

---

# The Complexity of Correlated Equilibria in Generalized Games

---

**Martino Bernasconi**  
Bocconi University  
martino.bernasconi@unibocconi.it

**Matteo Castiglioni**  
Politecnico di Milano  
matteo.castiglioni@polimi.it

**Andrea Celli**  
Bocconi University  
andrea.celli2@unibocconi.it

**Gabriele Farina**  
Massachusetts Institute of Technology  
gfarina@mit.edu

## Abstract

Correlated equilibria—and their generalizations known as  $\Phi$ -equilibria—are a fundamental object of study in game theory, offering a more tractable alternative to Nash equilibria in multi-player settings. While computational aspects of equilibrium computation are well-understood in some settings, fundamental questions are still open in *generalized games*, that is, games in which the set of strategies allowed to each player depends on the other players’ strategies. These classes of games model fundamental settings in economics, and have been a cornerstone of economics research since the seminal paper of [Arrow and Debreu \[1954\]](#). Recently, there has been growing interest, both in economics and in computer science, in studying correlated equilibria in generalized games. It is known that finding a social welfare maximizing correlated equilibrium in generalized games is NP-hard. However, the existence of efficient algorithms to find *any* equilibrium remains an important open question. In this paper, we answer this question in the negative, showing that this problem is PPAD-complete.

## 1 Introduction

Game Theory is a fundamental tool to study the interaction between rational agents. From an application point of view, it has found important applications such as training GANs [[Goodfellow et al., 2020](#)], auctions [[Nisan et al., 2007](#)], and superhuman performance in strategic games [[Brown and Sandholm, 2019](#), [FAIR et al., 2022](#)]. In a series of papers, originating from [Papadimitriou \[1994\]](#) and culminating in [Daskalakis et al. \[2009\]](#), [Chen et al. \[2009\]](#), the problem of computing a Nash equilibrium in multiplayer games was proven to be complete for the complexity class PPAD, which contains problems for which no efficient algorithm is known. On the other hand, correlated equilibria [[Aumann, 1987](#)], coarse-correlated equilibria [[Moulin and Vial, 1978](#)], and their generalization known as  $\Phi$ -equilibria [[Greenwald and Jafari, 2003](#)], are a relaxation of Nash equilibria, and provide a computationally tractable alternative to Nash in most instances.

The seminal result of [Nash \[1951\]](#) for the existence of equilibrium in non-cooperative games was extended to *generalized games* (also known as pseudo-games or abstract economies) by [Arrow and Debreu \[1954\]](#) and [Rosen \[1965\]](#). In these games, the set of feasible strategies of one player depends on the strategies played by the other players.

Coupled constraints between players model various applications involving finite resources, where consumption depends on the joint actions of the players. For example, generalized games have been applied to dynamic pricing [[Adida and Perakis, 2010](#)], electricity markets [[Hobbs and Pang, 2007](#)], and Fisher markets [[Brânzei et al., 2014](#), [Conitzer et al., 2022](#)]. A particularly relevant example of

joint constraints is bidding in repeated auctions under budget constraints, where budget consumption depends on the joint bids of all players. It is well known that when all players follow no-regret strategies, their average play converges to a correlated equilibrium [Cesa-Bianchi and Lugosi, 2006]. In online advertising, where many companies deploy automated bidders that act as “cost-aware” regret minimizers on behalf of advertisers [Balseiro and Gur, 2019, Aggarwal et al., 2019], it becomes particularly important to ask whether similar convergence guarantees hold for “cost-aware” no-regret algorithms (such as the algorithms presented in Mannor et al. [2009], Mahdavi et al. [2012]).

Thus, an important and natural question is to determine if there are efficient algorithms that compute correlated equilibria in generalized games. While Bernasconi et al. [2023] proved the computational infeasibility of computing a social welfare maximizing correlated equilibrium in generalized games, the complexity of computing *any* equilibrium remains open (see also Zhang et al. [2025]). In this paper, we resolve this open question.

## 1.1 Contributions and Overview of Techniques

The main contribution of this paper is to show that the problem of computing an approximate  $\Phi$ -equilibrium in generalized games (as formalized in the CONstrained- $\Phi$ -EQUILIBRIUM problem in Section 3) is PPAD-complete. The hardness holds for approximation factors that are inverse-polynomial in the instance both in (i) games with two players, and (ii) games with an arbitrary number of players even for a constant number of actions and constraints per player. By providing this characterization, our paper puts to rest a recent line of inquiry in the literature [Bernasconi et al., 2023, Zhang et al., 2025].

At a fundamental level, the question of the complexity of correlated equilibria in generalized games boils down to the tension between two opposing forces:

1. On the one hand, the introduction of *correlation* —as opposed to the *independence* between the play of different players required by Nash equilibria— is a major driver of computational feasibility. For instance, while Nash equilibria are believed to be computationally intractable, correlated solution concepts [Aumann, 1987] are typically computable in polynomial time, even in games with exponentially many actions [Papadimitriou and Roughgarden, 2005, Farina and Pipis, 2024, Daskalakis et al., 2024].
2. On the other hand, the introduction of coupled constraints often has adverse effects on computation [Daskalakis et al., 2021, Papadimitriou et al., 2023, Bernasconi et al., 2024, Anagnostides et al., 2025], even when the constraints are *simple*, *e.g.*, when they define a polytope in the joint strategy space.

Thus, a natural and fundamental question that tries to resolve this tension is the following:

*Which of the two forces prevails? Does correlation remain beneficial in the face of constraints? Or are constraints so detrimental to push the complexity of correlated equilibria in generalized games to the realm of intractability?*

We resolve this question by showing that correlation is not a sufficient relaxation in the presence of coupled constraints. We prove this result by a reduction from the problem of computing Nash equilibria in polymatrix games. At an intuitive level, we could say that the computation of linearly constrained correlated equilibria is indistinguishable from the computation of unconstrained Nash (*i.e.*, independent) equilibria. In other words, the combination of correlation and constraints can simulate independence without constraints.

At the technical level, our main tool is to introduce a “left” and “right” copy of each player in the polymatrix instance, thus creating two teams of players. Each player on one team then plays only with players on the opposing team in a team zero-sum game. We then exploit the constraints to couple the players of the left team to the players of the right team by asking that their marginals coincide. This allows us to essentially remove the terms that are linear in the correlated strategy and only keep the ones that are linear in the marginals (which now resemble the ones of Nash equilibria).

One additional implication of our result is that the quasi-polynomial time algorithm for computing approximate  $\Phi$ -equilibria of Bernasconi et al. [2023] is tight. Indeed, under the Exponential Time Hypothesis for PPAD (introduced by Rubinfeld [2017]), we show that the problem of computing

a (coarse) correlated equilibria in generalized games requires quasi-polynomial time, even with a constant number of players and a constant violation of the incentive compatibility constraints.

Furthermore, we emphasize that although numerous works have established positive convergence guarantees for regret minimization in the unconstrained setting [e.g., Hart and Mas-Colell, 2000], our results demonstrate that no efficient no-regret algorithms can, in general, converge to constrained correlated equilibria. This impossibility has broad implications for systems of constrained no-regret learners, as it precludes convergence unless specific properties of the problem are exploited. A concrete example that has recently attracted significant theoretical and empirical attention is that of Internet advertising platforms, where automated bidding agents are typically operated by the platform itself (see, e.g., [Agarwal et al., 2014, Balseiro and Gur, 2019, Balseiro et al., 2021, Paes Leme et al., 2024]). In such environments, convergence to an equilibrium is often desirable since it promotes stability, predictability, and alignment with the advertisers’ individual incentives. Our result, however, shows that this goal cannot be achieved in a “black-box” manner. To design autobidders with convergence guarantees, one must develop algorithms that exploit the specific structural properties of the problem, since no general-purpose approach can ensure convergence in this setting.

## 2 Preliminaries

### 2.1 Multi-player Games and $\Phi$ -equilibrium

We consider an  $n$ -player game. Each player has a set  $A = [\ell]$  of actions, and for each player  $i \in [n]$  and each tuple of actions  $\mathbf{a} \in \mathcal{A} := A^n$ , the utility of player  $i$  is  $u_i(\mathbf{a}) \in [0, 1]$ .<sup>1</sup> A  $\Phi$ -equilibrium is defined on correlated strategies  $z \in \Delta(\mathcal{A})$ . In particular, a mediator samples actions  $\mathbf{a} = (a_1, \dots, a_n) \sim z$  and communicates to each player  $i$  its own action  $a_i$ . Conditioned on the action recommendation, each player can now decide how to deviate. In particular, each player can pick a function  $\phi \in \Phi$  that prescribes its deviation from a set of functions  $\Phi$ . More specifically, each function  $\phi \in \Phi$  maps each action  $a \in A$  into a randomized action  $\phi(a) \in \Delta(A)$ . Note that each function  $\phi$  can be represented by a  $\ell \times \ell$ -dimensional right stochastic matrix on which each row corresponds to a specific action, i.e.  $\phi(a, b)$  is the probability of playing  $b$  when the player is recommended action  $a$ . By applying a deviation  $\phi$  to a correlated strategy  $z$  the  $i$ -th player induces a distribution  $\phi \circ_i z$  on  $\mathcal{A}$ . Formally, for all  $\mathbf{a} = (a_i, \mathbf{a}_{-i}) \in \mathcal{A}$  we let

$$(\phi \circ_i z)(\mathbf{a}) = \sum_{b \in A} \phi(b, a_i) z(b, \mathbf{a}_{-i}).$$

Moreover, for a function  $F : \mathcal{A} \rightarrow \mathbb{R}$  and a distribution  $z \in \Delta(\mathcal{A})$ , with abuse of notation we write  $F(z)$  to denote the expected value of  $F$  under the distribution  $z$ , i.e.  $\sum_{\mathbf{a} \in \mathcal{A}} F(\mathbf{a}) z(\mathbf{a})$ . For instance, we denote the expected utility of player  $i$  under distribution  $z$  as  $u_i(z) := \sum_{\mathbf{a} \in \mathcal{A}} z(\mathbf{a}) u_i(\mathbf{a})$ .

