Learning Sparse Polymatrix Games in Polynomial Time and Sample Complexity

Asish Ghoshal
aghoshal@purdue.edu

Jean Honorio
jhonorio@purdue.edu

Department of Computer Science
Purdue University, West Lafayette, IN, U.S.A. - 47907.

Abstract

We consider the problem of learning sparse polymatrix games from observations of strategic interactions. We show that a polynomial time method based on $\ell_{1,2}$-group regularized logistic regression recovers a game, whose Nash equilibria are the ϵ-Nash equilibria of the game from which the data was generated (true game), in $\mathcal{O}\left(m^4 d^4 \log(pd)\right)$ samples of strategy profiles — where m is the maximum number of pure strategies of a player, p is the number of players, and d is the maximum degree of the game graph. Under slightly more stringent separability conditions on the payoff matrices of the true game, we show that our method learns a game with the exact same Nash equilibria as the true game. We also show that $\Omega\left(d \log(pm)\right)$ samples are necessary for any method to consistently recover a game, with the same Nash-equilibria as the true game, from observations of strategic interactions. We verify our theoretical results through simulation experiments.

1 Introduction and Related Work

Motivation. Many complex real-world data can be thought of as resulting from the behavior of a large number of self-interested trying to myopically or locally maximize some utility. Over the past several decades, non-cooperative game theory has emerged as a powerful mathematical framework for reasoning about such strategic interactions between self-interested agents. Traditionally, research in game theory has focused on

computing the *Nash equilibria* (NE) (c.f. [BSK06] and [JLB11]) — which characterizes the stable outcome of the overall behavior of self-interested agents — *correlated equilibria* (c.f. [KKLO03]), and other solution concepts given a description of the game. Computing the *price of anarchy* (PoA) for graphical games, which in a sense quantifies the *inefficiency of equilibria*, is also of tremendous interest (c.f. [BZR11]). The aforementioned problems of computing the NE, correlated equilibria and PoA can be thought of as *inference problems* in graphical games, and require a description of the game, i.e., the payoffs of the players. In many real-world settings, however, only the behavior of the agents are observed, in which case inferring the latent payoffs of the players from observations of behavioral data becomes imperative. *This problem of learning a game from observations of behavioral data, i.e., recovering the structure and parameters of the player payoffs such that the Nash equilibria of the game, in some sense, approximates the Nash equilibria of the true game, is the primary focus of the paper.*

Recovering the underlying game from behavioral data is an important tool in exploratory research in political science and behavioral economics, and recent times have seen a surge of interest in such problems (c.f. [IO14, HO15, GH16, GJ16, GH17]). For instance, in political science, [IO14] identified the *most influential* senators in the U.S congress — a small coalition of senators whose collective behavior forced every other senator to a unique choice of action — by learning a *linear influence game* from congressional voting records. [GJ16] showed that a *tree-structured polymatrix game* [1] learned from U.S. Supreme Court data was able to recover the known ideologies of the justices. However, many open problems remain in this area of active research. One such problem is whether there exists efficient (polynomial time) methods for learning polymatrix games [Jan68] from noisy observations of

---

[1][GJ16] call their game a *potential game* even though the formulation of their game is similar to ours.

strategic interactions. This is the focus of the current paper.

Related Work. Various methods have been proposed for learning games from data. [HO15] proposed a maximum-likelihood approach to learn "linear influence games" — a class of parametric graphical games with linear payoffs. However, in addition to being exponential time, the maximum-likelihood approach of [HO15] also assumed a specific observation model for the strategy profiles. [GH16] proposed a polynomial time algorithm, based on ℓ_1-regularized logistic regression, for learning linear influence games. They again assumed the specific observation model proposed by [HO15] in which the strategy profiles (or joint actions) were drawn from a mixture of uniform distributions: one over the pure-strategy Nash equilibria (PSNE) set, and the other over the complement of the PSNE set. [GH17] obtained necessary and sufficient conditions for learning linear influence games under arbitrary observation model. Finally, [GJ16] use a discriminative, max-margin based approach, to learn tree structured polymatrix games. However, their method is exponential time and they show that learning polymatrix games is NP-hard under this max-margin setting, even when the class of graphs is restricted to trees. Furthermore, all the aforementioned works, with the exception of [GJ16], consider binary strategies only. In this paper, we propose a polynomial time algorithm for learning polymatrix games, which are non-parametric graphical games where the pairwise payoffs between players are characterized by matrices (or pairwise potential functions). In this setting, each player has a finite number of pure-strategies.

Our Contributions. We propose an $\ell_{1,2}$ group-regularized logistic regression method to learn polymatrix games, which has been considered by [GJ16] and is a generalization of linear influence games considered by [GH17]. We make no assumptions on the latent payoff functions and show that our polynomial time algorithm recovers an ε-Nash equilibrium of the true game [2], with high probability, if the number of samples is $\mathcal{O}\left(m^4 d^4 \log(pd)\right)$, where p is the number of players, d is the maximum degree of the game graph and m is the maximum number of pure-strategies of a player. Under slightly more stringent separability conditions on the payoff functions of the underlying game, we show that our method recovers the Nash equilibria set exactly. We further generalize the observation model from [GH17] in the sense that we allow strategy profiles

---

[2]By the phrase "recovering the Nash equilibria" we mean that we learn a game with the same Nash equilibria as the true game. We use this phrase elsewhere in the paper for brevity.

in the non-Nash equilibria set to have zero measure. This should be compared with the results of [GJ16] who show that learning tree-structured polymatrix games is NP-hard under a max-margin setting. We also obtain necessary conditions on learning polymatrix games and show that $\Omega\left(d \log(pm)\right)$ samples are required by any method for recovering the PSNE set of a polymatrix game from observations of strategy profiles.

Finally, we conclude this section by referring the reader to the work of [JRVS11] who analyze $\ell_{1,2}$-regularized logistic regression for learning undirected graphical models. However, our setting differs from that of learning discrete graphical models in many ways. First, unlike discrete graphical models, where the underlying distribution over the variables is described by a potential function that factorizes over the cliques of the graph, we make no assumptions whatsoever on the generative distribution of data. Further, we are interested in recovering the PSNE set of a game, since the graph structure in generally unidentifiable from observational data, whereas [JRVS11] obtain guarantees on the graph structure of the discrete graphical model. As a result, our theoretical analysis and proofs differ significantly from those of [JRVS11].

2 Notation and Problem Formulation

In this section, we introduce our notation and formally define the problem of learning polymatrix games from behavioral data.

