TOWARDS CONVERGENCE TO NASH EQUILIBRIA IN TWO-TEAM ZERO-SUM GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Contemporary applications of machine learning in two-team e-sports and the superior expressivity of multi-agent generative adversarial networks raise important and overlooked theoretical questions regarding optimization in two-team games. Formally, two-team zero-sum games are defined as multi-player games where players are split into two competing sets of agents, each experiencing a utility identical to that of their teammates and opposite to that of the opposing team. We focus on the solution concept of Nash equilibria (NE). We first show that computing NE for this class of games is *hard* for the complexity class CLS. To further examine the capabilities of online learning algorithms in games with full-information feedback, we propose a benchmark of a simple —yet nontrivial— family of such games. These games do not enjoy the properties used to prove convergence for relevant algorithms. In particular, we use a dynamical systems perspective to demonstrate that gradient descent-ascent, its optimistic variant, optimistic multiplicative weights update, and extra gradient fail to converge (even locally) to a Nash equilibrium. On a brighter note, we propose a first-order method that leverages control theory techniques and under some conditions enjoys last-iterate local convergence to a Nash equilibrium. We also believe our proposed method is of independent interest for general min-max optimization.

1 INTRODUCTION

Online learning shares an enduring relationship with game theory that has a very early onset dating back to the analysis of *fictitious play* by (Robinson, 1951) and Blackwell's *approachability theorem* (Blackwell, 1956). A key question within this context is whether self-interested agents can arrive at a game-theoretic *equilibrium* in an *independent* and *decentralized* manner with only *limited feedback* from their environment. Learning dynamics that converge to different notions of equilibria are known to exist for two-player zero-sum games (Robinson, 1951; Arora et al., 2012; Daskalakis et al., 2011), potential games (Monderer & Shapley, 1996), near-potential games (Anagnostides et al., 2022b), socially concave games (Golowich et al., 2020), and extensive form games (Anagnostides et al., 2022a). We try to push the boundary further and explore whether equilibria —in particular, Nash equilibria— can be reached by agents that follow decentralized learning algorithms in two-team zero-sum games.

Team competition has played a central role in the development of game theory Marschak (1955); von Stengel & Koller (1997); Bacharach (1999); Gold (2005), economics Marschak (1955); Gottinger (1974), and evolutionary biology Nagylaki (1993); Nowak et al. (2004). Recently, competition among teams has attracted the interest of the machine learning community due to the advances that multi-agent systems have accomplished: *e.g.*, multi-GAN's (Hoang et al., 2017; Hardy et al., 2019) for generative tasks, adversarial regression with multiple learners (Tong et al., 2018), or AI agents competing in e-sports (*e.g.,* CTF (Jaderberg et al., 2019) or Starcraft Vinyals et al. (2019)) as well as card games (Moravčík et al., 2017; Brown & Sandholm, 2018; Bowling et al., 2015).

Our class of games. We turn our attention to *two-team zero-sum games* a quite general class of min-max optimization problems that include bilinear games and a wide range of nonconvex-nonconcave games as well. In this class of games, players fall into two teams of size n, m and submit their own randomized strategy vectors independently. We note that the games that we focus on are

not restricted to team games in the narrow sense of the term "team" as we use it in sports, games, and so on; the players play independently and do not have a central coordinating authority. Rather, for the purpose of this paper, *teams* are constituted by agents that merely enjoy the same utility function. This might already hint that the solution concept that we engage with is the *Nash equilibrium* (NE). Another class of games that is captured by this framework is the class of *adversarial potential games.* In these games, the condition that all players of the same team experience the same utility is weakened as long as there exists a *potential function* that can track differences in the utility of each player when they unilaterally deviate from a given strategy profile (see Appendix A.2 for a formal definition). A similar setting has been studied in the context of nonatomic games (Babaioff et al., 2009).

Positive duality gap. In two-player zero-sum games, *i.e.*, $n = m = 1$, min-max (respectively max-min) strategies are guaranteed to form a Nash equilibrium due to Von Neumann's minmax theorem Von Neumann (1928); ultimately endowing the game with a unique value. The challenges arise for the case of $n, m > 1$; Schulman & Vazirani (2019b) prove that, in general, two-team games do not have a unique value. They do so by presenting a family of team games with a positive duality gap, together with bounds concerning this gap. These bounds quantify the effect of exchanging the order of commitment to their strategy either between the teams as a whole or the individual players.

Solution concept. In this work, we examine the solution concept of Nash equilibrium (NE). Under a Nash equilibrium, no player can improve their utility by unilaterally deviating. The main downside of a NE for team games is the fact that such an equilibrium can be arbitrarily suboptimal for the team Basilico et al. (2017a).

This is one of the reasons that the solution concept of team-maxmin equilibrium with a coordination device (TMECor) has dominated contemporary literature of team games, especially in regard to applications Farina et al. (2018); Zhang et al. (2020); Cacciamani et al. (2021). Under a TMECor, players are allowed to communicate before the game and decide upon combinations of strategies to be played during the game using an external source of randomness.

The undeniable advantage of a TMECor is that the expected utility of the team under it is greater than the expected utility under a NE (Basilico et al., 2017a). Nevertheless, this favorable property of TMECor can by no means render the study of NE irrelevant. In fact, the study of NE is always of independent interest within the literature of algorithmic game theory —especially questions corresponding to computational complexity. Moreover, there exist settings in which *ex ante* coordination cannot be expected to be possible or even sensible; for example in (i) environments where the external sources of randomness are unreliable or nonexistent or visible to the adversarial team, (ii) games in which players cannot know in advance who they share a common utility with, (iii) *adversarial potential games*. These games can model naturally occurring settings such as (a) security games with multiple uncoordinated defenders versus multiple similarly uncoordinated attackers, (b) the load balancing "game" between telecommunication service providers trying to minimize the maximum delay of service experienced by their customers versus the service users that try to individually utilize the maximum amount of broadband possible, and (c) the *weak selection* model of evolutionary biology where a species as a whole is a team, the genes of its population are the players and the alleles of each gene are in turn the actions of a player; the allele frequencies are independent across genes (Nagylaki, 1993; Nowak et al., 2004; Mehta et al., 2015).

Concluding, we could not possibly argue for a single correct solution concept for two-team games; there is no silver bullet. In contrast, one has to assess which is the most fitting based on the constraints of a given setting. A Nash equilibrium is a cornerstone concept of game theory and examining its properties in different games is always important.

The optimization point of view. As we opt for the solution concept of NE, we first note that computing local-NE in general nonconvex-nonconcave games is PPAD-complete Daskalakis et al. (2009; 2021). Thus, all well-celebrated online learning, first-order methods like gradient descent-ascent Lin et al. (2020); Daskalakis & Panageas (2019), its optimistic Popov (1980); Chiang et al. (2012); Sridharan & Tewari (2010), optimistic multiplicative weights update Sridharan (2012), and the extra gradient method Korpelevich (1976) would require an exponential number of steps in the parameters of the problem in order to compute an approximate NE under the oracle optimization model of Nemirovskij & Yudin (1983) . Additionally, in the continuous time regime, similar classes

of games exhibit behaviors antithetical to convergence like cycling, recurrence, or chaos Vlatakis-Gkaragkounis et al. (2019). Second, even if a regret notion within the context of team-competition could be defined, no-regret dynamics are guaranteed to converge only to the set of coarse correlated equilibria (CCE) (Fudenberg, 1991; Hannan, 2016). CCE is a weaker equilibrium notion whose solutions could potentially be exclusively supported on strictly dominated strategies, even for simple symmetric two-player games (See also Viossat & Zapechelnyuk (2013)).

Surely, the aforementioned intractability remarks for the general case of nonconvex-nonconcave min-max problems provide a significant insight. But, they cannot *per se* address the issue of computing Nash equilibria when the game is equipped with a particular structure, *i.e.*, that of two-team zero-sum games. In fact, our paper addresses the following questions:

Can we get provable convergence guarantees to NE of decentralized first-order methods in two-team zero-sum games?

Our results. First, with regards to computational complexity, we prove that computing an approximate (and possibly mixed) NE in two-team zero-sum games is CLS-hard (Theorem 3.1); *i.e.*, it is computationally harder than finding pure NE in a congestion game or computing an approximate fixed point of gradient descent.

Second, regarding online learning for equilibrium computation, we prove that a number of established, decentralized, first-order methods are not fit for the purpose and fail to converge even asymptotically. Specifically, we present a simple —yet nontrivial— family of two-team zero-sum games (with each team consisting of two players) where projected gradient descent-ascent (GDA), optimistic gradient descent-ascent (OGDA), optimistic multiplicative weights update (OMWU), and the extra gradient method (EG) fail to locally converge to a mixed NE (Theorem 3.3). More broadly, in the case of GDA in nondegenerate team games with unique mixed NE, one could acquire an even stronger result for any high-dimensional configuration of actions and players (Theorem 3.2). To the best of our knowledge, the described family of games is the first-of-its-kind in which all these methods provably fail to converge at the same time.

Third, we propose a novel first-order method inspired by adaptive control (Bazanella et al., 1997; Hassouneh et al., 2004). In particular, we use a technique that manages to stabilize unstable fixed point of a dynamical system without prior knowledge of their position and without introducing new ones. It is important to note that this method is a modification of GDA that uses a stabilizing feedback which maintains the decentralized nature of GDA.

Finally, in Section 4 we provide a series of experiments in simple two-team zero-sum GAN's. We also show that multi-GAN architectures achieve better performance than the single-agent ones, relative to the network capacity, when they are trained on synthetic or real-world datasets like CIFAR10.

Existing algorithms for NE in multiplayer games. The focus of the present paper is examining algorithms for the setting of *repeated games* (Cesa-Bianchi & Lugosi, 1999, Chapter 7). If we do not restrict ourselves to this setting, there are numerous centralized algorithms (Lipton et al., 2003; Berg & Sandholm, 2017) and heuristics (Gemp et al., 2021) that solve the problem of computing Nash equilibria in general multi-player games .

2 Preliminaries

Our setting. A *two-team game* in normal form is defined as a tuple $\Gamma(\mathcal{N}, \mathcal{M}, \mathcal{A}, \mathcal{B}, \{U_A, U_B\})$. The tuple is defined by

(i) a finite set of $n = |\mathcal{N}|$ *players* belonging to team A, as well as a finite set of $m = |\mathcal{M}|$ *players* belonging to team B;

(ii) a finite set of *actions* (or *pure strategies*) $\mathcal{A}_i = \{\alpha_1, \dots, \alpha_{n_i}\}$ per player $i \in \mathcal{N}$; where $\mathcal{A} := \prod_i \mathcal{A}_i$ denotes the ensemble of all possible action profiles of team A, and respectively, a finite set of *actions* (or *pure strategies*) $\mathcal{B}_i = \{\beta_1, \dots, \beta_{n_i}\}$ per player $i \in \mathcal{M}$, where $\mathcal{B} := \prod_i \mathcal{B}_i$.

(iii) a utility function for team A, $U_A : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ (resp. $U_B : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ for team B)

We also use $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ to denote the strategy profile of team A players and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)$ the strategy profile of team B players.

Finally, each team's *payoff* function is denoted by $U_A, U_B : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$, where the *individual utility* of a player is identical to her teammates, i.e., $U_i = U_A$ & $U_j = U_B$ $\forall i \in \mathcal{N}$ and $j \in \mathcal{M}$. In this general context, players could also submit *mixed strategies*, i.e, probability distributions over actions. Correspondingly, we define the product distributions $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$, $\boldsymbol{y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_m)$ as team A and B's strategies respectively, in which $\boldsymbol{x}_i \in \Delta(\mathcal{A}_i)$ and $\boldsymbol{y}_j \in \Delta(\mathcal{B}_j)$. Conclusively, we will write $\mathcal{X} := \prod_{i\in\mathcal{N}} \mathcal{X}_i = \prod_{i\in\mathcal{N}} \Delta(\mathcal{A}_i), \mathcal{Y} := \prod_{i\in\mathcal{M}} \mathcal{Y}_i = \prod_{i\in\mathcal{M}} \Delta(\mathcal{B}_i)$ the space of mixed strategy profiles of teams A, B. A two-team game is called *two-team zero-sum* if $U_B = -U_A = U$ which is the main focus of this paper. Moreover, we assume that the game is *succinctly representable* and satisfies the *polynomial expectation property* Daskalakis et al. (2006). This means that given a mixed strategy profile, the utility of each player can be computed in polynomial time in the number of agents, the sum of the number of strategies of each player, and the bit number required to represent the mixed strategy profile.

A *Nash equilibrium* (NE) is a strategy profile $(\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathcal{X} \times \mathcal{Y}$ such that

$$\begin{cases} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \leq U(\boldsymbol{x}_i, \boldsymbol{x}^*_{-i}, \boldsymbol{y}^*), \ \forall \boldsymbol{x}_i \in \mathcal{X}_i \\ U(\boldsymbol{x}^*, \boldsymbol{y}^*) \geq U(\boldsymbol{x}^*, \boldsymbol{y}_j, \boldsymbol{y}^*_{-j}), \ \forall \boldsymbol{y}_j \in \mathcal{Y}_j \end{cases} \quad [1] \tag{NE}$$

A first approach to computing NE in Two-Team Zero-Sum games. Due to the multilinearity of the utility and the existence of a duality gap, the linear programming method used in two-player zero-sum games cannot be used to compute a Nash equilibrium. For the goal of computing Nash equilibrium in two-team zero-sum games, we have experimented with a selection of online learning, first-order methods that have been utilized with varying success in the setting of the two-person zero-sum case. Namely, we analyze the following methods: (i) gradient descent-ascent (GDA) (ii) optimistic gradient descent-ascent (OGDA) (iii) extra gradient method (EG) (iv) optimistic multiplicative weights update method (OMWU). For their precise definitions, we refer to Appendix B.

The below folklore fact will play a key role hereafter.

Fact 2.1. Any fixed point of the aforementioned discrete-time dynamics (apart from OMWU) on the utility function necessarily corresponds to a Nash Equilibrium of the game.

Hence, an important test for the asymptotic behavior of GDA, OGDA, EG, and OMWU methods is to examine whether these methods stabilize around their fixed points which effectively constitute the Nash equilibria of the game. In Section 3.2, we show that in the absence of pure Nash equilibria, all the above methods fail to stabilize on their fixed points even for a simple class of two-team game with $(n = 2, m = 2)$. Consequently, they fail to converge to the mixed Nash equilibrium of the game.

The presence of these results demonstrates the need for a different approach that lies outside the scope of traditional optimization techniques. Inspired by the applications of washout filters to stabilize unknown fixed points and the adaptive control generalizations of the former, we design a new variant of GDA "vaned" with a feedback loop dictated by a pair of two matrices. In contrast to the aforementioned conventional methods, our proposed technique surprisingly accomplishes asymptotic last-iterate convergence to its fixed point, *i.e.*, the mixed Nash equilibria of the team game.

$(\mathbf{K}, \mathbf{P})$-vaned GDA Method. After concatenating the vectors of the minimizing and the maximizing agents — $\boldsymbol{z}^{(k)} = (\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)})$— our method for appropriate matrices $\mathbf{K}, \mathbf{P}$ reads:

$$\begin{cases} \boldsymbol{z}^{(k+1)} = \Pi_{\mathcal{Z}} \left\{ \boldsymbol{z}^{(k)} + \eta \left(\begin{smallmatrix} -\nabla_{\boldsymbol{x}} f(\boldsymbol{z}^{(k)}) \\ \nabla_{\boldsymbol{y}} f(\boldsymbol{z}^{(k)}) \end{smallmatrix} \right) + \eta \mathbf{K} (\boldsymbol{z}^{(k)} - \boldsymbol{\theta}^{(k)}) \right\} \\ \boldsymbol{\theta}^{(k+1)} = \Pi_{\mathcal{Z}} \left\{ \boldsymbol{\theta}^{(k)} + \eta \mathbf{P} (\boldsymbol{z}^{(k)} - \boldsymbol{\theta}^{(k)}) \right\} \end{cases} \tag{KPV-GDA}$$

Intuitively, the additional variable $\boldsymbol{\theta}^{(k)}$ holds an estimate of the fixed point, and through the feedback law $\eta \mathbf{K}(\boldsymbol{z}^{(k)} - \boldsymbol{\theta}^{(k)})$ the vector $\boldsymbol{z}$ stabilizes around that estimate which slowly moves towards the real fixed point of the GDA dynamic.

2.1 Two Illustrative Examples

Our first example plays a dual role: first, it demonstrates how two-team min-max competition can capture the formulation of multi-agent GAN architectures; second, it hints at the discrepancy be-

[1] We are using here the shorthand $\boldsymbol{x}_{-i}$ (or $\boldsymbol{y}_{-i}$) to highlight the strategy of all agents $\mathcal{N}$ (or $\mathcal{M}$) but i.