An  $\epsilon$ - $\Phi$ -equilibrium is a correlated distribution over strategies such that

$$u_i(z) \geq u_i(\phi \circ_i z) - \epsilon \quad \forall i \in [n], \phi \in \Phi.$$

We call a  $\Phi$ -equilibrium a 0- $\Phi$ -equilibrium. In Appendix A we formally write the sets of functions  $\Phi_{\text{CE}}$  and  $\Phi_{\text{CCE}}$ , which define correlated (CE) and coarse-correlated equilibria (CCE), respectively. Intuitively, CCEs deviations are deviations that cannot depend on the recommended action, while CE deviations are a larger set where deviations could be different based on the recommended action.

### 2.2 Constrained $\Phi$ -equilibria

To model situations in which there are common shared resources, we introduce costs that depend on the joint actions of all players. In particular, each player  $i \in [n]$  has  $m$  costs  $\{C_i^j(\mathbf{a})\}_{j \in [m]} \in [-1, 1]^m$  associated with each tuple of action  $\mathbf{a} \in \mathcal{A}$ . For any  $\nu \geq 0$ , we define as  $\mathcal{S}_i^\nu$  the set of correlated strategies that are  $\nu$ -safe (in expectation) for the  $i$ -th player, i.e.,

$$\mathcal{S}_i^\nu := \left\{ z \in \Delta(\mathcal{A}) : C_i^j(z) \leq \nu \quad \forall j \in [m] \right\}.$$

Intuitively,  $\mathcal{S}_i^\nu$  guarantees that the expected cost of player  $i$  is at most  $\nu$  for each of its resources. Moreover, we define the set  $\mathcal{S}^\nu = \bigcap_{i \in [n]} \mathcal{S}_i^\nu$  as the set of strategies that are  $\nu$ -safe for all players. For

<sup>1</sup>Given an integer  $i \in \mathbb{N}$ , we denote with  $[i]$  the set  $\{1, \dots, i\}$ .

each  $z \in \mathcal{S}^\nu$  and player  $i \in [n]$  we can define the set of *safe-deviations*, *i.e.*,

$$\Phi_i^{\mathcal{S}}(z) := \{\phi \in \Phi : (\phi \circ_i z) \in \mathcal{S}_i\}.$$

To guarantee that the game is well-defined and the existence of solutions, we make the following mild assumption. This assumption requires that all agents always have a deviation that ensures that cost constraints are satisfied. Formally, for all  $z \in \Delta(\mathcal{A})$  and  $i \in [n]$ , it holds that  $\Phi_i^{\mathcal{S}}(z) \neq \emptyset$ . This assumption is required to guarantee that players always have safe strategies, and it is common in the literature. [Bernasconi et al. \[2023\]](#) proved existence (but not PPAD-membership) with a stronger notion related to strict feasibility, which was then removed in subsequent works [[Boufous et al., 2024](#), [Ni et al., 2025](#)]. This assumption is generally satisfied, as it is reasonable to let players have a void action with zero utility and zero cost, irrespective of other players' strategies. We say that a correlated strategy  $z \in \mathcal{S}^\nu$  is a Constrained  $(\epsilon, \nu)$ - $\Phi$ -equilibrium if it holds that  $u_i(z) \geq u_i(\phi \circ_i z) - \epsilon$  for all players  $i \in [n]$  and deviations in  $\phi \in \Phi_i^{\mathcal{S}}(z)$ . Intuitively, ignoring the technicalities related to the relaxation  $\nu$ , a correlated strategy  $z$  is a constrained  $\epsilon$ - $\Phi$ -equilibrium if  $z$  is safe for each player and each player cannot earn more than  $\epsilon$  by applying a safe deviation to the correlated strategy  $z$ .

### 2.3 Polymatrix and team games

Polymatrix games are a type of multiplayer games in which the  $n$  players interact only with a subset of other players and the game between any two players is a bi-matrix game. Formally, we can introduce the following computational problem:

#### POLYMATRIX

**Input:** a graph  $G = (V, E)$ , a matrix  $A^{i,j} \in [0, 1]^{k \times k}$  for each  $(i, j) \in E$ , and an approximation  $\epsilon > 0$ .

**Output:** vectors  $x_i \in \Delta([k])$ ,  $i \in V$  such that:

$$\sum_{j \in V: (i,j) \in E} x_i^\top A^{i,j} x_j \geq \sum_{j \in V: (i,j) \in E} \tilde{x}_i^\top A^{i,j} x_j - \epsilon \quad \forall i \in V, \tilde{x}_i \in \Delta([k]).$$

In a breakthrough result [Rubinstein \[2015\]](#) (later improved by [Deligkas et al. \[2022\]](#)) proved constant inapproximability of POLYMATRIX for constant number of actions and degree of the graph.

**Theorem 2.1** ([Rubinstein \[2015, Theorem 1\]](#)). *There exists a constant  $\epsilon^*$  such that POLYMATRIX is PPAD-complete, even when the graph has degree 3 and each player has 2 actions.*

An important subclass of polymatrix games are *two-team zero-sum games*, in which  $V$  can be partitioned into two teams such that: (i) each edge between players of different teams is a zero-sum game, and (ii) each edge between players of the same team is a coordination game. If one of the teams has no internal coordination edges, these games are called *two-team polymatrix zero-sum game with independent adversaries* [[Hollender et al., 2025](#)]. Here we also consider the case in which there are no coordination games in between both teams, *i.e.*, the graph  $G$  is fully bipartite. We call these *two-team polymatrix zero-sum game with independent teams*. These games are special cases of zero-sum polymatrix games that admit LP-based polynomial-time algorithms [[Cai et al., 2016](#)].

### 2.4 PPAD-hardness and Total Problems

The class PPAD is an important subclass of total search problems *i.e.*, TFNP, which includes search problems with solutions that can be checked in polynomial time. PPAD was introduced by [Papadimitriou \[1994\]](#) to capture problems whose totality is guaranteed by parity-like arguments and, equivalently, by fixed-point arguments such as Brouwer's. PPAD-complete problems are unlikely to be solvable in polynomial time. This is also supported by several lower bounds based on cryptographic assumptions [[Bitansky et al., 2015](#), [Garg et al., 2016](#), [Choudhuri et al., 2019](#)].

Moreover, for ease of presentation, we use a more manageable class of total problems, as we allow for promise problems, *i.e.*, problems whose totality is guaranteed only if the instance satisfies some additional properties (promises) whose validity is not clear how to check in polynomial time.<sup>2</sup>

<sup>2</sup>Formally, to guarantee the totality of our problems we should accept as solutions a violation of the promise (see, [Fearnley et al. \[2020, 2022\]](#), [Hollender \[2021\]](#) for an in-depth treatment of promise-preserving reductions.

### 3 Reduction from POLYMATRIX to CONSTRAINED- $\Phi$ -EQUILIBRIUM

We define a computational search problem associated with constrained  $\Phi$ -equilibria.

#### CONSTRAINED- $\Phi$ -EQUILIBRIUM

**Input:**  $\epsilon' > 0$ ,  $\nu > 0$ , an utility function  $u_i : \mathcal{A} \rightarrow [0, 1]$  and  $m$  cost functions  $C_i^j : \mathcal{A} \rightarrow [-1, 1]$ ,  $j \in [m]$  for each player  $i$ , and a set of possible deviations  $\Phi$  (given as a list of linear inequalities, defining a polytope of right stochastic matrices) with  $A \in \mathbb{R}^{k \times 2\ell}$  and  $b \in \mathbb{R}^k$ .

**Output:** a vector  $z \in \mathcal{S}^\nu$  such that  $u_i(z) \geq u_i(\phi \circ_i z) - \epsilon' \quad \forall i \in [n], \phi \in \Phi_i^S(z)$ .

**Promise:** It is promised that  $\Phi_i^S(z) \neq \emptyset$  for all  $z \in \Delta(\mathcal{A})$  and  $i \in [n]$ .

Formally, to fix a representation of the output, we think of  $z$  as a mixture of product distributions. Notice that such a representation is essential to guarantee a representation polynomial in the support and independent of  $|\mathcal{A}|$ .

In the following, we will provide a reduction from POLYMATRIX to CONSTRAINED- $\Phi$ -EQUILIBRIUM.

**Intuitive high-level idea of the proof** To give a high-level intuition of the reduction, we sketch a simplified reduction from the problem of computing Nash equilibria in two-player general-sum games. Given a general-sum game  $A, B \in [0, 1]^{k \times k}$ , we can write a feasibility LP for coarse correlated equilibria as follows:<sup>3</sup>

$$\text{Find } z \text{ such that: } \begin{cases} A(z) \geq A(\tilde{x} \otimes z_{-1}) \quad \forall \tilde{x} \in \Delta([k]) & (1a) \\ B(z) \geq B(z_{-2} \otimes \tilde{y}) \quad \forall \tilde{y} \in \Delta([k]) & (1b) \\ z \in \Delta([k]^2) \end{cases}$$

If we were to add a constraint forcing  $z$  to be a product distribution, *i.e.*,  $z = x \otimes y$ , then Equation (1) would turn into a feasibility problem whose solutions are Nash equilibria of  $A, B$ .<sup>4</sup> This (non-linear) constraint is exactly what makes computing Nash equilibria hard. Thus, ideally, we would like to impose constraints to force the correlated strategy to be a product distribution. However, the constraints we can employ are linear in  $z$ , and we cannot follow this idea directly.

