
The Complexity of Symmetric Equilibria in Min-Max Optimization and Team Zero-Sum Games*

Ioannis Anagnostides¹, Ioannis Panageas², Tuomas Sandholm^{1,3}, and Jingming Yan²

¹Carnegie Mellon University

²University of California, Irvine

³Strategy Robot, Inc.

³Strategic Machine, Inc.

³Optimized Markets, Inc.

{ianagnos,sandholm}@cs.cmu.edu, {ipanagea,jingmy1}@uci.edu

Abstract

We consider the problem of computing stationary points in min-max optimization, with a focus on the special case of Nash equilibria in (two-)team zero-sum games. We first show that computing ϵ -Nash equilibria in 3-player *adversarial* team games—wherein a team of 2 players competes against a *single* adversary—is CLS-complete, resolving the complexity of Nash equilibria in such settings. Our proof proceeds by reducing from *symmetric* ϵ -Nash equilibria in *symmetric*, identical-payoff, two-player games, by suitably leveraging the adversarial player so as to enforce symmetry—without disturbing the structure of the game. In particular, the class of instances we construct comprises solely polymatrix games, thereby also settling a question left open by Hollender, Maystre, and Nagarajan (2024).

Moreover, we establish that computing *symmetric* (first-order) equilibria in *symmetric* min-max optimization is PPAD-complete, even for quadratic functions. Building on this reduction, we show that computing symmetric ϵ -Nash equilibria in symmetric, 6-player (3 vs. 3) team zero-sum games is also PPAD-complete, even for $\epsilon = \text{poly}(1/n)$. As a corollary, this precludes the existence of symmetric dynamics—which includes many of the algorithms considered in the literature—converging to stationary points. Finally, we prove that computing a *non-symmetric* $\text{poly}(1/n)$ -equilibrium in symmetric min-max optimization is FNP-hard.

1 Introduction

We consider computing local equilibria in constrained min-max optimization problems of the form

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}, \mathbf{y}), \quad (1)$$

where $\mathcal{X} \subseteq \mathbb{R}^{d_x}$ and $\mathcal{Y} \subseteq \mathbb{R}^{d_y}$ are convex and compact constraint sets, and $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ is a smooth objective function. Tracing all the way back to Von Neumann’s celebrated minimax theorem [von Neumann, 1928] and the inception of game theory, such problems are attracting renewed interest in recent years propelled by a variety of modern machine learning applications, such as generative modeling [Goodfellow et al., 2014], reinforcement learning [Daskalakis et al., 2020, Bai and Jin, 2020, Wei et al., 2021], and adversarial robustness [Madry et al., 2018, Cohen et al., 2019, Bai et al., 2021, Carlini et al., 2019]. Another prominent class of problems encompassed by (1) concerns computing Nash equilibria in *(two-)team zero-sum games* [Zhang et al., 2023, 2021, Basilico et al., 2017, von Stengel

*The authors are ordered alphabetically.

and Koller, 1997, Carminati et al., 2023, Orzech and Rinard, 2023, Farina et al., 2018, Zhang and An, 2020, Celli and Gatti, 2018, Schulman and Vazirani, 2017], which is a primary focus of this paper.

Perhaps the most natural solution concept—guaranteed to always exist—pertaining to (1), when f is nonconvex-nonconcave, is a pair of strategies $(\mathbf{x}^*, \mathbf{y}^*)$ such that both players (approximately) satisfy the associated *first-order* optimality conditions [Tsaknakis and Hong, 2021, Jordan et al., 2023, Ostrovskii et al., 2021, Nouiehed et al., 2019], as formalized in the definition below.

Definition 1.1. A point $(\mathbf{x}^*, \mathbf{y}^*) \in \mathcal{X} \times \mathcal{Y}$ is an ϵ -first-order Nash equilibrium of (1) if

$$\langle \mathbf{x} - \mathbf{x}^*, \nabla_{\mathbf{x}} f(\mathbf{x}^*, \mathbf{y}^*) \rangle \geq -\epsilon \quad \text{and} \quad \langle \mathbf{y} - \mathbf{y}^*, \nabla_{\mathbf{y}} f(\mathbf{x}^*, \mathbf{y}^*) \rangle \leq \epsilon \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}.$$

Definition 1.1 can be equivalently recast as a variational inequality (VI) problem: if $\mathbf{z} := (\mathbf{x}, \mathbf{y})$ and $F : \mathbf{z} \mapsto F(\mathbf{z}) := (\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}), -\nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}))$, we are searching for a point $\mathbf{z}^* \in \mathcal{Z} := \mathcal{X} \times \mathcal{Y}$ such that $\langle \mathbf{z} - \mathbf{z}^*, F(\mathbf{z}^*) \rangle \geq -2\epsilon$ for all $\mathbf{z} \in \mathcal{Z}$. Yet another equivalent definition is instead based on approximate *fixed points* of gradient descent/ascent (GDA); namely, Definition 1.1 amounts to bounding the gradient mappings

$$\|\mathbf{x}^* - \Pi_{\mathcal{X}}(\mathbf{x}^* - \nabla_{\mathbf{x}} f(\mathbf{x}^*, \mathbf{y}^*))\| \leq \epsilon', \|\mathbf{y}^* - \Pi_{\mathcal{Y}}(\mathbf{y}^* + \nabla_{\mathbf{y}} f(\mathbf{x}^*, \mathbf{y}^*))\| \leq \epsilon' \quad (\text{Fixed points of GDA})$$

for some approximation parameter $\epsilon' > 0$ that is (polynomially) dependent on $\epsilon > 0$, where $\|\cdot\|$ is the (Euclidean) ℓ_2 norm and $\Pi(\cdot)$ is the projection operator. Other definitions that differentiate between the order of play between players—based on the notion of a Stackelberg equilibrium—have also been considered in the literature [Jin et al., 2020].

The complexity of min-max optimization is well-understood in certain special cases, such as when f is convex-concave (e.g., Korpelevich [1976], Mertikopoulos et al. [2019], Cai et al. [2022], Choudhury et al. [2023], Gorbunov et al. [2022], and references therein), or more broadly, nonconvex-concave [Lin et al., 2020, Xu et al., 2023, Luo et al., 2020]. However, the complexity of general min-max optimization problems, when the objective function f is nonconvex-nonconcave, has remained wide open despite intense efforts in recent years. Daskalakis et al. [2021] made progress by establishing certain hardness results targeting the more challenging setting in which there is a *joint* (that is, coupled) set of constraints. In fact, it turns out that their lower bounds apply even for linear-nonconcave objective functions (cf. Bernasconi et al. [2024]), showing that their hardness result is driven by the presence of joint constraints—indeed, under uncoupled constraints, many efficient algorithms attaining Definition 1.1 (for linear-nonconcave problems) have been documented in the literature. In the context of min-max optimization, the most well-studied setting posits that players have independent constraints; this is the primary focus of our paper.

1.1 Our results

We establish new complexity lower bounds in min-max optimization for computing equilibria in the sense of Definition 1.1; our main results are gathered in Table 1.

Table 1: The main results of this paper. NE stands for Nash equilibrium and FONE for first-order Nash equilibrium (Definition 1.1). We also abbreviate symmetric to “sym.” (second column).

Class of problems	Eq. concept	Complexity	Even for
Adversarial team games	ϵ -NE	CLS-complete (Theorem 1.3)	3-player (2 vs 1), polymatrix
Symmetric min-max	sym. ϵ -FONE	PPAD-complete (Theorem 1.7)	polymatrix, team 0-sum, $\epsilon = 1/n^c$
	non-sym. ϵ -FONE	NP-hard (Theorem 1.6)	quadratics, $\epsilon = 1/n^c$

Adversarial team games We first examine an important special case of (1): *adversarial team games* [von Stengel and Koller, 1997]. Here, a team of n players with identical interests is competing against a single adversarial player. (In such settings, Definition 1.1 captures precisely the Nash equilibria of the game.) The computational complexity of this problem was placed by Anagnostides et al. [2023] in the complexity class CLS—which stands for continuous local search [Daskalakis and Papadimitriou, 2011]. Further, by virtue of a result of Babichenko and Rubinstein [2021], computing Nash equilibria in adversarial team games when $n \gg 1$ is CLS-complete. In the context of this prior work, an important question left open by Anagnostides et al. [2023] concerns the case where n is a small constant, a regime not captured by the hardness result of Babichenko and Rubinstein [2021]

pertaining to identical-interest games—in such games, one can simply identify the strategy leading to the highest payoff, which is tractable when n is small.

We show that even when $n = 2$, computing an ϵ -Nash equilibrium in adversarial team games is CLS-complete. (The case where $n = 1$ amounts to two-player zero-sum games, known to be in P.)

Theorem 1.2. *Computing an ϵ -Nash equilibrium in 3-player (that is, 2 vs. 1) adversarial team games is CLS-complete.*

Coupled with earlier results, [Theorem 1.2](#) completely characterizes the complexity landscape for computing Nash equilibria in adversarial team games.

Our proof is based on a recent hardness result of [Ghosh and Hollender \[2024\]](#) (cf. [Tewolde et al. \[2025\]](#)), who proved that computing a *symmetric* ϵ -Nash equilibrium in a symmetric two-player game with identical payoffs is CLS-complete. The key idea in our reduction is that one can leverage the adversarial player so as to enforce symmetry between the team players, without affecting the equilibria of the original game; the basic gadget underpinning this reduction is analyzed in [Section 3.1](#).

Incidentally, our CLS-hardness reduction hinges on a *polymatrix* adversarial team game, thereby addressing another open question left recently by [Hollender et al. \[2025\]](#).

Theorem 1.3. *[Theorem 1.2](#) holds even when one restricts to polymatrix, 3-player adversarial team games.*

We complement the above hardness result by further characterizing the complexity of deciding whether an adversarial team game admits a unique (approximate) Nash equilibrium ([Theorem 3.5](#)).

Symmetric min-max optimization As we have seen, symmetry plays a key role in the proof of [Theorem 1.2](#), but that result places no restrictions on whether the equilibrium is symmetric or not—this is indeed the crux of the argument. The next problem we consider concerns computing *symmetric* equilibria in *symmetric* min-max optimization problems, in the following natural sense.

Definition 1.4 (Symmetric min-max optimization). A function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ is called *antisymmetric* if $\mathcal{X} = \mathcal{Y}$ and

$$f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x}) \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}.$$

Furthermore, a point $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$ is called *symmetric* if $\mathbf{x} = \mathbf{y}$. The associated min-max optimization problem is called symmetric if the underlying function f is *antisymmetric*.²

Symmetric zero-sum games are ubiquitous in the literature and in practical applications alike. Many popular recreational games used for AI benchmarking, such as poker and battleship, are symmetric when roles are assigned at random; the symmetry assumption is particularly natural as it ensures that no player has an *a priori* advantage before the game begins.

The study of symmetric equilibria has a long history in the development of game theory, propelled by Nash’s pathbreaking PhD thesis [[Nash, 1950](#)] (cf. [Gale et al. \[1951\]](#)), and has remained a popular research topic ever since [[Tewolde et al., 2025](#), [Emmons et al., 2022](#), [Garg et al., 2018](#), [Ghosh and Hollender, 2024](#), [Mehta, 2014](#)]. Classic examples in game theory, including rock–paper–scissors and matching pennies, are also symmetric; these games were already discussed in the original work of [von Neumann \[1928\]](#).

It is not hard to see that symmetric min-max optimization problems, in the sense of [Definition 1.4](#), always admit *symmetric* first-order Nash equilibria. What is more, we show that computing such a symmetric equilibrium is in the complexity class PPAD [[Papadimitriou, 1994](#)]; this is based on an argument of [Etessami and Yannakakis \[2010\]](#), and complements [Daskalakis et al. \[2021\]](#), who proved that the problem of computing approximate fixed points of gradient descent/ascent—which they refer to as GDAFIXEDPOINT—lies in PPAD. In a celebrated series of work, it was shown that PPAD captures the complexity of computing Nash equilibria in finite games [[Daskalakis et al., 2009](#), [Chen et al., 2009](#)]. In this context, we establish that PPAD also characterizes the complexity of computing symmetric first-order Nash equilibria in symmetric min-max optimization problems:

Theorem 1.5. *Computing a symmetric $1/n^c$ -approximate first-order Nash equilibrium in symmetric n -dimensional min-max optimization is PPAD-complete for any constant $c > 0$.*

²The nomenclature of this definition is consistent with the usual terminology in the context of (two-player) zero-sum games: a symmetric zero-sum game is one in which the underlying game matrix $\mathbf{A}$ is antisymmetric (that is, skew-symmetric), so that $\langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle = -\langle \mathbf{y}, \mathbf{A}\mathbf{x} \rangle$ for all $(\mathbf{x}, \mathbf{y})$.

Barring major complexity breakthroughs, [Theorem 1.5](#) precludes the existence of algorithms with complexity polynomial in the dimension and $1/\epsilon$, where $\epsilon > 0$ measures the precision (per [Definition 1.1](#)), under the symmetry constraint of [Definition 1.4](#). This stands in contrast to (nonconvex) minimization problems, wherein gradient descent converges to stationary points at a rate of $\text{poly}(1/\epsilon)$; even in the regime where $\epsilon = 1/\exp(n)$, computing a stationary point of a smooth function is in CLS [[Daskalakis and Papadimitriou, 2011](#)], which is a subclass of PPAD [[Fearnley et al., 2023](#)]. In fact, our reduction also rules out the existence of polynomial-time algorithms even when $\epsilon = \Theta(1)$ under some well-believed complexity assumptions ([Corollary 4.3](#)).

The proof of [Theorem 1.5](#) is elementary, and is based on the PPAD-hardness of computing *symmetric* Nash equilibria in *symmetric* two-player games. Importantly, our reduction gives an immediate, and significantly simpler, proof ([Theorem 4.4](#)) of the PPAD-hardness result of [Daskalakis et al. \[2021\]](#), while being applicable even with respect to quadratic and anti-symmetric functions defined on a product of simplexes.

Independent and concurrent work [Bernasconi et al. \[2024\]](#) also considerably simplified the proof of [Daskalakis et al. \[2021\]](#). Our hardness result hinges on the intermediate problem of finding a Nash equilibrium in symmetric two-player games [[Chen et al., 2009](#)], whereas [Bernasconi et al. \[2024\]](#) showed their hardness result via the problem of finding a Nash equilibrium in (multi-player) polymatrix two-action games. The main qualitative difference between the two is that ours applies to simplex domains while the result of [Bernasconi et al. \[2024\]](#) to box domains. The basic idea of both reductions then is that one can enforce the symmetry constraint $\mathbf{x} \approx \mathbf{y}$ via coupled constraints.

As a byproduct of [Theorem 1.5](#) and the result of [Bernasconi et al. \[2024\]](#), it follows that any *symmetric* dynamics—whereby both players follow the same online algorithm, as formalized in [Definition 4.5](#)—cannot converge to a first-order Nash equilibrium in polynomial time, subject to $\text{PPAD} \neq \text{P}$ ([Theorem 4.6](#)). This already captures many natural dynamics for which prior papers in the literature (e.g., [Kalogiannis et al. \[2023b\]](#)) have painstakingly shown lack of convergence; [Theorem 4.6](#) provides a complexity-theoretic justification for such prior results, while precluding a much broader family of algorithms.

The complexity of non-symmetric equilibria Remaining on symmetric min-max optimization, one natural question arising from [Theorem 1.5](#) concerns the complexity of *non-symmetric* equilibria—defined as having distance at least $\delta > 0$. Unlike their symmetric counterparts, non-symmetric first-order Nash equilibria are not guaranteed to exist. In fact, we establish the following result.

Theorem 1.6. *For a symmetric min-max optimization problem, constants $c_1, c_2 > 0$, and $\epsilon = n^{-c_1}$, it is NP-hard to distinguish between the following two cases under the promise that one of them holds:*

- any ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies $\|\mathbf{x}^* - \mathbf{y}^*\| \leq n^{-c_2}$, and
- there is an ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ such that $\|\mathbf{x}^* - \mathbf{y}^*\| \geq \Omega(1)$.

The main technical piece is [Theorem 4.7](#), which concerns symmetric, identical-interest, two-player games. It significantly refines the hardness result of [McLennan and Tourky \[2010\]](#) by accounting even for $\text{poly}(1/n)$ -Nash equilibria.

Team zero-sum games Finally, building on the reduction of [Theorem 1.5](#) coupled with the gadget behind [Theorem 1.2](#), we establish similar complexity results for team zero-sum games, which generalize adversarial team games by allowing the presence of multiple adversaries. In particular, a *symmetric* two-team zero-sum game and a *symmetric* equilibrium thereof are in accordance with [Definition 1.4](#)—no symmetry constraints are imposed within the same team, but only across teams. We obtain a result significantly refining [Theorem 1.5](#).

Theorem 1.7. *Computing a symmetric $1/n^c$ -Nash equilibrium in symmetric, 6-player (3 vs. 3) team zero-sum polymatrix games is PPAD-complete for some constant $c > 0$.*

Unlike our reduction in [Theorem 1.5](#) that comprises quadratic terms, the crux in team zero-sum games is that one needs to employ solely multilinear terms. The basic idea is to again use the gadget underpinning [Theorem 1.2](#), which enforces symmetry without affecting the equilibria of the game, thereby (approximately) reproducing the objective function that establishes [Theorem 1.5](#).

It is interesting to note that the class of polymatrix games we construct to prove [Theorem 1.7](#) belongs to a certain family introduced by [Cai and Daskalakis \[2011\]](#): one can partition the players into 2

groups so that any pairwise interaction between players of the same group is a coordination game, whereas any pairwise interaction across groups is a zero-sum game. Cai and Daskalakis [2011] showed that computing a Nash equilibrium is PPAD-hard in the more general case where there are 3 groups of players. While the complexity of that problem under 2 groups remains wide open, Theorem 1.7 shows PPAD-hardness for computing *symmetric* Nash equilibria in such games.

Taken together, our results bring us closer to characterizing the complexity of computing equilibria in min-max optimization.