Polymatrix games. A p-player *polymatrix game* is a graphical game where the set of nodes of the graph denote players and the edges correspond to two-player games. We will denote the graph by $G = ([p], E)$, where $[p] \stackrel{\text{def}}{=} \{1, \ldots, p\}$ is the vertex set and $E \subseteq [p] \times [p]$ is set of directed edges. An edge $(i, j) \in E$ denotes the directed edge $i \leftarrow j$. Each player i has a set of pure-strategies or actions $\mathcal{A}_i$, and the set of pure-strategy profiles or joint actions of all the p players is denoted by $\mathcal{A} = \times_{i \in [p]} \mathcal{A}_i$. We will denote $\mathcal{A}_{-i} \stackrel{\text{def}}{=} \times_{j \in -i} \mathcal{A}_j$. With each edge $(i, j) \in E$ is associated a payoff matrix $u^{i,j} : \mathcal{A}_i \times \mathcal{A}_j \to \mathbb{R}$, such that $u^{i,j}(x_i, x_j)$ gives the finite payoff of the i-th player (with respect to the j-th player), when player i plays $x_i \in \mathcal{A}_i$ and player j plays $x_j \in \mathcal{A}_j$. We assume that $(i, j) \in E$, if and only if $u^{i,j}(\cdot, \cdot) \neq 0$. Given a strategy profile $\mathbf{x} \in \mathcal{A}$, the total payoff, or simply the payoff, of the i-th player is given by the following potential function:

$$u^i(x_i, \mathbf{x}_{-i}; G) = u^{i,i}(x_i) + \sum_{j \in \mathcal{N}_i} u^{i,j}(x_i, x_j), \qquad (1)$$

where $\mathcal{N}_i(G) \stackrel{\text{def}}{=} \{j \in [p] | (i, j) \in E\}$ is the set of neighbors of i in the graph G, and $u^{i,i} : \mathcal{A}_i \to \mathbb{R}$

gives the (finite) *individual payoff* of i for playing x_i. We will denote the number of neighbors of player i by $d_i \overset{\text{def}}{=} |\mathcal{N}_i(G)|$, and the maximum degree of the graph G by $d = \max\{d_1, \ldots, d_p\}$. A polymatrix game $\mathcal{G} = (G, \mathcal{U})$ is then completely defined by a graph $G = ([p], E)$ and a collection of potential functions $\mathcal{U}(G) = \{u^i : \mathcal{A}_{-i} \to \mathbb{R}\}_{i \in [p]}$, where each of the payoff functions $u^i(\cdot; G)$ decomposes according to (1). Finally, we will also assume that the number of strategies of each player, $m_i \overset{\text{def}}{=} |\mathcal{A}_i|$, is non-zero and $\mathcal{O}(1)$ with respect to p and d, and that $m \overset{\text{def}}{=} \max\{m_i\}$.

Nash equilibria of polymatrix games. The pure-strategy Nash equilibria (PSNE) set for the game $\mathcal{G} = (G, \mathcal{U})$ is given by the set of strategy profiles where no player has any incentive to unilaterally deviate from its strategy given the strategy profiles of its neighbors, and is defined as follows:

$$\mathcal{NE}(\mathcal{G}) = \left\{ \mathbf{x} \in \mathcal{A} \;\middle|\; x_i \in \operatorname*{argmax}_{a \in \mathcal{A}_i} u^i(a, \mathbf{x}_{-i}) \right\}. \quad (2)$$

The set of ε-Nash equilibria of the game $\mathcal{G}$ are those strategy profiles where each player can gain at most ε payoff by deviating from its strategy, and is defined as follows:

$$\varepsilon\text{-}\mathcal{NE}(\mathcal{G}) = \Big\{ \mathbf{x} \in \mathcal{A} \mid u^i(x_i, \mathbf{x}_{-i}) \geq u^i(a, \mathbf{x}_{-i}) - \varepsilon,$$
$$\forall a \in \mathcal{A}_i \text{ and } \forall i \in [p] \Big\}. \quad (3)$$

Observation model. Without getting caught up in the dynamics of gameplay — something that is difficult to observe or reason about in real-world scenarios — we abstract the learning problem as follows. Assume that we are given "noisy" observations of strategy profiles, or joint actions, $\mathcal{D} = \{\mathbf{x}^{(l)} \in \mathcal{A}\}_{l \in [n]}$ drawn from a game $\mathcal{G} = (G, \mathcal{U})$. The noise process models our uncertainty over the individual actions of the players due to observation noise, for instance, when we observe the actions through a noisy channel, or due to the unobserved dynamics of gameplay during which equilibrium is reached. By "observations drawn from a game" we simply mean that there exists a distribution $\mathcal{P}$, from which the strategy profiles are drawn, satisfying the following condition:

$$\forall \mathbf{x}, \mathbf{x}' \text{ such that } \mathbf{x} \in \mathcal{NE}(\mathcal{G}) \text{ and } \mathbf{x}' \in \mathcal{A} \setminus \mathcal{NE}(\mathcal{G}):$$
$$\mathcal{P}(\mathbf{x}) > \mathcal{P}(\mathbf{x}').$$

The above condition ensures that the signal level is more than the noise level. This should be compared with the observation model of [GH17], who assume that $\forall \mathbf{x}' \in \mathcal{A} \setminus \mathcal{NE}(\mathcal{G}), \mathcal{P}(\mathbf{x}') > 0$. Our observation model thus encompasses specific observation models considered in prior literature [HO15, GH16]: the global

and local noise model. The global noise model is parameterized by a constant $q \in (\mathcal{NE}(\mathcal{G})/|\mathcal{A}|, 1)$ such that the probability of observing a strategy profile $\mathbf{x} \in \mathcal{A}$ is given by a mixture of two uniform distributions:

$$\mathcal{P}_g(\mathbf{x}; \mathcal{G}) = q \frac{\mathbf{1}\left[\mathbf{x} \in \mathcal{NE}(\mathcal{G})\right]}{|\mathcal{NE}(\mathcal{G})|} + (1 - q) \frac{\mathbf{1}\left[\mathbf{x} \notin \mathcal{NE}(\mathcal{G})\right]}{|\mathcal{A}| - |\mathcal{NE}(\mathcal{G})|}. \quad (4)$$

In the local noise model, we observe strategy profiles $\mathbf{x}$ from the PSNE set with each entry (strategy) corrupted independently. Therefore, in the local noise model we have the following distribution over strategy profiles:

$$\mathcal{P}_l(\mathbf{x}; \mathcal{G}) = \frac{1}{|\mathcal{NE}(\mathcal{G})|} \times$$
$$\sum_{\mathbf{y} \in \mathcal{NE}(\mathcal{G})} \prod_{i=1}^{p} (q_i)^{\mathbf{1}[x_i = y_i]} \left(\frac{1 - q_i}{m_i - 1}\right)^{\mathbf{1}[x_i \neq y_i]}, \quad (5)$$

with $q_i > 0.5$ for all $i \in [p]$.

In essence, we assume that we observe multiple "stable outcomes" of the game, which may or may-not be in equilibria. Treating the outcomes of the game as "samples" observed across multiple "plays" of the same game is a recurring theme in the literature for learning games (c.f. [HO15], [GH16], [GH17], [GJ16]).

The learning problem then corresponds to recovering a game $\widehat{\mathcal{G}} = (\widehat{G}, \widehat{\mathcal{U}})$ from $\mathcal{D}$ such that $\mathcal{NE}(\widehat{\mathcal{G}}) = \mathcal{NE}(\mathcal{G})$ with high probability. *Given that computing a single Nash equilibria is PPAD-complete [DGP09], any efficient learning algorithm must learn the game without explicitly computing or enumerating the Nash equilibria of the game.* It has also been shown that even computing an ε-Nash equilibria is hard under the exponential time hypothesis for PPAD [Rub16]. We also emphasize that we do not observe any information about the latent player payoffs, and neither do we impose any restrictions on the payoffs for obtaining our ε-Nash equilibria guarantees. Also, note that in our definition of the learning problem, we do not impose any restriction on the "closeness" of the recovered graph $\widehat{G}$ to the true graph G. This is because multiple graphs G can give rise to the same PSNE set under different payoff functions and thus be *unidentifiable* from observations of joint actions alone (see section 4.4.1 of [HO15] for a counter example.)