By observing Problem 1, we can see that the right-hand sides of the two incentive compatibility (IC) constraints (1a) and (1b), are linear only in the marginals. Thus, it could be convenient: (i) impose the costs to the  $x$  player such that the marginal of the  $x$  player copies the marginal of the  $y$  player, *i.e.*,  $z_{-1} = z_{-2}$  (note that this is a linear constraint in  $z$ ), (ii) take  $A = -B$  so to erase the terms linear in  $z$  when summing the two IC constraints of Equation (1).

It is easy to see that, due to (i), we also have that the deviations  $\tilde{x}$  admissible to the  $x$  player are just  $\{z_{-1}\}$ . On the other hand, there are no costs on the second player, and thus the deviations  $\tilde{y}$  are free. Define  $h = z_{-1} = z_{-2}$ . Then summing (1a) and (1b) we obtain

$$A(z) + B(z) \geq A(h \otimes h) + B(h \otimes \tilde{y}) \xrightarrow{\text{Due to (ii)}} h^\top B h \geq h^\top B \tilde{y} \quad \forall \tilde{y} \in \Delta([k]).$$

It is easy to choose  $B$  such that it is PPAD-hard to find a solution to the right-hand side inequality above. For example, we can choose  $B = \begin{bmatrix} 0 & \tilde{A} \\ \tilde{B}^\top & 0 \end{bmatrix}$  for another two-player general-sum game

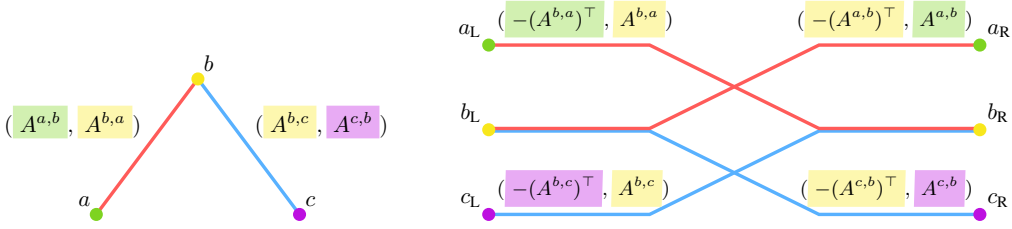
$\tilde{A}, \tilde{B}$ . This choice introduces two additional players for each player of the original game, essentially creating two opposing teams playing a zero-sum game.

The above construction is informal and skips many of the technical challenges, such as the fact that we are only allowed to use constraints that are satisfied approximately. Also, we derive a more general result and broader consequences by reducing from POLYMATRIX instead of computing Nash in two-player games (see Section 3.2). This is indeed one central part of the formal reduction that we will present in the next section.

Given a graph  $G = (V, E)$ , we denote with  $\deg(G)$  the maximum degree of graph  $G$ . Then, the main result of this section reads as follows:

<sup>3</sup>We denote with  $z_{-i}$  the marginalization of  $z$  with respect to the  $i$ -th player.

<sup>4</sup>Note that  $A(x \otimes y) = x^\top A y$ .



(a) Example of a POLYMATRIX instance with three players  $V = \{a, b, c\}$ . (b) Corresponding instance of CONSTRAINED- $\Phi$ -EQUILIBRIUM with  $\mathcal{N}_L = \{a_L, b_L, c_L\}$  and  $\mathcal{N}_R = \{a_R, b_R, c_R\}$ .

Figure 1: On the edges, we reported the matrix corresponding to the utilities of the players. Colors indicate the row player of the bi-matrix game and the player associated with the utility that is associated with the matrix.

**Theorem 3.1.** *There is a polynomial-time reduction from instances  $G = (V, E)$  of POLYMATRIX with  $n$  players,  $k$  actions per player, and approximation  $\epsilon$  to instances of CONSTRAINED- $\Phi$ -EQUILIBRIUM with approximation  $\epsilon' = O(\frac{\epsilon}{n})$  and  $\nu = O(\frac{\epsilon}{nk \deg(G)})$  and CCE deviations.*

### 3.1 Proof of Theorem 3.1

**Construction** Given an instance of POLYMATRIX with graph  $G = (V, E)$ , matrices  $\{A^{i,j}\}_{(i,j) \in E}$  and approximation  $\epsilon$ , we build an instance of CONSTRAINED- $\Phi$ -EQUILIBRIUM as follows:

- ▶ Set of players  $\mathcal{N} = \mathcal{N}_L \cup \mathcal{N}_R$ , where  $\mathcal{N}_L = \{i_L : i \in V\}$  and  $\mathcal{N}_R = \{i_R : i \in V\}$ ;
- ▶ For each player in  $\mathcal{N}$  the set of actions is  $A = [k]$  as in the original POLYMATRIX instance;
- ▶ For every  $i \in V$ , the utility of a player  $i_L \in \mathcal{N}_L$ , under an action profile  $\mathbf{a} \in A^{|\mathcal{N}|}$ , is defined as  $u_{i_L}(\mathbf{a}) = -\sum_{j:(i,j) \in E} (A^{j,i})^\top(a_{i_L}, a_{j_R})$ , while the utility of a player  $i_R \in \mathcal{N}_R$  under action profile  $\mathbf{a}$  is  $u_{i_R}(\mathbf{a}) = \sum_{j:(i,j) \in E} A^{i,j}(a_{i_R}, a_{j_L})$ ;
- ▶ The set of deviations is  $\Phi_{\text{CCE}}$  (see Appendix A for a definition);
- ▶ For each player  $i_L$  there are  $2k$  costs. For each  $j \in [2k]$ , let

$$C_{i_L}^j(\mathbf{a}) = \begin{cases} 2\mathbb{I}(j \leq k) - 1 & \text{if } a_{i_L} = j \pmod{k} \text{ and } a_{i_R} \neq j \pmod{k} \\ 2\mathbb{I}(j > k) - 1 & \text{if } a_{i_R} \neq j \pmod{k} \text{ and } a_{i_L} = j \pmod{k} \\ 0 & \text{otherwise} \end{cases}$$

For more intuition on the construction, we refer to Figure 1.

**Properties of the instance** The instance of CONSTRAINED- $\Phi$ -EQUILIBRIUM resulting from this construction has twice as many players as the original instance, and the same degree number of actions per player  $k$  as the original instance. Moreover, the CONSTRAINED- $\Phi$ -EQUILIBRIUM instance guarantees that each player has at most  $2k$  costs.

The game resulting from the reduction is a two-team polymatrix zero-sum game with independent teams. Indeed, we implicitly defined a graph  $\tilde{G} = (\mathcal{N}, \tilde{E})$  which has edges only between players  $\mathcal{N}_L$  and  $\mathcal{N}_R$ . More specifically, an edge is included between  $i_L \in \mathcal{N}_L$  and  $j_R \in \mathcal{N}_R$  if and only if there was an edge between  $i$  and  $j$  in the original POLYMATRIX instance. Moreover, the game is zero-sum. Indeed, for each couple of players  $i_L \in \mathcal{N}_L$  and  $j_R \in \mathcal{N}_R$ , the utilities are given by payoff matrices  $\tilde{A}^{i_L, j_R} = -(A^{j,i})^\top$  and  $\tilde{A}^{j_R, i_L} = A^{i,j}$ , respectively. Finally, the costs of player  $i_L$  only depend on the actions of player  $i_R$ .

Summing over the utilities of all players on each side, we get the following lemma:

**Lemma 3.2.** *For each tuple  $\mathbf{a} \in \mathcal{A}$ , define the utility of the “left team” as  $u_L(\mathbf{a}) = \sum_{i \in V} u_{i_L}(\mathbf{a})$  and the utility of the “right team” as  $u_R(\mathbf{a}) = \sum_{i \in V} u_{i_R}(\mathbf{a})$ . Then  $u_L(\mathbf{a}) = -u_R(\mathbf{a})$ .*

**Marginalization** Before delving into the correctness of the reduction, we need to introduce some additional notation that helps bridge between correlated equilibria and Nash equilibria. For any

subset of players  $\tilde{\mathcal{N}} \subseteq \mathcal{N}$  and any tuple of actions  $\mathbf{a} \in \mathcal{A}$ , we denote with  $\mathbf{a}_{\tilde{\mathcal{N}}}$  the actions pertaining to players in  $\tilde{\mathcal{N}}$ . We also denote with  $-\tilde{\mathcal{N}} = \mathcal{N} \setminus \tilde{\mathcal{N}}$  the complement of  $\tilde{\mathcal{N}}$ .

Now, we can define the marginalization of the correlated strategy  $z$  restricted to a set of players  $\tilde{\mathcal{N}}$  by summing over the actions of all other players. Formally, we write

$$m_{\tilde{\mathcal{N}}}(z|\mathbf{a}_{\tilde{\mathcal{N}}}) = \sum_{\mathbf{a}_{-\tilde{\mathcal{N}}} \in A^{|\mathcal{N} \setminus \tilde{\mathcal{N}}|}} z(\mathbf{a}_{\tilde{\mathcal{N}}}, \mathbf{a}_{-\tilde{\mathcal{N}}}).$$

We can then consider  $m_{\tilde{\mathcal{N}}}(z)$  as a vector indexed by all the tuples  $\mathbf{a}_{\tilde{\mathcal{N}}} \in A^{|\tilde{\mathcal{N}}|}$  which, clearly, is a distribution over  $A^{|\tilde{\mathcal{N}}|}$ .