1.2 Further related work

Adversarial team games have been the subject of much research tracing back to the influential work of von Stengel and Koller [1997], who introduced the concept of a *team maxmin equilibrium (TME)*; a TME can be viewed as the best Nash equilibrium for the team. Notwithstanding its intrinsic appeal, it turns out that computing a TME is FNP-hard [Borgs et al., 2010]. Indeed, unlike two-player zero-sum games, team zero-sum games generally exhibit a *duality gap*—characterized in the work of Schulman and Vazirani [2017].

This realization has shifted the focus of contemporary research to exploring more permissive solution concepts. One popular such relaxation is *TMECor*, which enables team players to *ex ante* correlate their strategies [Zhang et al., 2023, 2021, Basilico et al., 2017, Carminati et al., 2022, Farina et al., 2018, Zhang and An, 2020, Celli and Gatti, 2018]. Yet, in the context of extensive-form games, computing a TMECor remains intractable; Zhang et al. [2023] provided an exact characterization of its complexity. Team zero-sum games can be thought of as two-player zero-sum games but with *imperfect recall*, and many natural problems immediately become hard without perfect recall (e.g., Tewolde et al. [2023]). Parameterized algorithms have been developed for computing a TMECor based on some natural measure of shared information [Zhang et al., 2023, Carminati et al., 2022]. Beyond adversarial team games, Carminati et al. [2023] recently explored *hidden-role* games, wherein there is uncertainty regarding which players belong in the same team, a feature that often manifests itself in popular recreational games—and used certain cryptographic primitives to solve them.

In contrast, this paper focuses on the usual Nash equilibrium concept, being thereby orthogonal to the above line of work. One drawback of Nash equilibria in adversarial team games is that the (worst-case) value of the team can be significantly lower compared to TME [Basilico et al., 2017]. On the other hand, Anagnostides et al. [2023] showed that ϵ -Nash equilibria in adversarial team games admit an FPTAS, which stands in stark contrast to TME, and indeed, Nash equilibria in general games [Daskalakis et al., 2009, Chen et al., 2009]. This was further strengthened by Kalogiannis et al. [2023a, 2024] for computing ϵ -Nash equilibria in adversarial team *Markov* games—the natural generalization to Markov (aka. stochastic) games.

Related to Definition 1.1 is the natural notion of a *local* min-max equilibrium [Daskalakis and Panageas, 2018, Daskalakis et al., 2021]. It is easy to see that any local min-max equilibrium—with respect to a sufficiently large neighborhood of (x^*, y^*) —must satisfy Definition 1.1 [Daskalakis et al., 2021]. Unlike first-order Nash equilibria, local min-max equilibria are not guaranteed to exist.

Finally, Mehta et al. [2015] showed that in two-player symmetric games, deciding whether a non-symmetric Nash equilibrium exists is NP-hard, which directly relates to our Theorem 1.6.

2 Preliminaries

Notation We use boldface lowercase letters, such as $\mathbf{x}, \mathbf{y}, \mathbf{z}$, to represent vectors, and boldface capital letters, such as $\mathbf{A}, \mathbf{C}$, for matrices. We denote by x_i the i th coordinate of a vector $\mathbf{x} \in \mathbb{R}^n$. We use the shorthand notation $[n] := \{1, 2, \dots, n\}$. $\Delta^n := \{\mathbf{x} \in \mathbb{R}_{\geq 0}^n : \sum_{i=1}^n x_i = 1\}$ is the probability simplex on $\mathbb{R}^n$. For $i \in [n]$, $\mathbf{e}_i \in \Delta^n$ is the i th unit vector. $\langle \cdot, \cdot \rangle$ denotes the inner product. For a vector $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is its Euclidean norm. For $m \leq n$, $\mathbf{x}_{[1..m]} \in \mathbb{R}^m$ is the vector containing the first m coordinates of $\mathbf{x}$. We sometimes use the $O(\cdot)$, $\Theta(\cdot)$, $\Omega(\cdot)$ notation to suppress absolute constants. A continuously differentiable function f is L -smooth if its gradient is L -Lipschitz continuous with respect to $\|\cdot\|_2$; that is, $\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{x}')\|_2 \leq L\|\mathbf{x} - \mathbf{x}'\|_2$ for all $\mathbf{x}, \mathbf{x}'$.

Two-player games In a two-player game, represented in normal-form game, each player has a finite set, let $[n]$, of actions. Under a pair of actions $(i, j) \in [n] \times [n]$, the utility of the *row* player is given by $\mathbf{R}_{i,j}$, where $\mathbf{R} \in \mathbb{Q}^{n \times n}$ is the payoff matrix of the row player. Further, we let $\mathbf{C} \in \mathbb{Q}^{n \times n}$ be the payoff matrix of the column player. Players are allowed to randomize by selecting mixed strategies—points in Δ^n . Under a pair of mixed strategies $(\mathbf{x}, \mathbf{y}) \in \Delta^n \times \Delta^n$, the expected utility of the players is given by $\langle \mathbf{x}, \mathbf{R}\mathbf{y} \rangle$ and $\langle \mathbf{x}, \mathbf{C}\mathbf{y} \rangle$, respectively. The canonical solution concept in such games is the *Nash equilibrium* [Nash, 1951], which is recalled below.

Definition 2.1. A pair of strategies $(\mathbf{x}^*, \mathbf{y}^*)$ is an ϵ -Nash equilibrium of $(\mathbf{R}, \mathbf{C})$ if

$$\langle \mathbf{x}^*, \mathbf{R}\mathbf{y}^* \rangle \geq \langle \mathbf{x}, \mathbf{R}\mathbf{y}^* \rangle - \epsilon \quad \text{and} \quad \langle \mathbf{x}^*, \mathbf{C}\mathbf{y}^* \rangle \geq \langle \mathbf{x}^*, \mathbf{C}\mathbf{y} \rangle - \epsilon \quad \forall (\mathbf{x}, \mathbf{y}) \in \Delta^n \times \Delta^n.$$

Symmetric two-player games One of our reductions is based on *symmetric* two-player games, meaning that $\mathbf{R} = \mathbf{C}^\top$. A basic fact is that any symmetric game admits a *symmetric* Nash equilibrium $(\mathbf{x}^*, \mathbf{x}^*)$. Further, computing a Nash equilibrium in a general game can be reduced to computing a symmetric Nash equilibrium in a symmetric game [Nisan et al., 2007, Theorem 2.4]. In conjunction with the hardness result of Chen et al. [2009], we state the following consequence.

Theorem 2.2 (Chen et al., 2009). *Computing a symmetric $1/n^c$ -Nash equilibrium in a symmetric two-player game is PPAD-hard for any constant $c > 0$.*

Team zero-sum games A (two-)team zero-sum game is a multi-player game—represented in normal form for the purposes of this paper—in which the players’ utilities have a certain structure; namely, we can partition the players into two (disjoint) subsets, such that each player within the same team shares the same utility, whereas players in different teams have opposite utilities—under any possible combination of actions. An *adversarial* team game is a specific type of team zero-sum game wherein one team consists of a single player. As in Definition 2.1 for two-player games, an ϵ -Nash equilibrium is a tuple of strategies such that no unilateral deviation yields more than an ϵ additive improvement in the utility of the deviator.

3 Complexity of adversarial team games

We begin by examining equilibrium computation in adversarial team games.

3.1 CLS-completeness for 3-player games

Computing ϵ -Nash equilibria in adversarial team games was placed in CLS by Anagnostides et al. [2023], but whether CLS tightly characterizes the complexity of that problem remained open—that was only known when the number of players is large, so that the hardness result of Babichenko and Rubinstein [2021] can kick in. Our reduction here answers this question in the affirmative.

We rely on a recent hardness result of Ghosh and Hollender [2024] concerning symmetric, two-player games with identical payoffs. We summarize their main result below.

Theorem 3.1 (Ghosh and Hollender, 2024). *Computing an ϵ -Nash equilibrium in a symmetric, identical-payoffs, two-player game is CLS-complete.*

Now, let $\mathbf{A} \in \mathbb{Q}^{n \times n}$ be the common payoff matrix of a two-player game, which satisfies $\mathbf{A} = \mathbf{A}^\top$ so that the game is symmetric. Without loss of generality, we will assume that $\mathbf{A}_{i,j} \leq -1$ for all $i, j \in [n]$. We denote by $\mathbf{A}_{\min}$ and $\mathbf{A}_{\max}$ the minimum and maximum entry of $\mathbf{A}$, respectively (which satisfy $\mathbf{A}_{\max}, \mathbf{A}_{\min} \leq -1$). The basic idea of our proof is to suitably use the adversarial player so as to force the other two players to play roughly the same strategy (Lemma 3.2), while (approximately) maintaining the structure of the game (Lemma 3.3). The formal proofs are in Section A.1.

Definition of the adversarial team game Based on $\mathbf{A}$, we construct a 3-player adversarial team game as follows. The utility function of the adversary reads

$$u(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \langle \mathbf{x}, \mathbf{A}\mathbf{y} \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \sum_{i=1}^n (z_i(x_i - y_i) + z_{n+i}(y_i - x_i)) + z_{2n+1}|\mathbf{A}_{\min}|. \quad (2)$$

The adversary selects a strategy $\mathbf{z} \in \Delta^{2n+1}$, while the team players, who endeavor to minimize (2), select strategies $\mathbf{x} \in \Delta^n$ and $\mathbf{y} \in \Delta^n$, respectively. (While the range of the utilities in (2) grows with $1/\epsilon$, normalizing to $[-1, 1]$ maintains all of the consequences by suitably adjusting the approximation.)

The first important lemma establishes that, in equilibrium, $\mathbf{x} \approx \mathbf{y}$. The basic argument proceeds as follows. By construction of (2), the adversary would be able to secure a large payoff whenever there is a coordinate $i \in [n]$ such that $|x_i - y_i| \gg 0$ —by virtue of the second term in (2). But that cannot happen in equilibrium, for Player $\mathbf{x}$ (or symmetrically Player $\mathbf{y}$) can simply neutralize that term in the adversary’s utility by playing $\mathbf{x} = \mathbf{y}$.

Lemma 3.2 (Equilibrium forces symmetry). *Consider an ϵ^2 -Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ of the adversarial team game (2) with $\epsilon^2 \leq 1/2$. Then, $\|\mathbf{x}^* - \mathbf{y}^*\|_\infty \leq 2\epsilon$.*

Having established that $\mathbf{x} \approx \mathbf{y}$, the next step is to make sure that the adversarial player does not distort the original game by much. In particular, we need to make sure that the effect of the second term in (2) is negligible. We do so by showing that $z_{2n+1} \approx 1$ (Lemma 3.3).

The argument here is more subtle; roughly speaking, it goes as follows. Suppose that $z_i \gg 0$ or $z_{n+i} \gg 0$ for some $i \in [n]$. Since Player $\mathbf{z}$ is approximately best responding, it would then follow that $|y_i^* - x_i^*| \gg 0$ —otherwise Player $\mathbf{z}$ would prefer to switch to action $2n+1$. But, if $|y_i^* - x_i^*| \gg 0$, Player $\mathbf{x}$ could profitably deviate by reallocating probability mass by either removing from or adding to i (depending on whether $y_i^* - x_i^* > 0$), which leads to a contradiction.

Lemma 3.3 (Most probability mass in a_{2n+1}). *Given any ϵ^2 -Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ of the adversarial team game (2) with $\epsilon \leq 1/10$, $z_j \geq 9\epsilon$ for all $j \in [2n]$. In particular, $z_{2n+1} \geq 1 - 18n\epsilon$.*

By combining Lemmas 3.2 and 3.3, we can complete the reduction from symmetric two-player games with common payoffs to 3-player adversarial team games, as stated below.

Theorem 3.4. *Given any ϵ^2 -Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ in the adversarial team game (2), with $\epsilon \leq 1/10$, $(\mathbf{y}^*, \mathbf{y}^*)$ is a symmetric $(21n+1)|\mathbf{A}_{\min}|\epsilon$ -Nash equilibrium of the symmetric, two-player game $(\mathbf{A}, \mathbf{A})$ (that is, $\mathbf{A} = \mathbf{A}^\top$).*

3.2 The complexity of determining uniqueness

Another natural question concerns the complexity of determining whether an adversarial team game admits a unique Nash equilibrium. Our next theorem establishes NP-hardness for a version of that problem that accounts for approximate Nash equilibria.

Theorem 3.5. *For polymatrix, 3-player adversarial team games, constants $c_1, c_2 > 0$, and $\epsilon = n^{-c_1}$, it is NP-hard to distinguish between the following two cases under the promise that one of them holds:*

- any two ϵ -Nash equilibria have ℓ_1 -distance at most n^{-c_2} , and
- there are two ϵ -Nash equilibria that have ℓ_1 -distance $\Omega(1)$.

We will discuss more about the proof of this theorem later in Section 4.2 when we examine the complexity of computing non-symmetric equilibria in symmetric min-max optimization problems. It is also interesting to point out that an adversarial team game can have a unique Nash equilibrium supported on *irrational* numbers, as we show in Section A.2.

4 Complexity of equilibria in symmetric min-max optimization

This section characterizes the complexity of computing symmetric first-order Nash equilibria (Definition 1.1) in symmetric min-max optimization problems in the sense of Definition 1.4; namely, when $f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$ for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$ and $\mathcal{X} = \mathcal{Y}$.

4.1 Problem definitions and hardness results for symmetric equilibria

Given a continuously differentiable function $f : \mathcal{D} \rightarrow \mathbb{R}$, we set $F_{\text{GDA}} : \mathcal{D} \rightarrow \mathcal{D}$ to be

$$F_{\text{GDA}}(\mathbf{x}, \mathbf{y}) := \prod_{\mathcal{D}} [\mathbf{x} - \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}), \mathbf{y} + \nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})] \text{ for } (\mathbf{x}, \mathbf{y}) \in \mathcal{D},$$

the norm of which measures the fixed-point gap and corresponds to the update rule of GDA with stepsize equal to one; we recall that Player $\mathbf{x}$ is the minimizer, while Player $\mathbf{y}$ is the maximizer. The domain $\mathcal{D}$ is a compact subset of $\mathbb{R}^d$ for some $d \in \mathbb{N}$. Moreover, the projection operator $\prod$ is applied

jointly on $\mathcal{D}$.³ When $\mathcal{D}$ can be expressed as a *Cartesian* product $\mathcal{X} \times \mathcal{Y}$, the domain set is called *uncoupled* (and the projection can be done independently), otherwise it is called *coupled* (or *joint*).

We begin by introducing the problem of computing fixed points of gradient descent/ascent (GDA) for domains expressed as the Cartesian product of polytopes, modifying the computational problem GDAFIXEDPOINT introduced by Daskalakis et al. [2021].

GDAFIXEDPOINT Problem.

INPUT:

- Precision parameter $\epsilon > 0$ and smoothness parameter L ,
- Polynomial-time Turing machine $\mathcal{C}_f$ evaluating a L -smooth function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ and its gradient $\nabla f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^d$, where $\mathcal{X} = \{\mathbf{x} : \mathbf{A}_x \mathbf{x} \leq \mathbf{b}_x\}$ and $\mathcal{Y} = \{\mathbf{y} : \mathbf{A}_y \mathbf{y} \leq \mathbf{b}_y\}$ are nonempty, bounded polytopes described by input matrices $\mathbf{A}_x \in \mathbb{R}^{m_x \times d_x}$, $\mathbf{A}_y \in \mathbb{R}^{m_y \times d_y}$ and vectors $\mathbf{b}_x \in \mathbb{R}^{m_x}$, $\mathbf{b}_y \in \mathbb{R}^{m_y}$, with $d := d_x + d_y$.

OUTPUT: A point $(\mathbf{x}^*, \mathbf{y}^*) \in \mathcal{X} \times \mathcal{Y}$ such that $\|(\mathbf{x}^*, \mathbf{y}^*) - F_{GDA}(\mathbf{x}^*, \mathbf{y}^*)\|_2 \leq \epsilon$.

Based on GDAFIXEDPOINT, we introduce the problem SYMGDAFIXEDPOINT, which captures the problem of computing *symmetric* (approximate) fixed points of GDA for symmetric min-max optimization problems. We define our computational problems as promise problems.

SYMGDAFIXEDPOINT Problem.

INPUT:

- Precision parameter $\epsilon > 0$ and smoothness parameter L ,
- Polynomial-time Turing machine $\mathcal{C}_f$ evaluating a L -smooth, antisymmetric function $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ and its gradient $\nabla f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^{2d}$, where $\mathcal{X} = \{\mathbf{x} : \mathbf{A} \mathbf{x} \leq \mathbf{b}\}$ is a nonempty, bounded polytope described by an input matrix $\mathbf{A} \in \mathbb{R}^{m \times d}$ and vector $\mathbf{b} \in \mathbb{R}^m$.

OUTPUT: A point $(\mathbf{x}^*, \mathbf{x}^*) \in \mathcal{X} \times \mathcal{X}$ such that $\|(\mathbf{x}^*, \mathbf{x}^*) - F_{GDA}(\mathbf{x}^*, \mathbf{x}^*)\|_2 \leq \epsilon$.

We start by showing that SYMGDAFIXEDPOINT also lies in PPAD; the fact that GDAFIXEDPOINT is in PPAD—even under coupled domains—was shown to be the case by Daskalakis et al. [2021]. The detailed proof is included in the appendix.

Lemma 4.1. SYMGDAFIXEDPOINT is a total search problem and lies in PPAD.

Having established that SYMGDAFIXEDPOINT belongs in PPAD, we now state the first main hardness result of this section.

Theorem 4.2 (Complexity for symmetric equilibrium). SYMGDAFIXEDPOINT is PPAD-complete, even for quadratic functions.

The basic idea of the proof is to consider the objective

$$f(\mathbf{x}, \mathbf{y}) := \frac{1}{2} \langle \mathbf{y}, \mathbf{A} \mathbf{y} \rangle - \frac{1}{2} \langle \mathbf{x}, \mathbf{A} \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{C} \mathbf{x} \rangle, \quad (3)$$

where $\mathbf{A}$ is symmetric and $\mathbf{C}$ is skew-symmetric. Theorem 4.2 then follows from some elementary calculations, as we show in Section A.3.