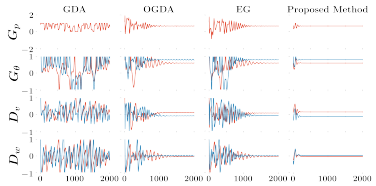


Figure 1: Parameter training of the configuration under different algorithms

tween the results of optimization methods, since —as we will see— GDA will not converge to the Nash equilibrium/ground-truth distribution. Generally, the solution that is sought after is the min max solution of the objective function Goodfellow et al. (2014) which are NP-hard to compute in the general case Borgs et al. (2008); nevertheless, applications of GANs have proven that first-order stationary points of the objective function suffice to produce samples of very good quality.

2.1.1 Learning a mixture of gaussians with multi-agent GAN's

Consider the case of $\mathcal{O}$, a mixture of gaussian distribution with two components, $C_1 \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{I})$ and $C_2 \sim \mathcal{N}(-\boldsymbol{\mu}, \mathbf{I})$ and mixture weights π_1, π_2 to be positive such that $\pi_1 + \pi_2 = 1$ and $\pi_1, \pi_2 \neq \frac{1}{2}$.

To learn the distribution above, we utilize an instance of a *Team*-WGAN in which there exists a generating team of agents $G_p : \mathbb{R} \to \mathbb{R}, G_{\boldsymbol{\theta}} : \mathbb{R}^n \to \mathbb{R}^n$, and a discriminating team of agents $D_{\boldsymbol{v}} : \mathbb{R}^n \to \mathbb{R}, D_{\boldsymbol{w}} : \mathbb{R}^n \to \mathbb{R}$, all described by the following equations:

$$\begin{array}{rl} \text{Generators:} & G_p(\zeta) = p + \zeta \ , \ G_\theta(\boldsymbol{\xi}) = \boldsymbol{\xi} + \boldsymbol{\theta} \\ \text{Discriminators:} & D_{\boldsymbol{v}}(\boldsymbol{y}) = \langle \boldsymbol{v}, \boldsymbol{y} \rangle \ , D_{\boldsymbol{w}}(\boldsymbol{y}) = \sum_i w_i y_i^2 \end{array} \tag{1}$$

The generating agent G_θ maps random noise $\boldsymbol{\xi} \sim \mathcal{N}(0, \mathbf{I})$ to samples while generating agent $G_p(\zeta)$, utilizing an independent source of randomness $\zeta \sim \mathcal{N}(0, 1)$, probabilistically controls the sign of the output of the generator G_θ. The probability of ultimately generating a sample $\boldsymbol{y} = \boldsymbol{\xi} + \boldsymbol{\theta}$ is in expectation equal to p, while the probability of the sample being $\boldsymbol{y} = -\boldsymbol{z} - \boldsymbol{\theta}$ is equal to $1 - p$.

On the other end, there stands the discriminating team of $D_{\boldsymbol{v}}, D_{\boldsymbol{w}}$. Discriminators, $D_v(\boldsymbol{y}), D_w(\boldsymbol{y})$ map any given sample $\boldsymbol{y}$ to a scalar value accounting for the "realness" or "fakeness" of it – negative meaning fake, positive meaning real. The discriminators are disparate in the way they measure the realness of samples as seen in their definitions.

We follow the formalism of the Wasserstein GAN to form the optimization objective:

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \left\{ \mathbb{E}_{\boldsymbol{y}\sim\mathcal{O}}\Big[D_{\boldsymbol{v}}(\boldsymbol{y}) + D_{\boldsymbol{w}}(\boldsymbol{y})\Big] - \mathbb{E}_{\substack{\boldsymbol{\xi}\sim\mathcal{N}(0,\mathbf{I}),\\ \zeta\sim\mathcal{N}(0,1)}} \left[\begin{array}{c} G_p(\zeta)\cdot\Big(D_{\boldsymbol{v}}\big(G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)+D_{\boldsymbol{w}}\big(G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)\Big) \\ + \\ \big(1-G_p(\zeta)\big)\cdot\Big(D_{\boldsymbol{v}}\big(-G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)+D_{\boldsymbol{w}}\big(-G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)\Big) \end{array} \right] \right\} \tag{2}$$

Equation (2) yields the simpler form:

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} (\pi_1 - \pi_2)\boldsymbol{v}^T\boldsymbol{\mu} - 2p\boldsymbol{v}^T\boldsymbol{\theta} + \boldsymbol{v}^T\boldsymbol{\theta} + \sum_i^n w_i(\mu_i^2 - \theta_i^2) \tag{3}$$

It is easy to check that Nash equilibria of (2) must satisfy:

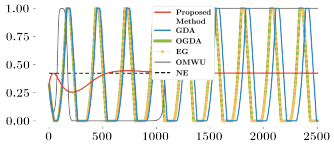
$$\left\{ \begin{array}{rcrl} \boldsymbol{\theta} & = & \boldsymbol{\mu}, & p = 1 - \pi_2 = \pi_1 \\ \boldsymbol{\theta} & = & -\boldsymbol{\mu}, & p = 1 - \pi_1 = \pi_2. \end{array} \right\}$$

Figure 1 demonstrates both GDA's failure and OGDA, EG, and our proposed method, KPV-GDA succeeding to converge to the above Nash equilibria and simultaneously discovering the mixture of the ground-truth.

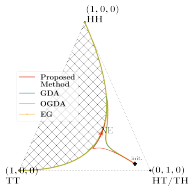
2.1.2 Multiplayer Matching Pennies

Interestingly enough, there are non-trivial instances of two-team competition settings in which even OGDA and EG fail to converge. Such is the case for a team version of the well-known game of

matching pennies. The game can be shortly described as such: "*coordinate with your teammates to play a game of matching pennies against the opposing team, coordinate not and pay a penalty*". (We note that this game is a special case of the game presented in Section 3.3.) As we can see in Figures 2a and 2b, this multiplayer generalized matching pennies game constitutes an excellent benchmark on which all traditional gradient flow discretizations fail under the perfect competition setting. Interestingly, we are not aware of a similar example in min-max literature and it has been our starting point for seeking new optimization techniques inspired by Control theory. Indeed, the KPV-GDA variation with $(\mathbf{K}, \mathbf{P}) = (-1.1 \cdot \mathbf{I}, 0.3 \cdot \mathbf{I})$ achieves to converge to the unique mixed Nash Equilibrium of the game. In the following sections, we provide theorems that explain formally the behavior of the examined dynamics.



(a) Generalized matching pennies under different algorithms. For the precise definition of the game, we refer to appendix C.4



(b) Projected Trajectory of Team A under different algorithms. The sketched surface is not part of the feasible team strategy profiles (product of distributions).

3 Main Results

In this section we will prove that computing a Nash equilibrium in two-team zero-sum games is computationally hard and thus getting a polynomial-time algorithm that computes a Nash equilibrium is unlikely. Next, we will demonstrate the shortcomings of an array of commonly used online learning, first-order methods, and then we will provide a novel, decentralized, first-order method that locally converges to NE under some conditions.

3.1 On the hardness of computing NE in two-team zero-sum games

As promised, our first statement characterizes the hardness of NE computation in two-team zero-sum games:

Theorem 3.1 (CLS-hard)**.** *Computing a Nash equilibrium in a succinctly represented two-team zero-sum games is* CLS*-hard.*

The main idea of the proof of Theorem 3.1 relies on a reduction of approximating Nash equilibria in congestion games, which has been shown to be complete for the CLS complexity class. The class CLS contains the problem of continuous optimization. We defer the proof of the above theorem to the paper's supplement.

3.2 Failure of common online, first-order methods

The negative computational complexity result we proved for two-team zero-sum games (Theorem 3.1) does not preclude the prospect of attaining algorithms (learning first-order methods) that converge to Nash equilibria. Unfortunately, we prove that these methods cannot guarantee convergence to Nash equilibria in two-team zero-sum games in general.

In this subsection, we are going to construct a family of two-team zero-sum games with the property that the dynamics of GDA, OGDA, OMWU, and EG are unstable on Nash equilibria. This result is indicative of the challenges that lie in min-max optimization of two-team zero-sum games and the reason that provable, nonasymptotic convergence guarantees of online learning have not yet been established.

Before defining our benchmark game, we prove an important theorem which states that GDA does not converge to mixed Nash equilibria. This fact is a stepping stone in constructing the family of team-zero sum games later. We present the proof of all of the below statements in detail in the paper's appendix (Appendix B).

Weakly-stable Nash equilibrium. (Kleinberg et al., 2009; Mehta et al., 2015) Consider the set of Nash equilibria with the property that if any single randomizing agent of one team is forced to play any strategy in their current support with probability one, all other agents of the same team must remain indifferent between the strategies in their support. This type of Nash equilibria is called weakly-stable. We note that pure Nash equilibria are trivially weakly-stable. It has been shown that mixed Nash equilibria are not weakly-stable in generic games [2]

We can show that Nash equilibria that are not weakly-stable Nash are actually unstable for GDA. Moreover, through standard dynamical systems machinery, that the set of initial conditions that converges to Nash equilibria that are not weakly-stable should be of Lebesgue measure zero. Formally, we prove that:

Theorem 3.2 (Non weakly-stable Nash are unstable)**.** *Consider a two-team zero-sum game with utility function of Team B ($\boldsymbol{y}$ vector) being $U(\boldsymbol{x}, \boldsymbol{y})$ and Team A ($\boldsymbol{x}$ vector) being $-U(\boldsymbol{x}, \boldsymbol{y})$. Moreover, assume that $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is a Nash equilibrium of full support that is not weakly-stable. It follows that the set of initial conditions so that GDA converges to $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is of measure zero for step size $\eta < \frac{1}{L}$ where L is the Lipschitz constant of ∇U.*

3.3 Generalized Matching Pennies (GMP)

Inspired by Theorem 3.2, in this section we construct a family of team zero-sum games so that GDA, OGDA, OMWU, and EG methods fail to converge (if the initialization is a random point in the simplex, the probability of convergence of the aforementioned methods is zero). The intuition is to construct a family of games, each of which has only mixed Nash equilibria (that are not weakly-stable), *i.e.*, the constructed games should lack pure Nash equilibria; using Theorem 3.2, it would immediately imply our claim for GDA. It turns out that OGDA, OMWU, and EG also fail to converge for the same family.

Definition of GMP. Consider a setting with two teams (Team A, Team B), each of which has $n = 2$ players. Inspired by the standard matching pennies game and the game defined in Schulman & Vazirani (2019a), we allow each agent i to have two strategies/actions that is $S = \{H, T\}$ for both teams with 2^4 possible strategy profiles. In case all the members of a Team choose the same strategy say H or T then the Team "agrees" to play H or T (otherwise the Team "does not agree").

	HH	HT/TH	TT
HH	$1, -1$	$\omega, -\omega$	$-1, 1$
HT/TH	$-\omega, \omega$	$0, 0$	$-\omega, \omega$
TT	$-1, 1$	$\omega, -\omega$	$1, -1$

Thus, in the case that both teams "agree", the payoff of each team is actually the payoff for the two-player matching pennies. If one team "agrees" and the other does not, the team that "agrees" enjoys payoff $\omega \in (0, 1)$ and the other team suffers a penalty ω. If both teams fail to "agree", both teams get payoff zero. Let x_i with $i \in \{1, 2\}$ be the probability that agent i of Team A chooses H and $1 - x_i$ the probability that she chooses T. We also denote $\boldsymbol{x}$ the vector of probabilities for Team A. Similarly, we denote y_i for $i \in \{1, 2\}$ be the probability that agent i of Team B chooses H and $1 - y_i$ the probability that she chooses T and $\boldsymbol{y}$ the probability vector.

Properties of GMP. An important remark on the properties of our presented game is due. Existing literature tackles settings with (i) (weak-)monotonocity (Mertikopoulos et al., 2019; Diakonikolas et al., 2021), (ii) cocoercivity (Zhu & Marcotte, 1996), (iii) zero-duality gap (Von Neumann, 1928), (iv) unconstrained solution space (Golowich et al., 2020) . Our game is carefully crafted and – although it has a *distinct structure* and is nonconvex-nonconcave only due to *multilinearity*– satisfies none of the latter properties. This makes the (local) convergence of our proposed method even more surprising. (See also Appendix B.6.)

The first fact about the game that we defined is that for $\omega \in (0, 1)$, there is only one Nash equilibrium $(\boldsymbol{x}^*, \boldsymbol{y}^*)$, which is the uniform, *i.e.*, $x_1^* = x_2^* = y_1^* = y_2^* = \frac{1}{2}$ for all agents i.

Lemma 3.1 (GMP has a unique Nash)**.** The Generalized Matching Pennies game exhibits a unique Nash equilibrium which is $(\boldsymbol{x}^*, \boldsymbol{y}^*) = ((\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}))$.

[2] Roughly speaking, generic games where we add small Gaussian noise to perturb slightly every payoff only so that we preclude any payoff ties. In these games, all Nash equilibria in all but a measure-zero set of games exhibit the property that all pure best responses are played with positive probability.

Remark 1. *The fact that the game we defined has a unique Nash equilibrium that is in the interior of* $[0,1]^4$ *is really crucial for our negative convergence results later in the section as we will show that it is not a weakly-stable Nash equilibrium and the negative result about GDA will be a corollary due to Theorem 3.2. We also note that if* $\omega = 1$ *then there are more Nash equilibria, in particular the* $(\mathbf{0},\mathbf{0}), (\mathbf{1},\mathbf{0}), (\mathbf{0},\mathbf{1}), (\mathbf{1},\mathbf{1})$ *which are pure.*

The following Theorem is the main (negative) result of this section.

Theorem 3.3 (GDA, OGDA, EG, and OMWU fail)**.** *Consider GMP game with* $\omega \in (0,1)$. *Assume that* $\eta_{GDA} < \frac{1}{4}$, $\eta_{OGDA} < \min(\omega, \frac{1}{8})$, $\eta_{EG} < \frac{\omega}{2}$, *and* $\eta_{OMWU} < \min\left(\frac{1}{4}, \frac{\omega}{2}\right)$ *(bound on the stepsize for GDA, OGDA, OMWU, and EG methods respectively). It holds that the set of initial conditions so that GDA, OGDA, OMWU, and EG converge (stabilize to any point) is of measure zero.*

Remark 2. *Theorem 3.3 formally demonstrates that the behavior of algorithms mentioned in Section 2.1.2 are not a result of "bad parametrization", and in fact, the probability that any of them converges to the NE is equal to the probability that the initialization of the variables coincides with the NE (Lebesgue measure zero).*

Remark 3 (Average iterate also fails)**.** *One might ask what happens when we consider average iterates instead of the last iterate. It is a well-known fact Syrgkanis et al. (2015) that the average iterate of no-regret algorithms converges to coarse correlated equilibria (CCE) so we expect that the average iterate stabilizes. Nevertheless, CCE might not be Nash equilibria. Indeed we can construct examples in which the average iterate of GDA, OGDA, OMWU, and EG experimentally fail to stabilize to Nash equilibria. In particular, we consider a slight modification of GMP; players and strategies are the same but the payoff matrix has changed and can be found reads*

	HH	HT/TH	TT
HH	$2, -2$	$\frac{1}{2}, -\frac{1}{2}$	$-2, 2$
HT/TH	$-\frac{1}{2}, \frac{1}{2}$	$0, 0$	$-\frac{1}{2}, \frac{1}{2}$
TT	$-1, 1$	$\frac{1}{2}, -\frac{1}{2}$	$1, -1$

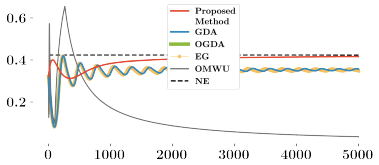


Figure 3: GDA, OGDA, OMWU, & EG fail to converge to a Nash Equilibrium even in average

Figure 3 shows that the average iterates of GDA, OGDA, OMWU, and EG stabilize to points that are not Nash equilibria. We note that since our method (see next subsection) converges locally, the average iterate should converge locally to a Nash equilibrium.