**Effects on the costs** We now prove the main properties related to costs. Intuitively, every safe strategy guarantees that the marginals of players  $i_L$  are close to the marginals of the players  $i_R$  for each  $i \in V$ . In particular, each action  $j$  has two associated costs  $C_{i_L}^j$  and  $C_{i_L}^{j+k}$  that force the marginal of the first player to mimic the marginal of the second player. Namely,

**Lemma 3.3.** *Given a  $z \in \mathcal{S}^\nu$ , and  $i \in V$ , it holds that  $\|m_{i_L}(z) - m_{i_R}(z)\|_\infty \leq \nu$ . Moreover, given a  $z \in \Delta(A^{|\mathcal{N}|})$  and an  $i \in V$ , the set of safe deviations of player  $i_L$  is  $\Phi_{i_L}^S(z) = \{x_{i_L} : \|x_{i_L} - m_{i_R}(z)\|_\infty \leq 0\}$ , which guarantees  $\Phi_{i_L}^S(z) \neq \emptyset$ .*

**Correctness** We recall that  $k = |A|$  is the number of actions of each player and  $n = |V|$  is the number of players. Take any  $z \in \mathcal{S}^\nu$  that is a solution to CONstrained- $\Phi$ -EQUILIBRIUM, combining it with Lemma 3.3, we get that for all  $i \in V$  it holds:

$$u_{i_L}(z) \geq u_{i_L}(x_{i_L} \otimes m_{\mathcal{N} \setminus i_L}(z)) - \epsilon' \quad \forall x_{i_L} \in \Delta(A) : \|x_{i_L} - m_{i_R}(z)\|_\infty \leq 0 \quad (2)$$

and

$$u_{i_R}(z) \geq u_{i_R}(x_{i_R} \otimes m_{\mathcal{N} \setminus i_R}(z)) - \epsilon' \quad \forall x_{i_R} \in \Delta(A). \quad (3)$$

Notice that we can write explicitly the right-hand side of Equation (2) as

$$u_{i_L}(x_{i_L} \otimes m_{\mathcal{N} \setminus i_L}(z)) = - \sum_{j:(i,j) \in E} x_{i_L}^\top (A^{j,i})^\top m_{j_R}(z) = - \sum_{j:(i,j) \in E} m_{j_R}(z)^\top A^{j,i} x_{i_L},$$

We then specialize Equation (2) for  $x_{i_L} = m_{i_R}(z)$  and sum over all  $i \in V$ . This results in the following inequality (where we also swapped the identity of  $i$  and  $j$  in the last equality):

$$u_L(z) \geq - \sum_{(i,j) \in E} m_{j_R}(z)^\top A^{j,i} m_{i_R}(z) - n\epsilon' = - \sum_{(i,j) \in E} m_{i_R}(z)^\top A^{i,j} m_{j_R}(z) - n\epsilon'. \quad (4)$$

Similarly, we analyze Equation (3) and use the first statement of Lemma 3.3:

$$\begin{aligned} u_R(z) &\geq \sum_{i \in V} u_{i_R}(x_{i_R} \otimes m_{\mathcal{N} \setminus i_R}(z)) - n\epsilon' \\ &= \sum_{(i,j) \in E} x_{i_R}^\top A^{i,j} m_{j_L}(z) - n\epsilon' \\ &= \sum_{(i,j) \in E} x_{i_R}^\top A^{i,j} m_{j_R}(z) + \sum_{i \in V} \sum_{j:(i,j) \in E} x_{i_R}^\top A^{i,j} (m_{j_L}(z) - m_{j_R}(z)) - n\epsilon' \\ &\geq \sum_{(i,j) \in E} x_{i_R}^\top A^{i,j} m_{j_R}(z) - n(\nu k \deg(G) + \epsilon') \quad \forall \{x_{i_R}\}_{i \in V} \in \Delta(A)^n. \end{aligned} \quad (5)$$

Summing Equation (4) and Equation (5), and using the fact that  $u_R(z) = -u_L(z)$  from Lemma 3.2, we obtain that

$$\sum_{(i,j) \in E} m_{i_R}(z)^\top A^{i,j} m_{j_R}(z) \geq \sum_{(i,j) \in E} x_{i_R}^\top A^{i,j} m_{j_R}(z) - n(\nu k \deg(G) + 2\epsilon') \quad \forall \{x_{i_R}\}_{i \in V} \in \Delta(A)^n.$$

Which, for all players  $i \in V$  can be specialized to

$$\sum_{j:(i,j) \in E} m_{i_R}(z)^\top A^{i,j} m_{j_R}(z) \geq \sum_{j:(i,j) \in E} x_{i_R}^\top A^{i,j} m_{j_R}(z) - n(\nu k \deg(G) + 2\epsilon') \quad \forall x_{i_R} \in \Delta(A),$$

by setting  $x_{j_R} = m_{j_R}(z)$  for all  $j \neq i$ .

Thus, taking  $h_i = m_{i_R}(z)$  for all players  $i \in V$ , and  $\epsilon' = \epsilon/4n$  and  $\nu = \epsilon/(2nk \deg(G))$  we prove that  $h_i$  is a solution of POLYMATRIX with approximation  $\epsilon$ .

### 3.2 Implications of the PPAD-hardness

The class of deviations used in the reduction of Theorem 3.1 is  $\Phi_{\text{CCE}}$ , which is a subset of the deviations of  $\Phi_{\text{CE}}$ , thus, our hardness works also for correlated equilibrium in generalized games.

We now discuss some special cases in which our problem is PPAD-hard. POLYMATRIX is PPAD-hard even when  $\epsilon, \deg(G), k = O(1)$  (see Theorem 2.1). Thus, Theorem 3.1 implies that CONSTRAINED- $\Phi$ -EQUILIBRIUM is hard when each player has  $m = O(1)$  constraints and actions, when the slackness  $\nu$  and approximation  $\epsilon$  are  $O(n^{-1})$ .

Moreover, the instances resulting from our reduction form a bipartite graph and, more specifically, a two-team zero-sum polymatrix game with independent teams, as defined in Section 2.3. Importantly, for this case of two-team zero-sum polymatrix game with no constraints, there are known LP-based polynomial-time algorithms [Cai et al., 2016]. It is an interesting open question whether it is possible to prove hardness for constant approximations, arbitrary number of players, constant number of actions per player, and constant number of constraints per player.

We can also consider the bi-matrix instance of POLYMATRIX in which we have only 2 players and an arbitrary number of actions  $k$ . This shows that our problem is PPAD-hard even with a constant number of players when  $\epsilon' = O(1)$  and  $\nu = O(k^{-1})$ .

Moreover, assuming the Exponential Time Hypothesis for PPAD (namely that solving the canonical problem END-OF-THE-LINE requires exponential time), Rubinfeld [2017] showed that it must take at least  $k^{\log^{1-o(1)}(k)}$  time to find a constant approximation for any  $k \times k$  bi-matrix game. This imposes a quasi-polynomial lower bound on the CONSTRAINED- $\Phi$ -EQUILIBRIUM problem for constant  $\epsilon$  and  $\nu = \text{poly}(k^{-1})$ . On the other hand, Bernasconi et al. [2023, Corollary 4.3] provides a quasi-polynomial time algorithm for constant approximation and a constant number of players in  $k^{\log(k)}$  time, thus essentially closing the gap.

## 4 PPAD-membership

We now introduce a computational problem related to quasi-variational inequalities that will serve as the basis for our membership reduction.

### QUASIVI

**Inputs:**  $G, L, \epsilon > 0, \nu > 0$ , a circuit implementing a  $G$ -Lipschitz function  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and two circuits implementing a  $L$ -Lipschitz continuous matrix valued function  $A : [0, 1]^d \rightarrow \mathbb{R}^{n \times d}$  and a  $L$ -Lipschitz continuous vector valued function  $b : \mathbb{R}^n \rightarrow \mathbb{R}^d$  defining the correspondence  $Q_\nu(\tilde{z}) := \{z \in [0, 1]^d : A(\tilde{z})z \leq b(\tilde{z}) + \nu 1_d\}$ .<sup>a</sup>

**Output:** a point  $z \in Q_\nu(z)$  such that  $F(z)^\top(\tilde{z} - z) \geq -\epsilon$ , for all  $\tilde{z} \in Q_\nu(z)$ .

**Promise:** The correspondence is promised to satisfy  $Q_0(z) \neq \emptyset$  for all  $z \in [0, 1]^d$ , the function  $F$  is  $G$ -Lipschitz and the correspondence  $Q_0(z)$  is  $L$ -Lipschitz.

<sup>a</sup>We say that a matrix-valued function  $A : [0, 1]^d \rightarrow \mathbb{R}^{n \times d}$  is  $L$ -Lipschitz if the function  $\tilde{z} \mapsto A(\tilde{z})z$  is  $L$ -Lipschitz for every  $z \in [0, 1]^d$  in the  $\ell_2$ -norm.

**Theorem 4.1** (Bernasconi et al. [2024, Theorem 3.4]). QUASIVI  $\in$  PPAD.

To prove that CONSTRAINED- $\Phi$ -EQUILIBRIUM is total and more specifically in PPAD we reduce it to QUASIVI. Our idea is to restrict to Nash-like equilibria with CCE deviations, *i.e.*, restrict to product strategies  $\times_{i \in [n]} p_i \in \Delta(A)^n$ , where  $p_i \in \Delta(A)$ . We will show that these are a subset of constrained  $\Phi$ -equilibria for any set of deviations  $\Phi$ .

