^{(t)}) > \epsilon$ for some player $i \in \llbracket n \rrbracket$. Then, by (38) we conclude that

$$\sum_{i=1}^n \text{Reg}_i^{(T)} \leq 4LD_{\mathcal{X}}^2 - \frac{1}{8LD_{\mathcal{X}}^2} \epsilon^2 T \leq -\frac{1}{16LD_{\mathcal{X}}^2} \epsilon^2 T,$$

since $T \geq \frac{64L^2 D_{\mathcal{X}}^4}{\epsilon^2} \geq \frac{64L^2 D_{\mathcal{X}}^4}{\epsilon^2}$. As a result, by (λ, μ) -smoothness, it follows that

$$\sum_{i=1}^n \text{Reg}_i^{(T)} \geq \lambda \text{OPT}_{\mathcal{G}} T - (1 + \mu) \sum_{t=1}^T \text{SW}(\mathbf{x}^{(t)}),$$

in turn implying that

$$\frac{1}{T} \sum_{t=1}^T \text{SW}(\mathbf{x}^{(t)}) \geq \rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} - \frac{1}{T} \sum_{i=1}^n \text{Reg}_i^{(T)} \geq \rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} + \frac{\epsilon^2}{16(\mu + 1)LD_{\mathcal{X}}^2}.$$

As a result, there is a time $t^* \in \llbracket T \rrbracket$ such that $\text{SW}(\mathbf{x}^{(t^*)}) \geq \rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} + \frac{\epsilon^2}{16(\mu + 1)LD_{\mathcal{X}}^2}$. In the contrary case, if there is a time $t^* \in \llbracket T \rrbracket$ such that $\text{BRGAP}_i(\mathbf{x}^{(t^*)}) \leq \epsilon$ for any player $i \in \llbracket n \rrbracket$, it follows that $\text{SW}(\mathbf{x}^{(t^*)}) \geq \text{PoA}_{\mathcal{G}}^{\epsilon} \cdot \text{OPT}_{\mathcal{G}}$ (by definition). This concludes the proof. \square

In particular, we note that Corollary 4.3 stated earlier in the main body is an immediate consequence of Theorem 4.1 under Condition 4.2 since $\text{rPoA}_{\mathcal{G}} \geq \rho_{\mathcal{G}}$. It is also worth noting that for the special case of normal-form games, one can state Theorem 4.1 so that the first term in the right-hand side of (3) reads $\rho_{\mathcal{G}}(\lambda, \mu) \cdot \text{OPT}_{\mathcal{G}} + \frac{\epsilon^2}{32(\mu + 1)L}$; thus, for a broad class of games (see Lemmas A.7 to A.9), the improvement over the smoothness bound is an absolute constant when ϵ and μ are also bounded by absolute constants. Relatedly, we should note that Theorem 6.2 applies in the regime where μ is polynomially bounded (we are not aware of any application of smoothness where this is not the case).

Remark A.22. It is direct to see that Item 2 of Theorem 4.1 can be refined so that there is a set $S \subseteq \llbracket T \rrbracket$, with $|S| \geq (1 - \gamma)T$, so that

$$\frac{1}{|S|} \sum_{t \in S} \text{SW}(\mathbf{x}^{(t)}) \geq \sup_{\epsilon \geq \epsilon_0} \min \left\{ \rho_{\mathcal{G}} \cdot \text{OPT}_{\mathcal{G}} + \frac{\gamma \epsilon^2}{16(\mu + 1)LD_{\mathcal{X}}^2}, \text{PoA}_{\mathcal{G}}^{\epsilon} \cdot \text{OPT}_{\mathcal{G}} \right\}.$$

Remark A.23. Theorem 4.1 can also be refined using the primal-dual framework of Nadav and Roughgarden [75] discussed in Appendix A.3, with the caveat that the variables $\{z_i\}_{i=1}^n$ need to

again have a bounded pairwise ratio. It is interesting to note, however, that using CGD enables making non-trivial conclusions even when a subset of the variables $\{z_i\}_{i=1}^n$ are zero in an optimal solution. In particular, when the quantity $\sum_{i=1}^n z_i \text{Reg}_i^{(T)}$ is nonnegative, we can readily draw the non-trivial conclusion that all players in the set $\{i \in [n] : z_i \neq 0\}$ are eventually best responding (by virtue of Corollary A.21). It is not at all clear if such conclusions apply to OGD as well. In other words, using CGD one can essentially replace PoA_G^ϵ in Theorem 4.1 by the price of anarchy with respect to a solution concept in which a *particular subset* of players are best responding. Shedding light to this type of guarantee necessitates understanding the implications of having some of the variables $\{z_i\}_{i=1}^n$ being zero in an optimal solution of the LP (21). Relatedly, even if we constraint $z_1 = z_2 = \dots = z_n = z$, can we characterize the games for which $z \approx 0$? We explained earlier in Remark A.10 that such pathologies (typically) occur in constant-sum games, for which questions concerning social welfare are trivial.