For symmetric first-order Nash equilibria, our argument establishes PPAD-hardness for any $\epsilon \leq 1/n^c$, where $c > 0$ (as claimed in Theorem 1.5). Moreover, leveraging the hardness result of Rubinfeld [2016], we can also immediately obtain constant inapproximability under the so-called *exponential-time hypothesis* (ETH) for PPAD—which postulates that any algorithm for solving END OF A LINE, the prototypical PPAD-complete problem, requires $2^{\tilde{\Omega}(n)}$ time.

Corollary 4.3. Computing an $\Theta(1)$ -approximate first-order Nash equilibrium in symmetric n -dimensional min-max optimization requires $n^{\tilde{\Omega}(\log n)}$ time, assuming ETH for PPAD.

³This is the “safe” version of GDA because it ensures that the mapping always lies in $\mathcal{D}$. One could also project independently on $\mathcal{D}(\mathbf{y}) = \{\mathbf{x}' : (\mathbf{x}', \mathbf{y}) \in \mathcal{D}\}$ and $\mathcal{D}(\mathbf{x}) = \{\mathbf{y}' : (\mathbf{x}, \mathbf{y}') \in \mathcal{D}\}$; see Daskalakis et al. [2021] for further details and the polynomial equivalence for finding fixed points for both versions.

The argument of [Theorem 4.2](#) can be slightly modified to imply the main result of [Daskalakis et al. \[2021\]](#)—with simplex instead of box constraints—as stated below.

Theorem 4.4 (PPAD-hardness for coupled domains). *The problem GDAXEDPOINT is PPAD-hard when the domain is a joint polytope, even for quadratic functions.*

The main idea is to add coupled constraints in order to *force symmetry*: $-\delta \leq x_i - y_i \leq \delta$ for all $i \in [n]$, where, if ϵ is the approximation accuracy, δ is of order $\Theta(\epsilon^{1/4})$. Compared to the equilibrium studied in [Daskalakis et al. \[2021\]](#), the symmetric equilibrium considered in our work is stronger in that it accounts for all deviations, not merely ones on the coupled feasibility set. We present the proof of [Theorem 4.4](#) in [Section A.3](#).

Hardness results for symmetric dynamics Another interesting consequence of [Theorem 4.2](#) is that it precludes convergence under a broad class of algorithms in general min-max optimization.

Definition 4.5 (Symmetric learning algorithms for min-max). Let $T \in \mathbb{N}$. A deterministic, polynomial-time learning algorithm $\mathcal{A}$ proceeds as follows for any time $t \in [T]$. It outputs a strategy as a function of the history $\mathcal{H}^{(t)}$ it has observed so far (where $\mathcal{H}^{(1)} := \emptyset$), and then receives as feedback $\mathbf{g}^{(t)}$. It then updates $\mathcal{H}^{(t+1)} := (\mathcal{H}^{(t)}, \mathbf{g}^{(t)})$. A *symmetric* learning algorithm in min-max optimization consists of Player $\mathbf{x}$ employing algorithm $\mathcal{A}$ with history $\mathcal{H}_x^{(t)} := (\nabla_{\mathbf{x}} f(\mathbf{x}^{(t)}, \mathbf{y}^{(t)}))_{t=1}^T$, and Player $\mathbf{y}$ employing the *same* algorithm with history $\mathcal{H}_y^{(t)} := (-\nabla_{\mathbf{y}} f(\mathbf{x}^{(t)}, \mathbf{y}^{(t)}))_{t=1}^T$.

Note that a consequence of the above definition is that both players initialize from the same strategy. Many natural and well-studied algorithms in min-max optimization adhere to [Definition 4.5](#). Besides the obvious example of gradient descent/ascent, we mention extragradient descent/ascent, optimistic gradient descent/ascent, and optimistic multiplicative weights—all assumed to be executed simultaneously. A simple non-example is *alternating* gradient descent/ascent [[Wibisono et al., 2022](#), [Bailey et al., 2020](#)], wherein players do not update their strategies simultaneously.

Theorem 4.6. *No symmetric learning algorithm (per [Definition 4.5](#)) can converge to ϵ -first-order Nash equilibria in min-max optimization in polynomial time when $\epsilon = 1/n^c$, unless $\text{PPAD} = \text{P}$.*

This is a consequence of our argument in [Theorem 4.2](#): under [Definition 4.5](#) and the min-max optimization problem (3), it follows inductively that $\mathbf{x}^{(t)} = \mathbf{y}^{(t)}$ and $\mathcal{H}_x^{(t)} = \mathcal{H}_y^{(t)}$ for all $t \in [T]$. But computing a symmetric first-order Nash equilibrium is PPAD-hard when $\epsilon = 1/n^c$ ([Theorem 4.2](#)).

Assuming that $\text{P} \neq \text{PPAD}$, [Theorem 4.6](#), and in particular its instantiation in team zero-sum games ([Theorem 1.7](#)), significantly generalizes some impossibility results shown by [Kalogiannis et al. \[2023b\]](#) concerning certain algorithms, such as optimistic gradient descent/ascent—our hardness result goes much further, precluding any algorithm subject to [Definition 4.5](#), albeit being conditional.

4.2 The complexity of non-symmetric fixed points

An immediate question raised by [Theorem 4.2](#) concerns the computational complexity of finding *non-symmetric* fixed points of GDA for symmetric min-max optimization problems. Since totality is not guaranteed, unlike SYMGDAFIXEDPOINT, we cannot hope to prove membership in PPAD. In fact, we show that finding a non-symmetric fixed point of GDA is FNP-hard. To do so, we first define formally the computational problem of interest.

NONSYMGDAFIXEDPOINT Problem.

INPUT:

- Parameters $\epsilon, \delta > 0$ and Lipschitz constant L and
- Polynomial-time Turing machine $\mathcal{C}_f$ evaluating a L -smooth antisymmetric function $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ and its gradient $\nabla f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^{2d}$, where $\mathcal{X} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is a nonempty, bounded polytope described by a matrix $\mathbf{A} \in \mathbb{R}^{m \times d}$ and vector $\mathbf{b} \in \mathbb{R}^m$.

OUTPUT: A point $(\mathbf{x}^*, \mathbf{y}^*) \in \mathcal{X} \times \mathcal{X}$ such that $\|\mathbf{x}^* - \mathbf{y}^*\|_2 \geq \delta$ and $\|(\mathbf{x}^*, \mathbf{y}^*) - F_{\text{GDA}}(\mathbf{x}^*, \mathbf{y}^*)\|_2 \leq \epsilon$ if it exists, otherwise return NO.

We establish that NONSYMGDAFIXEDPOINT is FNP-hard. Our reduction builds on the hardness result of [McLennan and Tourky \[2010\]](#)—in turn based on earlier work by [Gilboa and Zemel \[1989\]](#),

Conitzer and Sandholm [2008]—which we significantly refine in order to account for $\text{poly}(1/n)$ -Nash equilibria. Our result, which forms the basis for Theorem 1.6 and Theorem 3.5, is summarized below.

Theorem 4.7. *For symmetric, identical-interest, two-player games, constants $c_1, c_2 > 0$, and $\epsilon = n^{-c_1}$, it is NP-hard to distinguish between the following two cases under the promise that one of them holds:*

- any two symmetric ϵ -Nash equilibria have ℓ_1 -distance at most n^{-c_2} , and
- there are two symmetric ϵ -Nash equilibria that have ℓ_1 -distance $\Omega(1)$.

The proof of Theorem 1.6 now follows by considering the antisymmetric function $f(x, y) := y^\top B y - x^\top B x$ for a suitable matrix B (defined per the hard instance from Theorem 4.7 based on k -clique). FNP-hardness follows similarly by considering a search version of maximum clique.

Finally, the proof of Theorem 3.5 that was claimed earlier follows immediately by combining Theorem 4.7 with the reduction of Section 3.1, and in particular, Lemmas 3.2 and 3.3.

4.3 Team zero-sum games

Our previous hardness result concerning symmetric min-max optimization problems does not have any immediate implications for (normal-form) team zero-sum games since the class of hard instances we constructed earlier contains a quadratic term. Our next result provides such a hardness result by combining the basic gadget we introduced in Section 3.1 in the context of adversarial team games; the basic pieces of the argument are similar to the ones we described in Section 3.1, and so the proof is deferred to Section A.5. Our goal is to prove the following.

Theorem 1.7. *Computing a symmetric $1/n^c$ -Nash equilibrium in symmetric, 6-player (3 vs. 3) team zero-sum polymatrix games is PPAD-complete for some constant $c > 0$.*

Let us describe the class of 3 vs. 3 team zero-sum games upon which our hardness result is based on. Based on (2), we define the auxiliary function

$$\delta : \Delta^n \times \Delta^n \times \Delta^{2n+1} \ni (x, y, z) \mapsto \frac{|\mathbf{A}_{\min}|}{\epsilon} \sum_{i=1}^n (z_i(x_i - y_i) + z_{n+i}(y_i - x_i)) + |\mathbf{A}_{\min}| z_{2n+1}.$$

In what follows, the 3 players of the one team will be identified with (x, y, z) , while the 3 players of the other team with $(\hat{x}, \hat{y}, \hat{z})$. We define the utility of the latter team to be

$$u(x, y, z, \hat{x}, \hat{y}, \hat{z}) = \langle x, \mathbf{A}y \rangle - \langle \hat{x}, \mathbf{A}\hat{y} \rangle + \langle x, \mathbf{C}\hat{x} \rangle + \delta(x, y, z) - \delta(\hat{x}, \hat{y}, \hat{z}), \quad (4)$$

where $\mathbf{A}$ is symmetric and $\mathbf{C}$ is skew-symmetric. The rest of the argument follows Section 3.1.

5 Conclusion and open problems

We have provided a number of new complexity results concerning min-max optimization in general, and team zero-sum games in particular (see Table 1). There are many interesting avenues for future research. The complexity of computing first-order Nash equilibria (equivalently, the GDAFIXEDPOINT problem) remains wide open, but our hardness results suggest a possible approach: as we have seen, in symmetric min-max optimization, computing *either* symmetric or non-symmetric equilibria is intractable, so it would be enough if one could establish this using the same underlying function—that is, somehow combine our two reductions into one. It would also be interesting to see whether our hardness results can be extended to more structured min-max optimization problems, such as adversarial training and GANs.

Acknowledgments

I.P. and J.Y. were supported by NSF grant CCF-2454115. I.P. would like to acknowledge ICS research award and a start-up grant from UCI. Part of this work was done while I.P. and J.Y. were visiting Archimedes Research Unit. T.S. is supported by the Vannevar Bush Faculty Fellowship ONR N00014-23-1-2876, National Science Foundation grants RI-2312342 and RI-1901403, ARO award W911NF2210266, and NIH award A240108S001. We are grateful to Alexandros Hollender for many valuable discussions.

References

- Ioannis Anagnostides, Fivos Kalogiannis, Ioannis Panageas, Emmanouil-Vasileios Vlatakis-Gkaragkounis, and Stephen McAleer. Algorithms and complexity for computing Nash equilibria in adversarial team games. In *Conference on Economics and Computation (EC)*, 2023.
- Yakov Babichenko and Aviad Rubinstein. Settling the complexity of Nash equilibrium in congestion games. In *Symposium on Theory of Computing (STOC)*, 2021.
- Tao Bai, Jinqi Luo, Jun Zhao, Bihan Wen, and Qian Wang. Recent advances in adversarial training for adversarial robustness. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2021.
- Yu Bai and Chi Jin. Provable self-play algorithms for competitive reinforcement learning. In *International Conference on Machine Learning (ICML)*, Proceedings of Machine Learning Research, 2020.
- James P. Bailey, Gauthier Gidel, and Georgios Piliouras. Finite regret and cycles with fixed step-size via alternating gradient descent-ascent. In *Conference on Learning Theory (COLT)*, 2020.
- 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 (AAAI)*, 2017.
- Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Gabriele Farina. On the role of constraints in the complexity of min-max optimization, 2024.
- Marie Louisa Tølbøll Berthelsen and Kristoffer Arnsfelt Hansen. On the computational complexity of decision problems about multi-player Nash equilibria. In Dimitris Fotakis and Evangelos Markakis, editors, *Algorithmic Game Theory*, pages 153–167, Cham, 2019. Springer International Publishing. ISBN 978-3-030-30473-7.
- Christian Borgs, Jennifer T. Chayes, Nicole Immorlica, Adam Tauman Kalai, Vahab S. Mirrokni, and Christos H. Papadimitriou. The myth of the folk theorem. *Games and Economic Behavior*, 70(1): 34–43, 2010.
- Yang Cai and Constantinos Daskalakis. On minmax theorems for multiplayer games. In *Symposium on Discrete Algorithms (SODA)*, 2011.
- Yang Cai, Argyris Oikonomou, and Weiqiang Zheng. Finite-time last-iterate convergence for learning in multi-player games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- Nicholas Carlini, Anish Athalye, Nicolas Papernot, Wieland Brendel, Jonas Rauber, Dimitris Tsipras, Ian J. Goodfellow, Aleksander Madry, and Alexey Kurakin. On evaluating adversarial robustness, 2019.
- Luca Carminati, Federico Cacciamani, Marco Ciccone, and Nicola Gatti. A marriage between adversarial team games and 2-player games: Enabling abstractions, no-regret learning, and subgame solving. In *International Conference on Machine Learning (ICML)*, 2022.
- Luca Carminati, Brian Hu Zhang, Gabriele Farina, Nicola Gatti, and Tuomas Sandholm. Hidden-role games: Equilibrium concepts and computation. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, 2023.
- Andrea Celli and Nicola Gatti. Computational results for extensive-form adversarial team games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.
- Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player Nash equilibria. *J. ACM*, 56(3):14:1–14:57, 2009.
- Sayantan Choudhury, Eduard Gorbunov, and Nicolas Loizou. Single-call stochastic extragradient methods for structured non-monotone variational inequalities: Improved analysis under weaker conditions. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

- Jeremy Cohen, Elan Rosenfeld, and Zico Kolter. Certified adversarial robustness via randomized smoothing. In *International Conference on Machine Learning (ICML)*, pages 1310–1320, 2019.
- Vincent Conitzer and Tuomas Sandholm. New complexity results about Nash equilibria. *Games and Economic Behavior*, 63(2):621–641, 2008.
- Constantinos Daskalakis and Ioannis Panageas. The limit points of (optimistic) gradient descent in min-max optimization. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2018.
- Constantinos Daskalakis and Christos H. Papadimitriou. Continuous local search. In *Symposium on Discrete Algorithms (SODA)*, 2011.
- Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM J. Comput.*, 39(1):195–259, 2009.
- Constantinos Daskalakis, Dylan J Foster, and Noah Golowich. Independent policy gradient methods for competitive reinforcement learning. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2020.
- Constantinos Daskalakis, Stratis Skoulakis, and Manolis Zampetakis. The complexity of constrained min-max optimization. In *Symposium on Theory of Computing (STOC)*, 2021.
- Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, and Stuart Russell. For learning in symmetric teams, local optima are global Nash equilibria. In *International Conference on Machine Learning (ICML)*, 2022.
- Kousha Etessami and Mihalis Yannakakis. On the complexity of Nash equilibria and other fixed points. *SIAM J. Comput.*, 39(6):2531–2597, 2010.
- Gabriele Farina, Andrea Celli, Nicola Gatti, and Tuomas Sandholm. Ex ante coordination and collusion in zero-sum multi-player extensive-form games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2018.
- John Fearnley, Paul Goldberg, Alexandros Hollender, and Rahul Savani. The complexity of gradient descent: $\text{CLS} = \text{PPAD} \cap \text{PLS}$. *J. ACM*, 70(1):7:1–7:74, 2023.
- D. Gale, H. W. Kuhn, and A. W. Tucker. *On Symmetric Games*, pages 81–88. Princeton University Press, 1951.
- Jugal Garg, Ruta Mehta, Vijay V. Vazirani, and Sadra Yazdanbod. $\exists$ r-completeness for decision versions of multi-player (symmetric) Nash equilibria. *ACM Trans. Econ. Comput.*, 6(1), 2018.
- Saeed Ghadimi and Guanghui Lan. Accelerated gradient methods for nonconvex nonlinear and stochastic programming. *Mathematical Programming*, 156(1):59–99, 2016.
- Abheek Ghosh and Alexandros Hollender. The complexity of symmetric bimatrix games with common payoffs. In *Conference on Web and Internet Economics (WINE)*, 2024.
- Itzhak Gilboa and Eitan Zemel. Nash and correlated equilibria: Some complexity considerations. *Games and Economic Behavior*, 1(1):80–93, 1989.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, 2014.
- Eduard Gorbunov, Adrien B. Taylor, and Gauthier Gidel. Last-iterate convergence of optimistic gradient method for monotone variational inequalities. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- Alexandros Hollender, Gilbert Maystre, and Sai Ganesh Nagarajan. The complexity of two-team polymatrix games with independent adversaries. In *International Conference on Learning Representations (ICLR)*, 2025.

- Chi Jin, Praneeth Netrapalli, and Michael I. Jordan. What is local optimality in nonconvex-nonconcave minimax optimization? In *International Conference on Machine Learning (ICML)*, 2020.
- Michael I. Jordan, Tianyi Lin, and Manolis Zampetakis. First-order algorithms for nonlinear generalized Nash equilibrium problems. *Journal of Machine Learning Research*, 24:38:1–38:46, 2023.
- Fivos Kalogiannis, Ioannis Anagnostides, Ioannis Panageas, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Vaggos Chatziafratis, and Stelios Andrew Stavroulakis. Efficiently computing Nash equilibria in adversarial team markov games. In *International Conference on Learning Representations (ICLR)*, 2023a.
- Fivos Kalogiannis, Ioannis Panageas, and Emmanouil-Vasileios Vlatakis-Gkaragkounis. Towards convergence to Nash equilibria in two-team zero-sum games. *ICLR*, 2023b.
- Fivos Kalogiannis, Jingming Yan, and Ioannis Panageas. Learning equilibria in adversarial team markov games: A nonconvex-hidden-concave min-max optimization problem. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- Galina M Korpelevich. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.
- Tianyi Lin, Chi Jin, and Michael I. Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In *International Conference on Machine Learning (ICML)*, 2020.
- Luo Luo, Haishan Ye, Zhichao Huang, and Tong Zhang. Stochastic recursive gradient descent ascent for stochastic nonconvex-strongly-concave minimax problems. In *Advances in Neural Information Processing Systems*, 2020.
- Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations (ICLR)*, 2018.
- Andrew McLennan and Rabee Tourky. Simple complexity from imitation games. *Games and Economic Behavior*, 68(2):683–688, 2010.
- Ruta Mehta. Constant rank bimatrix games are ppad-hard. In *Symposium on Theory of Computing (STOC)*, 2014.
- Ruta Mehta, Vijay V Vazirani, and Sadra Yazdanbod. Settling some open problems on 2-player symmetric Nash equilibria. In *International Symposium on Algorithmic Game Theory*, pages 272–284. Springer, 2015.
- Panayotis Mertikopoulos, Bruno Lecouat, Houssam Zenati, Chuan-Sheng Foo, Vijay Chandrasekhar, and Georgios Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *International Conference on Learning Representations (ICLR)*, 2019.
- John Nash. *Non-cooperative games*. PhD thesis, Princeton University, 1950.
- John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.
- Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, USA, 2007.
- Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D. Lee, and Meisam Razaviyayn. Solving a class of non-convex min-max games using iterative first order methods. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- Idan Orzech and Martin C. Rinard. Correlated vs. uncorrelated randomness in adversarial congestion team games, 2023.
- Dmitrii M. Ostrovskii, Andrew Lowy, and Meisam Razaviyayn. Efficient search of first-order nash equilibria in nonconvex-concave smooth min-max problems. *SIAM J. Optim.*, 31(4):2508–2538, 2021.