**Proposition 4.2.** CONSTRAINED- $\Phi$ -EQUILIBRIUM  $\in$  PPAD.

*Proof.* The proof hinges on showing that for any instance of CONSTRAINED- $\Phi$ -EQUILIBRIUM with CCE deviations: (i) we can build an instance of QUASIVI in polynomial time and (ii) from the solution of QUASIVI, we can build a solution to the original CONSTRAINED- $\Phi$ -EQUILIBRIUM instance which is a product distribution. We show that considering CCE deviations is enough, since these deviations are the most expressive when applied to product distributions. Indeed, the image



of any product distribution under the set of deviations  $\Phi_{\text{CCE}}$  is a superset of the image under any possible set of phi deviations. Formally, given a player  $i \in [n]$ , a product distribution  $p = \otimes_{i \in [n]} p_i$ , a set of deviations  $\Phi$ , and a deviation  $\phi \in \Phi$ , there exists a marginal  $\tilde{p}_i$  of the  $i$ -th player such that  $\phi \circ_i p = \tilde{p}_i \otimes p_{-i}$ . In particular, we can take  $\tilde{p}_i \in \Delta(A)$  such that  $\tilde{p}_i(a_i) = \sum_{b \in A} \phi(b, a_i) p_i(b)$  for each  $a_i \in A$ . Then, a CCE deviation can replicate the marginal  $\tilde{p}_i$  by taking  $\phi \in \Phi_{\text{CCE}}$  such that each row of the stochastic matrix is  $\tilde{p}_i$ . This also proves that there exists a safe CCE deviations for every product distribution, from the existence of a safe deviation of the original set of deviations  $\Phi$ .

Indeed, any instance of **CONSTRAINED- $\Phi$ -EQUILIBRIUM** is guaranteed by assumption to have the set of feasible deviations non-empty for each correlated distribution. As a direct consequence of our previous observation, we have that the set of feasible deviations in CCE is not empty when considering only product distributions  $z$ . This can be proven by observing that the strategies obtained by applying deviations from any set  $\Phi$  to product distributions form a subset of the strategies obtained by applying CCE deviations.

**Construction** To define an instance of **QUASIVI**, we define a correspondence  $Q : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$  and an operator  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying the appropriate conditions of the **QUASIVI** problem. Namely, we need the correspondence  $Q(\tilde{z})$  to be linear for every  $\tilde{z}$  and with Lipschitz coefficients. Moreover, we also require that the operator  $F$  is Lipschitz.<sup>5</sup>

Since we only consider product distributions,  $\ell n$  numbers suffice to uniquely specify the distribution. Thus, we can “flatten” the product strategies  $p = \otimes_{i \in [n]} p_i$  into  $z = [p_1^\top | \dots | p_n^\top]^\top$  (equivalently  $\tilde{z}$  is the flattening of  $\tilde{p}$ ) and consider the correspondence  $Q(\tilde{z}) = \{z \in [0, 1]^{\ell n} : A(\tilde{z})z \leq b(\tilde{z})\}$ , that encodes both the costs and the simplex constraints. Formally:  $A(\tilde{z}) = \text{diag}(A_1(\tilde{z}), \dots, A_n(\tilde{z})) \in \mathbb{R}^{n(m+2) \times \ell n}$  and  $b(\tilde{z}) = [b_1(\tilde{z})^\top | \dots | b_n(\tilde{z})^\top]^\top \in \mathbb{R}^{n(m+2)}$ , where  $A_i(z)$  is a block matrix  $A_i(\tilde{z}) = \begin{bmatrix} D_i(\tilde{z}) \\ 1_\ell^\top \\ -1_\ell^\top \end{bmatrix} \in \mathbb{R}^{(m+2) \times \ell}$ ,  $b_i(\tilde{z}) = \begin{bmatrix} 0_m \\ 1 \\ -1 \end{bmatrix} \in \mathbb{R}^{m+2}$ , and  $D_i(\tilde{z}) \in \mathbb{R}^{m \times \ell}$  is such that

$$[0_{m \times \ell}, \dots, D_i(\tilde{z}), \dots, 0_{m \times \ell}]z \leq 0_m \iff C_i^j(p_i \otimes \tilde{p}_{-i}) \leq 0 \forall j \in [m].$$

The explicit definition of  $D_i(\tilde{z})$  is cumbersome due to the flattening notation and we defer it to Appendix C.1.

It is easy to check that  $z \in Q(z)$  if the corresponding “unflattened” product strategies  $p$  satisfy the simplex and cost constraints for each player. The next claim shows that the instances created by our reduction satisfy the promises needed by the **QUASIVI** problem.

**Claim 4.3.** *The functions  $\tilde{z} \mapsto A(\tilde{z})z$  and  $\tilde{z} \mapsto b(\tilde{z})^\top z$  defining the correspondence are  $L$ -Lipschitz for every  $z \in [0, 1]^{\ell n}$  where  $L$  has a representation polynomial in the size of the instance.*

Now, we build the operator  $F$  of the QVI by stacking the gradients of the utilities. For any product distribution  $p = \otimes_{i \in [n]} p_i$  and flattening  $z$  we define, with slight abuse of notation,  $F(z) := (-\nabla_{p_1} u_1(p), \dots, -\nabla_{p_n} u_n(p))$ . The following claims that the operator  $F$  satisfies the properties required by the **QUASIVI** problem. This shows that the constructed instance satisfies all the promises of the **QUASIVI** problem.

**Claim 4.4.** *The operator  $F : [0, 1]^{\ell n} \rightarrow [0, 1]^{\ell n}$  is  $G$ -Lipschitz where  $G$  has a representation polynomial in the size of the instance.*

**Correctness** For every vector  $r \in [0, 1]^{\ell n}$ , define  $r^i = r_{[\ell(i-1), \dots, \ell i - 1]}$ , which effectively divides (and “unflattens”) the vector  $r$  in  $n$  components of length  $\ell$  each corresponding to the strategies of the  $i$ -th player. We will use this notation to reconstruct a solution to **CONSTRAINED- $\Phi$ -EQUILIBRIUM** from a solution  $z$  of **QUASIVI**. Indeed, take a solution  $z$  of the **QUASIVI** problem, instantiated with  $\epsilon'$  and  $\nu'$  later to be defined. We now claim that the renormalization  $p = \otimes_{i \in [n]} p_i$  (where

<sup>5</sup>Crucially, the Lipschitz constants are inputs of the **QUASIVI** problem, and thus, we need to make sure that they have a polynomial representation in the original **CONSTRAINED- $\Phi$ -EQUILIBRIUM** instance.

$p_i = z^i / \|z^i\|_1$ ) is a solution to CONstrained- $\Phi$ -EQUILIBRIUM.<sup>6</sup> The only tricky part is simply handling the fact that all the constraints are only approximately satisfied.

**Claim 4.5.** *Let  $\nu' = \min(\frac{\epsilon}{2}, \frac{\nu^2}{2n})$  and  $\epsilon' = \frac{\epsilon}{2}(1 - n\nu')$ , then  $p$  is such that*

$$u_i(p) \geq u_i(\tilde{p}_i \otimes p_{-i}) - \epsilon' \quad \forall \tilde{p}_i \in \Phi_i^S(p)$$

and  $p \in \mathcal{S}^{\nu'}$ .

This concludes the proof, as it shows that  $p$  (which is a product distribution) is a CONstrained- $\Phi$ -EQUILIBRIUM for CCE deviations, with approximation  $\epsilon$  and slackness  $\nu$  and thus a CONstrained- $\Phi$ -EQUILIBRIUM for any set  $\Phi$ .  $\square$

## Acknowledgments

The work of MB and AC was partially funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

MB and AC are supported by an ERC grant (Project 101165466 — PLA-STEER) and by MUR - PRIN 2022 project 2022R45NBB funded by the Next Generation EU program. MC is partially supported by the FAIR (Future Artificial Intelligence Research) project PE0000013, funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, investment 1.3, line on Artificial Intelligence) and the EU Horizon project ELIAS (European Lighthouse of AI for Sustainability, No. 101120237). GF is supported in part by the National Science Foundation award CCF-2443068, ONR grant N000142512296, and an AI2050 Early Career Fellowship.

The authors also thank Tingting Ni for helpful discussion and suggestions.

## References

- E. Adida and G. Perakis. Dynamic pricing and inventory control: Uncertainty and competition. *Operations Research*, 58(2):289–302, 2010.
- D. Agarwal, S. Ghosh, K. Wei, and S. You. Budget pacing for targeted online advertisements at linkedin. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1613–1619, 2014.
- G. Aggarwal, A. Badanidiyuru, and A. Mehta. Autobidding with constraints. In *Web and Internet Economics: 15th International Conference, WINE 2019*, pages 17–30. Springer, 2019.
- I. Anagnostides, I. Panageas, T. Sandholm, and J. Yan. The complexity of symmetric equilibria in min-max optimization and team zero-sum games. *arXiv preprint arXiv:2502.08519*, 2025.
- K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954.
- R. J. Aumann. Correlated equilibrium as an expression of bayesian rationality. *Econometrica: Journal of the Econometric Society*, pages 1–18, 1987.
- S. Balseiro, Y. Deng, J. Mao, V. Mirrokni, and S. Zuo. Robust auction design in the auto-bidding world. *Advances in Neural Information Processing Systems*, 34:17777–17788, 2021.
- S. R. Balseiro and Y. Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.
- M. Bernasconi, M. Castiglioni, A. Marchesi, F. Trovo, and N. Gatti. Constrained phi-equilibria. In *International Conference on Machine Learning*, pages 2184–2205. PMLR, 2023.

<sup>6</sup>Note that  $p$  is a valid probability distribution over actions as  $\|p_i\|_1 = 1$  by construction, and  $p_i$  (as well as  $z^i$ ) has only positive components since the relaxation  $\nu > 0$  required by the QUASIVI does not relax the hypercube constraints.