3.4 Our proposed method

The aforementioned results prove that the challenging goal of computing two-team zero-sum games calls for an expansion of existing optimization techniques. The mainstay of this effort and our positive result is the KPV-GDA method defined in (KPV-GDA) which is inspired by techniques of adaptive control literature. The first statement we make is that KPV-GDA stabilizes around any Nash equilibrium for appropriate choices of matrices $\mathbf{K}, \mathbf{P}$:

Theorem 3.4 (KPV-GDA stabilizes)**.** *Consider a team zero-sum game so that the utility of Team* B *is* $U(\boldsymbol{x}, \boldsymbol{y})$ *and hence the utility of Team* A *is* $-U(\boldsymbol{x}, \boldsymbol{y})$ *and a Nash equilibrium* $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ *of the game. Moreover, we assume*

$$\begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} \text{ is invertible.}$$

For any fixed step size $\eta > 0$, *we can always find matrices* K, P *so that KPV-GDA method defined in* (KPV-GDA) *converges locally to* $(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

This is an existential theorem and cannot be generally useful in practice. Further, this dynamic would not be necessarily uncoupled and the design of matrices $\mathbf{K}$ and $\mathbf{P}$ could necessitate knowledge of the NE we are trying to compute. Instead, our next statement provides sufficient conditions under which a simple parametrization of matrices $\mathbf{K}, \mathbf{P}$ results in an uncoupled, converging dynamic:

Theorem 3.5. *Consider a two-team zero-sum game so that the utility of Team B is $U(\boldsymbol{x}, \boldsymbol{y})$, the utility of Team A is $-U(\boldsymbol{x}, \boldsymbol{y})$, and a Nash equilibrium $(\boldsymbol{x}^*, \boldsymbol{y}^*)$. Moreover, let*

$$\mathbf{H} := \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix}.$$

and E be the set of eigenvalues ρ of $\mathbf{H}$ with real part positive, that is $E = \{$Eigenvalues of matrix $\mathbf{H}$, $\rho : Re(\rho) > 0\}$. We assume that $\mathbf{H}$ is invertible and moreover

$$\beta = \min_{\rho \in E} \frac{Re(\rho)^2 + Im(\rho)^2}{Re(\rho)} > \max_{\rho \in E} Re(\rho) = \alpha. \tag{4}$$

We set $\mathbf{K} = k \cdot \mathbf{I}$, $\mathbf{P} = p \cdot \mathbf{I}$. There exist small enough step size $\eta > 0$ and scalar $p > 0$ and for any $k \in (-\beta, -\alpha)$ so that (KPV-GDA) with chosen $\mathbf{K}, \mathbf{P}$ converges locally to $(\boldsymbol{x}^*, \boldsymbol{y}^*)$.[3]

4 Experiments

In this section, we perform a series of experiments to further motivate the study of two-team zero-sum games, especially in the context of multi-agent generative adversarial networks (multi-GANs). A multi-agent generative adversarial network (multi-GAN) (Arora et al., 2017; Hoang et al., 2017; 2018; Zhang et al., 2018; Tang, 2020; Hardy et al., 2019; Albuquerque et al., 2019) is a generative adversarial network (GAN) that leverages multiple "agents" (generators and/or discriminators) in order achieve statistical and computational benefits. In particular, Arora et al. formally proved the expressive superiority of multi-generator adversarial network architectures something that we empirically verify in Section 4. In this direction, researchers strive to harness the efficacy of distributed processing by utilizing shallower networks that can collectively learn more diverse datasets[4].

At first, the superiority of multi-GANs might appear to contrast our theoretical findings; but in reality, the superiority comes from the quality of solutions that are attainable from multi-agent architectures (*expressivity*) and the fact that hardness (*complexity*) translates to rates of convergence but not non-convergence. Single agent GANs quickly converge to critical points that are not guaranteed to capture the distribution very well. In figure 4 we see the fast convergence of a single-agent GAN to solutions of bad quality versus the convergence of a multi-GAN to an obviously better solution. Due to space constraints, we defer further discussion of the experiments at Section D.1.

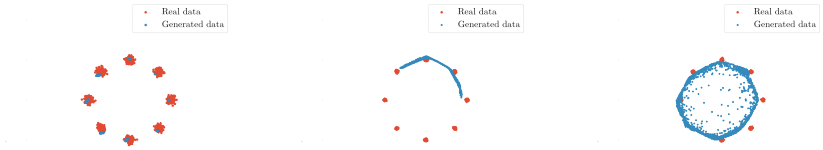


Figure 4: From left to right: (i) Each generator of MGAN learns one mode of the mixture of 8 gaussians, (ii) Mode Collapse of single-agent GANs, (iii) Single-agent GAN can't discriminate between the modes.

5 Conclusions and Open Problems

In this work we study the wide class of nonconvex-nonconcave games that express the two-team competition, inspired broadly by the structure of the complex competition between multi-agent generators and discriminators in GAN's. Furthermore, in this setting of *two-team zero-sum games*, we have presented a number of negative results about the problem of computing a Nash equilibrium. Moreover, through a simple family of games that we construct, we prove the inability of commonly used methods for min-max optimization such as GDA, OGDA, OMWU, and EG to converge both in average and in the last iterate to Nash Equilibria which comes to a stark contrast to recent literature that is concerned with simpler games. We have also presented an optimization method (called KPV-GDA) that manages to stabilize around Nash equilibria.

[3] As long as aforementioned conditions are satisfied, (KPV-GDA) locally converges in any nonconvex-nonconcave game. Indeed, GMP with any ω satisfies the sufficient conditions of 3.5. See also, Appendix B.7.

[4] Indeed, it is preferable from a computational standpoint to back-propagate through two equally sized neural networks rather than through a single one that would be twice as deep (Tang, 2020).

REFERENCES

Isabela Albuquerque, João Monteiro, Thang Doan, Breandan Considine, Tiago Falk, and Ioannis Mitliagkas. Multi-objective training of generative adversarial networks with multiple discriminators. In *International Conference on Machine Learning*, pp. 202–211. PMLR, 2019.

Ioannis Anagnostides, Gabriele Farina, Christian Kroer, Andrea Celli, and Tuomas Sandholm. Faster no-regret learning dynamics for extensive-form correlated and coarse correlated equilibria. *arXiv preprint arXiv:2202.05446*, 2022a.

Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, and Tuomas Sandholm. On last-iterate convergence beyond zero-sum games. *arXiv preprint arXiv:2203.12056*, 2022b.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 214–223, 2017.

Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

Sanjeev Arora, Rong Ge, Yingyu Liang, Tengyu Ma, and Yi Zhang. Generalization and equilibrium in generative adversarial nets (gans). In *International Conference on Machine Learning*, pp. 224–232. PMLR, 2017.

Moshe Babaioff, Robert Kleinberg, and Christos H Papadimitriou. Congestion games with malicious players. *Games and Economic Behavior*, 67(1):22–35, 2009.

Yakov Babichenko and Aviad Rubinstein. Settling the complexity of nash equilibrium in congestion games. In Samir Khuller and Virginia Vassilevska Williams (eds.), *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pp. 1426–1437. ACM, 2021.

Michael Bacharach. Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, 53(2):117–147, 1999. ISSN 1090-9443. doi: https://doi.org/10.1006/reec.1999.0188. URL https://www.sciencedirect.com/science/article/pii/S1090944399901886.

Nicola Basilico, Andrea Celli, Giuseppe De Nittis, and Nicola Gatti. Computing the team–maxmin equilibrium in single–team single–adversary team games. *Intelligenza Artificiale*, 11(1):67–79, 2017a.

Nicola Basilico, Andrea Celli, Giuseppe De Nittis, and Nicola Gatti. Team-maxmin equilibrium: efficiency bounds and algorithms. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017b.

A. S. Bazanella, P. V. Kokotovic, and A. S. e Silva. On the control of dynamic systems with unknown operating point. In *1997 European Control Conference (ECC)*, pp. 3434–3439, 1997. doi: 10.23919/ECC.1997.7082644.

Alexandre S. Bazanella, Petar V. Kokotovic, and Aguinaldo S. E Silva. On the control of dynamic systems with unknown operating point. *International Journal of Control*, 73(7): 600–605, 2000. doi: 10.1080/002071700219443. URL https://doi.org/10.1080/002071700219443.

Kimmo Berg and Tuomas Sandholm. Exclusion method for finding nash equilibrium in multiplayer games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

Christian Borgs, Jennifer Chayes, Nicole Immorlica, Adam Tauman Kalai, Vahab Mirrokni, and Christos Papadimitriou. The myth of the folk theorem. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pp. 365–372, 2008.

Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015.

Noam Brown and Tuomas Sandholm. Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

Federico Cacciamani, Andrea Celli, Marco Ciccone, and Nicola Gatti. Multi-agent coordination in adversarial environments through signal mediated strategies. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 269–278, 2021.

Yongcan Cao, Wenwu Yu, Wei Ren, and Guanrong Chen. An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial informatics*, 9(1): 427–438, 2012.

Andrea Celli and Nicola Gatti. Computational results for extensive-form adversarial team games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

N. Cesa-Bianchi and G. Lugosi. On prediction of individual sequences. *Annals of Statistics*, pp. 1865–1895, 1999.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Conference on Learning Theory*, pp. 6–1. JMLR Workshop and Conference Proceedings, 2012.

Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. *Advances in Neural Information Processing Systems*, 31, 2018.

Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. *Innovations in Theoretical Computer Science*, 2019.

Constantinos Daskalakis and Christos Papadimitriou. Continuous local search. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pp. 790–804. SIAM, 2011.

Constantinos Daskalakis, Alex Fabrikant, and Christos H Papadimitriou. The game world is flat: The complexity of nash equilibria in succinct games. In *International Colloquium on Automata, Languages, and Programming*, pp. 513–524. Springer, 2006.

Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

Constantinos Daskalakis, Alan Deckelbaum, and Anthony Kim. Near-optimal no-regret algorithms for zero-sum games. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pp. 235–254. SIAM, 2011.

Constantinos Daskalakis, Stratis Skoulakis, and Manolis Zampetakis. The complexity of constrained min-max optimization. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1466–1478, 2021.

Jelena Diakonikolas, Constantinos Daskalakis, and Michael I Jordan. Efficient methods for structured nonconvex-nonconcave min-max optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 2746–2754. PMLR, 2021.

Ishan Durugkar, Ian Gemp, and Sridhar Mahadevan. Generative multi-adversarial networks. *arXiv preprint arXiv:1611.01673*, 2016.

Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. The complexity of pure nash equilibria. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pp. 604–612, 2004.

Gabriele Farina, Andrea Celli, Nicola Gatti, and Tuomas Sandholm. Ex ante coordination and collusion in zero-sum multi-player extensive-form games. *Advances in Neural Information Processing Systems*, 31, 2018.

John Fearnley, Paul W Goldberg, Alexandros Hollender, and Rahul Savani. The complexity of gradient descent: CLS = PPAD∩PLS. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 46–59, 2021.

Lampros Flokas, Emmanouil-Vasileios Vlatakis-Gkaragkounis, and Georgios Piliouras. Solving min-max optimization with hidden structure via gradient descent ascent. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 2021, virtual*, 2021.

Drew Fudenberg. Jean tirole game theory, 1991.

Ian Gemp, Rahul Savani, Marc Lanctot, Yoram Bachrach, Thomas Anthony, Richard Everett, Andrea Tacchetti, Tom Eccles, and János Kramár. Sample-based approximation of nash in large many-player games via gradient descent. *arXiv preprint arXiv:2106.01285*, 2021.

Natalie Gold. Introduction: Teamwork in theory and in practice. In *Teamwork*, pp. 1–21. Springer, 2005.

Noah Golowich, Sarath Pattathil, and Constantinos Daskalakis. Tight last-iterate convergence rates for no-regret learning in multi-player games. *Advances in neural information processing systems*, 33:20766–20778, 2020.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pp. 2672–2680, 2014.

Hans W Gottinger. J. marschak and roy radner," economic theory of teams"(book review). *Theory and Decision*, 5(3):349, 1974.

James Hannan. 4. approximation to rayes risk in repeated play. In *Contributions to the Theory of Games (AM-39), Volume III*, pp. 97–140. Princeton University Press, 2016.

Corentin Hardy, Erwan Le Merrer, and Bruno Sericola. Md-gan: Multi-discriminator generative adversarial networks for distributed datasets. In *2019 IEEE international parallel and distributed processing symposium (IPDPS)*, pp. 866–877. IEEE, 2019.

Munther A Hassouneh, Hsien-Chiarn Lee, and Eyad H Abed. Washout filters in feedback control: Benefits, limitations and extensions. In *Proceedings of the 2004 American control conference*, volume 5, pp. 3950–3955. IEEE, 2004.

Quan Hoang, Tu Dinh Nguyen, Trung Le, and Dinh Phung. Multi-generator generative adversarial nets. *arXiv preprint arXiv:1708.02556*, 2017.

Quan Hoang, Tu Dinh Nguyen, Trung Le, and Dinh Phung. Mgan: Training generative adversarial nets with multiple generators. In *International conference on learning representations*, 2018.

Max Jaderberg, Wojciech M Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castaneda, Charles Beattie, Neil C Rabinowitz, Ari S Morcos, Avraham Ruderman, et al. Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science*, 364(6443):859–865, 2019.

Jeongbin Kim, Thomas R Palfrey, and Jeffrey R Zeidel. A theory of games played by teams of players. 2019.

R. Kleinberg, G. Piliouras, and É. Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *STOC*, 2009.

GM Korpelevich. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.

Jason D. Lee, Ioannis Panageas, Georgios Piliouras, Max Simchowitz, Michael I. Jordan, and Benjamin Recht. First-order methods almost always avoid strict saddle points. *Math. Program.*, 176(1-2):311–337, 2019. doi: 10.1007/s10107-019-01374-3. URL https://doi.org/10.1007/s10107-019-01374-3.

Joel Z Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel. Multi-agent reinforcement learning in sequential social dilemmas. *arXiv preprint arXiv:1702.03037*, 2017.

Dan Li, Dacheng Chen, Baihong Jin, Lei Shi, Jonathan Goh, and See-Kiong Ng. Mad-gan: Multivariate anomaly detection for time series data with generative adversarial networks. In *International Conference on Artificial Neural Networks*, pp. 703–716. Springer, 2019.

Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In *International Conference on Machine Learning*, pp. 6083–6093. PMLR, 2020.

Richard J Lipton, Evangelos Markakis, and Aranyak Mehta. Playing large games using simple strategies. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pp. 36–41, 2003.

Jakob Marschak. Elements for a theory of teams. *Management science*, 1(2):127–137, 1955.

H Brendan McMahan, Geoffrey J Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 536–543, 2003.

R. Mehta, I. Panageas, and G. Piliouras. Natural selection as an inhibitor of genetic diversity: Multiplicative weights updates algorithm and a conjecture of haploid genetics. In *ITCS*, 2015.

Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2703–2717. SIAM, 2018.

Panayotis Mertikopoulos, Houssam Zenati, Bruno Lecouat, Chuan-Sheng Foo, Vijay Chandrasekhar, and Georgios Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *ICLR'19-International Conference on Learning Representations*, pp. 1–23, 2019.

Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisỳ, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

Mohammad Sal Moslehian. Ky fan inequalities. *CoRR*, abs/1108.1467, 2011.

Thomas Nagylaki. The evolution of multilocus systems under weak selection. *Genetics*, 134(2): 627–647, 1993.

Arkadij Semenovič Nemirovskij and David Borisovich Yudin. Problem complexity and method efficiency in optimization. 1983.

Martin A Nowak, Akira Sasaki, Christine Taylor, and Drew Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428(6983):646–650, 2004.

Christos H Papadimitriou. The complexity of the lin–kernighan heuristic for the traveling salesman problem. *SIAM Journal on Computing*, 21(3):450–465, 1992.

Leonid Denisovich Popov. A modification of the arrow-hurwicz method for search of saddle points. *Mathematical notes of the Academy of Sciences of the USSR*, 28(5):845–848, 1980.

Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pp. 296–301, 1951.

R.W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

Tim Roughgarden. Intrinsic robustness of the price of anarchy. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 513–522, 2009.

Leonard J. Schulman and Umesh V. Vazirani. The duality gap for two-team zero-sum games. *Games Econ. Behav.*, 115:336–345, 2019a. doi: 10.1016/j.geb.2019.03.011. URL https://doi.org/10.1016/j.geb.2019.03.011.

Leonard J Schulman and Umesh V Vazirani. The duality gap for two-team zero-sum games. *Games and Economic Behavior*, 115:336–345, 2019b.

Hassam Ullah Sheikh, Mina Razghandi, and Ladislau Boloni. Learning distributed cooperative policies for security games via deep reinforcement learning. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, volume 1, pp. 489–494. IEEE, 2019.

Karthik Sridharan. Learning from an optimization viewpoint. *arXiv preprint arXiv:1204.4145*, 2012.

Karthik Sridharan and Ambuj Tewari. Convex games in banach spaces. *COLT 2010*, pp. 1, 2010.

Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems*, pp. 2989–2997, 2015.