- Christos H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences*, 48(3):498–532, 1994.
- Aviad Rubinstein. Settling the complexity of computing approximate two-player Nash equilibria. In *Symposium on Foundations of Computer Science (FOCS)*, 2016.
- Leonard Schulman and Umesh V Vazirani. The duality gap for two-team zero-sum games. In *Innovations in Theoretical Computer Science Conference (ITCS)*, 2017.
- Ke Sun. Some properties of the Nash equilibrium in 2×2 zero-sum games, 2022.
- Emanuel Tewolde, Caspar Oesterheld, Vincent Conitzer, and Paul W. Goldberg. The computational complexity of single-player imperfect-recall games. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2023.
- Emanuel Tewolde, Brian Hu Zhang, Caspar Oesterheld, Tuomas Sandholm, and Vincent Conitzer. Computing game symmetries and equilibria that respect them. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2025.
- Ioannis C. Tsaknakis and Mingyi Hong. Finding first-order Nash equilibria of zero-sum games with the regularized nikaido-isoda function. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2021.
- John von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928.
- Bernhard von Stengel and Daphne Koller. Team-maxmin equilibria. *Games and Economic Behavior*, 21(1):309–321, 1997.
- Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Last-iterate convergence of decentralized optimistic gradient descent/ascent in infinite-horizon competitive markov games. In *Conference on Learning Theory (COLT)*, 2021.
- Andre Wibisono, Molei Tao, and Georgios Piliouras. Alternating mirror descent for constrained min-max games. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2022.
- Zi Xu, Huiling Zhang, Yang Xu, and Guanghui Lan. A unified single-loop alternating gradient projection algorithm for nonconvex-concave and convex-nonconcave minimax problems. *Mathematical Programming*, 201(1):635–706, 2023.
- Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. Team belief DAG: generalizing the sequence form to team games for fast computation of correlated team max-min equilibria via regret minimization. In *International Conference on Machine Learning (ICML)*, 2023.
- Youzhi Zhang and Bo An. Converging to team-maxmin equilibria in zero-sum multiplayer games. In *International Conference on Machine Learning (ICML)*, 2020.
- Youzhi Zhang, Bo An, and Jakub Cerný. Computing ex ante coordinated team-maxmin equilibria in zero-sum multiplayer extensive-form games. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2021.

A Omitted proofs

This section contains the proofs omitted from the main body.

A.1 Proofs from Section 3.1

We begin by stating and proving a simple auxiliary lemma.

Lemma A.1. *Let $(\mathbf{x}_i^*, \mathbf{x}_{-i}^*)$ be an ϵ^2 -Nash equilibrium of a normal-form game, and a_j any action of player i . If $u_i(a_k, \mathbf{x}_{-i}^*) \leq u_i(a_j, \mathbf{x}_{-i}^*) - c$ for some $c > 0$ and $k \in [n]$, then $x_i^*(a_k) \leq \epsilon^2/c$.*

Proof. For the sake of contradiction, suppose that $x_i^*(a_k) > \frac{\epsilon^2}{c}$ for some $k \in [n]$ such that $u_i(a_k, \mathbf{x}_{-i}^*) \leq u_i(a_j, \mathbf{x}_{-i}^*) - c$. Consider the strategy $\Delta^n \ni \mathbf{x}'_i = \mathbf{x}_i^* + x_i^*(a_k)\mathbf{e}_j - x_i^*(a_k)\mathbf{e}_k$. Then, we have

$$\begin{aligned} u_i(\mathbf{x}'_i, \mathbf{x}_{-i}^*) - u_i(\mathbf{x}_i^*, \mathbf{x}_{-i}^*) &= x_i^*(a_k)u_i(a_j, \mathbf{x}_{-i}^*) - x_i^*(a_k)u_i(a_k, \mathbf{x}_{-i}^*) \\ &\geq cx_i^*(a_k) \\ &> \epsilon^2. \end{aligned}$$

That is, deviating to $\mathbf{x}'_i$ yields a utility benefit strictly larger than ϵ^2 , which contradicts the assumption that $(\mathbf{x}_i^*, \mathbf{x}_{-i}^*)$ is an ϵ^2 -Nash equilibrium. $\square$

We move on to the proof of Lemma 3.2.

Proof of Lemma 3.2. For the sake of contradiction, suppose that $x_i^* - y_i^* > 2\epsilon$ for some $i \in [n]$ (the case where $y_i^* - x_i^* > 2\epsilon$ is symmetric, and can be treated analogously). Player z could then choose action a_i (with probability 1), which secures a utility of

$$u(\mathbf{x}^*, \mathbf{y}^*, a_i) = \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \cdot (x_i^* - y_i^*) > \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + 2|\mathbf{A}_{\min}|,$$

since $x_i^* - y_i^* > 2\epsilon$. At the same time, Player z could choose action a_{2n+1} , which secures a utility of $u(\mathbf{x}^*, \mathbf{y}^*, a_{2n+1}) = \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + |\mathbf{A}_{\min}|$. So,

$$u(\mathbf{x}^*, \mathbf{y}^*, a_i) - u(\mathbf{x}^*, \mathbf{y}^*, a_{2n+1}) \geq |\mathbf{A}_{\min}|.$$

Applying Lemma A.1,

$$z_{2n+1}^* \leq \frac{\epsilon^2}{|\mathbf{A}_{\min}|} \leq \epsilon^2. \quad (5)$$

Also, using the fact that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is an ϵ^2 -Nash equilibrium,

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*) &\geq u(\mathbf{x}^*, \mathbf{y}^*, a_i) - \epsilon^2 \\ &\geq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + 2|\mathbf{A}_{\min}| - \epsilon^2 \\ &\geq \mathbf{A}_{\min} + 2|\mathbf{A}_{\min}| - \epsilon^2 \\ &= |\mathbf{A}_{\min}| - \epsilon^2, \end{aligned} \quad (6)$$

since we have assumed that $\mathbf{A}_{\min} < 0$. Now, consider the deviation of Player x (from strategy $\mathbf{x}^*$) to $\mathbf{x}' := \mathbf{y}^*$. Then, $u(\mathbf{x}', \mathbf{y}^*, \mathbf{z}^*) = \langle \mathbf{y}^*, \mathbf{A}\mathbf{y}^* \rangle + z_{2n+1}^* \cdot |\mathbf{A}_{\min}|$. Thus, combining with (6) and (5),

$$\begin{aligned} u(\mathbf{x}', \mathbf{y}^*, \mathbf{z}^*) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*) &\leq \langle \mathbf{y}^*, \mathbf{A}\mathbf{y}^* \rangle + z_{2n+1}^* \cdot |\mathbf{A}_{\min}| - |\mathbf{A}_{\min}| + \epsilon^2 \\ &\leq \langle \mathbf{y}^*, \mathbf{A}\mathbf{y}^* \rangle + \epsilon^2 \cdot |\mathbf{A}_{\min}| - |\mathbf{A}_{\min}| + \epsilon^2 \\ &\leq \mathbf{A}_{\max} - |\mathbf{A}_{\min}| + \epsilon^2(|\mathbf{A}_{\min}| + 1) \\ &\leq -1 < -\epsilon^2, \end{aligned} \quad (7)$$

where we used that $\mathbf{A}_{\max}, \mathbf{A}_{\min} \leq -1$ and $\epsilon^2 \leq 1/2$. But (7) contradicts the fact that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is an ϵ^2 -Nash equilibrium since deviating to $\mathbf{x}'$ yields a utility improvement (equivalently, decrease in cost) strictly larger than ϵ^2 . This completes the proof. $\square$

We proceed with the proof of [Lemma 3.3](#).

Proof of Lemma 3.3. From [Lemma 3.2](#) it holds that $\|\mathbf{x}^* - \mathbf{y}^*\|_\infty \leq 2\epsilon$. Let $i \in [n]$. We assume that i is such that $x_i^* - y_i^* \geq 0$; the contrary case is symmetric. We consider two cases. First, suppose that $|x_i^* - y_i^*| \leq \epsilon/2$. Then, we have

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}^*, a_{2n+1}) - u(\mathbf{x}^*, \mathbf{y}^*, a_i) &\geq |\mathbf{A}_{\min}| - \frac{|\mathbf{A}_{\min}|}{\epsilon}(x_i^* - y_i^*) \\ &\geq \frac{1}{2}|\mathbf{A}_{\min}| \geq \frac{1}{2}. \end{aligned}$$

Thus, by [Lemma A.1](#), we conclude that $z_i \leq 2\epsilon^2$. Similarly,

$$u(\mathbf{x}^*, \mathbf{y}^*, a_{2n+1}) - u(\mathbf{x}^*, \mathbf{y}^*, a_{n+i}) \geq |\mathbf{A}_{\min}| - \frac{|\mathbf{A}_{\min}|}{\epsilon}(y_i^* - x_i^*) \geq |\mathbf{A}_{\min}| \geq 1,$$

since $x_i^* - y_i^* \geq 0$. Again, [Lemma A.1](#) implies that $z_{n+i} \leq \epsilon^2$.

It thus suffices to treat the case where $|x_i^* - y_i^*| > \epsilon/2$ (assuming that $x_i^* - y_i^* \geq 0$). It follows that there exists $j \in [n]$ such that $x_j^* - y_j^* < 0$. In addition, we observe that $u(\mathbf{x}^*, \mathbf{y}^*, a_{2n+1}) = \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + |\mathbf{A}_{\min}| \geq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + 1$, whereas $u(\mathbf{x}^*, \mathbf{y}^*, a_j) < \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle$ and $u(\mathbf{x}^*, \mathbf{y}^*, a_{n+i}) < \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle$. As a result, [Lemma A.1](#) implies that $z_{n+i}^* \leq \epsilon^2$ and $z_j^* \leq \epsilon^2$.

Now, consider the deviation

$$\Delta^n \ni \mathbf{x}' = \mathbf{x}^* + (y_i^* - x_i^*)\mathbf{e}_i + (x_i^* - y_i^*)\mathbf{e}_j;$$

that is, $\mathbf{x}'$ is the strategy that results from $\mathbf{x}$ by reallocating $(x_i^* - y_i^*)$ probability mass from action a_i to action a_j . The difference $u(\mathbf{x}', \mathbf{y}^*, \mathbf{z}^*) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ can be expressed as

$$\begin{aligned} &\langle \mathbf{x}' - \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} (z_i^*(x_i' - x_i^*) + z_j^*(x_j' - x_j^*) + z_{n+i}^*(x_i^* - x_i') + z_{n+j}^*(x_j^* - x_j')) \\ &= \langle \mathbf{x}' - \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} (z_i^*(y_i^* - x_i^*) + z_j^*(x_i^* - y_i^*) + z_{n+i}^*(x_i^* - y_i^*) + z_{n+j}^*(y_i^* - x_i^*)) \\ &\leq 4\epsilon|\mathbf{A}_{\min}| + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(z_i^* \cdot \left(-\frac{\epsilon}{2} \right) + \epsilon^2 \cdot 2\epsilon + \epsilon^2 \cdot 2\epsilon \right) \end{aligned} \quad (8)$$

$$\leq -\left(\frac{1}{2}z_i^* - 4\epsilon^2 - 4\epsilon \right) |\mathbf{A}_{\min}|, \quad (9)$$

where (8) uses the following:

- $y_i^* - x_i^* \leq -\epsilon/2$;
- $x_i^* - y_i^* \leq 2\epsilon$ ([Lemma 3.2](#));
- $z_{n+i}^* \leq \epsilon^2$ and $z_j^* \leq \epsilon^2$; and
- $\langle \mathbf{x}' - \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle \leq \|\mathbf{x}' - \mathbf{x}^*\|_1 \|\mathbf{A}\mathbf{y}^*\|_\infty \leq 2|x_i^* - y_i^*| |\mathbf{A}_{\min}| \leq 4\epsilon |\mathbf{A}_{\min}|$ (since $\mathbf{A}$ has negative entries);

Moreover, given that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is assumed to be an ϵ^2 -Nash equilibrium, we have

$$-u(\mathbf{x}', \mathbf{y}^*, \mathbf{z}^*) + u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*) \leq \epsilon^2. \quad (10)$$

(The utility of Player $\mathbf{x}$ is given by $-u$.) Combining (10) and (9),

$$\left(\frac{1}{2}z_i^* - 4\epsilon^2 - 4\epsilon \right) |\mathbf{A}_{\min}| \leq \epsilon^2 \quad (11)$$

$$\Rightarrow z_i^* \leq 8\epsilon + 10\epsilon^2 \leq 9\epsilon \text{ for } \epsilon \leq \frac{1}{10}. \quad (12)$$

In summary, when $x_i^* - y_i^* \geq 0$, we have shown that $z_i^* \leq 9\epsilon$ and $z_{n+i}^* \leq \epsilon^2$. The case where $y_i^* - x_i^* \geq 0$ can be treated similarly. $\square$

We continue with the proof of [Theorem 3.4](#), which combines [Lemmas 3.2](#) and [3.3](#) to complete the CLS-hardness reduction of [Section 3.1](#).

Theorem 3.4. *Given any ϵ^2 -Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ in the adversarial team game (2), with $\epsilon \leq 1/10$, $(\mathbf{y}^*, \mathbf{y}^*)$ is a symmetric $(21n + 1)|\mathbf{A}_{\min}| \epsilon$ -Nash equilibrium of the symmetric, two-player game $(\mathbf{A}, \mathbf{A})$ (that is, $\mathbf{A} = \mathbf{A}^\top$).*

Proof. Since $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is an ϵ^2 -Nash equilibrium, we have that for any for any deviation $\mathbf{y}' \in \Delta^n$ of Player $\mathbf{y}$,

$$\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle \leq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}' \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n z_i(x_i^* - y'_i) + z_{n+i}(y'_i - x_i^*) \right) + \epsilon^2. \quad (13)$$

Moreover, by considering a deviation of Player $\mathbf{x}$ again to $\mathbf{y}'$,

$$\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle \leq \langle \mathbf{y}', \mathbf{A}\mathbf{y}^* \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n z_i(y'_i - y_i^*) + z_{n+i}(y_i^* - y'_i) \right) + \epsilon^2 \quad (14)$$

Adding (13) and (14), and using the fact that $\mathbf{A}$ is a symmetric matrix,

$$\begin{aligned} 2\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle &\leq \langle \mathbf{y}', \mathbf{A}(\mathbf{x}^* + \mathbf{y}^*) \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n z_i(x_i^* - y_i^*) + z_{n+i}(y_i^* - x_i^*) \right) + 2\epsilon^2 \\ &\leq \langle \mathbf{y}', \mathbf{A}(2\mathbf{y}^*) \rangle + 2\epsilon n |\mathbf{A}_{\min}| + \frac{|\mathbf{A}_{\min}|}{\epsilon} (2n \cdot 9\epsilon \cdot 2\epsilon) + 2\epsilon^2 \end{aligned} \quad (15)$$

$$\leq 2\langle \mathbf{y}', \mathbf{A}\mathbf{y}^* \rangle + (38n + 2)|\mathbf{A}_{\min}|\epsilon, \quad (16)$$

where in (15) we use [Lemmas 3.2](#) and [3.3](#). Also,

$$\begin{aligned} \langle \mathbf{y}^*, \mathbf{A}\mathbf{y}^* \rangle &= \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \langle \mathbf{y}^* - \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle \\ &\leq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + 2\epsilon n |\mathbf{A}_{\min}|. \end{aligned} \quad (17)$$

Finally, combining (16) and (17), we conclude that for any $\mathbf{y}' \in \Delta^n$,

$$\langle \mathbf{y}^*, \mathbf{A}\mathbf{y}^* \rangle \leq \langle \mathbf{y}', \mathbf{A}\mathbf{y}^* \rangle + (21n + 1)|\mathbf{A}_{\min}|\epsilon.$$

This concludes the proof. $\square$

We now restate [Theorem 1.2](#), which establishes the main complexity implication of [Theorem 3.1](#).

Theorem 1.2. *Computing an ϵ -Nash equilibrium in 3-player (that is, 2 vs. 1) adversarial team games is CLS-complete.*

Proof. CLS-hardness follows directly from [Theorem 3.4](#) and [Theorem 3.1](#) (due to [Ghosh and Hollender \[2024\]](#)). The inclusion was shown by [Anagnostides et al. \[2023\]](#). $\square$

A.2 Irrational Nash equilibria in adversarial team games

We next describe an interesting property for adversarial team games. Namely, similar to general-sum 3-player games [\[Nash, 1950\]](#), there exist adversarial team games that admit a unique irrational Nash equilibrium, as stated below.

Proposition A.2 ([Berthelsen and Hansen, 2019](#)). *There exists a 3-player adversarial team game with a unique Nash equilibrium that is supported on irrationals.*

Although the paper of [Berthelsen and Hansen \[2019\]](#) provides such an instance, no proof is given. Here, as a complement to their work, we provide a relatively simple and general way to analyze the irrational Nash equilibrium in 3-player adversarial team games.