- M. Bernasconi, M. Castiglioni, A. Celli, and G. Farina. On the role of constraints in the complexity of min-max optimization. *arXiv preprint arXiv:2411.03248*, 2024.
- N. Bitansky, O. Paneth, and A. Rosen. On the cryptographic hardness of finding a nash equilibrium. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1480–1498. IEEE, 2015.
- O. Boufous, R. El-Azouzi, M. Touati, E. Altman, and M. Bouhtou. Constrained correlated equilibria. In *2024 60th Annual Allerton Conference on Communication, Control, and Computing*, pages 1–8. IEEE, 2024.
- S. Brânzei, Y. Chen, X. Deng, A. Filos-Ratsikas, S. Frederiksen, and J. Zhang. The fisher market game: Equilibrium and welfare. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.
- N. Brown and T. Sandholm. Superhuman ai for multiplayer poker. *Science*, 365(6456):885–890, 2019.
- Y. Cai, O. Candogan, C. Daskalakis, and C. Papadimitriou. Zero-sum polymatrix games: A generalization of minmax. *Mathematics of Operations Research*, 41(2):648–655, 2016.
- N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- X. Chen, X. Deng, and S.-H. Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the ACM (JACM)*, 56(3):1–57, 2009.
- A. R. Choudhuri, P. Hubáček, C. Kamath, K. Pietrzak, A. Rosen, and G. N. Rothblum. Finding a nash equilibrium is no easier than breaking fiat-shamir. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1103–1114, 2019.
- V. Conitzer, C. Kroer, E. Sodomka, and N. E. Stier-Moses. Multiplicative pacing equilibria in auction markets. *Operations Research*, 70(2):963–989, 2022.
- C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.
- C. Daskalakis, S. Skoulakis, and M. Zampetakis. The complexity of constrained min-max optimization. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1466–1478, 2021.
- C. Daskalakis, G. Farina, M. Fishelson, C. Pipis, and J. Schneider. Efficient learning and computation of linear correlated equilibrium in general convex games. *arXiv preprint arXiv:2412.20291*, 2024.
- A. Deligkas, J. Fearnley, A. Hollender, and T. Melissourgos. Pure-circuit: Strong inapproximability for ppad. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 159–170. IEEE, 2022.
- M. F. A. R. D. T. FAIR, A. Bakhtin, N. Brown, E. Dinan, G. Farina, C. Flaherty, D. Fried, A. Goff, J. Gray, H. Hu, et al. Human-level play in the game of diplomacy by combining language models with strategic reasoning. *Science*, 378(6624):1067–1074, 2022.
- G. Farina and C. Pipis. Polynomial-time computation of exact  $\phi$ -equilibria in polyhedral games. In *Advances in Neural Information Processing Systems*, volume 37. Curran Associates, Inc., 2024.
- J. Fearnley, S. Gordon, R. Mehta, and R. Savani. Unique end of potential line. *Journal of Computer and System Sciences*, 114:1–35, 2020.
- J. Fearnley, P. Goldberg, A. Hollender, and R. Savani. The complexity of gradient descent:  $\text{CLS} = \text{PPAD} \cap \text{PLS}$ . *Journal of the ACM*, 70(1):1–74, 2022.
- S. Garg, O. Pandey, and A. Srinivasan. Revisiting the cryptographic hardness of finding a nash equilibrium. In *Annual International Cryptology Conference*, pages 579–604. Springer, 2016.

- I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.
- A. Greenwald and A. Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop, COLT/Kernel 2003, Washington, DC, USA, August 24-27, 2003. Proceedings*, pages 2–12. Springer, 2003.
- S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.
- B. F. Hobbs and J.-S. Pang. Nash-cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. *Operations Research*, 55(1):113–127, 2007.
- A. Hollender. *Structural results for total search complexity classes with applications to game theory and optimisation*. PhD thesis, University of Oxford, 2021.
- A. Hollender, G. Maystre, and S. G. Nagarajan. The complexity of two-team polymatrix games with independent adversaries. In *International Conference on Learning Representations ICLR*, 2025.
- M. Mahdavi, R. Jin, and T. Yang. Trading regret for efficiency: online convex optimization with long term constraints. *The Journal of Machine Learning Research*, 13(1):2503–2528, 2012.
- S. Mannor, J. N. Tsitsiklis, et al. Online learning with sample path constraints. *Journal of Machine Learning Research*, 10(3), 2009.
- H. Moulin and J. P. Vial. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7:201–221, 1978.
- J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.
- T. Ni, A. Maddux, and M. Kamgarpour. On characterization and existence of a constrained correlated equilibria in markov games. *arXiv preprint arXiv:2507.03502*, 2025.
- N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.
- R. Paes Leme, G. Piliouras, J. Schneider, K. Spendlove, and S. Zuo. Complex dynamics in auto-bidding systems. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 75–100, 2024.
- C. Papadimitriou, E.-V. Vlatakis-Gkaragkounis, and M. Zampetakis. The computational complexity of multi-player concave games and kakutani fixed points. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 1045–1045, 2023.
- C. 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.
- C. H. Papadimitriou and T. Roughgarden. Computing equilibria in multi-player games. In *SODA*, volume 5, pages 82–91. Citeseer, 2005.
- J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.
- A. Rubinfeld. Inapproximability of nash equilibrium. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 409–418, 2015.
- A. Rubinfeld. Settling the complexity of computing approximate two-player nash equilibria. *ACM SIGecom Exchanges*, 15(2):45–49, 2017.
- B. H. Zhang, I. Anagnostides, E. Tewolde, R. E. Berker, G. Farina, V. Conitzer, and T. Sandholm. Expected variational inequalities. *ICML 2025-42st International Conference on Machine Learning*, 2025.

## 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: The introduction and abstract accurately describe the main results of the paper.

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: Limitations of the model are discussed in the paper.

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 the claims are supported by formal proof.

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: This is mainly a theoretical paper, and no experiments are performed.

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: This is mainly a theoretical paper, and no experiments are performed.

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, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA]

Justification: This is mainly a theoretical paper, and no experiments are performed.

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: This is mainly a theoretical paper, and no experiments are performed.

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: This is mainly a theoretical paper, and no experiments are performed.

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 authors conformed 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: This is mainly a theoretical paper, and there are no direct negative societal impacts.

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: There is no model or data released with the paper.

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](https://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: The 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: The 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: The 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 rigor, or originality of the research, declaration is not required.

Answer: [NA]

Justification: the core method development in this research 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.

## A Examples of $\Phi$ -equilibria

We now introduce two notable examples of sets  $\Phi$ , namely Correlated and Coarse Correlated equilibria.

Correlated equilibria (CEs) are obtained by considering all possible deviation strategies, *i.e.*,

$$\Phi_{\text{CE}} = \left\{ \phi \in [0, 1]^{\ell \times \ell} : \sum_{b \in A} \phi(a, b) = 1 \quad \forall a \in A \right\}.$$

This set models a player that can observe its own recommendation and deviate to any other action with some probability.

Another important class of equilibria is Coarse Correlated Equilibria (CCEs) that are defined by the set

$$\Phi_{\text{CCE}} = \left\{ \phi \in [0, 1]^{\ell \times \ell} : \phi(a, b) = \phi(a', b) \quad \forall a, a' \in A \wedge \sum_{b \in A} \phi(a, b) = 1 \quad \forall a \in A \right\}.$$

This models a player whose deviations are forced to be equal for each recommended action  $a \in A$ . Intuitively, the player has to decide their own deviation strategy before seeing the recommended action  $a$ . This greatly simplifies the possible deviation since  $\phi \in \Phi_{\text{CCE}}$  can simply be identified with the marginals it induces.

## B Missing proof from Section 3 (hardness)

**Lemma 3.2.** *For each tuple  $\mathbf{a} \in \mathcal{A}$ , define the utility of the “left team” as  $u_{\text{L}}(\mathbf{a}) = \sum_{i \in V} u_{i_{\text{L}}}(\mathbf{a})$  and the utility of the “right team” as  $u_{\text{R}}(\mathbf{a}) = \sum_{i \in V} u_{i_{\text{R}}}(\mathbf{a})$ . Then  $u_{\text{L}}(\mathbf{a}) = -u_{\text{R}}(\mathbf{a})$ .*

*Proof.* The statement follows from straightforward calculations

$$\begin{aligned} u_{\text{L}}(\mathbf{a}) &= \sum_{i \in V} u_{i_{\text{L}}}(\mathbf{a}) \\ &= \sum_{i \in V} \sum_{j: (i,j) \in E} \tilde{A}^{i_{\text{L}}, j_{\text{R}}}(a_{i_{\text{L}}}, a_{j_{\text{R}}}) \\ &= - \sum_{i \in V} \sum_{j: (i,j) \in E} (A^{j,i})^{\top}(a_{i_{\text{L}}}, a_{j_{\text{R}}}) && (\tilde{A}^{i_{\text{L}}, j_{\text{R}}} = -(A^{j,i})^{\top}) \\ &= - \sum_{i \in V} \sum_{j: (i,j) \in E} (A^{i,j})^{\top}(a_{j_{\text{L}}}, a_{i_{\text{R}}}) && \text{(By swapping the sum and the identity of } i \text{ and } j) \\ &= - \sum_{i \in V} \sum_{j: (i,j) \in E} A^{i,j}(a_{i_{\text{R}}}, a_{j_{\text{L}}}) \\ &= - \sum_{i \in V} \sum_{j: (i,j) \in E} \tilde{A}^{i_{\text{R}}, j_{\text{L}}}(a_{i_{\text{R}}}, a_{j_{\text{L}}}) && (\tilde{A}^{i_{\text{R}}, j_{\text{L}}} = A^{i,j}) \\ &= - \sum_{i \in V} u_{i_{\text{R}}}(\mathbf{a}) \\ &= -u_{\text{R}}(\mathbf{a}), \end{aligned}$$

concluding the proof. □

**Lemma 3.3.** *Given a  $z \in S^{\nu}$ , and  $i \in V$ , it holds that  $\|m_{i_{\text{L}}}(z) - m_{i_{\text{R}}}(z)\|_{\infty} \leq \nu$ . Moreover, given a  $z \in \Delta(A^{|\mathcal{N}|})$  and an  $i \in V$ , the set of safe deviations of player  $i_{\text{L}}$  is  $\Phi_{i_{\text{L}}}^{\text{S}}(z) = \{x_{i_{\text{L}}} : \|x_{i_{\text{L}}} - m_{i_{\text{R}}}(z)\|_{\infty} \leq 0\}$ , which guarantees  $\Phi_{i_{\text{L}}}^{\text{S}}(z) \neq \emptyset$ .*