3 Method

In this section, we describe our method for learning polymatrix games from observational data. The individual and pairwise payoffs can be equivalently written, in linear form, as follows:

$$u^{i,i}(x_i) = (\boldsymbol{\theta}^{i,0})^T \mathbf{f}^{i,0}(x_i),$$
$$u^{i,j}(x_i, x_j) = (\boldsymbol{\theta}^{i,j})^T \mathbf{f}^{i,j}(x_i, x_j),$$

where for $j \in \mathcal{N}_i$, $\mathbf{f}^{i,j}(x_i, x_j) = (\mathbf{1}[x_i = a, x_j = b])_{a \in \mathcal{A}_i, b \in \mathcal{A}_j}$ and $\boldsymbol{\theta}^{i,j} = (\theta^{i,j}_{a,b})_{a \in A_i, b \in A_j}$, $\mathbf{f}^{i,0}(x_i) = (\mathbf{1}[x_i = a])_{a \in A_i}$ and $\boldsymbol{\theta}^{i,0} = (\theta^{i,0}_a)_{a \in A_i}$. Note that $\mathbf{f}^{i,j} \in \{0,1\}^{(m_i m_j)}$, $\boldsymbol{\theta}^{i,j} \in \mathbb{R}^{(m_i m_j)} \neq \mathbf{0}$, $\mathbf{f}^{i,0}(x_i) \in \{0,1\}^{m_i}$, and $\boldsymbol{\theta}^{i,0} \in \mathbb{R}^{m_i}$. Let

$$\boldsymbol{\theta}^i \overset{\text{def}}{=} (\boldsymbol{\theta}^{i,0}, \boldsymbol{\theta}^{i,1}, \ldots, \boldsymbol{\theta}^{i,i-1}, \boldsymbol{\theta}^{i,i+1}, \ldots, \boldsymbol{\theta}^{i,p}),$$

$$\mathbf{f}^i(x_i, \mathbf{x}_{-i}) \overset{\text{def}}{=} (\mathbf{f}^{i,0}(x_i), \mathbf{f}^{i,1}(x_i, x_1), \ldots, \mathbf{f}^{i,i-1}(x_i, x_{i-1}),$$
$$\mathbf{f}^{i,i+1}(x_i, x_{i+1}), \ldots, \mathbf{f}^{i,p}(x_i, x_p)), \qquad (6)$$

with $\boldsymbol{\theta}^{i,j} = \mathbf{0}$ for $j > 0 \wedge j \notin \mathcal{N}_i$, and $\boldsymbol{\theta}^i \in \mathbb{R}^{(m_i + \sum_{j \in -i} m_i m_j)}, \mathbf{f}^i(x_i, \mathbf{x}_{-i}) \in \{0,1\}^{(m_i + \sum_{j \in -i} m_i m_j)}$. Thus the payoff for the i-th player can be written, in linear form, as:

$$u^i(x_i, \mathbf{x}_{-i}) = (\boldsymbol{\theta}^i)^T \mathbf{f}^i(x_i, \mathbf{x}_{-i}). \qquad (7)$$

The learning problem then corresponds to learning the parameters $\boldsymbol{\theta}^i$ for each player i. The sparsity pattern of $\boldsymbol{\theta}^i$ identifies the neighbors of i. The way this differs from the binary strategies considered by [GH17] is that the parameters $\boldsymbol{\theta}^i$ have a *group-sparsity* structure, i.e., for all $j > 0 \wedge j \notin \mathcal{N}_i$ the entire group of parameters $\boldsymbol{\theta}^{i,j}$ is zero. In order to ensure that the payoffs are finite, we will assume that the parameters for the i-th player belong to the set $\Theta^i \overset{\text{def}}{=} \{\mathbf{y} \in \mathbb{R}^{(m_i + \sum_{j \in -i} m_i m_j)} \mid \|\mathbf{y}\|_\infty < \infty\}$.

Our approach for estimating the parameters $\boldsymbol{\theta}^i$ is to perform one-versus-rest multinomial logistic regression with $\ell_{1,2}$ group-sparse regularization. In more detail, we obtain estimators $\widehat{\boldsymbol{\theta}}^i$ by solving the following optimization problem for each $i \in [p]$:

$$\widehat{\boldsymbol{\theta}}^i = \underset{\boldsymbol{\theta} \in \Theta^i}{\operatorname{argmin}} L^i(\mathcal{D}; \boldsymbol{\theta}) + \lambda \|\boldsymbol{\theta}\|_{1,2}, \qquad (8)$$

$$L^i(\mathcal{D}; \boldsymbol{\theta}) = \frac{1}{n} \sum_{l=1}^n \ell^i(\mathbf{x}^{(l)}; \boldsymbol{\theta}), \qquad (9)$$

$$\ell^i(\mathbf{x}; \boldsymbol{\theta}) = -\log \left(\frac{\exp(\boldsymbol{\theta}^T \mathbf{f}^i(x_i, \mathbf{x}_{-i}))}{\sum_{a \in \mathcal{A}_i} \exp(\boldsymbol{\theta}^T \mathbf{f}^i(a, \mathbf{x}_{-i}))} \right), \quad (10)$$

where $\|\boldsymbol{\theta}\|_{1,2} = \sum_{j \in [p]} \|\boldsymbol{\theta}_j\|_2$, with $\boldsymbol{\theta}_j$ being the j-th group of $\boldsymbol{\theta}$. When referring to a block of a matrix or vector we will use bold letters, e.g, $\boldsymbol{\theta}_j$ denotes the j-th group or block of $\boldsymbol{\theta}$, while θ_j denotes the j-th element of $\boldsymbol{\theta}$. In general, we define the $\ell_{a,b}$ group structured norm as follows: $\|\boldsymbol{\theta}\|_{a,b} = \|(\|\boldsymbol{\theta}_1\|_b, \ldots, \|\boldsymbol{\theta}_p\|_b)\|_a$. Also, when using group structured norms, we will use the group structure as shown in (6), i.e., we will assume that there are p groups and, in the context of the i-th player, the sizes of the groups are: $\{m_i, m_i m_1, \ldots, m_i m_{i-1}, m_i m_{i+1}, \ldots, m_i m_p\}$. Finally, we will define the support set of $\boldsymbol{\theta}^i$ as the set of all indices corresponding to the active groups, i.e.,

$S_i = \{(j,k) | j \in \{0\} \cup \mathcal{N}_i$ and $k \in [m_i]$ for $j = 0, k \in [m_i m_j]$ for $j > 0\}$, where j can be thought of as indexing the groups, while k can be thought of as the indexing the elements within the j-th group. Thus, $|S_i| = m_i + \sum_{j \in \mathcal{N}_i} m_i m_j$.