We consider a 3-player adversarial team game in which the utility function of the adversary $u : \{1, 2\} \times \{1, 2\} \times \{1, 2\} : (\mathbf{x}, \mathbf{y}, \mathbf{z}) \mapsto \mathbb{R}$ reads

$z \backslash (x, y)$				
	(1, 1)	(1, 2)	(2, 1)	(2, 2)
1	1	3	$\frac{99}{100}$	$-\frac{1}{100}$
2	$\frac{9}{10}$	$-\frac{1}{10}$	1	3

The proof of this result makes use of a characterization of Nash equilibria in 2×2 two-player zero-sum games, stated below; for the proof, we refer to, for example, [Sun \[2022, Theorem 1.2\]](#).

Lemma A.3. *Let $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ such that*

$$(\mathbf{A}_{1,1} - \mathbf{A}_{1,2})(\mathbf{A}_{2,2} - \mathbf{A}_{2,1}) > 0 \text{ and } (\mathbf{A}_{1,1} - \mathbf{A}_{2,1})(\mathbf{A}_{2,2} - \mathbf{A}_{1,2}) > 0.$$

Then, the two-player zero-sum game $\min_{x \in \Delta^2} \max_{z \in \Delta^2} \langle x, \mathbf{A}z \rangle$ admits a unique (exact) Nash equilibrium with value

$$v := \frac{\mathbf{A}_{1,1}\mathbf{A}_{2,2} - \mathbf{A}_{1,2}\mathbf{A}_{2,1}}{\mathbf{A}_{1,1} - \mathbf{A}_{1,2} - \mathbf{A}_{2,1} + \mathbf{A}_{2,2}}.$$

Furthermore, the unique Nash equilibrium (x^, z^*) satisfies*

$$x^* = \left(\frac{\mathbf{A}_{2,2} - \mathbf{A}_{2,1}}{\mathbf{A}_{1,1} - \mathbf{A}_{1,2} - \mathbf{A}_{2,1} + \mathbf{A}_{2,2}}, \frac{\mathbf{A}_{1,1} - \mathbf{A}_{1,2}}{\mathbf{A}_{1,1} - \mathbf{A}_{1,2} - \mathbf{A}_{2,1} + \mathbf{A}_{2,2}} \right)$$

$$z^* = \left(\frac{\mathbf{A}_{2,2} - \mathbf{A}_{1,2}}{\mathbf{A}_{1,1} - \mathbf{A}_{1,2} - \mathbf{A}_{2,1} + \mathbf{A}_{2,2}}, \frac{\mathbf{A}_{1,1} - \mathbf{A}_{2,1}}{\mathbf{A}_{1,1} - \mathbf{A}_{1,2} - \mathbf{A}_{2,1} + \mathbf{A}_{2,2}} \right).$$

Proof of Proposition A.2. By construction of the adversarial team game, the mixed extension of the utility can be expressed as

$$x_1 z_1 (y_1 + 3y_2) + x_1 z_2 \left(\frac{9}{10} y_1 - \frac{1}{10} y_2 \right) + x_2 z_1 \left(\frac{99}{100} y_1 - \frac{1}{100} y_2 \right) + x_2 z_2 (3y_2 + y_1).$$

Suppose that we fix $y \in \Delta^2$. Then, Players x and y are engaged in a (two-player) zero-sum game with payoff matrix

$$\mathbf{A}(y) := \begin{bmatrix} 1 + 2y_2 & \frac{9}{10} - y_2 \\ \frac{99}{100} - y_2 & 1 + 2y_2 \end{bmatrix}. \quad (18)$$

We now invoke [Lemma A.3](#). Indeed, we have

$$(\mathbf{A}(y)_{1,1} - \mathbf{A}(y)_{1,2})(\mathbf{A}(y)_{2,2} - \mathbf{A}(y)_{2,1}) = \left(3y_2 + \frac{1}{10} \right) \left(3y_2 + \frac{1}{100} \right) > 0 \quad (19)$$

and

$$(\mathbf{A}(y)_{1,1} - \mathbf{A}(y)_{2,1})(\mathbf{A}(y)_{2,2} - \mathbf{A}(y)_{1,2}) = \left(3y_2 + \frac{1}{100} \right) \left(3y_2 + \frac{1}{10} \right) > 0; \quad (20)$$

that is, the precondition of [Lemma A.3](#) is satisfied, and so the value of (18) reads

$$v(y) = \min_{x \in \Delta^2} \max_{z \in \Delta^2} \langle x, \mathbf{A}(y)z \rangle = \frac{109 + 5890y_2 + 3000y_2^2}{110 + 6000y_2}. \quad (21)$$

It is easy to verify that v is a strictly convex function in $[0, 1]$, and admits a unique minimum corresponding to $y^* = \left(\frac{611-9\sqrt{3}}{600}, \frac{9\sqrt{3}-11}{600} \right)$, which is irrational. Now, suppose that (x^*, y^*, z^*) is a Nash equilibrium of the adversarial team game. We will first argue that (x^*, z^*) is the unique Nash equilibrium of $\mathbf{A}(y^*)$. Indeed, suppose that there exists $x' \in \Delta^2$ such that $\langle x', \mathbf{A}(y^*)z^* \rangle < \langle x^*, \mathbf{A}(y^*)z^* \rangle$, or equivalently, $u(x', y^*, z^*) < u(x^*, y^*, z^*)$; this is a contradiction since (x^*, y^*, z^*) is assumed to be a Nash equilibrium. Similar reasoning applies with respect to Player z . Thus, (x^*, z^*) is a Nash equilibrium of $\mathbf{A}(y^*)$, and thereby uniquely determined by y^* —by [Lemma A.3](#) coupled with (19) and (20). Furthermore, given the value of y^* , we get that $x^* = \left(\frac{3-\sqrt{3}}{6}, \frac{3+\sqrt{3}}{6} \right)$ and $z^* = \left(\frac{3+\sqrt{3}}{6}, \frac{3-\sqrt{3}}{6} \right)$. Now, consider the utility of Player y

when playing the first action a_1 or the second action a_2 ; plugging in the value of $\mathbf{x}^*$ and $\mathbf{z}^*$, we have $u(\mathbf{x}^*, e_1, \mathbf{z}^*) = \frac{578+9\sqrt{3}}{600}$ and $u(\mathbf{x}^*, e_2, \mathbf{z}^*) = \frac{578+9\sqrt{3}}{600}$. Since $u(\mathbf{x}^*, e_1, \mathbf{z}^*) = u(\mathbf{x}^*, e_2, \mathbf{z}^*)$, $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is a Nash equilibrium.

Moreover, suppose there exists another Nash equilibrium $(\mathbf{x}', \mathbf{y}', \mathbf{z}')$ that is different from $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$. As shown above, $(\mathbf{x}', \mathbf{z}')$ is the unique equilibrium of the zero-sum game induced by $\mathbf{y}'$. Thus, if we have two different Nash equilibria, it implies that $\mathbf{y}' \neq \mathbf{y}^*$. We consider the following three cases:

- First, let $\mathbf{y}'$ be a (fully) mixed strategy. Since $\mathbf{x}'$ and $\mathbf{z}'$ forms the unique NE in of $\mathbf{A}(\mathbf{y}')$, we have

$$\begin{aligned}\mathbf{x}' &= \left(\frac{1 + 300y'_2}{11 + 600y'_2}, \frac{10 + 300y'_2}{11 + 600y'_2} \right), \\ \mathbf{z}' &= \left(\frac{10 + 300y'_2}{11 + 600y'_2}, \frac{1 + 300y'_2}{11 + 600y'_2} \right).\end{aligned}$$

Further, for Player $\mathbf{y}$,

$$\begin{aligned}u(\mathbf{x}', e_1, \mathbf{z}') &= \frac{1199 + 130800y_2 + 3501000y_2^2}{10(11 + 600y_2)^2}, \\ u(\mathbf{x}', e_2, \mathbf{z}') &= \frac{589 + 196800y_2 + 5301000y_2^2}{10(11 + 600y_2)^2}.\end{aligned}$$

Since $\mathbf{y}'$ is a mixed strategy, we have $u(\mathbf{x}', e_1, \mathbf{z}') = u(\mathbf{x}', e_2, \mathbf{z}')$; solving the equality we get $\mathbf{y}' = \left(\frac{611-9\sqrt{3}}{600}, \frac{9\sqrt{3}-11}{600} \right)$, which contradicts the assumption that $\mathbf{y}' \neq \mathbf{y}^*$.

- If $\mathbf{y}' = (1, 0)$, we have $u(\mathbf{x}', e_1, \mathbf{z}') = \frac{1199}{1210}$ and $u(\mathbf{x}', e_2, \mathbf{z}') = \frac{589}{1210}$. Thus, it follows that by unilaterally deviating to play $(0, 1)$, Player $\mathbf{y}$ can decrease the utility of the adversary, contradicting the fact that $(\mathbf{x}', \mathbf{y}', \mathbf{z}')$ is a Nash equilibrium.
- Finally, suppose that $\mathbf{y}' = (0, 1)$. Similarly to the second case, we get $u(\mathbf{x}', e_1, \mathbf{z}') < u(\mathbf{x}', e_2, \mathbf{z}')$, which is a contradiction.

Thus, we conclude that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*)$ is the unique Nash equilibrium of the 3-player adversarial team game defined above, completing the proof. $\square$

A natural question arising from [Proposition A.2](#) concerns the complexity of determining whether an adversarial team game admits a unique Nash equilibrium. [Theorem 3.5](#)—that was presented earlier in the main body—establishes NP-hardness for a version of that problem that accounts for approximate Nash equilibria.

A.3 Proofs from Section 4.1

We continue with the proofs from [Section 4.1](#). We first apply Brouwer's fixed point theorem to show that symmetric min-max optimization problems always admit a symmetric equilibrium.

Lemma A.4. *Let $\mathcal{X}$ be a convex and compact set. Then, any L -smooth, antisymmetric function ([Definition 1.4](#)) $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ admits a symmetric first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{x}^*)$.*

Proof. We define the function $M : \mathcal{X} \rightarrow \mathcal{X}$ to be

$$M(\mathbf{x}') := \prod_{\mathbf{x}} \left[\mathbf{x}' - \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}) \Big|_{(\mathbf{x}, \mathbf{y}) = (\mathbf{x}', \mathbf{x}')} \right]. \quad (22)$$

Given that f is L -smooth, we conclude that $M(\mathbf{x}')$ is $(L+1)$ -Lipschitz, hence continuous. Therefore, from Brouwer's fixed point theorem, there exists an $\mathbf{x}^*$ so that $M(\mathbf{x}^*) = \mathbf{x}^*$. Moreover, the symmetry

of f implies that $\nabla_y f(x, y) \Big|_{(x, y) = (w, w)} = -\nabla_x f(x, y) \Big|_{(x, y) = (w, w)}$ for all $w \in \mathcal{X}$; as a result,

$$\begin{aligned} \mathbf{x}^* &= \prod_{\mathcal{X}} \left[\mathbf{x}^* - \nabla_x f(x, y) \Big|_{(x, y) = (\mathbf{x}^*, \mathbf{x}^*)} \right] \\ &= \prod_{\mathcal{X}} \left[\mathbf{x}^* + \nabla_y f(x, y) \Big|_{(x, y) = (\mathbf{x}^*, \mathbf{x}^*)} \right]. \end{aligned}$$

Therefore, $(\mathbf{x}^*, \mathbf{x}^*)$ is a first-order Nash equilibrium of the symmetric min-max problem with function f . $\square$

We now present the proof of [Lemma 4.1](#)

Proof of Lemma 4.1. We first define the function (as in [Lemma A.4](#)) $M : \mathcal{X} \rightarrow \mathcal{X}$ as

$$M(\mathbf{x}') := \prod_{\mathcal{X}} \left[\mathbf{x}' - \nabla_x f(x, y) \Big|_{(x, y) = (\mathbf{x}', \mathbf{x}')} \right],$$

where we recall that Π is the projection operator on $\mathcal{X}$. Assuming that the input function f is L -smooth, it follows that $M(\mathbf{x}')$ is $(L + 1)$ -Lipschitz. Furthermore, projecting on the polytope $\mathcal{X}$ takes polynomial time, and so M is polynomial-time computable. As a result, we can use [Etessami and Yannakakis \[2010, Proposition 2, part 2\]](#) (see also [Fearnley et al. \[2023, Proposition D.1\]](#)), where it was shown that finding an ϵ -approximate fixed point of a Brouwer function that is efficiently computable and continuous, when the domain is a bounded polytope, lies in PPAD. $\square$

We proceed with the proof of [Theorem 4.2](#)

Proof of Theorem 4.2. We P -time reduce the problem of finding approximate symmetric NE in two-player symmetric games to SYMGDAFIXEDPOINT. Given any two-player symmetric game with payoff matrices $(\mathbf{R}, \mathbf{R}^\top)$ of size $n \times n$, we set

$$\mathbf{A} := \frac{1}{2} (\mathbf{R} + \mathbf{R}^\top) \text{ (symmetric matrix) and } \mathbf{C} := \frac{1}{2} (\mathbf{R} - \mathbf{R}^\top) \text{ (skew-symmetric matrix)}. \quad (23)$$

We define the *quadratic*, antisymmetric function

$$f(x, y) := \frac{1}{2} \langle y, \mathbf{A}y \rangle - \frac{1}{2} \langle x, \mathbf{A}x \rangle + \langle y, \mathbf{C}x \rangle \quad (24)$$

with domain $\Delta^n \times \Delta^n$. Indeed, to see that f is antisymmetric, one can observe that

$$f(y, x) = \frac{1}{2} \langle x, \mathbf{A}x \rangle - \frac{1}{2} \langle y, \mathbf{A}y \rangle + \langle x, \mathbf{C}y \rangle = \frac{1}{2} \langle x, \mathbf{A}x \rangle - \frac{1}{2} \langle y, \mathbf{A}y \rangle - \langle y, \mathbf{C}^\top x \rangle = -f(x, y).$$

Assuming that all entries of $\mathbf{R}$ lie in $[-1, 1]$, it follows that the singular values of $\mathbf{A}$ and $\mathbf{C}$ are bounded by n . As a result f and $\nabla_x f = -\mathbf{A}x - \mathbf{C}y$, $\nabla_y f = \mathbf{A}y + \mathbf{C}x$ are polynomial time computable and continuous, and $\nabla_x f, \nabla_y f$ are L -Lipschitz for $L \leq 2n$, thus f is $4n$ -smooth.

We assume x is the minimizer and y is the maximizer, and let $(\mathbf{x}^*, \mathbf{x}^*)$ be an ϵ -approximate fixed point of GDA. We shall show that $(\mathbf{x}^*, \mathbf{x}^*)$ is an $4n\epsilon$ -approximate NE of the symmetric two-player game $(\mathbf{R}, \mathbf{R}^\top)$. Since $(\mathbf{x}^*, \mathbf{x}^*)$ is an ϵ -approximate fixed point of GDA, we can use [Lemma A.5](#) and obtain the following variational inequalities:

$$\max_{\mathbf{x}^* + \delta \in \Delta^n, \|\delta\|_2 \leq 1} \delta^\top (\mathbf{A}\mathbf{x}^* + \mathbf{C}\mathbf{x}^*) \leq \epsilon(2n + 1),$$

implying that (since the diameter of Δ^n is $\sqrt{2}$ in ℓ_2)

$$\langle x - \mathbf{x}^*, (\mathbf{A} + \mathbf{C})\mathbf{x}^* \rangle \leq \sqrt{2}\epsilon(2n + 1) \text{ for any } x \in \Delta^n. \quad (\text{VI for NE})$$

Now, we observe that (VI for NE) implies that $(\mathbf{x}^*, \mathbf{x}^*)$ is a $\sqrt{2}\epsilon(2n + 1)$ -approximate symmetric NE in the two-player symmetric game with payoff matrices $(\mathbf{A} + \mathbf{C}, \mathbf{A} - \mathbf{C})$ (recall Definition (2.1)). Since $\sqrt{2}\epsilon(2n + 1) \leq 4n\epsilon$ for $n \geq 2$, our claim follows.

By [Theorem 2.2](#) and [Lemma 4.1](#), we conclude that SYMGDAFIXEDPOINT is PPAD-complete, even for quadratic functions that are $O(n)$ -smooth, $O(n)$ -Lipschitz and $\epsilon \leq 1/n^{1+c}$, for any $c > 0$. $\square$

We next state a standard lemma that connects first-order optimality with the fixed-point gap of gradient ascent.

Lemma A.5 ([Ghadimi and Lan, 2016], Lemma 3 for $c = 1$). *Let $f(x)$ be an L -smooth function in $x \in \Delta^n$. Define the gradient mapping*

$$G(x) := \prod_{\Delta^n} \{x + \nabla_x f(x)\} - x.$$

If $\|G(x^)\|_2 \leq \epsilon$, that is, x^* is an ϵ -approximate fixed point of gradient ascent with stepsize equal to one, then*

$$\max_{x^* + \delta \in \Delta^n, \|\delta\|_2 \leq 1} \delta^\top \nabla_x f(x^*) \leq \epsilon(L + 1).$$

Similarly, the next lemma makes such a connection for min-max optimization problems with coupled constraints; it is mostly extracted from Daskalakis et al. [2021, Section B.2].