*Proof.* Consider the case  $j \leq k$ :

$$\begin{aligned}
C_{i_L}^j(z) &= \sum_{\mathbf{a} \in \mathcal{A}} C_{i_L}^j(\mathbf{a}) z(\mathbf{a}) \\
&= \sum_{\mathbf{a} \in A^n: a_{i_L}=j, a_{i_R} \neq j} z(\mathbf{a}) - \sum_{\mathbf{a} \in A^n: a_{i_L} \neq j, a_{i_R}=j} z(\mathbf{a}) \\
&= \sum_{\mathbf{a} \in \mathcal{A}: a_{i_L}=j} z(\mathbf{a}) - \sum_{\mathbf{a} \in \mathcal{A}: a_{i_R}=j} z(\mathbf{a}) \\
&= m_{i_L}(z|j) - m_{i_R}(z|j)
\end{aligned}$$

and thus  $C_{i_L}^j(z) \leq \nu$ , with  $j \leq k$  implies that  $m_{i_L}(z|j) - m_{i_R}(z|j) \leq \nu$ . On the other hand  $C_{i_L}^j(z) \leq 0$ , with  $j > k$  implies that  $m_{i_R}(z|j) - m_{i_L}(z|j) \leq 0$  which concludes the statement.  $\square$

## C Missing proofs and additional details from Section 4 (membership)

### C.1 Explicit definition of $A(\tilde{z})$

The correspondence  $Q : [0, 1]^{\ell n} \Rightarrow [0, 1]^{\ell n}$  is given by  $Q(\tilde{z}) = \{z : A(\tilde{z})z \leq b(\tilde{z})\}$  where

$$A(\tilde{z}) = \begin{bmatrix} \overbrace{D_1(\tilde{z})}^{\ell} & & & \\ \mathbf{1}_\ell^\top & & & \\ -\mathbf{1}_\ell^\top & & & \\ \hline & D_2(\tilde{z}) & & \\ & \mathbf{1}_\ell^\top & & \\ & -\mathbf{1}_\ell^\top & & \\ \hline & & \ddots & \\ \hline & & & D_n(\tilde{z}) \\ & & & \mathbf{1}_\ell^\top \\ & & & -\mathbf{1}_\ell^\top \end{bmatrix} \in \mathbb{R}^{n \cdot (m+2) \times \ell n}$$

where  $D_i(\tilde{z}) \in \mathbb{R}^{m \times \ell}$  is a matrix such that  $[0_{m \times \ell}, \dots, D_i(\tilde{z}), \dots, 0_{m \times \ell}]z \leq 0_m \iff C_i^j(\tilde{p})p_i \leq 0 \forall j \in [m]$ . In particular, for each  $i$ , only the components of  $z$  corresponding to the strategies of the  $i$ -th player matter and correspond to the strategy  $p_i$ . We define  $g$  the flattening function which takes a product distribution  $\tilde{p} = \bigotimes_{i \in [n]} \tilde{p}_i$  and returns the corresponding ‘‘unflattened’’ vector  $\tilde{z}$  while we call  $h$  its inverse. Thus, we can write that  $D_i(g(\tilde{p}))p_i \leq 0_m$  if and only if  $C_i^j(p_i \otimes \tilde{p}_{-i}) \leq 0$  for all  $j \in [m]$ . Notice that  $p_i \mapsto C_i^j(p_i \otimes \tilde{p}_{-i})$  is linear and can be written as  $C_i^j(p_i \otimes \tilde{p}_{-i}) = c_i^j(\tilde{p})^\top p_i$ , where  $c_i^j(\tilde{p}) \in \mathbb{R}^\ell$  and each component  $c_i^j(\tilde{p})_{\bar{a}}$ , indexed by  $\bar{a}$ , is given by  $\sum_{\mathbf{a} \in \mathcal{A}: a_i = \bar{a}} C_i^j(\mathbf{a}) \prod_{k \neq i} \tilde{p}_k(a_k)$ . Consequently, we can take

$$D_i(\tilde{z}) = \begin{bmatrix} c_i^1(h(\tilde{z})) \\ \vdots \\ c_i^m(h(\tilde{z})) \end{bmatrix} \in \mathbb{R}^{m \times \ell n}$$

### C.2 Proof of Claim 4.3 and Claim 4.4 from Proposition 4.2

**Claim 4.3.** *The functions  $\tilde{z} \mapsto A(\tilde{z})z$  and  $\tilde{z} \mapsto b(\tilde{z})^\top z$  defining the correspondence are  $L$ -Lipschitz for every  $z \in [0, 1]^{\ell n}$  where  $L$  has a representation polynomial in the size of the instance.*

*Proof.* First note that  $b(\tilde{z})$  does not depend on  $\tilde{z}$  and thus is trivially 0-Lipschitz. Thus, we only need to prove the statement about  $A(\tilde{z})$ . We recall the exact definition of  $A(\tilde{z})$  given in Appendix C.1.

We analyze the Jacobian of the  $D_i(\tilde{z})$ . Any entry of  $D_i(\tilde{z})$  would correspond to a cost  $j$  of the  $i$ -th player (rows) and to an action  $\hat{a} \in A$  (columns), and any component  $\ell$  of  $\tilde{z}$  would correspond to a player  $i'$  (possibly different from  $i$ ) and an action  $\bar{a} \in A$ . Thus, by defining  $\tilde{p} = h(\tilde{z})$ , we can compute the following

$$\frac{\partial c_i^j(h(\tilde{z}))_{\hat{a}}}{\partial \tilde{z}_\ell} = \frac{\partial c_i^j(\tilde{p})_{\hat{a}}}{\partial \tilde{p}_{i'}(\bar{a})} \frac{\partial \tilde{p}_{i'}(\bar{a})}{\partial \tilde{z}_\ell}$$

The second term is clearly 1, as the function  $h$  just rearranges the components  $\tilde{z}$ , while the first term is easily bounded as follows

$$\begin{aligned} \left| \frac{\partial c_i^j(\tilde{p})_{\hat{a}}}{\partial \tilde{p}_{i'}(\bar{a})} \right| &= \left| \frac{\partial \sum_{\mathbf{a} \in \mathcal{A}: a_i = \hat{a}} C_i^j(\mathbf{a}) \prod_{k \in [n], k \neq i} \tilde{p}_k(a_k)}{\partial \tilde{p}_{i'}(\bar{a})} \right| \\ &= \left| \sum_{\mathbf{a} \in \mathcal{A}: a_{i'} = \bar{a}, a_i = \hat{a}} C_i^j(\mathbf{a}) \prod_{k \in [n], k \neq i', k \neq i} \tilde{p}_k(a_k) \right| \leq \ell^n. \end{aligned}$$

The following elementary lemma lets us conclude the proof.

**Lemma C.1.** *Let  $M : \mathbb{R}^K \rightarrow \mathbb{R}^{m \times n}$  be a matrix valued function such that  $\left| \frac{\partial M_{i,j}(\tilde{z})}{\partial \tilde{z}_k} \right| \leq C$  for all  $i \in [m], j \in [n], k \in [K]$  then*

$$\|(M(\tilde{z}) - M(\tilde{z}'))z\| \leq Cm\sqrt{nK}\|\tilde{z} - \tilde{z}'\|,$$

for all  $z \in [0, 1]^K$ .

Indeed,

$$\begin{aligned} \|(A(\tilde{z}) - A(\tilde{z}'))z\| &\leq \ell^n(\ell n)\sqrt{\ell n \cdot n(m + \ell + 2)}\|\tilde{z} - \tilde{z}'\| \\ &\leq 2\ell^{n+2}n^2\sqrt{m}\|\tilde{z} - \tilde{z}'\| \end{aligned}$$

and thus  $L = \text{poly}(\ell^n, m, n)$  concluding the proof.  $\square$

**Claim 4.4.** *The operator  $F : [0, 1]^{\ell n} \rightarrow [0, 1]^{\ell n}$  is  $G$ -Lipschitz where  $G$  has a representation polynomial in the size of the instance.*

*Proof.* We can get a simple upper bound on the Lipschitz constant of  $F$  by bounding its gradient. In particular  $F(z) = (-\nabla_{p_1} u_1(p), \dots, -\nabla_{p_n} u_n(p))$ , where as usual  $z$  is the unrolling of the product distribution  $p = \bigotimes_{i=1}^n p_i$ . We can consider any component of  $F$ , which will correspond to some player  $i \in [n]$  and action  $\bar{a} \in A$ , and consider some component of  $z$  which will correspond to some player  $j \in [n]$  and some action  $\tilde{a} \in A$ . The component of  $F$  selected corresponds to  $-\frac{\partial u_i(p)}{\partial p_i(\bar{a})}$ . We can then consider the following:

$$-\frac{\partial^2 u_i(p)}{\partial p_j(\tilde{a}) \partial p_i(\bar{a})} = - \sum_{\mathbf{a} \in \mathcal{A}: a_i = \bar{a}, a_j = \tilde{a}} u_i(\mathbf{a}) \prod_{k \neq i, j} p_k(a_k)$$

and thus  $\left| \frac{\partial^2 u_i(p)}{\partial p_j(\tilde{a}) \partial p_i(\bar{a})} \right| \leq \ell^n$ . The mean value theorem trivially concludes the proof:

$$\|F(z) - F(z')\| \leq \|J_F(\xi)\| \cdot \|z - z'\|$$

for some  $\xi$  on the segment connecting  $z$  and  $z'$ . Now for any matrix  $M \in \mathbb{R}^{m \times n}$  it holds that  $\|M\| \leq \sqrt{mn} \cdot \sup_{i,j} |M_{i,j}|$  and thus  $\|J_F(\xi)\| \leq n\ell^{n+1} = G$ , concluding the proof.  $\square$

**Claim 4.5.** *Let  $\nu' = \min(\frac{\epsilon}{2}, \frac{\nu^2}{2n})$  and  $\epsilon' = \frac{\epsilon}{2}(1 - n\nu')$ , then  $p$  is such that*

$$u_i(p) \geq u_i(\tilde{p}_i \otimes p_{-i}) - \epsilon' \quad \forall \tilde{p}_i \in \Phi_i^S(p)$$

and  $p \in \mathcal{S}^\nu$ .