After estimating the parameters $\widehat{\boldsymbol{\theta}}^i$ for each $i \in [p]$, the payoff functions are simply estimated to be $\widehat{u}^i(x_i, \mathbf{x}_{-i}) = (\widehat{\boldsymbol{\theta}}^i)^T \mathbf{f}^i(x_i, \mathbf{x}_{-i})$. Finally, the graph $\widehat{G} = ([p], \widehat{E})$ is given by the group-sparsity structure of $\widehat{u}^i$s, i.e., $\widehat{u}^{i,j}(\cdot, \cdot) \neq 0 \implies (i,j) \in \widehat{E}$.

4 Sufficient Conditions

First, we obtain sufficient conditions on the number of samples n to ensure successful PSNE recovery. Since our theoretical results depend on certain properties of the Hessian of the loss function defined above, we introduce the Hessian matrix in this paragraph. Let $\mathbf{H}^i(\mathbf{x}; \boldsymbol{\theta})$ denote the Hessian of $\ell^i(\mathbf{x}; \boldsymbol{\theta})$. A little calculation shows that the (j,k)-th block of the Hessian matrix for the i-th player is given as:

$$\mathbf{H}^i_{j,k}(\mathbf{x}; \boldsymbol{\theta}) = \sum_{a \in \mathcal{A}_i} \sigma^i(a, \mathbf{x}_{-i}; \boldsymbol{\theta}) \mathbf{f}^{i,j}(a, x_j) (\mathbf{f}^{i,k}(a, x_k))^T -$$
$$\left\{ \left(\sum_{a \in \mathcal{A}_i} \sigma^i(a, \mathbf{x}_{-i}; \boldsymbol{\theta}) \mathbf{f}^{i,j}(a, x_j) \right) \times \right.$$
$$\left. \left(\sum_{a \in \mathcal{A}_i} \sigma^i(a, \mathbf{x}_{-i}; \boldsymbol{\theta}) \mathbf{f}^{i,k}(a, x_k) \right)^T \right\}, \qquad (11)$$

$$\sigma^i(x, \mathbf{x}_{-i}; \boldsymbol{\theta}) = \frac{\exp(\boldsymbol{\theta}^T \mathbf{f}^i(x, \mathbf{x}_{-i}))}{\sum_{a \in \mathcal{A}_i} \exp(\boldsymbol{\theta}^T \mathbf{f}^i(a, \mathbf{x}_{-i}))}, \qquad (12)$$

where we have overloaded the notation $\mathbf{f}^{i,j}(x_i, x_j)$ to also include $\mathbf{f}^{i,0}(x_i)$, i.e., we let $\mathbf{f}^{i,0}(x_i, x_0) \overset{\text{def}}{=} \mathbf{f}^{i,0}(x_i)$. We will denote the i-th expected Hessian matrix at any parameter $\boldsymbol{\theta} \in \Theta^i$ as $\mathbf{H}^i(\boldsymbol{\theta}) = \mathbb{E}_{\mathbf{x}}[\mathbf{H}^i(\mathbf{x}; \boldsymbol{\theta})]$, and the i-th Hessian matrix at the true parameter $\boldsymbol{\theta}^i$ as $\mathbf{H}^i(\boldsymbol{\theta}^i)$. We will also drop the superscript i from the i-th Hessian matrix, whenever clear from context. We will denote the finite sample version of $\mathbf{H}^i(\boldsymbol{\theta}^i)$ by $\mathbf{H}^i(\mathcal{D}, \boldsymbol{\theta}^i)$, i.e., $\mathbf{H}^i(\mathcal{D}, \boldsymbol{\theta}^i) = \frac{1}{n} \sum_{l=1}^n \mathbf{H}^i(\mathbf{x}^{(l)}, \boldsymbol{\theta}^i)$. Finally, we will denote the Hessian matrix restricted to the true support set S_i by: $\mathbf{H}^i(\cdot; \boldsymbol{\theta}^i_{S_i}) \in \mathbb{R}^{|S_i| \times |S_i|}$. In order to prove our main result, we will present a series of technical lemmas slowly building towards our main result. Detailed proofs of the lemmas are given in Appendix A.

The following lemma states that the i-th population Hessian is positive definite. Specifically, the i-th population Hessian evaluated at the true parameter $\boldsymbol{\theta}^i$, are positive definite with the minimum eigenvalue being $C_{\min}$. We prove the following lemma by showing that the loss function given by (10), when restricted to an arbitrary line, is strongly convex as long as the payoffs are finite.

Lemma 1 (Minimum eigenvalue of population Hessian). *For $\boldsymbol{\theta}^i \in \Theta^i$, $\lambda_{\min}(\mathbf{H}^i(\boldsymbol{\theta}^i)) \stackrel{\text{def}}{=} C_{\min} > 0$.*

Given that population Hessian matrices are positive-definite, we then show that the finite sample Hessian matrices, evaluated at any parameter $\boldsymbol{\theta}_{S_i}$, are positive definite with high probability. We use tools from random matrix theory developed by [Tro12] to prove the following lemma.

Lemma 2 (Minimum eigenvalue of finite sample Hessian). *Let $\boldsymbol{\theta} \in \Theta^i$ be any arbitrary vector and let $\lambda_{\min}(\mathbf{H}^i(\boldsymbol{\theta}_{S_i})) \stackrel{\text{def}}{=} \lambda_{\min} > 0$. Then, if the number of samples satisfies the following condition:*

$$n \geq \frac{8(d_i + 1)}{\lambda_{\min}} \log\left(\frac{m_i(1 + d_i m)}{\delta}\right),$$

then $\lambda_{\min}(\mathbf{H}^i(\mathcal{D}; \boldsymbol{\theta}_{S_i})) \geq \frac{\lambda_{\min}}{2}$ with probability at least $1 - \delta$ for some $\delta \in (0, 1)$.

Now that we have shown that the loss function given by (10) is strongly convex (Lemmas 1 and 2), we exploit strong convexity to control the difference between the true parameter and the estimator $\|\boldsymbol{\theta}^i - \widehat{\boldsymbol{\theta}^i}\|_{1,2}$. However, before proceeding further, we need to bound the $\ell_{\infty,2}$ norm of the gradient, as done in the following lemma. We prove the lemma by using McDiarmid's inequality to show that in each group the finite sample gradient concentrates around the expected gradient, and then use a union bound over all the groups to control the $\ell_{\infty,2}$ norm.

Lemma 3 (Gradient bound). *Let $\|\mathbb{E}_{\mathbf{x}}\left[\nabla \ell^i(\mathbf{x}; \boldsymbol{\theta}^i)\right]\|_{\infty,2} = \nu$, then we have that*

$$\|\nabla L^i(\mathcal{D}; \boldsymbol{\theta}^i)\|_{\infty,2} \leq \nu + \sqrt{\frac{2}{n} \log\left(\frac{2(d_i + 1)}{\delta}\right)},$$

with probability at least $1 - \delta$.