Lemma A.6. *Let $f(x, y)$ be a G -Lipschitz, L -smooth function defined in some polytope domain $\mathcal{D} \subseteq \Delta^n \times \Delta^n$ of diameter D . Define the mapping*

$$G(x, y) := \prod_{\mathcal{D}} \{x - \nabla_x f(x, y), y + \nabla_y f(x, y)\} - (x, y).$$

If $\|G(x^, y^*)\|_2 \leq \epsilon$, that is, (x^*, y^*) is an ϵ -approximate fixed point of (the safe version) of GDA with stepsize equal to one, then*

$$\langle x - x^*, \nabla_x f(x^*, y^*) \rangle \geq -\sqrt{\epsilon}K \text{ for } x \in \mathcal{D}(y^*) \text{ and } \langle y - y^*, \nabla_y f(x^*, y^*) \rangle \leq \sqrt{\epsilon}K \text{ for } y \in \mathcal{D}(x^*),$$

where $\mathcal{D}(x^) = \{y : (x^*, y) \in \mathcal{D}\}$, $\mathcal{D}(y^*) = \{x : (x, y^*) \in \mathcal{D}\}$ and $K = (L + 1)\sqrt{(G + 4\sqrt{2})}$.*

Proof. Let $(x_\Delta, y_\Delta) = (x^* - \nabla_x f(x^*, y^*), y^* + \nabla_y f(x^*, y^*))$. In Daskalakis et al. [2021, Claim B.2], it was shown that for all $(x, y) \in \mathcal{D}$, we have

$$\langle (x_\Delta, y_\Delta) - (x^*, y^*), (x, y) - (x^*, y^*) \rangle \leq (G + 2D)\epsilon.$$

Using the above inequality, it was concluded that (x^*, y^*) is an approximate fixed point of the “unsafe” version of GDA; specifically,

$$\left\| x^* - \prod_{\mathcal{D}(y^*)} \{x^* - \nabla_x f(x^*, y^*)\} \right\| \leq \sqrt{(G + 2D)\epsilon}$$

and

$$\left\| y^* - \prod_{\mathcal{D}(x^*)} \{y^* + \nabla_y f(x^*, y^*)\} \right\| \leq \sqrt{(G + 2D)\epsilon}.$$

We now use Lemma A.5 for both inequalities above, together the fact that $D = 2\sqrt{2}$, to conclude that

$$\langle x - x^*, \nabla_x f(x^*, y^*) \rangle \geq -\sqrt{(G + 4\sqrt{2})\epsilon}(L + 1) \text{ for } x \in \mathcal{D}(y^*),$$

and

$$\langle y - y^*, \nabla_y f(x^*, y^*) \rangle \leq \sqrt{(G + 4\sqrt{2})\epsilon}(L + 1) \text{ for } y \in \mathcal{D}(x^*).$$

□

We proceed to establish Theorem 4.4.

Proof of Theorem 4.4. The proof follows similar steps with Theorem 4.2, namely, we P -time reduce the problem of finding approximate symmetric NE in two-player symmetric games to GDAFIXEDPOINT with coupled domains. Given a two-player symmetric game with payoff matrices $(\mathbf{R}, \mathbf{R}^\top)$ of size $n \times n$, we set $\mathbf{A} := \frac{1}{2}(\mathbf{R} + \mathbf{R}^\top)$, $\mathbf{C} := \frac{1}{2}(\mathbf{R} - \mathbf{R}^\top)$ and define again the quadratic, antisymmetric function

$$f(x, y) := \frac{1}{2}\langle y, \mathbf{A}y \rangle - \frac{1}{2}\langle x, \mathbf{A}x \rangle + \langle y, \mathbf{C}x \rangle.$$

Moreover, given a parameter $\delta > 0$ (to be specified shortly), we define the joint domain of f to be

$$\mathcal{D} := \{(\mathbf{x}, \mathbf{y}) \in \Delta^n \times \Delta^n : -\delta \leq x_i - y_i \leq \delta \text{ for all } i \in [n]\}. \quad (\text{joint Domain})$$

Let $(\mathbf{x}^*, \mathbf{y}^*)$ be an ϵ -approximate fixed point of GDA. We will show that $(\frac{\mathbf{x}^* + \mathbf{y}^*}{2}, \frac{\mathbf{x}^* + \mathbf{y}^*}{2})$ is an $O(\epsilon^{1/4})$ -approximate (symmetric) NE of the game $(\mathbf{R}, \mathbf{R}^\top)$ for an appropriate choice of δ .

We set $\mathcal{D}(\mathbf{x}^*) = \{\mathbf{y} : (\mathbf{x}^*, \mathbf{y}) \in \mathcal{D}\}$ and $\mathcal{D}(\mathbf{y}^*) = \{\mathbf{x} : (\mathbf{x}, \mathbf{y}^*) \in \mathcal{D}\}$. In words, $\mathcal{D}(\mathbf{x}^*)$ and $\mathcal{D}(\mathbf{y}^*)$ capture the allowed deviations for $\mathbf{y}$ and $\mathbf{x}$ respectively. It also holds that f is G -Lipschitz continuous with $G = 4n$ and also $4n$ -smooth (using the same reasoning as in Theorem 4.2).

Since $(\mathbf{x}^*, \mathbf{y}^*)$ is an ϵ -approximate fixed point of GDA, using Lemma A.6, the following VIs must hold for some positive constant $K < 10$ and n sufficiently large:

$$\begin{aligned} \langle \mathbf{x} - \mathbf{x}^*, -\mathbf{A}\mathbf{x}^* + \mathbf{C}^\top \mathbf{y}^* \rangle &\geq -Kn^{3/2}\sqrt{\epsilon} \text{ for any } \mathbf{x} \in \mathcal{D}(\mathbf{y}^*) \text{ and} \\ \langle \mathbf{y} - \mathbf{y}^*, \mathbf{A}\mathbf{y}^* + \mathbf{C}\mathbf{x}^* \rangle &\leq Kn^{3/2}\sqrt{\epsilon} \text{ for any } \mathbf{y} \in \mathcal{D}(\mathbf{x}^*). \end{aligned} \quad (25)$$

Let $\overline{\mathcal{D}} = \{z \in \Delta^n : \|z - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}\|_\infty \leq \frac{\delta}{2}\}$. By triangle inequality, it follows that $\overline{\mathcal{D}} \subseteq \mathcal{D}(\mathbf{y}^*) \cap \mathcal{D}(\mathbf{x}^*)$. We express the VIs of (25) using a single variable z and common deviation domain:

$$\langle z - \mathbf{x}^*, -\mathbf{A}\mathbf{x}^* + \mathbf{C}^\top \mathbf{y}^* \rangle \geq -Kn^{3/2}\sqrt{\epsilon} \text{ and } \langle z - \mathbf{y}^*, \mathbf{A}\mathbf{y}^* + \mathbf{C}\mathbf{x}^* \rangle \leq Kn^{3/2}\sqrt{\epsilon} \text{ for any } z \in \overline{\mathcal{D}}.$$

Multiplying the first inequality by $-1/2$ and the second with $1/2$ and adding them up gives

$$\left\langle z - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}, (\mathbf{A} + \mathbf{C}) \frac{\mathbf{x}^* + \mathbf{y}^*}{2} \right\rangle \leq \frac{1}{4} \langle \mathbf{x}^* - \mathbf{y}^*, \mathbf{A}(\mathbf{x}^* - \mathbf{y}^*) \rangle + Kn^{3/2}\sqrt{\epsilon}. \quad (26)$$

Since $\mathbf{x}^*, \mathbf{y}^* \in \mathcal{D}$, it follows that $\langle \mathbf{x}^* - \mathbf{y}^*, \mathbf{A}(\mathbf{x}^* - \mathbf{y}^*) \rangle \leq n\|\mathbf{x}^* - \mathbf{y}^*\|_2^2 \leq n^2\delta^2$. Combining this fact with (26), we conclude that

$$\left\langle z - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}, (\mathbf{A} + \mathbf{C}) \frac{\mathbf{x}^* + \mathbf{y}^*}{2} \right\rangle \leq n^2\delta^2 + Kn^{3/2}\sqrt{\epsilon} \text{ for any } z \in \overline{\mathcal{D}}. \quad (\text{VImedian})$$

(VImedian) shows that by deviating from $(\frac{\mathbf{x}^* + \mathbf{y}^*}{2}, \frac{\mathbf{x}^* + \mathbf{y}^*}{2})$ to some z in $\overline{\mathcal{D}}$, the payoff cannot increase by more than $(n^2\delta^2 + Kn^{3/2}\sqrt{\epsilon})$ in the two-player symmetric game with matrices $(\mathbf{R}, \mathbf{R}^\top)$.

We consider any pure strategy e_j for $j \in [n]$. If $\|e_j - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}\|_\infty \leq \frac{\delta}{2}$ then $e_j \in \overline{\mathcal{D}}$ and it is captured by (VImedian). Suppose that $\|e_j - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}\|_\infty > \frac{\delta}{2}$ and consider the point $z' \in \overline{\mathcal{D}}$ on the line segment between e_j and $\frac{\mathbf{x}^* + \mathbf{y}^*}{2}$ that intersects the boundary of $\overline{\mathcal{D}}$. It holds that $e_j - \frac{\mathbf{x}^* + \mathbf{y}^*}{2} = c(z' - \frac{\mathbf{x}^* + \mathbf{y}^*}{2})$ for some positive $c \leq \frac{2}{\delta}$ (it cannot be larger because otherwise the infinity norm of the difference between e_j and $\frac{\mathbf{x}^* + \mathbf{y}^*}{2}$ would exceed one, which is impossible as they both belong to Δ^n). Therefore,

$$\left\langle e_j - \frac{\mathbf{x}^* + \mathbf{y}^*}{2}, (\mathbf{A} + \mathbf{C}) \frac{\mathbf{x}^* + \mathbf{y}^*}{2} \right\rangle \leq 2n^2\delta + \frac{2Kn^{3/2}\sqrt{\epsilon}}{\delta} \text{ for any pure strategy } j. \quad (27)$$

From (27), we conclude that $\frac{\mathbf{x}^* + \mathbf{y}^*}{2}$ is $(2n^2\delta + \frac{2Kn^{3/2}\sqrt{\epsilon}}{\delta})$ -approximate NE of the symmetric two-player game $(\mathbf{R}, \mathbf{R}^\top)$. We choose $\delta = \epsilon^{1/4}n^{-1/4}$ and we get that $\frac{\mathbf{x}^* + \mathbf{y}^*}{2}$ is an $O(n^{7/4}\epsilon^{1/4})$ -approximate NE for $(\mathbf{R}, \mathbf{R}^\top)$, and thus the hardness result holds for ϵ of order $O(\frac{1}{n^{7+c}})$ for any constant $c > 0$. We note that if instead of an ϵ -approximate fixed point of GDA, we were given an ϵ -approximate first-order NE, then the hardness result would hold for any ϵ of order $\frac{1}{n^c}$ with $c > 0$. $\square$



Figure 1: An example of matrix $\mathbf{A} = \mathbf{A}(G)$ (right) for graph G (left).

A.4 Proofs from Section 4.2

The main result we want to show is restated below.

Theorem 1.6. *For a symmetric min-max optimization problem, constants $c_1, c_2 > 0$, and $\epsilon = n^{-c_1}$, it is NP-hard to distinguish between the following two cases under the promise that one of them holds:*

- any ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies $\|\mathbf{x}^* - \mathbf{y}^*\| \leq n^{-c_2}$, and
- there is an ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ such that $\|\mathbf{x}^* - \mathbf{y}^*\| \geq \Omega(1)$.

Our reduction builds on the hardness result of [McLennan and Tourky \[2010\]](#)—in turn based on earlier work by [Gilboa and Zemel \[1989\]](#), [Conitzer and Sandholm \[2008\]](#)—which we significantly refine in order to account for $\text{poly}(1/n)$ -Nash equilibria. We begin by describing their basic approach. Let $G = ([n], E)$ be an n -node, undirected, unweighted graph, and construct

$$\mathbf{A}_{i,j} = \begin{cases} -1 & \text{if } i = j, \\ 0 & \text{if } \{i, j\} \in E, \\ -2 & \text{otherwise.} \end{cases} \quad (28)$$

(Figure 1 depicts an illustrative example.) Based on this matrix, [McLennan and Tourky \[2010\]](#) consider the symmetric, identical-payoff, two-player game $(\mathbf{A}, \mathbf{A})$ —by construction, $\mathbf{A} = \mathbf{A}^\top$, and so this game is indeed symmetric. They were able to show the following key property.

Lemma A.7 ([\[McLennan and Tourky, 2010\]](#)). *Let $C_k \subseteq [n]$ be a maximum clique of G with size k and $\mathbf{x}^* = \frac{1}{k} \sum_{i \in C_k} \mathbf{e}_i$. Then, $(\mathbf{x}^*, \mathbf{x}^*)$ is a Nash equilibrium of $(\mathbf{A}, \mathbf{A})$ that attains value $-\frac{1}{k}$. Furthermore, any symmetric Nash equilibrium not in the form described above has value at most $-\frac{1}{k-1}$.*

The idea now is to construct a new symmetric, identical-payoff game $(\mathbf{B}, \mathbf{B})$, for

$$\mathbf{B} := \begin{bmatrix} \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,n} & r \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{A}_{n,1} & \cdots & \mathbf{A}_{n,n} & r \\ r & \cdots & r & V \end{bmatrix}, \quad (29)$$

where $V := -\frac{1}{k}$ and $r = \frac{1}{2}(-\frac{1}{k} - \frac{1}{k-1}) = -\frac{2k-1}{2(k-1)k}$. Coupled with [Lemma A.7](#), this new game yields the following NP-hardness result.

Theorem A.8 ([\[McLennan and Tourky, 2010\]](#)). *It is NP-hard to determine whether a symmetric, identical-payoff, two-player game has a unique symmetric Nash equilibrium.*

Our goal here is to prove a stronger result, [Theorem 4.7](#), that characterizes the set of ϵ -Nash equilibria even for $\epsilon = 1/n^c$; this will form the basis for our hardness result in min-max optimization and adversarial team games. To do so, we first derive some basic properties of game (29).

Game (29) always admits the trivial (symmetric) Nash equilibrium $(\mathbf{e}_{n+1}, \mathbf{e}_{n+1})$. Now, consider any symmetric Nash equilibrium $(\mathbf{x}^*, \mathbf{x}^*)$ with $x_{n+1}^* \neq 1$. If $x_{n+1}^* = 0$, it follows that $(\mathbf{x}_{[1 \dots n]}^*, \mathbf{x}_{[1 \dots n]}^*)$ is a Nash equilibrium of $(\mathbf{A}, \mathbf{A})$, which in turn implies that G admits a clique of size k ; this follows from [Lemma A.7](#), together with the fact that $-1/(k-1) < r < -1/k$.

We now analyze the case where $x_{n+1}^* \in (0, 1)$. It then follows that $(\mathbf{x}_{[1 \dots n]}^*/(1-x_{n+1}^*), \mathbf{x}_{[1 \dots n]}^*/(1-x_{n+1}^*))$ is a (symmetric) Nash equilibrium of $(\mathbf{A}, \mathbf{A})$. Furthermore, the utility of playing action a_{n+1} is

$(1 - x_{n+1}^*)r + x_{n+1}^*V > r$. By Lemma A.7, it follows that $(\mathbf{x}_{[1 \dots n]}^*/(1 - x_{n+1}^*), \mathbf{x}_{[1 \dots n]}^*/(1 - x_{n+1}^*))$ has a value of V and G admits a clique of size k . As a result, the utility of playing any action a_i , with $i \in \text{supp}(\mathbf{x}^*)$ and $i \neq n+1$, is $(1 - x_{n+1}^*)V + x_{n+1}^*r$. At the same time, the utility of playing action a_{n+1} reads $(1 - x_{n+1}^*)r + x_{n+1}^*V$. Equating those two quantities, it follows that $x_{n+1}^* = 1/2$.

In summary, G contains a clique of size k if and only if game (29) admits a unique symmetric Nash equilibrium, which implies Theorem A.8. What is more, we have shown a stronger property. Namely, any symmetric Nash equilibrium of $(\mathbf{B}, \mathbf{B})$ has to be in one of the following forms:

1. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \mathbf{e}_{n+1}$;
2. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \frac{1}{k} \sum_{i \in C_k} \mathbf{e}_i$, where $C_k \subseteq [n]$ is a clique in G of size k ;
3. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \frac{1}{2} \mathbf{e}_{n+1} + \frac{1}{2k} \sum_{i \in C_k} \mathbf{e}_i$, where $C_k \subseteq [n]$ is a clique in G of size k .

In particular, the equilibria in Items 2 or 3—which exist iff G contains a clique of size k —are always far from the one in Item 1. However, this characterization only applies to exact Nash equilibria. In any two-player game Γ , when ϵ is sufficiently small with $\log(1/\epsilon) \leq \text{poly}(|\Gamma|)$, Etesami and Yannakakis [2010] have shown that any ϵ -Nash equilibrium is within ℓ_1 -distance δ from an exact one, and so the above characterization can be applied; unfortunately, this does not apply (for general games) in the regime we are interested, namely $\epsilon = \text{poly}(1/n)$.

We address this challenge by refining the result of McLennan and Tourky [2010]. Our main result, which forms the basis for Theorem 1.6 and Theorem 3.5, is recalled below.

Theorem A.9. *For symmetric, identical-interest, two-player games, constants $c_1, c_2 > 0$, and $\epsilon = n^{-c_1}$, it is NP-hard to distinguish between the following two cases:*

- any two symmetric ϵ -Nash equilibria have ℓ_1 -distance at most n^{-c_2} , and
- there are two symmetric ϵ -Nash equilibria that have ℓ_1 -distance $\Omega(1)$.

Our reduction proceeds similarly, but defines $\bar{\mathbf{A}}$ to be the adjacency matrix of G with $\delta \in (0, 1)$ in each diagonal entry. Using $\bar{\mathbf{A}}$, we show that we can refine Lemma A.7 of McLennan and Tourky [2010]. Before we state the key property we prove in Lemma A.11, we recall the following definition.

Definition A.10 (Well-supported NE). A symmetric strategy profile $(\mathbf{x}, \mathbf{x})$ is an ϵ -well-supported Nash equilibrium of the symmetric, identical-payoff game $(\bar{\mathbf{A}}, \bar{\mathbf{A}})$ if for all $i \in [n]$,

$$x_i > 0 \implies (\bar{\mathbf{A}}\mathbf{x})_i \geq \max_{j \in [n]} (\bar{\mathbf{A}}\mathbf{x})_j - \epsilon.$$

Lemma A.11. *Suppose that the maximum clique in G is of size k . For any symmetric ϵ -well-supported NE $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ of $(\bar{\mathbf{A}}, \bar{\mathbf{A}})$ not supported on a clique of size k , we have $u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \leq 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4} + 2\epsilon$.*

Equipped with this property, we will see shortly that a similar argument to the one described earlier concerning game (29) establishes Theorem 4.7. We proceed now with the proof of Theorem 1.6.