Shichang Tang. Lessons learned from the training of gans on artificial datasets. *IEEE Access*, 8: 165044–165055, 2020.

Liang Tong, Sixie Yu, Scott Alfeld, et al. Adversarial regression with multiple learners. In *International Conference on Machine Learning*, pp. 4946–4954. PMLR, 2018.

Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.

Yannick Viossat and Andriy Zapechelnyuk. No-regret dynamics and fictitious play. *Journal of Economic Theory*, 148(2):825–842, 2013.

Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, and Georgios Piliouras. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (eds.), *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pp. 10450–10461, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/6c7cd904122e623ce625613d6af337c4-Abstract.html.

John Von Neumann. Zur theorie der gesellschaftsspiele. *Math*, 100:295–320, 1928.

Bernhard von Stengel and Daphne Koller. Team-maxmin equilibria. *Games and Economic Behavior*, 21(1-2):309–321, 1997.

Hongyang Zhang, Susu Xu, Jiantao Jiao, Pengtao Xie, Ruslan Salakhutdinov, and Eric P Xing. Stackelberg gan: Towards provable minimax equilibrium via multi-generator architectures. *arXiv preprint arXiv:1811.08010*, 2018.

Youzhi Zhang and Bo An. Converging to team-maxmin equilibria in zero-sum multiplayer games. In *International Conference on Machine Learning*, pp. 11033–11043. PMLR, 2020.

Youzhi Zhang, Bo An, and Jakub Černỳ. Computing ex ante coordinated team-maxmin equilibria in zero-sum multiplayer extensive-form games. *arXiv preprint arXiv:2009.12629*, 2020.

Dao Li Zhu and Patrice Marcotte. Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. *SIAM Journal on Optimization*, 6(3):714–726, 1996.

A FURTHER RELATED WORK

The properties of team games have given rise to a large amount of literature which we cannot hope to review here; instead we refer the reader to (Kim et al., 2019; Gold, 2005) and references therein. Additionally, convergence properties of online learning, first-order methods (*e.g.* GDA, OGDA, OMWU, EG) in *team games* are not well understood.

Team games. For team games, the solution concept of *team-maxmin equilibria* (TME), a notion introduced by von Stengel & Koller (1997), is a rather well-understood solution concept. In this category of two-team zero-sum games, a team of n players plays defensively against a single *adversarial* player. TME's provide crucial information about robustness since they give the team players the highest payoff possible when they employ uncoordinated strategies. Unfortunately, TMEs correspond to the global optimum of a nonconvex nonlinear program which is NP-hard to compute and as such precludes the use of online learning, first-order algorithms for their computation. Basilico et al. (2017b) provided heuristics with iterative linear programming and support enumeration methods with proved approximation guarantees. For the case of extensive or large-sized normal form games, Basilico et al. (2017a); Celli & Gatti (2018) introduced a relaxed notion of *team-maxmin equilibrium with coordination device* (TMECor). In a TMECor, the team of players decides their joint strategies *ex-ante* against the adversary; that is, the team members are allowed to discuss and agree on tactics before the game starts, but they cannot communicate during the game.

While TMECor's share some properties with NE's in zero-sum two-player games (*e.g.*, exchangeability), the proposed algorithms by Celli & Gatti (2018) leverage tools like Mixed-Integer Linear Program (MILP) that involve a large number of integer variables. In order to tackle the large size of the program heuristically, Zhang et al. (2020); Zhang & An (2020) have proposed modified versions of the incremental strategy generation (ISG) algorithm (See McMahan et al. (2003)) for large-sized games.

Multi-GAN's. We cannot aspire to survey the literature of multi-GAN's in its entirety. Scratching only the surface, we mention that: (a) In MGAN Hoang et al. (2018) and MAD-GAN Li et al. (2019), we have a mixture of generative models with asymptotically infinite capacity networks. (b) In the seminal case of MIX+GAN Arora et al. (2017), an extra regularization term is typically added to discourage the weights of a mixture of generators being too far away from uniform. (c) In Stackelberg GAN, Zhang et al. (2018) exploit the leader-followers relation of discriminator vs generators to achieve smaller minimax gap and better Frechét metrics. For the case of *multiple-discriminators*: (d) In GMAN Durugkar et al. (2016), an increased number of agents are used to achieve higher quality images, while (e) MD-GAN Hardy et al. (2019) is designed to generalize over different datasets and multiple tasks that are spread on multiple workers.

A.1 DISCUSSION ABOUT DIFFERENT SOLUTION CONCEPTS

In this subsection, we would like to stress some differences between the different equilibrium notions (TMECor, TME, and NE). Initially, it is easy to see that TMECor similar to coarse correlated equilibria can achieve better social welfare but, as in the case of congestion games, their price of anarchy can become significantly worse Roughgarden (2009). In order to improve this, TMECor request prior knowledge of the game. For example, as Zhang et al. (2020) mention: *For example, in multiplayer poker games, a team may play against an adversary player, but they cannot communicate and discuss their strategy during the game due to the rule.*

However in many contemporary challenging AI models, the agents/players of each team do not necessarily have knowledge of the actual game. Thus, the only thing the team players can decide in advance is the decision rules they will follow. Thus, the main conceptual reason that we focus on NE's in comparison to TMECor is is because it is a way of reasoning about games in which players do not necessarily need to know in which team they fall into in advance and/or they cannot be expected to be able to coordinate their actions with players of common interest.

Problems of this kind are prevalent in modern ML applications like strategic conflict resolution (Leibo et al., 2017), coordination between autonomous vehicles (Cao et al., 2012) , or collaboration of agents in defensive escort teams (Sheikh et al., 2019). In this kind of games each agent is simulta-

neously working towards maximizing its own payoff *(local reward)* as well as the collective success of the team *(global reward)*.

Conclusively , the notion of Nash equilibrium seems to be a solution concept that needs to be investigated thoroughly in the setting of two-team zero-sum games.

A.2 ADVERSARIAL POTENTIAL GAMES.

Consider a normal form game with two sets of players $\mathcal{N}$ and $\mathcal{M}$. Following the conventions of Section 2, the minimizers $\mathcal{N}$ will be associated with cost functions $c_i,\ i \in [n]$, while the maximizers $\mathcal{M}$ will be associated with utilities $u_j,\ j \in [m]$. Further, there exists an underlying *potential function* $\Phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, so that

(i) for each minimizer $i \in \mathcal{N}$ and strategies $\boldsymbol{x}_i, \boldsymbol{x}_i' \in \mathcal{X}_i$,

$$c_i(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{y}) - c_i(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{y}) = \Phi(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{y}) - \Phi(\boldsymbol{x}_i, \boldsymbol{x}_{-i}, \boldsymbol{y}); \tag{5}$$

(ii) for each maximizer $j \in \mathcal{M}$ and strategies $\boldsymbol{y}_j, \boldsymbol{y}_j' \in \mathcal{Y}_j$,

$$u_j(\boldsymbol{x}, \boldsymbol{y}_j', \boldsymbol{y}_{-j}) - u_j(\boldsymbol{x}, \boldsymbol{y}_j, \boldsymbol{y}_{-j}) = \Phi(\boldsymbol{x}, \boldsymbol{y}_j', \boldsymbol{y}_{-j}) - \Phi(\boldsymbol{x}, \boldsymbol{y}_j, \boldsymbol{y}_{-j}). \tag{6}$$

While this setting is not a two-team zero-sum game, it can still be captured by one, namely $\Gamma(n, m, \Phi)$. More precisely, (5) implies that $c_i(\boldsymbol{x}, \boldsymbol{y}) = \Phi(\boldsymbol{x}, \boldsymbol{y}) + \Psi_i(\boldsymbol{x}_{-i}, \boldsymbol{y})$, for any $i \in \mathcal{N}$ and $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X} \times \mathcal{Y}$, where Ψ_i does not depend on $\boldsymbol{x}_i$, and (6) implies that $u_j(\boldsymbol{x}, \boldsymbol{y}) = \Phi(\boldsymbol{x}, \boldsymbol{y}) + \Psi_j'(\boldsymbol{x}, \boldsymbol{y}_{-j})$, where Ψ_j does not depend on $\boldsymbol{y}_j$. As a result, by the same reasoning we used in the proof of Theorem 3.1, we arrive at the following conclusion:

Proposition A.1. A strategy profile that is an ϵ-approximate Nash equilibrium with respect to the two-team zero-sum game $\Gamma(n, m, \Phi)$ will be an ϵ-approximate Nash equilibrium of the original adversarial potential game.

B GAME DYNAMICS

The following algorithms are ubiquitous in the literature of min-max optimization and have known variable success under different settings. Gradient descent-ascent (GDA) is the prototypical method upon which optimistic gradient descent (OGDA), extra gradient (EG), and our proposed method $(\mathbf{K}, \mathbf{P})$-vaned gradient descent-ascent (KPV-GDA) are built. Optimistic multiplicative weights update is the optimistic variant of the MWU.

B.1 GRADIENT DESCENT-ASCENT

$$\begin{cases} x_i^{(t+1)} = \Pi_{\mathcal{X}_i}\left\{x_i^{(t)} - \eta \nabla_{x_i} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)})\right\} \\ y_j^{(t+1)} = \Pi_{\mathcal{Y}_j}\left\{y_j^{(t)} + \eta \nabla_{y_j} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)})\right\} \end{cases}$$

B.2 OPTIMISTIC GRADIENT DESCENT-ASCENT

$$\begin{cases} x_i^{(t+1)} = \Pi_{\mathcal{X}_i}\left\{x_i^{(t)} - 2\eta \nabla_{x_i} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)}) + \eta \nabla_{x_i} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{(t-1)}\right\} \\ y_j^{(t+1)} = \Pi_{\mathcal{Y}_j}\left\{y_j^{(t)} + 2\eta \nabla_{y_j} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)}) - \eta \nabla_{y_j} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{(t-1)}\right\} \end{cases}$$

B.3 EXTRA GRADIENT METHOD

$$\begin{cases} x_i^{(t+\frac{1}{2})} = \Pi_{\mathcal{X}_i}\left\{x_i^{(t)} - \eta \nabla_{x_i} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)})\right\}, \quad x_i^{(t+1)} = \Pi_{\mathcal{X}_i}\left\{x_i^{(t)} - \eta \nabla_{x_i} f(\boldsymbol{x}^{(t+\frac{1}{2})}, \boldsymbol{y}^{(t+\frac{1}{2})})\right\} \\ y_j^{(t+\frac{1}{2})} = \Pi_{\mathcal{Y}_j}\left\{y_j^{(t)} + \eta \nabla_{y_j} f(\boldsymbol{x}^{(t)} \boldsymbol{y}^{(t)})\right\}, \quad y_j^{(t+1)} = \Pi_{\mathcal{Y}_j}\left\{y_j^{(t)} + \eta \nabla_{y_j} f(\boldsymbol{x}^{(t+\frac{1}{2})}, \boldsymbol{y}^{(t+\frac{1}{2})})\right\} \end{cases}$$

B.4 OPTIMISTIC MULTIPLICATIVE WEIGHTS UPDATE METHOD

$$\begin{cases} x_{i,k}^{(t+1)} = x_{i,k}^{(t)} \dfrac{\exp\left(-2\eta \frac{\partial}{\partial x_{i,k}} f(\boldsymbol{x}^{(t)}\boldsymbol{y}^{(t)}) + \eta \frac{\partial}{\partial x_{i,k}} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{(t-1)})\right)}{\sum_j x_j^{(t)} \exp\left(-2\eta \frac{\partial}{\partial x_{i,k}} f(\boldsymbol{x}^{(t)}\boldsymbol{y}^{(t)}) + \eta \frac{\partial}{\partial x_{i,k}} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{((t-1))})\right)} \\ y_{j,k}^{(t+1)} = y_{j,k}^{(t)} \dfrac{\exp\left(2\eta \frac{\partial}{\partial y_{j,k}} f(\boldsymbol{x}^{(t)}\boldsymbol{y}^{(t)}) - \eta \frac{\partial}{\partial y_{j,k}} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{(t-1)})\right)}{\sum_j y_j^{(t)} \exp\left(2\eta \frac{\partial}{\partial y_{j,k}} f(\boldsymbol{x}^{(t)}\boldsymbol{y}^{(t)}) - \eta \frac{\partial}{\partial y_{j,k}} f(\boldsymbol{x}^{((t-1))}), \boldsymbol{y}^{(t-1)})\right)} \end{cases}$$

B.5 $\mathbf{K}, \mathbf{P}$-VANED GRADIENT DESCENT-ASCENT

$$\begin{cases} \boldsymbol{z}^{(t+1)} = & \Pi_{\mathcal{Z}} \left\{ \boldsymbol{z}^{(t)} + \eta \binom{-\nabla_{\boldsymbol{x}} f(\boldsymbol{z}^{(t)})}{\nabla_{\boldsymbol{y}} f(\boldsymbol{z}^{(t)})} + \eta \mathbf{K}(\boldsymbol{z}^{(t)} - \boldsymbol{\theta}^{(t)}) \right\} \\ \boldsymbol{\theta}^{(t+1)} = & \Pi_{\mathcal{Z}} \left\{ \boldsymbol{\theta}^{(t)} + \eta \mathbf{P}(\boldsymbol{z}^{(t)} - \boldsymbol{\theta}^{(t)}) \right\} \end{cases}$$

Notations $x_i^{(t)}$ (or $y_j^{(t)}$) stand for the strategy vector of the i-th minimizing (or $\mathbf{J}$ -th maximizing) agent at time-step t; while $x_{i,k}$ and $y_{j,k}$ denote the k-th pure strategy of player i of team A and player $\mathbf{J}$ of team B respectively. The step size is denoted as η. Operators $\Pi_{\mathcal{X}_i}, \Pi_{\mathcal{Y}_j}, \Pi_{\mathcal{Z}}$ are the projection operators to the corresponding products of simplices. The projection operator is important since remaining inside of the set of feasible solutions is not guaranteed at every step except for the case of the OMWU method.

B.6 PROPERTIES OF GMP

We will formally demonstrate how a series of conditions needed for the convergence of the extra gradient method are not satisfied for GMP.

Variational Inequalities Let a saddle-point problem (which can be seen as a zero-sum game between agents $\boldsymbol{x}, \boldsymbol{y}$):

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{\boldsymbol{y} \in \mathcal{Y}} U(\boldsymbol{x}, \boldsymbol{y}) \tag{7}$$

Also, let an operator $F : \mathbb{R}^d \to \mathbb{R}^d$ that is L-Lipschitz continuous for a norm $\| \cdot \|_p$.

The saddle-point problem can be formulated in the form of the equivalent Stampacchia variational inequality (SVI), i.e. find a $\boldsymbol{z}^* \in \mathbb{R}^d$, $\boldsymbol{z}^* = (\boldsymbol{x}^*, \boldsymbol{y}^*)$, such that for every $\boldsymbol{z}$:

$$\langle F(\boldsymbol{z}^*) - \boldsymbol{z} - \boldsymbol{z}^* \rangle \geq 0. \tag{8}$$

Operator F is said to be monotone if for every $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{R}^d$:

$$\langle F(\boldsymbol{z}), F(\boldsymbol{w}), \boldsymbol{z} - \boldsymbol{w} \rangle \geq 0. \tag{9}$$

One more noteworthy variational inequality is the Minty Variational Inequality (MVI) with respect to a point $\boldsymbol{z}^*$ is defined as:

$$\langle F(\boldsymbol{z}), \boldsymbol{z} - \boldsymbol{z}^* \rangle \geq 0. \tag{10}$$

Mertikopoulos et al. (2019) prove the convergence of the extra gradient method by utilizing the monotonicity of F. In this case, the solutions to the MVI, $\boldsymbol{z}^*$, coincide with the solutions to the SVI. In particular F is taken to be $F(\boldsymbol{x}, \boldsymbol{y}) = \Big(\nabla_{\boldsymbol{x}} U(\boldsymbol{x}, \boldsymbol{y}), -\nabla_{\boldsymbol{y}} U(\boldsymbol{x}, \boldsymbol{y})\Big)$.

Diakonikolas et al. (2021) prove convergence under a weaker form of the MVI for operator F:

$$\langle F(\boldsymbol{z}), \boldsymbol{z} - \boldsymbol{z}^* \rangle \geq -\frac{\rho}{2} \|F(\boldsymbol{z})\|_{p^*}^2. \tag{11}$$

With ρ a constant in the interval $[0, \frac{1}{4L}]$. We note that with $\rho = 0$ we retrieve the MVI which we proved is not satisfied.