Note that the expected gradient at the parameter $\boldsymbol{\theta}^i$ does not vanish, i.e., $\|\mathbb{E}_{\mathbf{x}}\left[\nabla \ell^i(\mathbf{x}; \boldsymbol{\theta}^i)\right]\|_{\infty,2} = \nu$. This is because of the mismatch between the generating distribution $\mathcal{P}$ and the softmax distribution used for learning the parameters, as in (10). Indeed, if the data were drawn from a Markov random field, which induces a softmax distribution on the conditional distribution of node given the rest of the nodes, the parameter $\nu = 0$. However this is not the case for us. An unfortunate consequence of this is that, even with an infinite number of samples, our method will not be able to recover the parameters $\boldsymbol{\theta}^i$ exactly. Thus, without additional assumptions on the payoffs, our method only recovers the ε-Nash equilibrium of the game.

With the required technical results in place, we are now ready to bound $\|\boldsymbol{\theta}^i - \widehat{\boldsymbol{\theta}^i}\|_{1,2}$. Our analysis has two

steps. First, we bound the norm in the true support set, i.e., $\|\boldsymbol{\theta}^i_{S_i} - \widehat{\boldsymbol{\theta}}^i_{S_i}\|_{1,2}$. Then, we show that the norm of the difference between the true parameter and the estimator, outside the support set, is a constant factor (specifically 3) of the difference in the support set. For the first step with use a proof technique originally developed by [RBLZ08] in a different context, while the second step follows from matrix algebra and optimality of the estimator $\widehat{\boldsymbol{\theta}^i}$ for the problem (8).

The following technical lemma, which will be used later on in our proof to bound $\|\widehat{\boldsymbol{\theta}}^i_S - \boldsymbol{\theta}^i_S\|_{1,2}$, lower bounds the minimum eigenvalue of the i-th population Hessian at an arbitrary parameter $\boldsymbol{\theta} \in \Theta^i$, in terms of the minimum eigenvalue of the i-th population Hessian at the true parameter $\boldsymbol{\theta}^i$.

Lemma 4 (Minimum population eigenvalue at arbitrary parameter). *Let $\boldsymbol{\theta} \in \Theta^i$ be any vector. Then the minimum eigenvalue of i-th population Hessian matrix evaluated at $\boldsymbol{\theta}_{S_i}$ is lower bounded as follows:*

$$\lambda_{\min}(\mathbf{H}^i(\boldsymbol{\theta}_{S_i})) \geq \lambda_{\min}(\mathbf{H}^i(\boldsymbol{\theta}^i_{S_i}))$$
$$- \frac{1}{4}(d_i + 1)m^2 \|\boldsymbol{\theta}_{S_i} - \boldsymbol{\theta}^i_{S_i}\|_{1,2}.$$

Now, we are ready to bound the difference between the true parameter $\boldsymbol{\theta}^i$ and its estimator $\widehat{\boldsymbol{\theta}^i}$, in the true support set S_i.

Lemma 5 (Error of the i-th estimator on the support set). *If the regularization parameter and number of samples satisfy the following condition:*

$$\lambda \geq 2\left(\nu + \sqrt{\frac{2}{n} \log\left(\frac{2(d_i + 1)}{\delta}\right)}\right),$$

$$n > \frac{2}{N(m, d_i)} \log\left(\frac{2(d_i + 1)}{\delta}\right),$$

where $N(m, d_i) = \{C_{\min}/(36m^2(d_i+1)^2) - \nu\}^2$, and $C_{\min} \stackrel{\text{def}}{=} \lambda_{\min}(\mathbf{H}^i(\boldsymbol{\theta}^i_{S_i}))$; then with probability at least $1 - \delta$, for some $\delta \in (0, 1)$, we have:

$$\|\widehat{\boldsymbol{\theta}}^i_{S_i} - \boldsymbol{\theta}^i_{S_i}\|_{1,2} \leq \frac{6(d_i + 1)}{C_{\min}}\lambda. \tag{13}$$

Next, we bound the difference between the true parameter $\boldsymbol{\theta}^i$ and its estimator $\widehat{\boldsymbol{\theta}^i}$.

Lemma 6 (Error of the i-th parameter estimator). *Under the same conditions on the regularization parameter and number of samples as in Lemma 5 we have, with probability at least $1 - \delta$ for some $\delta \in (0, 1)$,*

$$\|\widehat{\boldsymbol{\theta}^i} - \boldsymbol{\theta}^i\|_{1,2} \leq \frac{24(d_i + 1)}{C_{\min}}\lambda.$$

Now that we have control over $\|\boldsymbol{\theta}^i - \widehat{\boldsymbol{\theta}^i}\|_{1,2}$ for all $i \in [p]$, we are ready to prove our main result concerning the

sufficient number of samples needed by our method to guarantee PSNE recovery with high probability.

Theorem 1. *Let $\mathcal{G} = (G, \mathcal{U})$, with $\mathcal{U} = \{u^i : \mathcal{A}_{-i} \to \mathbb{R}\}_{i \in [p]}$, be the true potential graphical game over p players and maximum degree d, from which the data set $\mathcal{D}$ is drawn. Let $\widehat{\mathcal{G}} = (\widehat{G}, \widehat{\mathcal{U}})$, with $\widehat{U} = \{\widehat{u}^i : \mathcal{A}_{-i} \to \mathbb{R}\}_{i \in [p]}$, be the game learned from the data set $\mathcal{D}$ by solving the optimization problem (8) for each $i \in [p]$. Then if the regularization parameter and the number of samples satisfy the condition:*

$$\lambda \geq 2\left(\nu + \sqrt{\frac{2}{n}\log\left(\frac{2p(d+1)}{\delta}\right)}\right),$$

$$n > \max\left\{\frac{2}{N(m,d)}\log\left(\frac{2p(d+1)}{\delta}\right),\right.$$
$$\left.\frac{8(d+1)}{C_{\min}}\log\left(\frac{m(1+dm)}{\delta}\right)\right\},$$

where $N(m,d) = \{C_{\min}/(36m^2(d+1)^2) - \nu\}^2$, then we have that the following hold with probability at least $1 - \delta$, for some $\delta \in (0,1)$:

(i) $\mathcal{NE}(\widehat{\mathcal{G}}) = \varepsilon\text{-}\mathcal{NE}(\mathcal{G})$, with $\varepsilon = \frac{48(d_i+1)}{C_{\min}}\lambda$.

(ii) Additionally, if the true game $\mathcal{G}$ satisfies the condition: $\forall i \in [p], \forall(x_i, \mathbf{x}_{-i}), (x_i', \mathbf{x}_{-i}) \in \mathcal{A}$ such that $(x_i, \mathbf{x}_{-i}) \in \mathcal{NE}(\mathcal{G}) \wedge (x_i', \mathbf{x}_{-i}) \notin \mathcal{NE}(\mathcal{G}) \implies u^i(x_i, \mathbf{x}_{-i}) > u^i(x_i', \mathbf{x}_{-i}) + \varepsilon$. Then, $\mathcal{NE}(\widehat{\mathcal{G}}) = \mathcal{NE}(\mathcal{G})$.