A.10 Beyond smoothness

In this subsection, we discuss an example in which the welfare predicted by the smoothness framework is far from the welfare obtained by learning algorithms such as OGD. We then suggest a natural direction that enables obtaining sharper predictions; we have already seen refined guarantees beyond the standard smoothness framework in Appendices A.3 and A.4, but here we explore a different direction.

Our example is again based on Shapley’s game, a bimatrix game in normal form defined with the matrices

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \quad (39)$$

Claim A.24. *Let \mathcal{G} be the bimatrix game defined in (39). Then, $\text{rPoA}_{\mathcal{G}} = 0$.*

Proof. Let $(\mathbf{x}_1^*, \mathbf{x}_2^*) := ((1, 0, 0), (0, 1, 0))$ be a welfare-maximizing joint action. By symmetry, the following argument will apply to any welfare-maximizing joint action. Consider $(\mathbf{x}_1, \mathbf{x}_2) = ((0, 1, 0), (0, 1, 0))$. Then, we have that $u_1(\mathbf{x}_1^*, \mathbf{x}_2) + u_2(\mathbf{x}_2^*, \mathbf{x}_1) = (\mathbf{x}_1^*)^\top \mathbf{A} \mathbf{x}_2 + \mathbf{x}_1^\top \mathbf{B} \mathbf{x}_2^* = 0 + 0$. As a result, the smoothness constraint corresponding to $(\mathbf{x}_1, \mathbf{x}_2)$ necessitates that $0 \geq \lambda$, which is incompatible with the constraint that $\lambda > 0$ (Definition 2.1). \square

In spite of the fact that $\text{rPoA}_{\mathcal{G}} = 0$, we see in Figure 2 (left) that OGD approaches close to the optimal social welfare 1. This can be explained by the fact that certain joint actions—in particular, those in the diagonal—are played with small probability under OGD (right-side of Figure 2). This motivates introducing a more refined notion of smoothness in which the smoothness constraints are only enforced on the joint actions that are visited with a non-negligible probability.

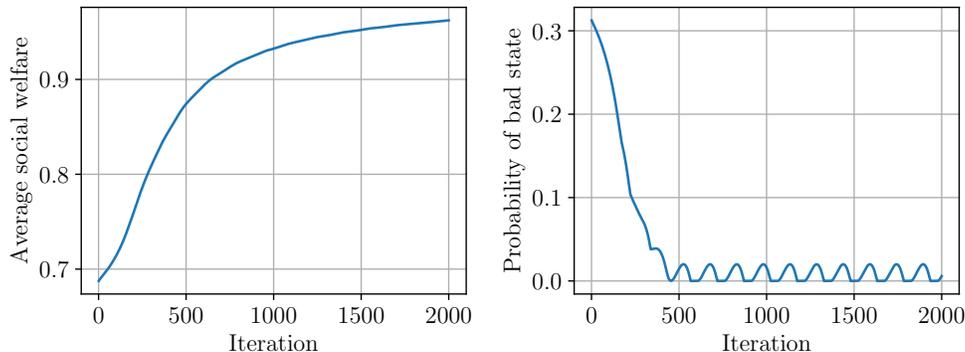


Figure 2: The behavior of OGD in the bimatrix game (39) with $\eta := 0.01$. On the left, we plot the average social welfare of the dynamics. On the right, we plot the probability of playing a joint action in the diagonal.

One class of games in which one can easily make such refinements concerns games solvable under iterated elimination of strictly dominated actions. As a concrete example, let

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (40)$$

Observation A.25. *Let \mathcal{G} be the game defined in (40). Then, $\text{rPoA}_{\mathcal{G}} = \frac{1}{2}$. Furthermore, if \mathcal{G}' is the game⁴ resulting from iterative removal of strictly dominated actions, then $\text{rPoA}_{\mathcal{G}'} = 1$.*

⁴Under elimination of strictly dominated actions, this is always uniquely defined.