Setting $\boldsymbol{z} = (\frac{1}{3}, \frac{2}{3}, \frac{1}{6}, \frac{5}{6}, \frac{2}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3})$ we see that the MVI is violated: $\langle F(\boldsymbol{z}), \boldsymbol{z} - \boldsymbol{z}^* \rangle = -\frac{\omega}{9}$.

As for the weak-MVI, we numerically verify that for $\omega = 1/2$ and every ρ in $\{\frac{k}{10^3} | k = 0, 1, ..., 10^3 \cdot \frac{1}{4L}\}$ that a $\boldsymbol{z}$ can be found such that is violates the inequality.

Hence, already known techniques (Mertikopoulos et al., 2018; Diakonikolas et al., 2021; Golowich et al., 2020) cannot tackle the problem of computing a Nash equilibrium in GMP. Additionally, Golowich et al. (2020) only analyze the unconstrained setting.

Cocoercivity Cocoerciviy is a condition that is stronger than the MVI. Hence, it follows that cocoercivity does not hold in GMP.

Hidden convex-concave games. Flokas et al. (2021) proved global convergence for gradient flow (continuous time) when a game is strictly hidden convex-concave. Our setting of GMP does not meet neither criteria.

B.7 GMP satisfies the sufficient conditions of Theorem 3.5

We can verify that the generalized matching pennies game satisfies the sufficient conditions of Theorem 3.5. Namely, the eigenvalues of H are the following

$$\rho \in \{-2\omega, 2\omega - 2i, 2\omega + 2i\}.$$

Hence, the set of the eigenvalues with a positive real part E is $E = \{2\omega - 2i, 2\omega + 2i\}$. As such, we see that indeed,

$$\beta = \min_{\rho \in E} \frac{\mathrm{Re}(\rho) + \mathrm{Im}(\rho)}{\mathrm{Re}(\rho)} = \frac{2\omega^2 + 2}{\omega} = 2\omega + \frac{2}{\omega} > 2\omega = \max_{\rho \in E} \mathrm{Re}(\rho) = \alpha.$$

For any choice of $k \in (-2\omega - \frac{2}{\omega}, -2\omega)$ the algorithm converges. For the case that $\omega = 1/2$, it suffices that $k \in (-5, -1)$.

C Proofs for Section 2.1

C.1 Derivation of the min-max objective in equation 3

By the definition of variance, we get that : $\mathrm{Var}\big[x_i\big] = \mathbb{E}[x_i^2] - \big(\mathbb{E}[x_i]\big)^2 \Leftrightarrow \mathbb{E}[x_i^2] = \mathrm{Var}\big[x_i\big] + \big(\mathbb{E}[x_i]\big)^2$. More precisely for any $\boldsymbol{x} \sim \mathcal{N}(\mu, \mathbf{I})$, we get $\mathbb{E}[x_i^2] = 1 + \mu_i^2$.

Additionally, after calculations, we can get that:

$$\begin{aligned}
&\mathbb{E}_{\xi \sim \mathcal{N}(0,\mathbf{I})}[(\xi_i + \theta_i)^2] = \mathrm{Var}\big[\xi_i + \theta_i\big] + \big(\mathbb{E}[\xi_i + \theta_i]\big)^2 \xRightarrow{\theta_i \text{ const.}} \\
&\mathbb{E}_{\xi \sim \mathcal{N}(0,\mathbf{I})}[(\xi_i + \theta_i)^2] = \mathrm{Var}\big[\xi_i\big] + \big(\mathbb{E}[\xi_i] + \mathbb{E}[\theta_i]\big)^2 \Rightarrow \\
&\mathbb{E}_{\xi \sim \mathcal{N}(0,\mathbf{I})}[(\xi_i + \theta_i)^2] = \mathrm{Var}\big[\xi_i\big] + \big(0 + \theta_i\big)^2 \Rightarrow \\
&\mathbb{E}_{\xi \sim \mathcal{N}(0,\mathbf{I})}[(\xi_i + \theta_i)^2] = 1 + \theta_i^2
\end{aligned}$$

Moreover, it is easy to check that for a mixture $\mathcal{D}(\mu_1, \sigma_1\mathbf{I}, \mu_2, \sigma_2\mathbf{I}, \pi_1, \pi_2)$ of two multi-dimensional normal distributions $\mathcal{N}(\mu_1, \sigma_1\mathbf{I}), \mathcal{N}(\mu_1, \sigma_1\mathbf{I})$ with corresponding weights π_1, π_2:

$$\mathrm{Var}_{x \sim \mathcal{D}}[x_i] = \pi_1\sigma_1^2 + \pi_2\sigma_2^2 + \pi_1\mu_{1,i}^2 + \pi_2\mu_{2,i}^2 - (\pi_1\mu_{1,i} + \pi_2\mu_{2,i})^2$$

Specifically for the case of $\mu_1 = \mu = -\mu_2$, this equivalent with:

$$\mathrm{Var}[x_i] = 1 + \mu_i^2 - (\pi_1\mu_i - \pi_2\mu_i)^2$$

Hence:

$$\mathbb{E}[x_i^2] = \mathrm{Var}[x_i] + (\pi_1\mu_i - \pi_2\mu_i)^2 = 1 + \mu_i^2$$

So we can apply the following manipulations in the objective (The notation $[x_i^2]$ stands for the vector that has as entries the values of the vector $\boldsymbol{x}$ squared):

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ \mathbb{E}_{\boldsymbol{y}\sim real}\Big[D_{\boldsymbol{v}}(\boldsymbol{y}) + D_{\boldsymbol{w}}(\boldsymbol{y})\Big] - \\ -\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I}),\zeta\sim\mathcal{N}(0,1)}\Big[G_p(\zeta)\cdot\Big(D_{\boldsymbol{v}}\big(-G_{\boldsymbol{\theta}}(\boldsymbol{y})\big) + D_{\boldsymbol{w}}\big(G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)\Big) + \\ + \big(1-G_p(\zeta)\big)\cdot\Big(D_{\boldsymbol{v}}\big(G_{\boldsymbol{\theta}}(\boldsymbol{y})\big) + D_{\boldsymbol{w}}\big(G_{\boldsymbol{\theta}}(\boldsymbol{y})\big)\Big)\Big] =$$

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ \mathbb{E}_{\boldsymbol{y}\sim real}\Big[\langle \boldsymbol{v},\boldsymbol{y}\rangle + \langle \boldsymbol{w},[y_i^2]\rangle\Big] - \\ -\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I}),\zeta\sim\mathcal{N}(0,1)}\Big[(p+\zeta)\cdot\Big(\langle \boldsymbol{v},-(\boldsymbol{z}+\boldsymbol{\theta})\rangle + \langle \boldsymbol{w},[(\theta_i+\xi_i)^2]\rangle\Big) + \\ + \big(1-(p+\zeta)\big)\cdot\Big(\langle \boldsymbol{v},-(\boldsymbol{z}+\boldsymbol{\theta})\rangle + \langle \boldsymbol{w},[\big(-(\theta_i+\xi_i)\big)^2]\rangle\Big)\Big] =$$

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ \mathbb{E}_{\boldsymbol{y}\sim real}\Big[\langle \boldsymbol{v},\boldsymbol{y}\rangle\Big] + \mathbb{E}_{\boldsymbol{y}\sim real}\Big[\langle \boldsymbol{w},[y_i^2]\rangle\Big] - \\ -\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I}),\zeta\sim\mathcal{N}(0,1)}\Big[(p+\zeta)\cdot\Big(\langle \boldsymbol{v},\boldsymbol{z}+\boldsymbol{\theta}\rangle + \langle \boldsymbol{w},[(\theta_i+\xi_i)^2]\rangle\Big) + \\ + \big(1-(p+\zeta)\big)\cdot\Big(\langle \boldsymbol{v},-(\boldsymbol{z}+\boldsymbol{\theta})\rangle + \langle \boldsymbol{w},[(\theta_i+\xi_i)^2]\rangle\Big)\Big] =$$

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ \Big\langle \boldsymbol{v},\mathbb{E}_{\boldsymbol{y}\sim real}\Big[\boldsymbol{y}\Big]\Big\rangle + \Big\langle \boldsymbol{w},\mathbb{E}_{\boldsymbol{y}\sim real}\Big[[y_i^2]\Big]\Big\rangle - \\ -p\cdot\Big(\Big\langle \boldsymbol{v},\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I})}\Big[(\boldsymbol{z}+\boldsymbol{\theta})\Big]\Big\rangle + \Big\langle \boldsymbol{w},\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I})}\Big[[(\theta_i+\xi_i)^2]\Big]\Big\rangle\Big) - \\ -\big(1-p\big)\Big(\cdot\Big\langle \boldsymbol{v},\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I})}\Big[-(\boldsymbol{z}+\boldsymbol{\theta})\Big]\Big\rangle + \Big\langle \boldsymbol{w},\mathbb{E}_{\xi\sim\mathcal{N}(0,\mathbf{I})}\Big[[(\theta_i+\xi_i)^2]\Big]\Big\rangle\Big) =$$

$$\max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ \Big\langle \boldsymbol{v},\Big(\pi_1\boldsymbol{\mu}+\pi_2(-\boldsymbol{\mu})\Big)\Big\rangle + \Big\langle \boldsymbol{w},[1+\mu_i^2]\Big\rangle - \\ -p\cdot\Big(\Big\langle \boldsymbol{v},\boldsymbol{\theta}\Big\rangle + \Big\langle \boldsymbol{w},[(\theta_i^2+1)]\Big]\Big\rangle\Big) - \big(1-p\big)\cdot\Big(\Big\langle \boldsymbol{v},-\boldsymbol{\theta}\Big\rangle + \Big\langle \boldsymbol{w},[(\theta_i^2+1)]\Big\rangle = \\ \max_{\boldsymbol{v},\boldsymbol{w}} \min_{\boldsymbol{\theta},p} \ (\pi_1-\pi_2)\boldsymbol{v}^T\boldsymbol{\mu} - 2p\boldsymbol{v}^T\boldsymbol{\theta} + \boldsymbol{v}^T\boldsymbol{\theta} + \sum_i w_i(\mu_i^2-\theta_i^2)$$

C.2 Proofs for Section 3

C.3 Proof of Theorem 3.1

We will reduce the problem of finding a Nash equilibrium in congestion games to the problem of finding a Nash equilibrium in two-team zero-sum games. The result then will follow since computing Nash equilibria in congestion games is CLS-hard Babichenko & Rubinstein (2021).

As a recent result (Fearnley et al., 2021) shows CLS is equal to the intersection of PLS and PPAD, two important classes of total problems. PPAD captures diverse problems in combinatorics and (non-)cooperative game theory, like the ε-approximation of a mixed Nash Equilibrium in a graphical game or the computation of market equilibria. PLS, for "Polynomial Local Search", captures problems of finding a local minimum of an objective function f, in contexts where any candidate solution x has a local neighborhood within which we can readily check for the existence of some other point having a lower value of f. Many diverse local optimization problems have been shown complete for PLS, attesting to its importance. Examples include searching for a local optimum of the TSP according to the Lin-Kernighan heuristic (Papadimitriou, 1992) and finding pure Nash equilibria in many-player congestion games (Fabrikant et al., 2004). The complexity class CLS ("Continuous Local Search") was introduced by Daskalakis and Papadimitriou (Daskalakis & Papadimitriou, 2011) to classify various important problems that lie in both PPAD and PLS. CLS is seen as a strong candidate for capturing the complexity of some of those important problems, like

the general versions of Banach's fixed point theorem, computation of KKT points, computation of gradient descent fixed points, etc.

In our reduction, a *congestion game* is defined by the tuple $(N; E; (S_i)_{i\in N}; (c_e)_{e\in E})$ where N is the set of *agents*, E is a set of *resources* (also known as *edges* or *facilities*), and each player i has a set S_i of subsets of E. Each strategy $s_i \in S_i$ is a set of edges (a *path*), and c_e is a cost (negative utility) function associated with facility e. For a strategy profile $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)$, the cost of player i is given by $c_i(\boldsymbol{s}) = \sum_{e\in s_i} c_e(\ell_e(\boldsymbol{s}))$, where $\ell_e(\boldsymbol{s})$ is the number of players using e in $\boldsymbol{s}$ (the load of edge e). It is a well-known result (Rosenthal, 1973) that congestion games exhibit a potential function $\Phi(\boldsymbol{s})$, that

$$\Phi(\boldsymbol{s}) = \sum_{e\in E} \sum_{j=1}^{\ell_e(\boldsymbol{s})} c_e(j)$$

with the property that if any agent i changes her strategy to s_i' it holds that

$$\Phi(s_i', \boldsymbol{s}_{-i}) - \Phi(s_i, \boldsymbol{s}_{-i}) = c_i(s_i', \boldsymbol{s}_{-i}) - c_i(s_i, \boldsymbol{s}_{-i}).$$

Succinct representation. We assume that the congestion game is *succinctly representable* (Daskalakis et al., 2006) and admits the *polynomial expectation property* (Daskalakis et al., 2006). That is, the expected cost of each player i, $c_i(\cdot)$, for any given mixed strategy profile $(s_1, \ldots, s_n)$ is computed in time polynomial in: (i) the number agents n, (ii) $\sum_i |S_i|$ where S_i is the finite set of agent i's strategies, (iii) the number of bits required to represent the mixed strategy profile $(s_1, \ldots, s_n)$.

Reduction. Consider a congestion game $(N; E; (S_i)_{i\in N}; (c_e)_{e\in E})$ with $n = |N|$ players and potential function Φ. We define a team zero-sum game as follows: Team A has n players, in which each agent i chooses strategies from S_i. Team B has n players, with each agent j having only one possible choice (singleton set of actions) call it d, i.e., these are dummy players. If players from Team A choose strategy profile $\boldsymbol{s}$ (Team B has only one choice) then they get utility $u_A(\boldsymbol{s}, d) = -\Phi(s)$. The utility members of Team B get is $u_B(\boldsymbol{s}, d) = \Phi(s)$.

Let $\boldsymbol{x}^* := (x_1^*, ..., x_n^*)$ and $(d, ..., d)$ be a (possibly mixed) Nash equilibrium in the team zero-sum game we defined. We shall show that $(x_1^*, ..., x_n^*)$ is a Nash equilibrium of the original congestion game and the reduction will be complete. Aiming for contradiction, suppose $(x_1^*, ..., x_n^*)$ is not a Nash equilibrium of the original congestion game. Then there exists an agent i that can deviate from strategy x_i^* to $\tilde{x}_i$ and decrease her expected cost. Hence we have that

$$\begin{aligned} 0 &< \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{x}^*}[c_i(\boldsymbol{s})] - \mathbb{E}_{\boldsymbol{s}\sim(\tilde{x}_i^*, \boldsymbol{x}_{-i}^*)}[c_i(\boldsymbol{s})] \\ &= \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{x}^*}[\Phi(\boldsymbol{s})] - \mathbb{E}_{\boldsymbol{s}\sim(\tilde{x}_i^*, \boldsymbol{x}_{-i}^*)}[\Phi(\boldsymbol{s})] \quad \text{(Property of potential).} \end{aligned}$$

We conclude that $\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{x}^*}[u_A(\boldsymbol{s}, d)] = -\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{x}^*}[\Phi(\boldsymbol{s})] < -\mathbb{E}_{\boldsymbol{s}\sim(\tilde{x}_i^*, \boldsymbol{x}_{-i}^*)}[\Phi(\boldsymbol{s})] = \mathbb{E}_{\boldsymbol{s}\sim(\tilde{x}_i^*, \boldsymbol{x}_{-i}^*)}[u_A(\boldsymbol{s}, d)]$ which is a contradiction since $(x_1^*, ..., x_n^*)$ is a Nash equilibrium for the team zero-sum game hence if player i deviates, her payoff (i.e., the payoff of her Team) should not increase.

C.4 Multiplayer Matching Pennies

Below, we present the exact definition of the (2x2) two-team of two-players matching pennies that we discuss in Section 3.2

		Team B		
		HH	HT/TH	TT
Team A	HH	$-1, 1$	$-1/2, 1/2$	$1, -1$
	HT/TH	$1/2, -1/2$	$0, 0$	$1/2, -1/2$
	TT	$1, -1$	$-1/2, 1/2$	$-1, 1$

C.5 PROOF OF THEOREM 3.2

Since $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is not weakly-stable, there exist players i, j from the same team (say B without loss of generality) and strategies k, l, l' so that if i is forced to play k, then $\mathbf{J}$'s best response is l and that gives larger payoff than another strategy l' in her support. Formally it holds that (by multi-linearity of U)

$$(\text{payoff if } i, j \text{ choose } k, l) \;\; \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl}} > \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl'}} \;\; (\text{payoff if } i, j \text{ choose } k, l'), \tag{12}$$

and also $\frac{\partial U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{jl}} = \frac{\partial U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{jl'}}$ (*). We shall show that $\nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*)$ has a strictly positive eigenvalue tangent in the product of simplices (note that if we were working with A, we would show that $\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*)$ has a strictly negative eigenvalue). Consider a vector of size $\sum_i |S_i|$, (where $|S_i|$ is the cardinality of the strategy space of agent i in Team B) which has 1 at coordinates $(i, k), (j, l)$, -1 at coordinate (j, l') and from which we subtract y_i^*; we denote by $\boldsymbol{v}$ the resulting vector. We shall show that $\boldsymbol{v}^\top \nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \boldsymbol{v} < 0$.