Proof. Note that $\|\mathbf{f}^i(x_i, \mathbf{x}_{-i})\|_{\infty,2} = \max\{\|\mathbf{f}^{i,0}(x_i)\|_2, \|\mathbf{f}^{i,1}(x_i, x_1)\|_2, \ldots, \|\mathbf{f}^{i,p}(x_i, x_p)\|_2\} = 1$, for any $\mathbf{x} \in \mathcal{A}$, since each binary vector $\mathbf{f}^{i,j}(x_i, x_j)$ has a single "1" at exactly one location. Then, from the Cauchy-Schwartz inequality, Lemma 6, and a union bound over all players, we have that:

$$(\forall \mathbf{x} \in \mathcal{A}, \forall i \in [p]) \; |\widehat{u}^i(x_i, \mathbf{x}_{-i}) - u^i(x_i, \mathbf{x}_{-i})|$$
$$= |(\widehat{\boldsymbol{\theta}}^i - \boldsymbol{\theta}^i)^T \mathbf{f}^i(x_i, \mathbf{x}_{-i})|$$
$$\leq \|\widehat{\boldsymbol{\theta}}^i - \boldsymbol{\theta}^i\|_{1,2}\|\mathbf{f}^i(x_i, \mathbf{x}_{-i})\|_{\infty,2}$$
$$= \|\widehat{\boldsymbol{\theta}}^i - \boldsymbol{\theta}^i\|_{1,2} \leq \frac{24(d_i+1)}{C_{\min}}\lambda = \frac{\varepsilon}{2}, \qquad (14)$$

with probability at least $1 - p\delta$. Now consider any $\mathbf{x} \in \mathcal{NE}(\widehat{\mathcal{G}})$ and any $i \in [p]$. Since $\mathbf{x} \in \mathcal{NE}(\widehat{\mathcal{G}})$, we have from (14), $(\forall x_i' \in \mathcal{A}_i)$:

$$u^i(x_i, \mathbf{x}_{-i}) + \varepsilon/2 \geq \widehat{u}^i(x_i, \mathbf{x}_{-i}) \geq \widehat{u}^i(x_i', \mathbf{x}_{-i})$$
$$\implies u^i(x_i, \mathbf{x}_{-i}) \geq \widehat{u}^i(x_i', \mathbf{x}_{-i}) - \varepsilon/2$$
$$\implies u^i(x_i, \mathbf{x}_{-i}) \geq u^i(x_i', \mathbf{x}_{-i}) - \varepsilon,$$

where the last line again follows from (14). This proves that $\mathcal{NE}(\widehat{\mathcal{G}}) \subseteq \varepsilon\text{-}\mathcal{NE}(\mathcal{G})$. Using exactly the same

arguments as above, we can also show that for any $\mathbf{x} \in \mathcal{NE}(\mathcal{G})$:

$$\widehat{u}^i(x_i, \mathbf{x}_{-i}) \geq \widehat{u}^i(x_i, \mathbf{x}_{-i}) - \varepsilon \qquad (\forall x_i' \in \mathcal{A}_i),$$

which proves that $\mathcal{NE}(\mathcal{G}) \subseteq \varepsilon\text{-}\mathcal{NE}(\widehat{\mathcal{G}})$. Thus we have that $\mathcal{NE}(\widehat{\mathcal{G}}) = \varepsilon\text{-}\mathcal{NE}(\mathcal{G})$, i.e., the set of joint strategy profiles $\mathbf{x} \in \mathcal{NE}(\widehat{\mathcal{G}})$ form an ε-Nash equilibrium set of the true game $\mathcal{G}$. This proves our first claim. For our second claim, consider any $(x_i, \mathbf{x}_{-i}) \in \mathcal{NE}(\mathcal{G})$ and $(x_i', \mathbf{x}_{-i}) \notin \mathcal{NE}(\mathcal{G})$. Then:

$$u^i(x_i, \mathbf{x}_{-i}) > u^i(x_i', \mathbf{x}_{-i}) + \varepsilon$$
$$\implies \widehat{u}^i(x_i, \mathbf{x}_{-i}) + \varepsilon/2 > \widehat{u}^i(x_i', \mathbf{x}_{-i}) - \varepsilon/2 + \varepsilon$$
$$\implies \widehat{u}^i(x_i, \mathbf{x}_{-i}) > \widehat{u}^i(x_i', \mathbf{x}_{-i})$$
$$\implies (x_i, \mathbf{x}_{-i}) \in \mathcal{NE}(\widehat{\mathcal{G}}) \wedge (x_i', \mathbf{x}_{-i}) \notin \mathcal{NE}(\widehat{\mathcal{G}}),$$

where the first line holds by assumption, and the second line again follows from (14). Thus we have that $\mathcal{NE}(\mathcal{G}) = \mathcal{NE}(\widehat{\mathcal{G}})$. By setting the probability of error $p\delta = \delta'$ for some $\delta' \in (0,1)$ we prove our claim. The second part of the lower bound on the number of samples is due to Lemma 2. $\qquad\square$

Remark 1. *The sufficient number of samples needed by our method to guarantee PSNE recovery, with probability at least $1 - \delta$, scales as $\mathcal{O}\left(m^4 d^4 \log(pd/\delta)\right)$. This should be compared with the results of [JRVS11] for learning undirected graphical models. They show that $\mathcal{O}\left(m^2 d^2 \log(m^2 p)\right)$ are sufficient for learning m-ary discrete graphical models. However, their sample complexity hides a constant K that is related to the maximum eigenvalue of the scatter matrix, which we have upper bounded by $m^2 d^2$ in our case, leading to a slightly higher sample complexity.*

Remark 2. *Note that as $n \to \infty$, the regularization parameter $\lambda \to 2\nu$, where ν is the maximum norm of the expected gradient at the true parameter $\boldsymbol{\theta}^i$ across all $i \in [p]$. Thus, even with an infinite number of samples, our method recovers the ε-Nash equilibria set of the true game with $\varepsilon \to \frac{96(d_i+1)\nu}{C_{\min}}$ as $n \to \infty$.*

5 Necessary Conditions

In this section, we obtain an information-theoretic lower bound on the number of samples needed to learn sparse polymatrix games. Let $\mathfrak{G}_{p,d,m}$ be set of polymatrix games over p players, with degree at most d, and maximum number of strategies per player being m. Our approach for doing so is to treat the inference procedure as a communication channel, where nature picks a game $\mathcal{G}^*$ from the set $\mathfrak{G}_{p,d,m}$ and then generates a data set $\mathcal{D}$ of n strategy profiles. A decoder $\psi : \mathcal{A}^n \to \mathfrak{G}_{p,d,m}$ then maps $\mathcal{D}$ to a game $\widehat{\mathcal{G}} \in \mathfrak{G}_{p,d,m}$. We wish to obtain

lower bounds on the number of samples required by any decoder ψ to recover the true game consistently. In this setting, we define the *minimax* estimation error as follows:

$$p_{\text{err}} = \min_{\psi} \sup_{\mathcal{G}^* \in \mathfrak{G}_{p,d,m}} \Pr \left\{ \mathcal{NE}(\psi(\mathcal{D})) \neq \mathcal{NE}(\mathcal{G}^*) \right\},$$

where the probability is computed over the data distribution. For obtaining necessary conditions on the sample complexity, we assume that the data distribution follows the global noise model described in (4). The following theorem prescribes the number of samples needed for learning sparse polymatrix games. Our proof of the theorem constitutes constructing restricted ensembles of "hard-to-learn" polymatrix games, from which nature picks a game uniformly at random and generates data. We then use the Fano's technique to lower bound the minimax error. The use of restricted ensembles is customary for obtaining information-theoretic lower bounds, c.f. [SW12, WWR10].