Proof of Theorem 1.6. It suffices to consider the antisymmetric function $f(\mathbf{x}, \mathbf{y}) := \mathbf{y}^\top \mathbf{B} \mathbf{y} - \mathbf{x}^\top \mathbf{B} \mathbf{x}$, where symmetric matrix $\mathbf{B}$ is defined as in (29), using our new matrix $\bar{\mathbf{A}}$ instead of $\mathbf{A}$ (see (32)). Any ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ of this (separable) min-max optimization problem induces, two symmetric ϵ -Nash equilibria—namely, $(\mathbf{x}^*, \mathbf{x}^*)$ and $(\mathbf{y}^*, \mathbf{y}^*)$ —in the symmetric, identical-interest, game $(\mathbf{B}, \mathbf{B})$. Using Theorem 4.7, the claim follows. $\square$

In what follows, our goal is to establish Theorem 4.7, which forms the basis for Theorems 1.6 and 3.5.

We consider the symmetric game $(\bar{\mathbf{A}}, \bar{\mathbf{A}}^\top)$, where $\bar{\mathbf{A}} \in \mathbb{R}^{n \times n}$ is a symmetric matrix. In particular, since $\bar{\mathbf{A}} = \bar{\mathbf{A}}^\top$, the two players in the game share the same payoff matrix. The payoff matrix $\bar{\mathbf{A}}$ is constructed based on an underlying graph $G = ([n], E)$ and a parameter $\delta \in (0, 1)$ as follows:

$$\bar{\mathbf{A}}_{i,j} = \begin{cases} \delta & \text{if } i = j, \\ 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma A.12. Let $(\hat{x}, \hat{x})$ be an ϵ -well-supported NE of the game where $\hat{x}$ is supported on a max clique of size k , denoted as C_k , then the value $u(\hat{x}, \hat{x})$ is at least $1 - \frac{1}{k} + \frac{\delta}{k} - (\frac{k-\delta}{1-\delta})\epsilon$ and $\|\hat{x} - x^*\|_\infty \leq \frac{k-\delta}{1-\delta}\epsilon$ where $x^* = \frac{1}{k} \sum_{i \in C_k} e_i$.

Proof. Since $\hat{x}$ is supported on C_k , let i be a coordinate that $\hat{x}^*$ puts the least probability mass on; that is, $i \in \operatorname{argmin}_{j \in C_k} \hat{x}_j$. Considering the utility of playing action a_i , we have

$$u(e_i, \hat{x}) = \hat{x}_i \cdot \delta + (1 - \hat{x}_i) \cdot 1 = 1 - \hat{x}_i + \hat{x}_i \delta.$$

Moreover, let $j \in C_k$ be a coordinate such that $\hat{x}_j \geq \frac{1}{k}$. It should hold that

$$u(e_j, \hat{x}) = 1 - (1 - \delta)\hat{x}_j \leq 1 - \frac{1}{k} + \frac{\delta}{k}.$$

Since $(\hat{x}, \hat{x})$ is an ϵ -well-supported NE, we have

$$\begin{aligned} u(e_j, \hat{x}) &\geq u(e_i, \hat{x}) - \epsilon \\ \Rightarrow 1 - \frac{1}{k} + \frac{\delta}{k} + \epsilon &\geq 1 - \hat{x}_i + \hat{x}_i \delta \\ \Rightarrow \hat{x}_i &\geq \frac{1}{k} - \frac{\epsilon}{1 - \delta}. \end{aligned}$$

Thus, for all coordinates $i \in C_k$, we have $\frac{1}{k} - \frac{\epsilon}{1-\delta} \leq \hat{x}_i \leq \frac{1}{k} + \frac{\epsilon}{1-\delta} \cdot (k-1)$. It then holds that

$$\begin{aligned} u(\hat{x}, \hat{x}) &\geq u(e_i, \hat{x}) - \epsilon \\ &\geq \left(\frac{1}{k} + \frac{\epsilon}{1-\delta} \cdot (k-1) \right) \cdot \delta + \left(1 - \frac{1}{k} - \frac{\epsilon}{1-\delta} \cdot (k-1) \right) - \epsilon \\ &\geq 1 - \frac{1}{k} + \frac{\delta}{k} - \left(\frac{k-\delta}{1-\delta} \right) \epsilon. \end{aligned}$$

□

Assumption A.13. For the rest of this subsection, we set the parameters as follows:

- $n \geq k \geq 10$;
- $\epsilon < \delta(1 - \delta)/6n^7$;
- $\delta := 1/2$.

(Using the symbolic value of δ is more convenient in our derivations below.)

Proof of Lemma A.11. We let $x^* = \frac{1}{k} \sum_{i \in C_k} e_i$. The proof considers the following three cases:

- The symmetric ϵ -well-supported NE $(\hat{x}, \hat{x})$ has support size less than k .
- The symmetric ϵ -well-supported NE $(\hat{x}, \hat{x})$ has support size greater than k .
- The symmetric ϵ -well-supported NE $(\hat{x}, \hat{x})$ has support size equal to k but is not supported on a clique.

We proceed to show that for any of the three cases above, we would not be able to have a symmetric ϵ -well-supported NE that achieves value greater than $1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4} + 2\epsilon$, which is a contradiction.

- For the first case, since the support has size less than k , we can find a coordinate $i \in [n]$ such that $\hat{x}_i \geq \frac{1}{k-1}$. Therefore, the value of playing that action is

$$u(e_i, \hat{x}) \leq 1 - \frac{1}{k-1} + \frac{\delta}{k-1}.$$

Since $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is a symmetric ϵ -well-supported NE, we have

$$\begin{aligned} u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) &\leq u(\mathbf{e}_i, \hat{\mathbf{x}}) + \epsilon \\ &\leq 1 - \frac{1}{k-1} + \frac{\delta}{k-1} + \epsilon \\ &\leq 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4} + 2\epsilon, \end{aligned}$$

where in the last step we use [Assumption A.13](#).

- For the second case, suppose $|\text{supp}(\hat{\mathbf{x}})| = m > k$. By [Lemma A.16](#), since the maximum clique size is k , we can find a set $\mathcal{S} \subseteq \text{supp}(\hat{\mathbf{x}})$ with at least $m - k + 1$ elements such that for each coordinate $i \in \mathcal{S}$, we can find a coordinate $j \in \mathcal{S}$ such that $\bar{\mathbf{A}}_{i,j} = \bar{\mathbf{A}}_{j,i} = 0$. Now we consider the utility of playing action a_i and a_j , we have

$$\begin{aligned} u(\mathbf{e}_i, \hat{\mathbf{x}}) &= \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{i,l} + \hat{x}_i \cdot \delta + \hat{x}_j \cdot \bar{\mathbf{A}}_{i,j}, \\ u(\mathbf{e}_j, \hat{\mathbf{x}}) &= \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{j,l} + \hat{x}_j \cdot \delta + \hat{x}_i \cdot \bar{\mathbf{A}}_{i,j}. \end{aligned}$$

Since $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is a symmetric ϵ -well-supported NE, we have

$$\begin{aligned} \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{i,l} + \hat{x}_i \cdot \delta + \hat{x}_j \cdot \bar{\mathbf{A}}_{i,j} &\geq \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{j,l} + \hat{x}_j \cdot \delta + \hat{x}_i \cdot \bar{\mathbf{A}}_{i,j} - \epsilon \\ \Rightarrow \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{i,l} - \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{j,l} &\geq \hat{x}_j \cdot \delta - \hat{x}_i \cdot \delta - \epsilon. \end{aligned}$$

Now, by moving all the probability mass from action j to action i , we form a new strategy $\mathbf{x}' = \hat{\mathbf{x}} + \hat{x}_j \cdot \mathbf{e}_i - \hat{x}_j \cdot \mathbf{e}_j$ such that

$$\begin{aligned} u(\mathbf{x}', \mathbf{x}') - u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) &= 2\hat{x}_j \left(\sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{i,l} - \sum_{l \in \text{supp}(\hat{\mathbf{x}}) - \{i,j\}} \hat{x}_l \cdot \bar{\mathbf{A}}_{j,l} \right) + 2\hat{x}_i \cdot \hat{x}_j \cdot \delta \\ &\geq 2\hat{x}_j \cdot ((\hat{x}_j - \hat{x}_i)\delta - \epsilon) + 2\hat{x}_i \cdot \hat{x}_j \cdot \delta \\ &\geq 2\hat{x}_j^2 \delta - 2\epsilon. \end{aligned} \tag{30}$$

Suppose there is a coordinate $i \in \mathcal{S}$ such that $\hat{x}_i \geq \frac{1}{nk^2}$; then, from (30), we have

$$\begin{aligned} u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) &\leq u(\mathbf{x}^*, \mathbf{x}^*) - 2 \cdot \left(\frac{1}{nk^2}\right)^2 \cdot \delta + 2\epsilon \\ &= 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4} + 2\epsilon. \end{aligned}$$

If $\hat{x}_i < \frac{1}{nk^2}$ for any $i \in \mathcal{S}$, then there exists a coordinate $l \notin \mathcal{S}$ such that $\hat{x}_l > \frac{1 - (m-k+1) \cdot \frac{1}{nk^2}}{m - (m-k+1)} \geq \frac{1}{k} + \frac{1}{k^2}$. Then, considering the utility when playing action a_l ,

$$\begin{aligned} u(\mathbf{e}_l, \hat{\mathbf{x}}) &\leq 1 \cdot \left(1 - \frac{1}{k} - \frac{1}{k^2}\right) + \delta \cdot \left(\frac{1}{k} + \frac{1}{k^2}\right) \\ &= 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{1}{k^2} + \frac{\delta}{k^2} \\ &< 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4}, \end{aligned} \tag{31}$$

where in (31) we used [Assumption A.13](#).

Since the l th action is played with positive probability and $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is an ϵ -well-supported NE, we have $u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \leq u(\mathbf{e}_l, \hat{\mathbf{x}}) + \epsilon < 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2 k^4} + 2\epsilon$.

- For the third case, since the support is not on a clique, there exists at least coordinates i, j such that $\hat{x}_i > 0, \hat{x}_j > 0$, and $\bar{\mathbf{A}}_{i,j} = \bar{\mathbf{A}}_{j,i} = 0$. Similarly as case two, if $\hat{x}_i \geq \frac{1}{nk^2}$ or $\hat{x}_j \geq \frac{1}{nk^2}$, then we have

$$\begin{aligned} u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) &\leq u(\mathbf{x}^*, \mathbf{x}^*) - 2 \cdot \left(\frac{1}{nk^2}\right)^2 \cdot \delta + 2\epsilon \\ &= 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2k^4} + 2\epsilon. \end{aligned}$$

If $\hat{x}_i < \frac{1}{nk^2}$ and $\hat{x}_j < \frac{1}{nk^2}$, then there exists a coordinate l such that $\hat{x}_l \geq \frac{1-2\cdot\frac{1}{nk^2}}{k-2} > \frac{1}{k} + \frac{1}{k^2}$. Same as (31), we conclude that $u(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is at most $1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2k^4} + 2\epsilon$.

The proof is complete. $\square$

We now construct a new symmetric identical payoff game $(\mathbf{B}, \mathbf{B})$, where $\mathbf{B}$ is defined as

$$\mathbf{B} = \begin{bmatrix} \bar{\mathbf{A}}_{1,1} & \cdots & \bar{\mathbf{A}}_{1,n} & r \\ \vdots & \ddots & \vdots & \vdots \\ \bar{\mathbf{A}}_{n,1} & \cdots & \bar{\mathbf{A}}_{n,n} & r \\ r & \cdots & r & V \end{bmatrix}. \quad (32)$$

Above, $V := 1 - \frac{1}{k} + \frac{\delta}{k}$ and $r := 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{\delta}{n^2k^4} + 3\epsilon$. Similarly to our discussion after [Theorem A.8](#), it follows that the symmetric (exact) Nash equilibria of this game can only be in one of the following forms:

1. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \mathbf{e}_{n+1}$;
2. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \frac{1}{k} \sum_{i \in C_k} \mathbf{e}_i$, where $C_k \subseteq [n]$ is a clique in G of size k ;
3. $(\mathbf{x}^*, \mathbf{x}^*)$ with $\mathbf{x}^* := \frac{1}{2} \mathbf{e}_{n+1} + \frac{1}{2k} \sum_{i \in C_k} \mathbf{e}_i$, where $C_k \subseteq [n]$ is a clique in G of size k .

We now show the following lemma.

Lemma A.14. *For any ϵ -well-supported NE $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ in game $(\mathbf{B}, \mathbf{B})$, it holds that $\|\hat{\mathbf{x}} - \mathbf{x}^*\|_\infty \leq 2n^6\epsilon$, where $(\mathbf{x}^*, \mathbf{x}^*)$ is an exact NE in one of the three cases above (1,2,3).*

Proof. First, we observe that since $V > r$, clearly $(\mathbf{e}_{n+1}, \mathbf{e}_{n+1})$ is a ϵ -well-supported NE; in this case, $\|\mathbf{x}' - \mathbf{x}^*\|_\infty = 0$. Furthermore, since $r = 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{\delta}{n^2k^4} + 3\epsilon$, the game does not attain any ϵ -NE with value less than $1 - \frac{1}{k} + \frac{\delta}{k} - \frac{\delta}{n^2k^4} + 2\epsilon$. Suppose the game admits a symmetric ϵ -well-supported NE $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ where $\hat{\mathbf{x}}$ is supported only on the first n actions. Since $u(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \geq 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{\delta}{n^2k^4} + 2\epsilon > 1 - \frac{1}{k} + \frac{\delta}{k} - \frac{2\delta}{n^2k^4} + 2\epsilon$, taking $\mathbf{x}^*$ as in (2), we conclude that $\|\hat{\mathbf{x}} - \mathbf{x}^*\|_\infty \leq \frac{k-1}{1-\delta}\epsilon < 2n^6\epsilon$ from [Lemma A.11](#).

We proceed to the case where there is a mixed symmetric ϵ -well-supported Nash $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ between the last action and the rest of actions such that $0 < \hat{x}_{n+1} < 1$. Denote $\hat{x}_{n+1} = \alpha$ and $\eta(\cdot)$ to denote the renormalization operation. Since $(\eta(\hat{\mathbf{x}}_{[1\dots n]}), \eta(\hat{\mathbf{x}}_{[1\dots n]}))$ is a symmetric strategy profile, from [McLennan and Tourky \[2010, Proposition 4\]](#), we conclude that there is at least one coordinate $i \in \text{supp}(\hat{\mathbf{x}}) - \{n+1\}$ such that

$$u_A(\mathbf{e}_i, \eta(\hat{\mathbf{x}}_{[1\dots n]})) \leq 1 - \frac{1}{k} + \frac{\delta}{k} = V,$$

where u_A is the utility when the payoff matrix is $\bar{\mathbf{A}}$. Since $(\eta(\hat{\mathbf{x}}_{[1\dots n]}), \eta(\hat{\mathbf{x}}_{[1\dots n]}))$ is an ϵ -well-supported NE, we have

$$\begin{aligned} u_B(\mathbf{e}_i, \hat{\mathbf{x}}) &\geq u_B(\mathbf{e}_{n+1}, \hat{\mathbf{x}}) - \epsilon \\ \Rightarrow (1 - \alpha) \cdot V + \alpha \cdot r &\geq (1 - \alpha)r + \alpha V - \epsilon \\ \Rightarrow \alpha &\leq \frac{1}{2} + \frac{\epsilon}{2(V - r)}, \end{aligned} \quad (33)$$

where u_B is the utility function when the payoff matrix is $\mathbf{B}$. Plugging in the value of V and r and using [Assumption A.13](#), we find that $\alpha \leq \frac{1}{2} + 2n^6\epsilon \leq \frac{2}{3}$.

Now, observe that for any action in the support other than the last action a_i , the utility of playing such action $u_B(a_i, \hat{\mathbf{x}}) = u_A(a_i, \hat{\mathbf{x}}_{[1 \dots n]}) + r\alpha$. Since $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is an ϵ -well-supported Nash Equilibrium, we have $u_B(a_i, \hat{\mathbf{x}}) \geq \max_j u_B(a_j, \hat{\mathbf{x}}) - \epsilon$ for all pairs $(i, j) \in \text{supp}(\hat{\mathbf{x}})$. Since $\alpha \leq \frac{2}{3}$, it follows that $u_A(a_i, \eta(\hat{\mathbf{x}}_{[1 \dots n]})) \geq \max_j u_A(a_j, \eta(\hat{\mathbf{x}}_{[1 \dots n]})) - 3\epsilon$ for any pairs $(i, j) \in \text{supp}(\hat{\mathbf{x}}) - \{n+1\}$. Thus, we conclude that $(\eta(\hat{\mathbf{x}}_{[1 \dots n]}), \eta(\hat{\mathbf{x}}_{[1 \dots n]}))$ forms a symmetric 3ϵ -well-supported Nash Equilibrium in game $(\bar{\mathbf{A}}, \bar{\mathbf{A}})$. Further, the value of playing the last action is $(1 - \alpha)r + \alpha V > r$, and so the only situation where there is a mixed Nash between the last action and the rest actions is when $u_A((\eta(\hat{\mathbf{x}}_{[1 \dots n]}), \eta(\hat{\mathbf{x}}_{[1 \dots n]}))) \geq r - \epsilon$. Therefore, by [Lemma A.12](#) and [Lemma A.11](#), we conclude that $\hat{\mathbf{x}}_{[1 \dots n]}$ is supported on a clique of size k . There exists at least one coordinate $i \in [n]$, with $0 < \hat{x}_i$ and $\hat{x}_i \geq \frac{1}{k}$, such that

$$u_A(e_i, \eta(\hat{\mathbf{x}}_{[1 \dots n]})) \geq 1 - \frac{1}{k} + \frac{\delta}{k} = V.$$

Since $(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is an ϵ -well-supported Nash, we have $u(e_i, \hat{\mathbf{x}}) \leq u(e_{n+1}, \hat{\mathbf{x}}) + \epsilon$, and so this gives

$$(1 - \alpha) \cdot V + \alpha \cdot r \leq (1 - \alpha) \cdot r + \alpha \cdot V + \epsilon \quad (34)$$

$$\Rightarrow \alpha \geq \frac{1}{2} - \frac{\epsilon}{2(V - r)}. \quad (35)$$

Using [Assumption A.13](#) and combining with (33),

$$\frac{1}{2} - 2n^6\epsilon \leq \alpha \leq \frac{1}{2} + 2n^6\epsilon.$$

By taking $\mathbf{x}^*$ as in (3), we conclude that $\|\hat{\mathbf{x}} - \mathbf{x}^*\|_\infty \leq 2n^6\epsilon$. $\square$

Theorem A.15. *For any ϵ -NE $(\mathbf{x}, \mathbf{x})$ in game $(\mathbf{B}, \mathbf{B})$, it holds that $\|\mathbf{x} - \mathbf{x}^*\|_\infty \leq n^6\sqrt{\epsilon}$, where $(\mathbf{x}^*, \mathbf{x}^*)$ is an exact NE in one of the forms specified above (1,2,3).*

Proof. [Chen et al. \[2009, Lemma 3.2\]](#) showed that from any $\epsilon^2/8$ -NE $(\mathbf{x}, \mathbf{y})$ in any two player bimatrix game, one can construct (in polynomial time) an ϵ -well-supported NE $(\mathbf{x}', \mathbf{y}')$ such that $\|\mathbf{x} - \mathbf{x}'\|_\infty \leq \frac{\epsilon}{4}$ and $\|\mathbf{y} - \mathbf{y}'\|_\infty \leq \frac{\epsilon}{4}$. Setting $\epsilon' := \frac{\epsilon^2}{8}$ for the ϵ defined in [Lemma A.14](#), the proof follows. $\square$

The proof of [Theorem 4.7](#) follows directly by observing that having two symmetric ϵ -NE $(\mathbf{x}, \mathbf{x})$ and $(\mathbf{y}, \mathbf{y})$ such that $\|\mathbf{x} - \mathbf{y}\|_\infty > 2n^6\sqrt{\epsilon}$ would imply that the game $(\mathbf{B}, \mathbf{B})$ has two distinct exact NE, which in turn implies that there is a clique of size k in the graph.