By multilinearity of U it follows that $\frac{\partial^2 U}{\partial y_{js} \partial y_{js'}} = 0$ (**) for all agents $\mathbf{J}$ and strategies s, s' and the same is true for x variables (team A). We conclude that

$$\begin{aligned}
\frac{1}{2} \boldsymbol{v}^\top \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \boldsymbol{v} &= \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl}} - \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl'}} - \sum_{s \in S_i} y_{is}^* \left(\frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{is} \partial y_{jl}} - \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{is} \partial y_{jl'}} \right) \\
&= \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl}} - \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl'}} - \left(\frac{\partial U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{jl}} - \frac{\partial U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{jl'}} \right) \\
&\overset{(*)}{=} \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl}} - \frac{\partial^2 U(\boldsymbol{x}^*, \boldsymbol{y}^*)}{\partial y_{ik} \partial y_{jl'}} \overset{(12)}{>} 0.
\end{aligned}$$

Therefore $\nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*)$ has a positive eigenvalue and as a result

$$\mathbf{R} := \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \mathbf{0} \\ \mathbf{0} & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} \tag{13}$$

must have a positive and a negative eigenvalue (since the trace is zero).

We consider the Jacobian of the GDA dynamics at $(\boldsymbol{x}^*, \boldsymbol{y}^*)$. The corresponding matrix is the following:

$$\mathbf{J}_{\text{GDA}} = \mathbf{I} + \eta \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix}, \tag{14}$$

We will show that $\mathbf{J}_{\text{GDA}}$ has an eigenvalue (possible complex) with an absolute value greater than one. It suffices to show that $\mathbf{J}_{\text{GDA}} - \mathbf{I}$ has an eigenvalue with a positive real part (because then $\mathbf{J}_{\text{GDA}}$ would have an eigenvalue with a real part greater than 1 and hence magnitude greater than one). Due to (**), we get that $\mathbf{J}_{\text{GDA}} - \mathbf{I}$ has trace zero. To reach contradiction suppose that no eigenvalue of $\mathbf{J}_{\text{GDA}} - \mathbf{I}$ has a positive real part, then all eigenvalues of $\mathbf{J}_{\text{GDA}} - \mathbf{I}$ should be imaginary or zero. But an imaginary eigenvalue (that is not zero) also results in an eigenvalue with a magnitude greater than one for $\mathbf{J}_{\text{GDA}}$, therefore all eigenvalues of $\mathbf{J}_{\text{GDA}} - \mathbf{I}$ should be zero. We use Ky Fan inequalities which states that the sequence (in decreasing order) of the eigenvalues of $\frac{1}{2}(\mathbf{H} + \mathbf{H}^\top)$ majorizes the real part of the sequence of the eigenvalues of $\mathbf{H}$ (see (Moslehian, 2011, p. 4)) for any matrix $\mathbf{H}$. We choose $\mathbf{H} = \frac{1}{\eta} \cdot (\mathbf{J}_{\text{GDA}} - \mathbf{I})$. Since $\mathbf{R}$ has both a negative and a positive eigenvalue, we get that $\mathbf{H}$ has an eigenvalue with a negative real part. The claim follows since $\mathbf{H}$ has a trace equal to zero, thus it must have an eigenvalue with a positive real part as well.

We conclude that $\mathbf{J}_{\text{GDA}}$ has an eigenvalue with an absolute value greater than one. Using (Daskalakis & Panageas, 2018, Theorem 2.2), it occurs that the set of initial conditions so that GDA converge to $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is of measure zero (for the particular choice of the step size).

C.6 PROOF OF LEMMA 3.1

First, we start with the min-max optimization objective of the GMP game with $\omega = \frac{1}{2}$:

$$\min_{\boldsymbol{x}\in[0,1]^2} \max_{\boldsymbol{y}\in[0,1]^2} -x_1x_2y_1y_2 - (1-x_1)(1-x_2)(1-y_1)(1-y_2) + x_1x_2(1-y_1)(1-y_2) + \\ + (1-x_1)(1-x_2)y_1y_2 + \omega\left(1 - x_1x_2 - (1-x_1)(1-x_2)\right)\left(y_1y_2 + (1-y_1)(1-y_2)\right) - \\ - \omega\left(1 - y_1y_2 - (1-y_1)(1-y_2)\right)\left(x_1x_2 + (1-x_1)(1-x_2)\right)$$

or after simplification, it is equivalent with

$$\min_{\boldsymbol{x}\in[0,1]^2} \max_{\boldsymbol{y}\in[0,1]^2} (\omega+1)(x_1+x_2) + (1-\omega)(y_1+y_2) - (x_1+x_2)(y_1+y_2) - 2\omega x_1x_2 + 2\omega y_1y_2. \quad (15)$$

Remark 4. *We highlight that the min-max optimization objective is multilinear and the degree of each variable in every summand is at most one (total degree is 2). Moreover, we note that due to the nonconvexity-nonconcavity of the function above, the max-min is not equal to the min-max.*

Let $(x_1^*, x_2^*, y_1^*, y_2^*)$ be a Nash equilibrium. Assuming $x_1^*, x_2^*, y_1^*, y_2^* \in (0,1)$ from (15) and first-order conditions we get the system of equations

1. $\omega + 1 - y_1^* - y_2^* - 2\omega x_2^* = 0,$
2. $\omega + 1 - y_1^* - y_2^* - 2\omega x_1^* = 0,$
3. $1 - \omega - x_1^* - x_2^* + 2\omega y_2^* = 0,$
4. $1 - \omega - x_1^* - x_2^* + 2\omega y_1^* = 0.$

Combining the first two equations, we have $x_1^* = x_2^*$ and combining the last two it follows $y_1^* = y_2^*$. Dividing equation one by ω and subtracting three, it follows that $\frac{1}{\omega} - 2\frac{y_1^*}{\omega} + \omega - 2\omega y_1^* = 0$. Hence we conclude that $y_1^* = \frac{1}{2}$. As a result $y_2^* = \frac{1}{2}$ and substituting in first equation $x_1^* = x_2^* = \frac{1}{2}$.

- Assume now that $x_1^* = 0$ and $x_2^*, y_1^*, y_2^* \in (0,1)$. Following the same idea, now only equations 2, 3, 4 hold and instead of the first we have the constraint $\omega+1-y_1^*-y_2^*-2\omega x_2^* \geq 0$. From 3, 4 we conclude that $y_1^* = y_2^*$ and using 2 it holds that $y_1^* = y_2^* = \frac{1+\omega}{2}$. Using 3 follows that $1 - \omega - x_2^* + \omega(\omega+1) = 0$. Thus $x_2^* = 1 + \omega^2 > 1$ (this is not possible because $x_2^* \in [0,1]$).
- Consider the case that $x_1^* = 1$ and $x_2^*, y_1^*, y_2^* \in (0,1)$. Only equations 2, 3, 4 hold and instead of the first we have the constraint $\omega + 1 - y_1^* - y_2^* - 2\omega x_2^* \leq 0$. From 3, 4 we conclude that $y_1^* = y_2^*$ and using 2 it holds that $y_1^* = y_2^* = \frac{1-\omega}{2}$. Using 3 follows that $-\omega - x_2^* + \omega(1-\omega) = 0$. Thus $x_2^* = -\omega^2 < 0$ (this is not possible because $x_2^* \in [0,1]$). By symmetry the same happens when $x_2^* = 0$ or $x_2^* = 1$ and $x_1^*, y_1^*, y_2^* \in (0,1)$.
- Case $x_1^* = x_2^* = 0$ and $y_1^*, y_2^* \in (0,1)$. Using 3, 4 we get $y_1^* = y_2^* = \frac{\omega-1}{2\omega} < 0$ (this is not possible).
- Case $x_1^* = x_2^* = 1$ and $y_1^*, y_2^* \in (0,1)$. Using 3, 4 we get $y_1^* = y_2^* = \frac{\omega+1}{2\omega} > 1$ (this is not possible).
- Case $x_1^* = 0$ and $x_2^* = 1$ and $y_1^*, y_2^* \in (0,1)$. Using 3, 4 we get $y_1^* = y_2^* = \frac{1}{2}$. Moreover one becomes $\omega - 2\omega x_2^* \geq 0$ and two $\omega - 2\omega x_1^* \leq 0$, that is $x_1^* \geq \frac{1}{2}$ and $x_2^* \leq \frac{1}{2}$ (contradiction). The case $x_1^* = 1$ and $x_2^* = 0$ and $y_1^*, y_2^* \in (0,1)$ is symmetric.

Similarly one can consider the case where the $\boldsymbol{y}$ team plays pure and $x_1^*, x_2^* \in (0,1)$. One can also check that all possible pure strategy profiles are not Nash equilibria.

C.7 PROOF OF THEOREM 3.3

We split the proof into 3 parts. Before we start the proof, note that the Hessian of U (19) has infinity norm less than 4 (since $\omega \in (0,1)$), so U has gradient Lipschitz with $L \leq 4$. Thus for the rest of the proof for GDA, we choose $\eta_{\text{OGDA}} < \frac{1}{4}$.

GDA. For GDA the proof will be straightforward. We will show that $(x_1^*, x_2^*, y_1^*, y_2^*)$ is a

weakly Nash equilibrium. Then the claim about GDA will follow because of Lemma 3.1, Theorem 3.2 and Remark 2.1.

Assume that player x_1 fixes his strategy to $x_1 = 0.$ and y_1, y_2 keep their strategy $(\frac{1}{2}, \frac{1}{2})$. We shall show that x_2 is not indifferent in his support and would like to change his mixed strategy $x_2 = \frac{1}{2}$ to pure. When $x_1 = 0$ and $y_1 = y_2 = \frac{1}{2}$ the payoff of Team A (x variables) becomes $-\omega x_2 - 1 + \frac{\omega}{2}$. Since $\omega \in (0, 1)$, x_2 prefers to play $x_2 = 0$ (instead of $\frac{1}{2}$ she had). We conclude that $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is not a weakly-stable Nash equilibrium.

OGDA. The Jacobian of the update rule of OGDA dynamics (use the same machinery of Section 3 in Daskalakis & Panageas (2018)) is the following:

$$\mathbf{J}_{\text{OGDA}} = \begin{pmatrix} \mathbf{I} - 2\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{xx}}U & -2\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{xy}}U & \eta_{\text{OGDA}}\nabla^2_{\boldsymbol{xx}}U & \eta_{\text{OGDA}}\nabla^2_{\boldsymbol{xy}}U \\ 2\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{yx}}U & \mathbf{I} + 2\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{yy}}U & -\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{yx}}U & -\eta_{\text{OGDA}}\nabla^2_{\boldsymbol{yy}}U \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{16}$$

where U is the payoff of Team B (max) and $-U$ is the payoff of team A. Substituting for MPG payoff at Nash equilibrium, we have

$$\nabla^2_{\boldsymbol{xx}}U = \begin{pmatrix} 0 & -2\omega \\ -2\omega & 0 \end{pmatrix}, \nabla^2_{\boldsymbol{yy}}U = \begin{pmatrix} 0 & 2\omega \\ 2\omega & 0 \end{pmatrix}, \nabla^2_{\boldsymbol{xy}}U = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}.$$

The corresponding Jacobian matrix becomes:

$$\mathbf{J}_{\text{OGDA}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \end{pmatrix} + \eta_{\text{OGDA}} \begin{pmatrix} 0 & 4\omega & 2 & 2 & 0 & -2\omega & -1 & -1 \\ 4\omega & 0 & 2 & 2 & -2\omega & 0 & -1 & -1 \\ -2 & -2 & 0 & 4\omega & 1 & 1 & 0 & -2\omega \\ -2 & -2 & 4\omega & 0 & 1 & 1 & -2\omega & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{17}$$

It turns out that matrix (17) has a characteristic polynomial (computed with the help of Mathematica)

$$\begin{aligned} \pi(\lambda) = (4(1+\omega^2)\eta^2_{\text{OGDA}}(1-2\lambda)^2 + (\lambda-1)^2\lambda^2 - 4\omega\eta_{\text{OGDA}}\lambda(1-3\lambda+2\lambda^2)) \\ \cdot((\lambda-1)\lambda + 2\omega\eta_{\text{OGDA}}(2\lambda-1))^2. \end{aligned} \tag{18}$$

A root of $\pi(\lambda)$ is $\frac{1}{2}\left(1 + 4\eta_{\text{OGDA}}j + 4\omega\eta_{\text{OGDA}} + \sqrt{1 - 16\eta^2_{\text{OGDA}} + 32j\omega\eta^2_{\text{OGDA}} + 16\omega^2\eta^2_{\text{OGDA}}}\right)$, which is in absolute value greater than one for $0 < \eta_{\text{OGDA}} \leq \omega$ It is also easy to see the Hessian of U, that is:

$$\begin{pmatrix} \nabla^2_{\boldsymbol{xx}}U & \nabla^2_{\boldsymbol{xy}}U \\ \nabla^2_{\boldsymbol{yx}}U & \nabla^2_{\boldsymbol{yy}}U \end{pmatrix} = \begin{pmatrix} 0 & -2\omega & -1 & -1 \\ -2\omega & 0 & -1 & -1 \\ -1 & -1 & 0 & 2\omega \\ -1 & -1 & 2\omega & 0 \end{pmatrix} \tag{19}$$

is invertible for $\omega \in (0, 1)$ (this is true as the eigenvalues are $-2\omega, 2\omega, -2\sqrt{1+\omega^2}, 2\sqrt{1+\omega^2}$, non of which is zero). From Theorem 3.2 in Daskalakis & Panageas (2018) follows that the initial conditions so that OGDA converges to the Nash equilibrium is of measure zero for step size $\eta_{\text{OGDA}} < \frac{1}{2L} \leq \frac{1}{8}$ (where L is the Lipschitz constant of ∇U). We choose $\eta_{\text{OGDA}} \leq \min(\frac{1}{8}, \omega)$ and the claim follows.

EG. The Jacobian of EG dynamics computed at the Nash equilibrium (fixed point) is given below (made use of chain rule):

$$\begin{aligned} \mathbf{J}_{\text{EG}} = \mathbf{I} + \eta_{\text{EG}} \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} \\ + \eta^2_{\text{EG}} \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix}^2. \end{aligned} \tag{20}$$

We substitute with the values and we get

$$\mathbf{J}_{\mathrm{EG}} = \mathbf{I} + \eta_{\mathrm{EG}} \begin{pmatrix} 0 & 2\omega & 1 & 1 \\ 2\omega & 0 & 1 & 1 \\ -1 & -1 & 0 & 2\omega \\ -1 & -1 & 2\omega & 0 \end{pmatrix} + \eta_{\mathrm{EG}}^2 \begin{pmatrix} 4\omega^2 - 2 & -2 & 4\omega & 4\omega \\ -2 & 4\omega^2 - 2 & 4\omega & 4\omega \\ -4\omega & -4\omega & -2 + 4\omega^2 & -2 \\ -4\omega & -4\omega & -2 & -2 + 4\omega^2 \end{pmatrix}. \tag{21}$$

The eigenvalues of $\mathbf{J}_{\mathrm{EG}}$ are $1 + \eta_{\mathrm{EG}}(-2\omega + 4\eta_{\mathrm{EG}}\omega^2)$, $1 + \eta_{\mathrm{EG}}(-2\omega + 4\eta_{\mathrm{EG}}\omega^2)$, $1 + \eta_{\mathrm{EG}}\left(-4\eta_{\mathrm{EG}} + 2\omega + 4\eta_{\mathrm{EG}}\omega^2 - 2\sqrt{-1 - 8\eta_{\mathrm{EG}}\omega - 16\eta_{\mathrm{EG}}^2\omega^2}\right)$, and $1 + \eta_{\mathrm{EG}}\left(-4\eta_{\mathrm{EG}} + 2\omega + 4\eta_{\mathrm{EG}}\omega^2 + 2\sqrt{-1 - 8\eta_{\mathrm{EG}}\omega - 16\eta_{\mathrm{EG}}^2\omega^2}\right)$. If $-4\eta_{\mathrm{EG}} + 2\omega + 4\eta_{\mathrm{EG}}\omega^2 > 0$ then $\mathbf{J}_{\mathrm{EG}}$ will have an eigenvalue with absolute value greater than one. A sufficient condition for that is $\eta_{\mathrm{EG}} \leq \frac{\omega}{2}$. Finally, for the same choice of stepsizes we have that $\mathbf{J}_{\mathrm{EG}}$ is invertible for all $(\boldsymbol{x}, \boldsymbol{y})$, hence the EG dynamics is a local diffeomorphism. From Theorem 2 in arxiv version of (Lee et al., 2019) the claim for EG method follows.