Theorem 2. *If the number of samples* $n \leq \frac{\log(m^d - m)\binom{p}{d}}{2 \log 2} - 1$, *then estimation fails with* $p_{\text{err}} \geq 1/2$.

Proof. Consider the following restricted ensemble $\widetilde{\mathfrak{G}} \subset \mathfrak{G}_{p,d,m}$ of p-player polymatrix games with degree d, and the set of pure-strategies of each player being $\mathcal{A}_i = [m]$. Each $\mathcal{G} = (G, \mathcal{U}) \in \widetilde{\mathfrak{G}}_{p,d,m}$ is characterized by a set $\mathcal{I}$ of *influential players*, and a set $\mathcal{I}^c \stackrel{\text{def}}{=} [p] \setminus \mathcal{I}$ of *non-influential* players, with $|\mathcal{I}| = d$. The graph G is a complete (directed) bipartite graph from the set $\mathcal{I}$ to $\mathcal{I}^c$. After picking the graph structure G, nature fixes the strategies of the influential players to some $\mathbf{a} \in \{\mathbf{b} \in [m]^{|\mathcal{I}|} \mid \exists i, j \in \mathcal{I} \text{ such that } b_i \neq b_j\}$. Finally, the payoff matrices are chosen as follows:

$$u^{i,i}(x_i) = \mathbf{1}[x_i = a_i] \qquad (\forall i \in \mathcal{I})$$

$$u^{j,j}(x_j) = \frac{1}{(2x_j)} \qquad (\forall j \in \mathcal{I}^c)$$

$$u^{j,i}(x_j, x_i) = \mathbf{1}[x_j = x_i] \qquad (\forall i \in \mathcal{I} \wedge j \in \mathcal{I}^c).$$

Therefore, each $\mathcal{G} \in \widetilde{\mathfrak{G}}$ game has a exactly one unique Nash equilibrium where the influential players play $\mathbf{a}$ (decided by nature) and the non-influential players play $\mathfrak{maj}(\mathbf{a})$ — where $\mathfrak{maj}(\mathbf{a})$ returns the majority strategy among $\mathbf{a}$, and in case of a tie between two or more strategies it returns the numerically lowest strategy (recall that the pure-strategy set for each player is $[m]$). Thus we have that $|\widetilde{\mathfrak{G}}| = (m^d - m)\binom{p}{d}$. Nature picks a game $\mathcal{G}$ uniformly at random from $\widetilde{\mathfrak{G}}$ by randomly selecting a set of d players as "influential", and then selecting a strategy profile $\mathbf{a}$ uniformly at random for the influential players and setting the payoff matrices as described earlier. Nature then generates a dataset

$\mathcal{D}$ using the global noise model with parameter $q \in (1/m^p, 2/(m^p + 1)]$. Then from the Fano's inequality we have that:

$$p_{\text{err}} \geq 1 - \frac{I(\mathcal{D}; \mathcal{G}) + \log 2}{H(\mathcal{G})}, \qquad (15)$$

where $I(\cdot; \cdot)$ and $H(\cdot)$ denote mutual information and entropy respectively. The mutual information $I(\mathcal{D}; \mathcal{G})$ can be bounded, using a result by [Yu97], as follows:

$$I(\mathcal{D}; \mathcal{G}) \leq \frac{1}{|\widetilde{\mathfrak{G}}|^2} \sum_{\mathcal{G}_1 \in \widetilde{\mathfrak{G}}} \sum_{\mathcal{G}_2 \in \widetilde{\mathfrak{G}}} \mathbb{KL}\left(\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1} \| \mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_2} \right),$$
$$(16)$$

where $\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1}$ (respectively $\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_2}$) denotes the data distribution under $\mathcal{G}_1$ (respectively $\mathcal{G}_2$). The KL divergence term from 16 can be bounded as follows:
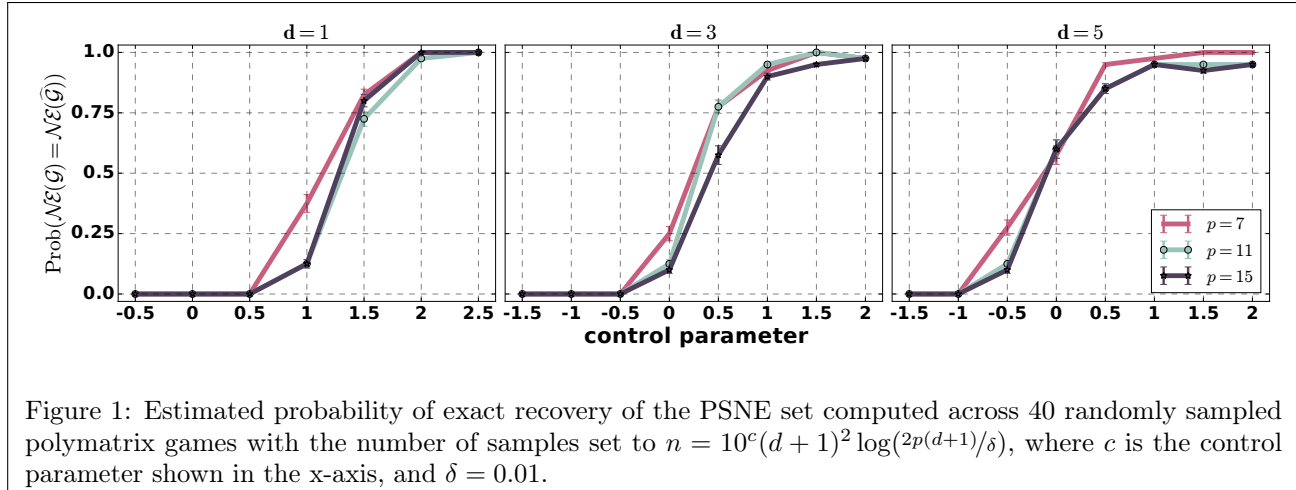
$$\mathbb{KL}\left(\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1} \| \mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_2} \right)$$
$$= n \sum_{\mathbf{x} \in \mathcal{A}} \mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1} \log \frac{\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1}}{\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_2}}$$
$$= n \left\{ \sum_{\mathbf{x} \in \mathcal{NE}(\mathcal{G}_1)} q \log \frac{q(m^p - 1)}{1 - q} + \right.$$
$$\left. \sum_{\mathbf{x} \in \mathcal{NE}(\mathcal{G}_2)} \frac{(1-q)}{m^p - 1} \log \frac{1 - q}{q(m^p - 1)} \right\}$$
$$= \frac{n(qm^p - 1)}{m^p - 1} \log \left(\frac{q(m^p - 1)}{1 - q} \right)$$
$$\leq n \log \left(\frac{q(m^p - 1)}{1 - q} \right) \leq n \log 2, \qquad (17)$$

where the first line follows from the fact that the samples are i.i.d , the second line follows from the fact the the distributions $\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_1}$ and $\mathcal{P}_{\mathcal{D}|\mathcal{G}=\mathcal{G}_2}$ assign the same probability to $\mathbf{x} \in \mathcal{A} \setminus (\mathcal{NE}(\mathcal{G}_1) \cup \mathcal{NE}(\mathcal{G}_2))$, and the last line follows from the fact that $q \in (1/m^p, 2/(m^p + 1)]$. Putting together (15), (16) and (17), we have that if

$$n \leq \frac{\log(m^d - m)\binom{p}{d}}{2 \log 2} - 1,$$

then $p_{\text{err}} \geq 1/2$. Since, learning the ensemble $\mathfrak{G}$ is at least as hard as learning a subset of $\mathfrak{G}$, our claim follows. $\qquad \square$