We next state and prove an auxiliary lemma that was used earlier.

Lemma A.16. *For any graph $G = (V, E)$ with n vertices, if the maximum clique has size k , then we can form a set $S \subseteq V$ of size at least $n - k + 1$ such that for any vertex $i \in S$, there exists a vertex $j \in S$ such that i and j are not connected.*

Proof. Suppose the largest set S we can form has cardinality $|S| < n - k + 1$, this implies there is a set $S' = V - S$ with at least $n - (n - k) = k$ vertices such that each vertex in S' is connected to all other vertices in G . However if this is the case, there is at least one vertex $v \notin S'$ that are connected to all vertices in S' . This contradicts the fact that maximum clique has size k . $\square$

We proceed now with the proof of [Theorem 1.6](#).

Proof of Theorem 1.6. It suffices to consider the antisymmetric function $f(\mathbf{x}, \mathbf{y}) := \mathbf{y}^\top \mathbf{B} \mathbf{y} - \mathbf{x}^\top \mathbf{B} \mathbf{x}$, where symmetric matrix $\mathbf{B}$ is defined as in (29), using our new matrix $\bar{\mathbf{A}}$ instead of $\mathbf{A}$ (see (32)). Any ϵ -first-order Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ of this (separable) min-max optimization problem induces, two symmetric ϵ -Nash equilibria—namely, $(\mathbf{x}^*, \mathbf{x}^*)$ and $(\mathbf{y}^*, \mathbf{y}^*)$ —in the symmetric, identical-interest, game $(\mathbf{B}, \mathbf{B})$. Using [Theorem 4.7](#), the claim follows. $\square$

A.5 Proofs from Section 4.3

We conclude with the missing proofs from Section 4.3. We first introduce two lemmas which are useful in later proofs.

The first key lemma, which mirrors Lemma 3.2, shows that, in an approximate Nash equilibrium of (4), $\mathbf{x} \approx \mathbf{y}$ and $\hat{\mathbf{x}} \approx \hat{\mathbf{y}}$. This is crucial as it allows us to construct—up to some small error—quadratic terms in the utility function, as in our hardness result for symmetric min-max optimization.

Lemma A.17. *Let $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ be an ϵ^2 -Nash equilibrium of (4) with $\epsilon^2 \leq 1/2$. Then, $\|\mathbf{x}^* - \mathbf{y}^*\|_\infty \leq 2\epsilon$ and $\|\hat{\mathbf{x}}^* - \hat{\mathbf{y}}^*\|_\infty \leq 2\epsilon$.*

Proof. For the sake of contradiction, suppose that $\|\mathbf{x}^* - \mathbf{y}^*\|_\infty > 2\epsilon$. Without loss of generality, let us further assume that there is some coordinate i such that $x_i^* - y_i^* > 2\epsilon$. The payoff difference for Player $\hat{\mathbf{z}}$ when playing the i th action compared to action a_{2n+1} reads

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, a_i) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, a_{2n+1}) &= \frac{|\mathbf{A}_{\min}|}{\epsilon} \cdot (x_i^* - y_i^*) - |\mathbf{A}_{\min}| \\ &> |\mathbf{A}_{\min}|. \end{aligned}$$

By Lemma A.1, it follows that $\hat{\mathbf{z}}_{2n+1}^* < \frac{\epsilon^2}{|\mathbf{A}_{\min}|} \leq \epsilon^2$. Moreover, given that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ is an ϵ^2 -Nash equilibrium, we have

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) &\geq u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, a_i) - \epsilon^2 \\ &= \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle - \langle \hat{\mathbf{x}}^*, \mathbf{A}\hat{\mathbf{y}}^* \rangle + \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle \\ &\quad + \frac{|\mathbf{A}_{\min}|}{\epsilon} (x_i^* - y_i^*) - \delta(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mathbf{z}) - \epsilon^2 \\ &\geq \mathbf{A}_{\min} - \langle \hat{\mathbf{x}}^*, \mathbf{A}\hat{\mathbf{y}}^* \rangle + \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle + 2|\mathbf{A}_{\min}| - \delta(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mathbf{z}) - \epsilon^2 \\ &= |\mathbf{A}_{\min}| - \langle \hat{\mathbf{x}}^*, \mathbf{A}\hat{\mathbf{y}}^* \rangle + \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle - \delta(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \mathbf{z}) - \epsilon^2. \end{aligned}$$

Now, considering the deviation of Player $\mathbf{y}$ to $\mathbf{y}' := \mathbf{x}^*$,

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}', \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) &\leq \langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle + \epsilon^2 |\mathbf{A}_{\min}| - |\mathbf{A}_{\min}| + \epsilon^2 \\ &\leq -2 + 2\epsilon^2 \\ &< -\epsilon^2, \end{aligned}$$

which contradicts the fact that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ is an ϵ^2 -Nash equilibrium. We conclude that $\|\mathbf{x}^* - \mathbf{y}^*\|_\infty \leq 2\epsilon$; the proof for the fact that $\|\hat{\mathbf{x}}^* - \hat{\mathbf{y}}^*\|_\infty \leq 2\epsilon$ follows similarly. $\square$

Next, following the argument of Lemma 3.3, we show that, in equilibrium, Players $\mathbf{z}$ and $\hat{\mathbf{z}}$ place most of their probability mass on action a_{2n+1} , thereby having only a small effect on the game between Players $\mathbf{x}$ and $\mathbf{y}$ vs. $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.

Lemma A.18. *Let $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ be an ϵ^2 -Nash equilibrium of (4) with $\epsilon \leq 1/10$. Then, $z_j, \hat{z}_j \leq 9\epsilon$ for all $j \in [2n]$.*

Proof. We will prove that $z_j \leq 9\epsilon$ for all $j \in [2n]$; the corresponding claim for Player $\hat{\mathbf{z}}$ follows similarly. Fix $i \in [n]$. Lemma A.17 shows that $|y_i^* - x_i^*| \leq 2\epsilon$. We shall consider two cases.

First, suppose that $|y_i^* - x_i^*| \leq \epsilon/2$. Then,

$$\begin{aligned} u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, a_{2n+1}) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, a_i) &\geq |\mathbf{A}_{\min}| - \frac{|\mathbf{A}_{\min}|}{\epsilon} \cdot (x_i^* - y_i^*) \\ &\geq \frac{|\mathbf{A}_{\min}|}{2} \geq \frac{1}{2}. \end{aligned}$$

By Lemma A.1, it follows that $\hat{z}_i \leq 2\epsilon^2$, and similar reasoning yields $\hat{z}_{n+i} \leq 2\epsilon^2$. On the other hand, suppose that $|y_i^* - x_i^*| > \epsilon/2$. Without loss of generality, we can assume that $y_i^* - x_i^* \geq 0$; the contrary case is symmetric. Since $\mathbf{x}^* \in \Delta^n$ and $\mathbf{y}^* \in \Delta^n$, there is some coordinate $j \in [n]$ such that

$y_j^* - x_j^* < 0$. As before, by [Lemma A.1](#), it follows that $\hat{z}_i^* \leq \epsilon^2$ and $\hat{z}_{n+j}^* \leq \epsilon^2$. Now, we consider the deviation

$$\Delta^n \ni \mathbf{y}' = \mathbf{y}^* + (x_i^* - y_i^*)\mathbf{e}_i + (y_i^* - x_i^*)\mathbf{e}_j.$$

Then, we have

$$\begin{aligned} & u(\mathbf{x}^*, \mathbf{y}', \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) \\ &= \langle \mathbf{x}^*, \mathbf{A}(\mathbf{y}' - \mathbf{y}^*) \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} (\hat{z}_i^*(x_i^* - y_i' - (x_i^* - y_i^*)) + \hat{z}_j^*(x_j^* - y_j' - (x_j^* - y_j^*))) \\ &\quad + \frac{|\mathbf{A}_{\min}|}{\epsilon} (\hat{z}_{n+i}^*(y_i' - x_i^* - (y_i^* - x_i^*)) + \hat{z}_{n+j}^*(y_j' - x_j^* - (y_j^* - x_j^*))) \\ &= \langle \mathbf{x}^*, \mathbf{A}(\mathbf{y}' - \mathbf{y}^*) \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} (\hat{z}_i^*(y_i^* - y_i') + \hat{z}_j^*(y_j^* - y_j') + \hat{z}_{n+i}^*(y_i' - y_i^*) + \hat{z}_{n+j}^*(y_j' - y_j^*)) \\ &\leq 4\epsilon|\mathbf{A}_{\min}| + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\epsilon^2 \cdot 2\epsilon + \hat{z}_{n+i}^* \cdot \left(-\frac{\epsilon}{2}\right) + \epsilon^2 \cdot 2\epsilon \right) \\ &\leq -\left(\frac{1}{2}\hat{z}_{n+i}^* - 4\epsilon^2 - 4\epsilon\right)|\mathbf{A}_{\min}|. \end{aligned}$$

At the same time, since $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ is an ϵ^2 -Nash equilibrium, we have

$$u(\mathbf{x}^*, \mathbf{y}', \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) - u(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*) \geq -\epsilon^2.$$

Thus,

$$\left(\frac{1}{2}\hat{z}_{n+i}^* - 4\epsilon^2 - 4\epsilon\right)|\mathbf{A}_{\min}| \leq \epsilon^2 \implies \hat{z}_{n+i}^* \leq 9\epsilon.$$

We conclude that $\hat{z}_i^* \leq \epsilon^2$ and $\hat{z}_{n+i}^* \leq 9\epsilon$. The case where $x_i^* - y_i^* \geq 0$ can be treated similarly. $\square$

Armed with those two basic lemmas, we are ready to complete the proof of [Theorem 1.7](#).

Proof of Theorem 1.7. Suppose that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}^*, \hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*, \hat{\mathbf{z}}^*)$ is an ϵ^2 -Nash equilibrium. We have that for any $\mathbf{y}' \in \Delta^n$,

$$\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle \leq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}' \rangle + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n \hat{z}_i^*(y_i^* - y_i') + \hat{z}_{n+i}^*(y_i' - y_i^*) \right) + \epsilon^2. \quad (36)$$

Moreover, considering a deviation from $\mathbf{x}^*$ to $\mathbf{y}'$,

$$\begin{aligned} \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle &\leq \langle \mathbf{y}', \mathbf{A}\mathbf{y}^* \rangle + \langle \mathbf{y}', \mathbf{C}\hat{\mathbf{x}}^* \rangle \\ &\quad + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n \hat{z}_i^*(y_i' - x_i^*) + \hat{z}_{n+i}^*(x_i^* - y_i') \right) + \epsilon^2. \end{aligned} \quad (37)$$

Summing (36) and (37),

$$\begin{aligned} 2\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle &\leq \langle \mathbf{y}', \mathbf{A}(\mathbf{x}^* + \mathbf{y}^*) \rangle + \langle \mathbf{y}', \mathbf{C}\hat{\mathbf{x}}^* \rangle \\ &\quad + \frac{|\mathbf{A}_{\min}|}{\epsilon} \left(\sum_{i=1}^n \hat{z}_i^*(y_i^* - x_i^*) + \hat{z}_{n+i}^*(x_i^* - y_i^*) \right) + 2\epsilon^2 \\ &\leq 2\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}' \rangle + \langle \mathbf{y}', \mathbf{C}\hat{\mathbf{x}}^* \rangle + 2\epsilon n|\mathbf{A}_{\min}| + \frac{|\mathbf{A}_{\min}|}{\epsilon} (2n \cdot 9\epsilon \cdot 2\epsilon) + 2\epsilon^2 \\ &\leq 2\langle \mathbf{x}^*, \mathbf{A}\mathbf{y}' \rangle + \langle \mathbf{y}', \mathbf{C}\hat{\mathbf{x}}^* \rangle + (38n + 2)|\mathbf{A}_{\min}|\epsilon. \end{aligned} \quad (38)$$

Moreover,

$$\langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle = \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + \langle \mathbf{x}^*, \mathbf{A}(\mathbf{x}^* - \mathbf{y}^*) \rangle \leq \langle \mathbf{x}^*, \mathbf{A}\mathbf{y}^* \rangle + 2|\mathbf{A}_{\min}|n\epsilon. \quad (39)$$

Combining (38) and (39), we get that for all $\mathbf{y}' \in \Delta^n$,

$$\langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle + \frac{1}{2}\langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle \leq \langle \mathbf{y}', \mathbf{A}\mathbf{x}^* \rangle + \frac{1}{2}\langle \mathbf{y}', \mathbf{C}\hat{\mathbf{x}}^* \rangle + (21n + 1)|\mathbf{A}_{\min}|\epsilon. \quad (40)$$

Similarly, we can show that for all $\hat{\mathbf{x}}' \in \Delta^n$,

$$\langle -\hat{\mathbf{x}}^*, \mathbf{A}\hat{\mathbf{x}}^* \rangle + \frac{1}{2} \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}^* \rangle \geq -\langle \hat{\mathbf{x}}', \mathbf{A}\hat{\mathbf{x}}^* \rangle + \frac{1}{2} \langle \mathbf{x}^*, \mathbf{C}\hat{\mathbf{x}}' \rangle - (21n+1)|\mathbf{A}_{\min}|\epsilon. \quad (41)$$

Taking $\mathbf{y}' = \hat{\mathbf{x}}'$ in (40) and summing with (41), we get that for all $\hat{\mathbf{x}}' \in \Delta^n$,

$$\langle \hat{\mathbf{x}}', \mathbf{A}\mathbf{x}^* \rangle + \langle \hat{\mathbf{x}}', \mathbf{A}\hat{\mathbf{x}}^* \rangle + \frac{1}{2} (\langle \hat{\mathbf{x}}', \mathbf{C}\hat{\mathbf{x}}^* \rangle + \langle \hat{\mathbf{x}}', \mathbf{C}\mathbf{x}^* \rangle) + (42n+2)|\mathbf{A}_{\min}|\epsilon \geq \langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle + \langle \hat{\mathbf{x}}^*, \mathbf{A}\hat{\mathbf{x}}^* \rangle,$$

where we used the fact that $\mathbf{C}$ is skew-symmetric. Now, using the fact that the Nash equilibrium is symmetric, so that $\mathbf{x}^* = \hat{\mathbf{x}}^*$, we have

$$\begin{aligned} \langle \hat{\mathbf{x}}', \mathbf{A}\mathbf{x}^* \rangle + \frac{1}{2} \langle \hat{\mathbf{x}}', \mathbf{C}\mathbf{x}^* \rangle + (21n+1)|\mathbf{A}_{\min}|\epsilon &\geq \langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle \\ &\geq \langle \mathbf{x}^*, \mathbf{A}\mathbf{x}^* \rangle + \frac{1}{2} \langle \mathbf{x}^*, \mathbf{C}\mathbf{x}^* \rangle. \end{aligned} \quad (42)$$

Setting $\mathbf{A} := -\frac{1}{2}(\mathbf{R} + \mathbf{R}^\top)$ and $\mathbf{C} := \mathbf{R}^\top - \mathbf{R}$, (42) shows that

$$\langle \hat{\mathbf{x}}', \mathbf{R}\mathbf{x}^* \rangle \leq \langle \mathbf{x}^*, \mathbf{R}\mathbf{x}^* \rangle + (21n+1)|\mathbf{A}_{\min}|\epsilon$$

for any $\hat{\mathbf{x}}' \in \Delta^n$; *ergo*, $(\mathbf{x}^*, \mathbf{x}^*)$ is a symmetric $(21n+1)|\mathbf{A}_{\min}|\epsilon$ -Nash equilibrium of the symmetric (two-player) game $(\mathbf{R}, \mathbf{R}^\top)$, and the proof follows from [Theorem 2.2](#). $\square$

NeurIPS Paper Checklist

1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: We provide proofs for all the claims made in the abstract and introduction.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: We formally state all assumptions needed for our results.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [\[Yes\]](#)

Justification: All assumptions are stated in the main body and the proofs are in the appendix.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
 - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
 - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
 - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
 - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [NA]

Justification: The paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: The research conducted in the paper conforms with the NeurIPS Code of Ethics.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: We do not foresee any direct societal impact.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The paper poses no such risks.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: The paper does not use existing assets.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset’s creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: This paper does not release new assets.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: This paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: This paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The development of the paper does not involve LLMs.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.