OMWU. The Jacobian of the update rule of OMWU dynamics (use the same machinery of Section 3 in Daskalakis & Panageas (2019)) is the following:

$$\mathbf{J}_{\mathrm{OMWU}} = \begin{pmatrix} 1 & \eta\omega & \eta & \eta & 0 & -\frac{\eta\omega}{2} & -\frac{\eta}{2} & -\frac{\eta}{2} \\ \eta\omega & 1 & \eta & \eta & -\frac{\eta\omega}{2} & 0 & -\frac{\eta}{2} & -\frac{\eta}{2} \\ -\eta & -\eta & 1 & \eta\omega & \frac{\eta}{2} & \frac{\eta}{2} & 0 & -\frac{\eta\omega}{2} \\ -\eta & -\eta & \eta\omega & 1 & \frac{\eta}{2} & \frac{\eta}{2} & -\frac{\eta\omega}{2} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{22}$$

It turns out that matrix (22) has a characteristic polynomial (retrieved with the help of Mathematica)

$$\pi(\lambda) = \frac{1}{16}\left((4 + \omega^2)\eta^2(1 - 2\lambda)^2 + 4(\lambda - 1)^2\lambda^2 - 4\omega\eta\lambda(1 - 3\lambda + 2\lambda^2)\right) \cdot \left(2(\lambda - 1)\lambda + \omega\eta(2\lambda - 1)\right)^2. \tag{23}$$

One root of the polynomial $\pi(\lambda)$ above is $\frac{1}{2}(1 + 2\eta j + \eta\omega + \sqrt{1 - 4\eta^2 + 4\eta^2\omega j + \eta^2\omega^2})$. As in the analysis of OGDA, it turns out that for $0 < \eta < \omega$ the aforementioned root has an absolute value greater than one.

Therefore we conclude that the Nash equilibrium where each agent plays $(\frac{1}{2}, \frac{1}{2})$ is repelling (and thus the fact that the set of initial conditions so that OMWU converges to that particular fixed point is of measure zero is derived from standard arguments from (Lee et al., 2019). To conclude the proof we need to exclude that OMWU will stabilize in other fixed points. Note that OMWU can stabilize to points that are not Nash equilibria (e.g., all pure strategy profiles are fixed points). To exclude such stabilization, the trick is that if OMWU stabilizes to a point, it should be (approximate coarse correlated equilibrium where the approximation depends on the size of the stepsize). However, none of the pure strategy profiles is a $\omega/2$- approximate coarse correlated equilibrium, so OMWU does not stabilize. Hence if we choose $0 < \eta < \omega/2$, OMWU does not stabilize.

C.8 Proof of Theorem 3.4

As previously, let a zero-sum game with payoff function $U(\boldsymbol{x}, \boldsymbol{y})$ for the maximizing team (it can either be a single player or more players in each team). The maximizing team controls $\boldsymbol{y}$ while the minimizing team controls $\boldsymbol{x}$.

It is easy to retrieve a black-box argument for the proof of Theorem 3.4 using the proof (Hassouneh et al., 2004). For the sake of completeness, we reprove the theorem from the bottom up in our own

vocabulary. Hassouneh et al. (2004) extend the proof (Bazanella et al., 2000) to the discrete-time case without using the singular-perturbation analysis that the latter use.

Let $\mathbf{A} = \mathbf{J}_{\text{GDA}} = \mathbf{I} + \eta \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix}$ be the Jacobian of GDA around a Nash equilibrium. and $\boldsymbol{z}$ be the concatenation of $\boldsymbol{x}$ and $\boldsymbol{y}$. We can locally represent the dynamic of GDA as a linear time-invariant dynamical system:

$$\boldsymbol{z}^{(t+1)} = \mathbf{A}\boldsymbol{z}^{(t)} + \mathbf{B}\boldsymbol{u}^{(t)} \tag{24}$$

$$\boldsymbol{u}^{(t+1)} = \mathbf{0} \tag{25}$$

Here, the feedback law, $\boldsymbol{u}$, is trivially zero, but we note it for the sake of clarity.

By assumption, $\mathbf{A} = \mathbf{J}_{\text{GDA}}$ is invertible, hence full-rank. Also, $\mathbf{B}$ has to be the identity matrix, $\mathbf{B} = \mathbf{I}$, as we want every player to be able to make decisions with access to their own estimation vector.

From a known theorem (the PBH controllability test) the pair of matrices $(\mathbf{A}, \mathbf{B})$, with $\mathbf{A} \in \mathbb{R}^{n \times n}$ is controllable if and only if, for every eigenvalue λ of $\mathbf{A}s$, the rank of the block matrix $[\mathbf{A} - \lambda\mathbf{I}, \mathbf{B}]$ is equal to n. As we know, stabilizability implies controllability. Hence, the pair of matrices $(\mathbf{A}, \mathbf{B}) = (\mathbf{J}_{\text{GDA}}, \mathbf{I})$ is stabilizable.

We can now write down the dynamic of (KPV-GDA) localized around a given (possibly Lyapunov unstable) Nash equilibrium as the following linear time-invariant system:

$$\begin{cases} \boldsymbol{z}^{(t+1)} & = & \underbrace{\mathbf{J}_{\text{GDA}}}_{\mathbf{A}} \boldsymbol{z}^{(t)} + \underbrace{\mathbf{I} \cdot \mathbf{K}}_{B \cdot \mathbf{K}} \boldsymbol{u}^{(t)} \\ \boldsymbol{\theta}^{(t+1)} & = & \boldsymbol{\theta}^{(t)} + \mathbf{P}\big(\boldsymbol{z}^{(t)} - \boldsymbol{\theta}^{(t)}\big) \\ \boldsymbol{u}^{(t+1)} & = & \boldsymbol{z}^{(t)} - \boldsymbol{\theta}^{(t)} \end{cases} . \tag{26}$$

This dynamic can be represented in a block matrix fashion as:

$$\mathbf{M} := \begin{pmatrix} \mathbf{J}_{\text{GDA}} + \mathbf{K} & -\mathbf{K} \\ \mathbf{P} & \mathbf{I} - \mathbf{P} \end{pmatrix}. \tag{27}$$

Knowing that eigenvalues are preserved under similarity transformations, we define two similarity transformations $\mathbf{T}_1, \mathbf{T}_2$ which we will apply consecutively on $\mathbf{M}$.

Let $\mathbf{Q}$ be any matrix. We define:

$$\mathbf{T}_1 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{-1} \end{pmatrix}, \quad \mathbf{T}_1^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix} \tag{28}$$

and:

$$\mathbf{T}_2 = \begin{pmatrix} \mathbf{I} & \mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \mathbf{T}_2^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{Q} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \tag{29}$$

By consecutively applying the latter transformations we get:

$$\begin{aligned} \mathbf{M}' =& \mathbf{T}_2\mathbf{T}_1\mathbf{M}\mathbf{T}_1^{-1}\mathbf{T}_2^{-1} = \\ =& \left(\begin{array}{c|c} \mathbf{J}_{\text{GDA}} + \mathbf{K} + \mathbf{Q} & -\mathbf{J}_{\text{GDA}}\mathbf{Q} - \mathbf{K}\mathbf{Q} - \mathbf{Q}^2 - \mathbf{K}\mathbf{P} + \mathbf{Q} - \mathbf{Q}\mathbf{P} \\ \hline \mathbf{I} & -\mathbf{Q} + \mathbf{I} - \mathbf{P} \end{array}\right) \end{aligned} \tag{30}$$

We now want to eliminate the top-left entry of $\mathbf{M}'$ in order to have an easy application of the Schur complement equations. We are free to do so if we can find an appropriate matrix $\mathbf{Q}$. Remember that $\mathbf{Q}$ can be any matrix.

First, we write matrices $\mathbf{P}$ and $\mathbf{Q}$ as (non-singular) perturbations of some appropriate matrices for $\epsilon > 0$ sufficiently small:

$$\mathbf{P} = \epsilon\mathbf{P}_1 \tag{31}$$

$$\mathbf{Q} = \mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2) \tag{32}$$

Setting the top-right entry of $\mathbf{M}'$ equal to $\mathbf{0}$:

$$-\mathbf{J}_{\text{GDA}}\mathbf{Q} - \mathbf{K}\mathbf{Q} - \mathbf{Q}^2 - \mathbf{K}\mathbf{P} + \mathbf{Q} - \mathbf{Q}\mathbf{P} = 0 \Rightarrow \tag{33}$$

$$-\mathbf{J}_{\text{GDA}}\Big(\mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2)\Big) - \mathbf{K}\Big(\mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2)\Big) - \Big(\mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2)\Big)^2 - \\ -\mathbf{K}\epsilon\mathbf{P}_1 + \Big(\mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2)\Big) - \Big(\mathbf{Q}_0 + \epsilon\mathbf{Q}_1 + O(\epsilon^2)\Big)\epsilon\mathbf{P}_1 = 0 \tag{34}$$

Gathering terms w.r.t. to the power of ϵ we get, for $O(1)$ terms:

$$(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I} + \mathbf{Q}_0)\mathbf{Q}_0 = \mathbf{0}. \tag{35}$$

And for $O(\epsilon)$ terms we get:

$$\mathbf{Q}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I}) + (\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1 = \mathbf{0}. \tag{36}$$

.

From equation 35 we get that

$$\mathbf{Q}_0 = \mathbf{0} \quad \text{or} \quad \mathbf{Q}_0 = -\mathbf{J}_{\text{GDA}} - \mathbf{K} + \mathbf{I}. \tag{37}$$

From equation 36, by observing that we can always select a $\mathbf{K}$ such that 1 is not an eigenvalue of $\mathbf{J}_{\text{GDA}} + \mathbf{K}$ we get that:

$$\mathbf{Q}_1 = -(\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1}. \tag{38}$$

We observe that $\mathbf{Q}_1$ is uniquely defined through a matrix inversion. Hence, from the Implicit Function Theorem applied to equation 33 we get that there exists a locally unique matrix $\mathbf{Q}$:

$$\mathbf{Q} = -\mathbf{J}_{\text{GDA}} - \mathbf{K} + \mathbf{I} - \epsilon(\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1} + O(\epsilon^2) \tag{39}$$

We return to matrix $\mathbf{M}'$ and substitute $\mathbf{Q}$ with what we calculated in order to get:

$$\mathbf{M}' = \left(\begin{array}{c|c} \mathbf{I} - \epsilon(\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1} + O(\epsilon^2) & \mathbf{0} \\ \hline \mathbf{I} & \mathbf{\Lambda} \end{array}\right) \tag{40}$$

With $\mathbf{\Lambda} = \mathbf{J}_{\text{GDA}} + \mathbf{K} + \epsilon\Big((\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1} - \mathbf{P}_1\Big) + O(\epsilon^2)$.

We have assumed $\eta \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{xy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{yx}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}}U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} = \mathbf{J}_{\text{GDA}} - \mathbf{I}$ to be invertible, hence in order for $\mathbf{M}'$ to be stable it suffices that:

- $\mathbf{I} - \epsilon(\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1} + O(\epsilon^2)$ is stable
- $\mathbf{J}_{\text{GDA}} + \mathbf{K}$ is stable with eigenvalues away from 1,
  so that term $\epsilon\Big((\mathbf{J}_{\text{GDA}} - \mathbf{I})\mathbf{P}_1(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})^{-1} - \mathbf{P}_1\Big)$ does not cause the eigenvalues of $\mathbf{J}_{\text{GDA}} + \mathbf{K}$ to get a a magnitude greater than 1 (not to become unstable).

For the first item it easy to see that an appropriate $\mathbf{P}_1$ always exists, that is, $\mathbf{P}_1 = (\mathbf{J}_{\text{GDA}} - \mathbf{I})^{-1}(\mathbf{J}_{\text{GDA}} + \mathbf{K} - \mathbf{I})$.

As for the second item, it follows from the fact that $(\mathbf{J}_{\text{GDA}}, \mathbf{I})$ is stabilizable as we have demonstrated.

C.9 PROOF OF THEOREM 3.5

We first compute the Jacobian of KPV-GDA dynamics (KPV-GDA) where we have eliminated one variable per agent (to avoid using the projection operator). The Jacobian has the following form:

$$\mathbf{J}_{\text{KPV-GDA}} = \mathbf{I} + \eta \begin{pmatrix} k \cdot \mathbf{I} + \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} & -k \cdot \mathbf{I} \\ p \cdot \mathbf{I} & -p \cdot \mathbf{I} \end{pmatrix}. \tag{41}$$

It suffices to show that the spectral radius of $\mathbf{J}_{\text{KPV-GDA}}$ is less than one for step size η small enough. We first show the claim below

Claim C.1. Let $\mathbf{J} = \mathbf{I} + \eta \mathbf{M}$ where $\mathbf{I}$ is the identity matrix. Suppose that $\mathbf{M}$ has all of its eigenvalues with a real part that is negative. Then, there exists an interval $(0, \eta_0)$ so that if $\eta \in (0, \eta_0)$ then $\mathbf{J}$ has all of its eigenvalues with an absolute value less than one.

Proof. Let $a_i + b_i \cdot j$ be an eigenvalue of $\mathbf{M}$ such that $a_i < 0$. Assume that $\eta < -\frac{2a_i}{a_i^2 + b_i^2}$, then the corresponding eigenvalue of $\mathbf{J}$ is $1 + \eta(a_i + b_i \cdot j)$, the magnitude of which is $(1 + \eta a_i)^2 + \eta^2 b_i^2 = 1 + 2\eta a_i + \eta^2(a_i^2 + b_i^2) < 1$ since $-2a_i > \eta(a_i^2 + b_i^2)$ (by assumption). Hence we can choose η_0 to be $\min_i \frac{-2a_i}{a_i^2 + b_i^2} > 0$. □

Using Claim C.1, we conclude that as long as the matrix $\mathbf{M}$ below has eigenvalues with real part negative then $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is attracting:

$$\mathbf{M} := \begin{pmatrix} k \cdot \mathbf{I} + \begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix} & -k \cdot \mathbf{I} \\ p \cdot \mathbf{I} & -p \cdot \mathbf{I} \end{pmatrix}. \tag{42}$$

By setting $\mathbf{M}_{11}, \mathbf{M}_{12}, \mathbf{M}_{21}, \mathbf{M}_{22}$ as the corresponding block matrices we compute the characteristic polynomial of $\mathbf{M}$. It holds

$$\det(\mathbf{M} - \lambda \mathbf{I}) = \det\left(\left[\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \hline \mathbf{M}_{21} & \mathbf{M}_{22} \end{array}\right] - \lambda \mathbf{I}\right) = \det\left(\left[\begin{array}{c|c} \mathbf{M}_{11} - \lambda \mathbf{I} & -k\mathbf{I} \\ \hline p\mathbf{I} & -p\mathbf{I} - \lambda \mathbf{I} \end{array}\right]\right) = \tag{43}$$

$$= \det\left(\left[\begin{array}{c|c} \mathbf{M}_{11} - \lambda \mathbf{I} - \frac{kp}{p+\lambda}\mathbf{I} & -k\mathbf{I} \\ \hline 0 & -p\mathbf{I} - \lambda \mathbf{I} \end{array}\right]\right) = \tag{44}$$

$$= \det\left(\mathbf{M}_{11} - \left(\lambda + \frac{kp}{p+\lambda}\right)\mathbf{I}\right) \det\left(-(p+\lambda)\mathbf{I}\right). \tag{45}$$

$\det(\mathbf{M} - \lambda \mathbf{I})$ has roots at $\lambda = -p$ and when

$$\det\left(\mathbf{M}_{11} - \left(\lambda + \frac{kp}{p+\lambda}\right)\mathbf{I}\right) = 0.$$