Remark 3. *From the above theorem we have that, the number of samples needed by any conceivable method, to recover the PSNE set consistently, is $\Omega(d \log(pm))$, assuming that $d = o(p)$. Therefore, the method based on $\ell_{1,2}$-regularized logistic regression is information-theoretically optimal in the number of players, for learning sparse polymatrix games.*

Figure 1: Estimated probability of exact recovery of the PSNE set computed across 40 randomly sampled polymatrix games with the number of samples set to $n = 10^c(d+1)^2 \log(2p(d+1)/\delta)$, where c is the control parameter shown in the x-axis, and $\delta = 0.01$.

6 Experiments

In order to validate our theoretical results, we performed various synthetic experiments by sampling a random polymatrix game, generating data from the sampled game, and then using our method to learn the game from the sampled data. We estimated the probability that our method learns the "correct" game, i.e., a game with the same PSNE set as the true game, across 40 randomly sampled games for each value of $p \in \{7, 11, 15\}$ and $d \in \{1, 3, 5\}$. The results are shown in Figure 1. We observe that the scaling of the sample complexity prescribed by Theorem 1 indeed holds in practice. The results show a phase transition behavior, where if the number of samples is less than $c(d+1)^2 \log(p(d+1)/\delta)$, for some constant c, then PSNE recovery fails with high probability, while if the number of samples is at least $C(d+1)^2 \log(p(d+1)/\delta)$, for some constant C, then PSNE recovery succeeds with high probability. More details about our synthetic experiments can be found in Appendix B.

We also evaluated our algorithm on real-world data sets containing (i) U.S. supreme court rulings, (ii) U.S. congressional voting records, and (iii) U.N. General Assembly roll-call votes.

Our algorithm recovers connected components corresponding to liberal and conservative blocs of justices within the Supreme Court of the U.S. The Nash equilibria consists of strategy profiles where all justices vote unanimously, as well as strategy profiles where the conservative and liberal blocs vote unanimously but in opposition to each other.

The game graph recovered from congressional voting records, groups Democrats and Republicans in separate components. Moreover, we observed that the connected components groups senators belonging to the same state or geographic region together. The recovered PSNE set sheds light on the voting patterns of senators — senators belonging to the same connected component vote (almost) identically on bills.

Finally, on the U.N. voting data set our method recovered connected components comprising of Arab League countries and Southeast Asian countries respectively. As was the case with the aforementioned data sets, the PSNE set grouped countries that vote almost identically on U.N. resolutions.

We were also able to compute the price of anarchy (PoA) for each data set, which quantifies the degradation of performance caused by selfish behavior of non-cooperative agents. For the two supreme court voting data sets, the PoA was 1.9 and 1.6 respectively. For the congressional voting data set the PoA was 2.6, while for the united nations voting data set the PoA was 3.0. More details and results from our real-world experiments can be found in Appendix C.

7 Concluding Remarks

We conclude this exposition with a discussion of potential avenues for future work. In this paper we considered the problem of learning a very general, and widely used, class of graphical games called polymatrix games, involving players with pure strategies. One can also consider *mixed strategies*, which would entail learning distributions, instead of "sets", under the framework of non-cooperative maximization of utility. Further, one can also consider other solution concepts like *correlated equilibria*.

References

[BSK06] Ben Blum, Christian R Shelton, and Daphne Koller. A continuation method for nash equilibria in structured games. *Journal of Artificial Intelligence Research*, 25:457–502, 2006.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[BZR11] Oren Ben-Zwi and Amir Ronen. Local and global price of anarchy of graphical games. *Theoretical Computer Science*, 412(12-14):1196–1207, 2011.

[DGP09] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[GH16] Asish Ghoshal and Jean Honorio. From behavior to sparse graphical games: Efficient recovery of equilibria. In *54th Annual Allerton Conference on Communication, Control, and Computing*, pages 1220–1227, 2016.

[GH17] Asish Ghoshal and Jean Honorio. Learning graphical games from behavioral data: Sufficient and necessary conditions. In *Artificial Intelligence and Statistics*, pages 1532–1540, 2017.

[GJ16] Vikas Garg and Tommi Jaakkola. Learning Tree Structured Potential Games. In *Advances in Neural Information Processing Systems 29*, pages 1552–1560, 2016.

[HO15] Jean Honorio and Luis Ortiz. Learning the structure and parameters of large-population graphical games from behavioral data. *Journal of Machine Learning Research*, 16:1157–1210, 2015.

[IO14] Mohammad T Irfan and Luis E Ortiz. On influence, stable behavior, and the most influential individuals in networks: A game-theoretic approach. *Artificial Intelligence*, 215:79–119, 2014.

[Jan68] E. Janovskaja. Equilibrium situations in multi-matrix games. *Litovskiĭ Matematicheskiĭ Sbornik*, 8:381–384, 1968.

[JLB11] Albert Xin Jiang and Kevin Leyton-Brown. Polynomial-time computation of exact correlated equilibrium in compact games. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 119–126. ACM, 2011.

[JN91] Charles R. Johnson and Peter Nylen. Monotonicity properties of norms. *Linear Algebra and its Applications*, 148:43–58, April 1991.

[JRVS11] Ali Jalali, Pradeep Ravikumar, Vishvas Vasuki, and Sujay Sanghavi. On Learning Discrete Graphical Models using Group-Sparse Regularization. In *AISTATS*, pages 378–387, 2011.

[KKLO03] Sham Kakade, Michael Kearns, John Langford, and Luis Ortiz. Correlated equilibria in graphical games. In *Proceedings of the 4th ACM Conference on Electronic Commerce*, pages 42–47. ACM, 2003.

[RBLZ08] Adam J. Rothman, Peter J. Bickel, Elizaveta Levina, and Ji Zhu. Sparse permutation invariant covariance estimation. *Electronic Journal of Statistics*, 2:494–515, 2008.

[Rub16] Aviad Rubinstein. Settling the complexity of computing approximate two-player Nash equilibria. *arXiv:1606.04550 [cs]*, June 2016. arXiv: 1606.04550.

[SW12] Narayana P Santhanam and Martin J Wainwright. Information-theoretic limits of selecting binary graphical models in high dimensions. *Information Theory, IEEE Transactions on*, 58(7):4117–4134, 2012.

[Tro12] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(May):389–434, 2012.

[WWR10] Wei Wang, Martin J Wainwright, and Kannan Ramchandran. Information-theoretic bounds on model selection for Gaussian Markov random fields. In *ISIT*, pages 1373–1377. Citeseer, 2010.

[Yu97] Bin Yu. *Festschrift for Lucien Le Cam: Research Papers in Probability and Statistics*, chapter Assouad, Fano, and Le Cam, pages 423–435. Springer New York, New York, NY, 1997.