*Proof.* First, observe that if  $\nu \geq 1$  then the costs bear no effects and PPAD-membership can be established by the PPAD-membership of Nash equilibria. Thus, we can assume w.l.o.g. that  $\nu \leq 1$ .

**Step 1: Safety** First, we claim that  $p$  is  $\nu$ -safe. Indeed,  $\|z^i\|_1 \in [1 - \nu', 1 + \nu']$  and for each  $i \in [n], j \in [m]$  and we can directly compute

$$\sum_{\mathbf{a} \in A^n} C_i^j(\mathbf{a}) \prod_{j \in [n]} z^j(a_j) \leq \nu',$$

then, we can divide the left and right hand side by  $\prod_{j \in [n]} \|z^j\|_1 \geq (1 - \nu')^n$  and, obtain that:

$$C_i^j(p) = \sum_{\mathbf{a} \in A^n} C_i^j(\mathbf{a}) \prod_{j \in [n]} p_j(a_j) \leq \frac{\nu'}{(1 - \nu')^n} \leq \frac{\nu'}{1 - n\nu'} \leq \nu,$$

where in the last inequality we used  $\nu' \leq \frac{\nu^2}{2n} \leq \frac{\nu}{1+n\nu}$  for all  $\nu \geq 0$  and  $n \geq 1$ . This shows that indeed  $p \in \mathcal{S}^\nu$ .

**Step 2: Simulating deviations** Now we prove that we can simulate safe deviations  $\tilde{p}_i$  with appropriate choices of  $\tilde{z}$ .

Take any safe deviation  $\tilde{p}_i \in \Delta(A)$  (i.e., it holds that  $C_i^j(\tilde{p}_i \otimes p_{-i}) \leq 0$ ), and consider  $\tilde{z}$  defined as  $\tilde{z} = [z^1, \dots, \tilde{z}^i, \dots, z^n]$ , where  $\tilde{z}^i = \tilde{p}_i$ . By the definition of the correspondence (see Appendix C.1 for notation and details on the explicit construction of the correspondence  $Q$ ) we have that  $\tilde{z} \in Q_{\nu'}(z)$ . Indeed, from the definition of  $D_i$  in Appendix C.1, we get that

$$\begin{aligned} [D_i(z)\tilde{z}^i]_j &= c_i^j(h(z))^\top \tilde{z}^i && \text{(by construction)} \\ &= \sum_{\bar{\mathbf{a}} \in A} \sum_{\mathbf{a} \in A^n, a_i = \bar{a}} C_i^j(\mathbf{a}) \prod_{k \neq i} z^k(a_k) \tilde{z}^i(\bar{a}) && \text{(def. of } c_i^j(h(z))) \\ &= \sum_{\mathbf{a} \in A^n, a_i} C_i^j(\mathbf{a}) \prod_{k \neq i} z^k(a_k) \tilde{z}^i(a_i) \\ &\leq 0 && (C_i^j(\tilde{p}_i \otimes p_{-i}) \leq 0) \end{aligned}$$

Moreover, it is easy to verify that  $1^\top \tilde{z}^i \in [1 - \nu', 1 + \nu']$  for all  $i \in [n]$ . This shows that  $\tilde{z} \in Q_{\nu'}(z)$ .

Formally, what we proved is that: for any player  $i \in [n]$  and for any safe deviation  $\tilde{p}_i$  such that  $C_i^j(\tilde{p}_i \otimes p_{-i}) \leq 0$ , there exists  $\tilde{z} \in Q_{\nu'}(z)$  which ‘‘simulates’’  $\tilde{p}_i$ . Moreover,  $\tilde{z}^j = z^j$  for all  $j \neq i$  and  $\tilde{z}^i = \tilde{p}_i$ .

**Step 3: Equilibrium conditions** Now, we also claim that  $p = \bigotimes_{i \in [n]} p_i$  satisfies the equilibrium constraints. Since  $z$  is a solution to QUASI-VI we have that for all  $\tilde{z} \in Q_{\nu'}(z)$  the following holds:

$$F(z)^\top (\tilde{z} - z) \geq -\epsilon' \implies - \sum_{i \in [n]} \sum_{\bar{\mathbf{a}} \in A} \sum_{\mathbf{a} \in A: a_i = \bar{a}} u_i(\mathbf{a}) \prod_{j \neq i} z^j(a_j) (\tilde{z}^i(\bar{a}) - z^i(\bar{a})) \geq -\epsilon',$$

which implies that:

$$\sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \neq i} z^j(a_j) (z^i(a_i) - \tilde{z}^i(a_i)) \geq -\epsilon',$$

once we specialize to  $\tilde{z}$  to the one built in step 2, since  $\tilde{z}^j = z^j$  for all  $j \neq i$ .

Now we can use the exact definition of  $\tilde{z}^i$  constructed in step 2, in the above equation, which leads to the following:

$$\sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \neq i} z^j(a_j) (z^i(a_i) - \tilde{z}^i(a_i)) = \sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \in [n]} z^j(a_j) - \sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \neq i} z^j(a_j) \tilde{p}_i(a_i) \geq -\epsilon'.$$

Rearranging it, we obtain

$$\sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \in [n]} z^j(a_j) \geq \sum_{\mathbf{a} \in A} u_i(\mathbf{a}) \prod_{j \neq i} z^j(a_j) \tilde{p}_i(a_i) - \epsilon'.$$

We can now divide each term by  $\gamma = \prod_{j \in [n]} \|z^j\|_1$ , which finally leads to

$$u_i(p) \geq \frac{u_i(\tilde{p}_i \otimes p_{-i})}{\|z^i\|_1} - \frac{\epsilon'}{\gamma} \geq u_i(\tilde{p}_i \otimes p_{-i}) - \left(\frac{\epsilon'}{\gamma} + \nu'\right)$$

and since  $\|z^j\|_1 \in [1 - \nu', 1 + \nu']$  for all  $j \in [n]$  this implies that  $\gamma \geq (1 - \nu')^n \geq 1 - n\nu'$  and thus  $u_i(p) \geq u_i(\tilde{p}_i \otimes p_{-i}) - \left(\frac{\epsilon'}{1 - n\nu'} + \nu'\right)$ .

Consider the first term  $\frac{\epsilon'}{1 - n\nu'}$ . The following inequalities hold:

$$\frac{\epsilon'}{1 - n\nu'} \leq \frac{\epsilon}{2}$$

since  $\epsilon' \leq \frac{\epsilon}{2}(1 - n\nu')$ .

Moreover,  $\nu' \leq \frac{\epsilon}{2}$  thus proving that  $u_i(p) \geq u_i(\tilde{p}_i \otimes p_{-i}) - \epsilon$  as desired.  $\square$

### C.3 Additional technical lemmas

**Lemma C.1.** Let  $M : \mathbb{R}^K \rightarrow \mathbb{R}^{m \times n}$  be a matrix valued function such that  $\left| \frac{\partial M_{i,j}(\tilde{z})}{\partial \tilde{z}_k} \right| \leq C$  for all  $i \in [m], j \in [n], k \in [K]$  then

$$\|(M(\tilde{z}) - M(\tilde{z}'))z\| \leq Cm\sqrt{nK}\|\tilde{z} - \tilde{z}'\|,$$

for all  $z \in [0, 1]^K$ .

*Proof.* Let  $\{m_i : [0, 1]^K \rightarrow \mathbb{R}^n\}_{i=1}^m$  be the functions defining the rows of  $M$  and  $h_i(\tilde{z}|z) = m_i(\tilde{z})^\top z$ . With this notation it is easy to check that  $\nabla_{\tilde{z}} h_i(\tilde{z}|z) = J_{m_i}(\tilde{z})^\top z$  and thus  $\|\nabla_{\tilde{z}} h_i(\tilde{z}|z)\| \leq \|J_{m_i}(\tilde{z})\| \|z\| \leq C\sqrt{mnK}$ .

By the mean value theorem, we have that for some  $\xi$  in the segment connecting  $\tilde{z}$  and  $\tilde{z}'$ , we have

$$\begin{aligned} |(m_i(\tilde{z}) - m_i(\tilde{z}'))^\top z| &\leq \|\nabla_{\tilde{z}} h(\xi|z)\| \cdot \|\tilde{z} - \tilde{z}'\| \\ &\leq C\sqrt{mnK}\|\tilde{z} - \tilde{z}'\| \end{aligned}$$

Finally,

$$\begin{aligned} \|(M(\tilde{z}) - M(\tilde{z}'))z\|^2 &= \sum_{i=1}^m ((m_i(\tilde{z}) - m_i(\tilde{z}'))^\top z)^2 \\ &\leq C^2 m^2 nK \|\tilde{z} - \tilde{z}'\|^2 \end{aligned}$$

concluding the proof.  $\square$