Therefore, if $\mathbf{M}$ has all eigenvalues with real part negative, we must have $p > 0$. Let ρ be an eigenvalue of $\begin{pmatrix} -\nabla^2_{\boldsymbol{xx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & -\nabla^2_{\boldsymbol{yx}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \nabla^2_{\boldsymbol{xy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) & \nabla^2_{\boldsymbol{yy}} U(\boldsymbol{x}^*, \boldsymbol{y}^*) \end{pmatrix}$. It must hold that

$$\lambda + \frac{kp}{p+\lambda} - k = \rho. \tag{46}$$

We need to investigate under what assumptions λ will have a real part negative. We expand (46) and we get

$$\lambda^2 + \lambda(p - k - \rho) - \rho p = 0. \tag{47}$$

We solve Equation (47) and we get

$$\lambda_{1,2} = \frac{-p + k + \rho \pm \sqrt{(p - k - \rho)^2 + 4\rho p}}{2}. \tag{48}$$

We provide sufficient conditions so that $\text{Re}(\lambda_{1,2}) < 0$. We consider the following cases:

- ρ is real and negative. In this case observe that for $p = 0$ we have that $\lambda_{1,2} = k + \rho$ and 0. Note that for $k < 0$, the first eigenvalue has real part negative and the second eigenvalue is zero. We compute the first derivative at $p = 0$ and this gives $\frac{1}{2}\left(-1 + \frac{-k+\rho}{|k+\rho|}\right) = \frac{\rho}{-k-\rho} < 0$. Therefore the $\lambda_{1,2}$ as a function of p is strictly decreasing at zero. We conclude that for p sufficiently small and positive, both eigenvalues of $\mathbf{M}$ will be real and negative. Hence if an eigenvalue of $\mathbf{H}$ is real and negative then for p sufficiently small positive and $k < 0$ both eigenvalues of $\mathbf{M}$ will be negative.
- ρ is complex with $|\text{Im}(\rho)| \neq 0$. In this case observe that for $p = 0$ we have that $\lambda_{1,2} = k + \rho$ and 0. If we choose $\min(0, -\text{Re}(\rho)) > k$, then the first eigenvalue will have real part negative. Now using the same idea as for the first case, we compute the first derivative of the real part of $\lambda_{1,2}$ as a function of p and we get

$$\frac{1}{2}\left(-1 \pm \frac{k^2 - \text{Re}(\rho)^2 - \text{Im}(\rho)^2}{(k + \text{Re}(\rho))^2 + \text{Im}(\rho)^2}\right).$$

  The expression above is negative as long as $-\text{Re}(\rho)^2 - \text{Im}(\rho)^2 - k\text{Re}(\rho) < 0$. The equation is trivially satisfied when $\text{Re}(\rho) < 0$ since k is chosen to be negative. Assume $\text{Re}(\rho) > 0$ (if it is zero then the above is trivially true since the imaginary part is non-zero). It occurs that the inequality above is true when $k > -\frac{|\rho|^2}{\text{Re}(\rho)}$. Since k is chosen to be smaller than $\min(0, -\text{Re}(\rho))$, it suffices to show that

$$-\text{Re}(\rho)) > -\frac{|\rho|^2}{\text{Re}(\rho)} \text{ is satisfied by our assumptions.}$$

  The above is true as long as $\text{Im}(\rho) \neq 0$. Let E be the set of eigenvalues ρ of $\mathbf{H}$ with real part positive (and non-zero imaginary part by assumption). There exists a choice for η, k, p, in which $\eta, p > 0$ and sufficiently small and k is negative and chosen to be

$$\min_{\rho \in E} \frac{|\rho|^2}{\text{Re}(\rho)} > -k > \max_{\rho \in E} \text{Re}(\rho).$$

D EXPERIMENTS

We elaborate on the experiment mentioned in the main body. Our experiment includes a dataset of a mixture of 2-D Gaussians with 8 modes. Our architecture includes 8 "shallow" generators and discriminators with 2 layers of 2-16-2 ReLU activations, compared with a "large" single-agent GAN with 4 layers of 2-128-256-1024-2 activations. Interestingly, the giant one fails in a double sense; It demonstrates both mode-collapsing and mode-drop phenomena without stabilizing. On the other hand, our architecture with a small number of neurons achieves to fit the data well.

The iterations illustrate the step at which the model roughly stops learning/converges. We observed no change after the demonstrated number of iterations.

Configuration	Generator(s)	Discriminator(s)	Typical outcome
Single GAN	# $1\times$ { 2-128-256-1024-2 Linear w/ ReLU	# $1\times$ { 2-128-256-1024-2 Linear w/ ReLU	Mode coll. in $\sim 1\times 10^3$ iters. Mode drop in $\sim 3\times 10^3$ iters.
Multi-GAN	# $8\times$ { 2-16-2 Linear w/ ReLU	# $8\times$ { 2-16-2 Linear w/ ReLU	Distr. learned in $\sim 3\times 10^3$ iters.

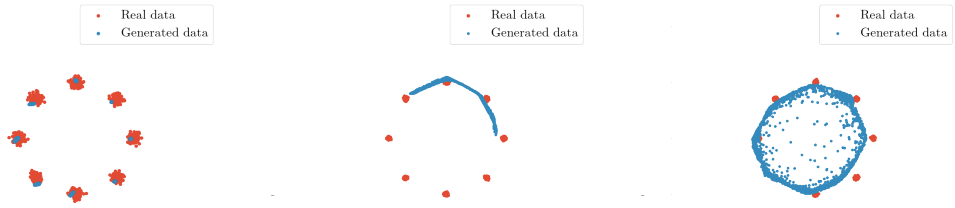


Figure 5: From left to right: (i) Each generator of MGAN learns one mode of a mixture of 8 gaussians, (ii) Mode Collapse of single-agent GAN's, (iii) Single-agent GAN can't discriminate between the modes.

D.1 MGAN VS. WGAN

CIFAR-10 is a well-established testbed for various GAN architectures. It contains 10 balanced classes of images of different objects.

Just to further illustrate the merits of multi-agent GAN's, we offer a selection of images generated by WGAN-GP (Arjovsky et al., 2017) and the MGAN (Hoang et al., 2017).

In the pictures provided by the WGAN-GP we see that across iterations, the generated samples tend to cover only certain classes of the dataset while it takes longer for the samples to become realistic.

On the other hand, the MGAN architecture from an early stage provides with diverse samples that tend to be more realistic from early on.

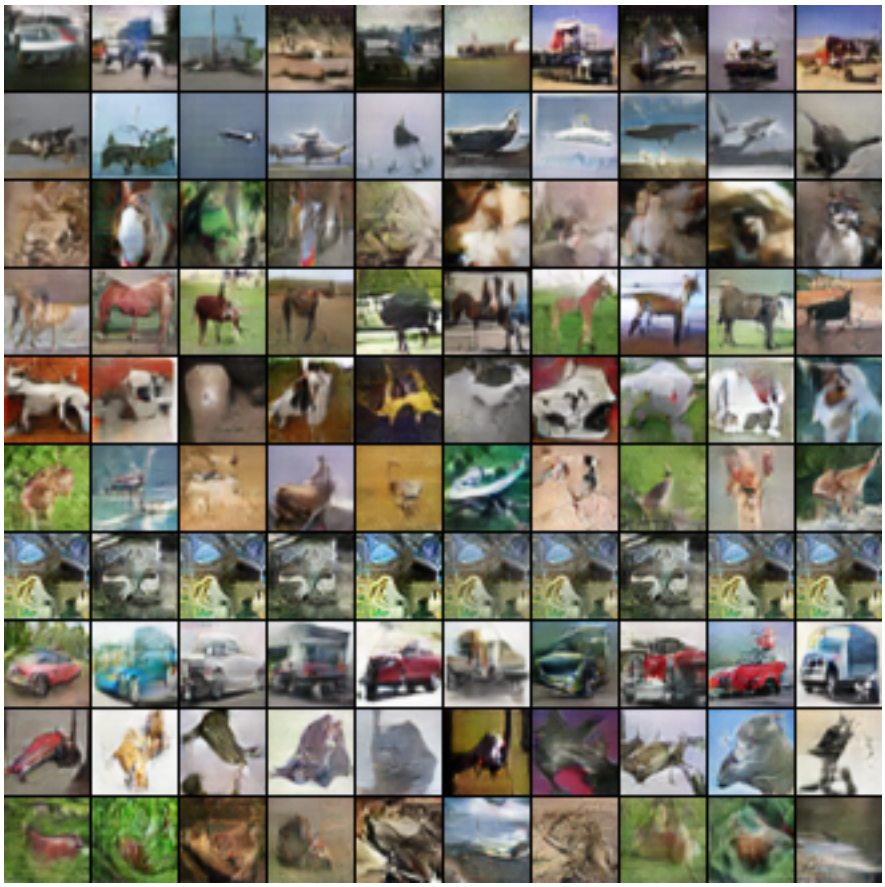


Figure 6: MGAN after 300 iterations

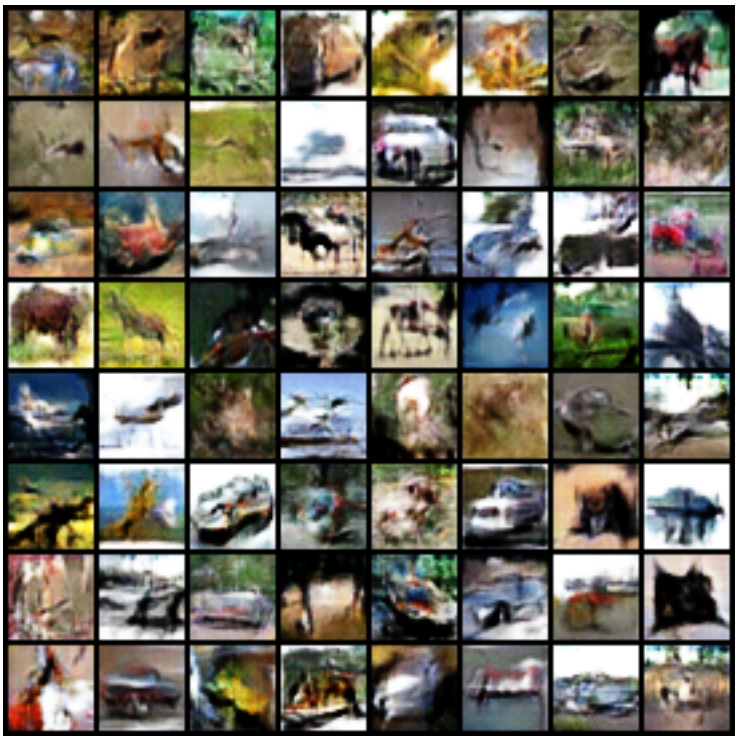


Figure 7: WGAN after 10000 iterations

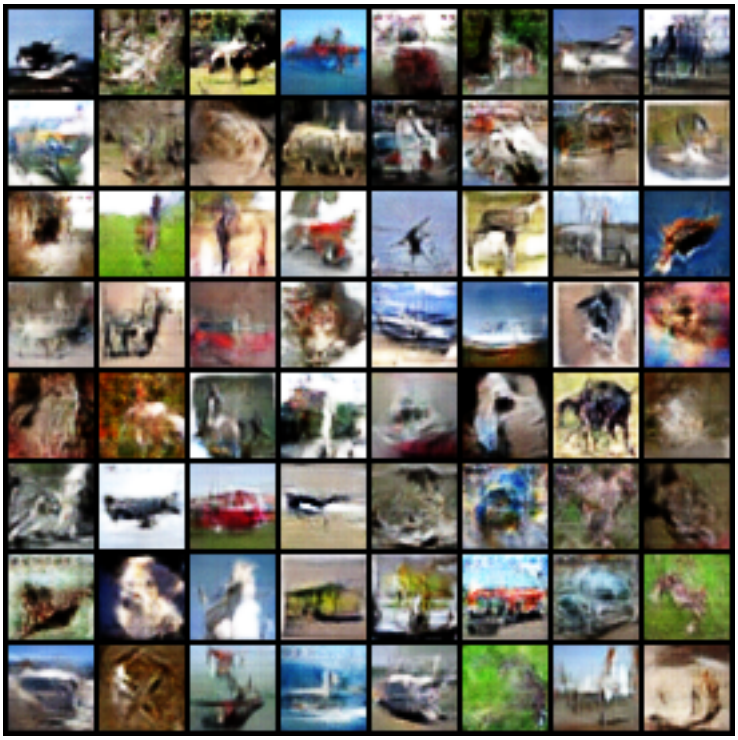


Figure 8: WGAN after 15000 iterations

Figure 9: WGAN after 25000 iterations

E EMPIRICAL PERFORMANCE OF THE PROPOSED METHOD

In this section, we perform a number of empirical experiments to assess the performance of our proposed method. As we highlighted, the conditions of Theorem 3.5 are only sufficient and not necessary. In the following figure, we plot the curves of the *Nash equilibrium gap* for 500 random two-team zero-sum games. In every such game, every player has 2 available strategies. We note that for all of these games we use a random initialization for the players' strategies and the same parameters $\eta = 0.05,\ k = -1.2,\ p = 0.02$ for our proposed method. It is notable that the Nash equilibrium gap vanishes for virtually all of these instances. This indicates that the conditions of Theorem 3.5 are only sufficient and a more complex analysis could show less restrictive conditions for convergence with the same simple parametrization.

For completeness, we include the definition of the Nash equilibrium gap:

Definition E.1 (NE-gap)**.** *We define the Nash equilibrium gap as the sum of the differences between the best-response and the strategy every player plays:*

$$\text{NE-gap}(\boldsymbol{x}) = \sum_{i=1}^{n} \max_{\boldsymbol{x}_i^*} \{u_i(\boldsymbol{x}_i^*, \boldsymbol{x}_{-i})\} - u_i(\boldsymbol{x}). \tag{49}$$

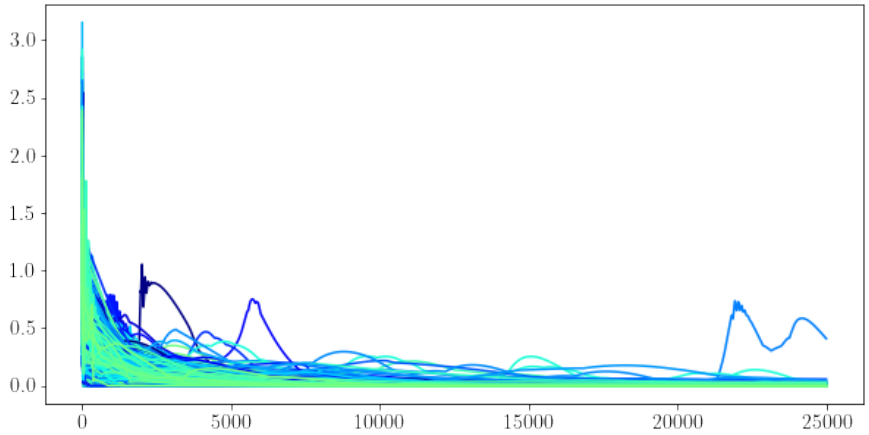


Figure 10: Nash equilibrium gap for 2-vs.-2 two-team zero-sum game with 2 strategies per player. KPV with parameters $\eta = 0.05,\ k = -1.2,\ p = 0.02$.

Additionally, we apply the same method with fixed parameters $\eta = 0.05,\ k = -1.05,\ p = 0.005$ to two-team zero-sum games with 3 strategies per player and 2 players in each team. Again, most random games seem to converge to a 0 NE-gap.

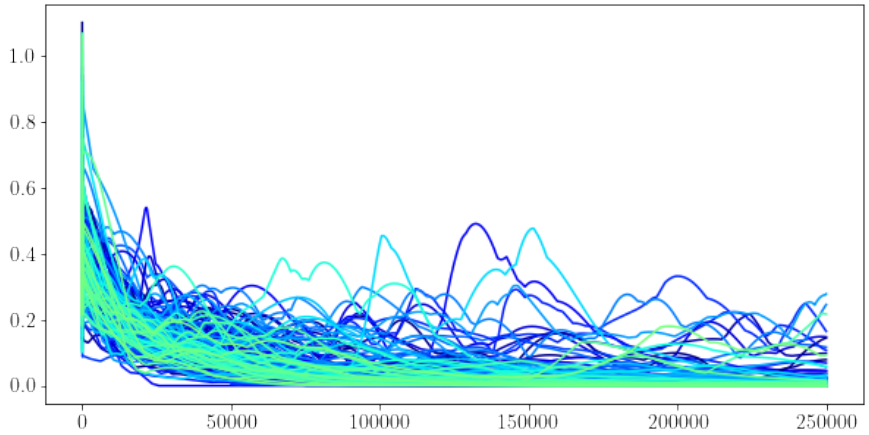


Figure 11: Nash equilibrium gap for 2-vs.-2 two-team zero-sum game with 3 strategies per player. KPV with parameters $\eta = 0.05,\ k = -1.05,\ p = 0.005$.