

# Differentially Private Equilibrium Finding in Polymatrix Games

Mingyang Liu<sup>1</sup>, Gabriele Farina<sup>1</sup>, Asuman Ozdaglar<sup>1</sup>

<sup>1</sup> LIDS, EECS, Massachusetts Institute of Technology

<sup>1</sup> {liumy19,gfarina,asuman}@mit.edu

## Abstract

We study equilibrium finding in polymatrix games under differential privacy constraints. To start, we show that high accuracy and asymptotically vanishing differential privacy budget (as the number of players goes to infinity) cannot be achieved simultaneously under either of the two settings: (i) We seek to establish equilibrium approximation guarantees in terms of Euclidean *distance* to the equilibrium set, and (ii) the adversary has access to all communication channels. Then, assuming the adversary has access to a constant number of communication channels, we develop a novel distributed algorithm that recovers strategies with simultaneously vanishing *Nash gap* (in expected utility, also referred to as *exploitability*) and *privacy budget* as the number of players increases.

## 1 Introduction

Many multi-agent settings involve a large number of players, each interacting with a smaller subset of players. Polymatrix games provide a tractable and rich model that can capture such settings (Janovskaja, 1968; Cai et al., 2016; Deligkas et al., 2017). In polymatrix games, players engage in pairwise interactions defined by an underlying graph in which nodes represent players and edges model the interaction between the two players they are connecting. The player chooses the same strategy in each pairwise game she is involved with and her utility is the sum of her utilities for each of these games.

In several settings, such as security games (De Nittis et al., 2018) and financial markets (Donmez et al., 2024), the players' utility functions may be sensitive and need to be kept private by the players as they update their behavior towards equilibrium. Typically, there are two approaches for finding equilibrium while keeping the utility functions private: centralized and distributed methods. In centralized methods, the players send their utility functions to a trusted central server (Kearns et al., 2014; Rogers and Roth, 2014; Cummings et al., 2015). When a trusted central server is infeasible, the only viable option for computing equilibrium is for the players to use a distributed algorithm where they make local computations and exchange information with their neighbors. Throughout the process, a malicious adversary may inspect the communication channel and infer the utility function, which might reveal sensitive information about individual preferences. To keep the computation secure and prevent the leakage of sensitive information, tools from differential privacy (DP) can be employed (e.g., Dwork (2006)).

Several authors have considered the question of differentially private equilibrium computation in games. However, previous work was either unable to achieve high accuracy and low differential privacy budget simultaneously (Ye et al., 2021; Wang et al., 2022), or left the adversary able to infer private information from the observations (Wang and Bašar, 2024; Wang and Nedić, 2024).

## Contributions and Techniques

In this paper, we study the problem of distributed, differentially private equilibrium computation in polymatrix games. We build on the notion of adjacent distributed optimization problem defined in [Huang et al. \(2015\)](#). In particular, given a polymatrix game  $\mathcal{G}$  we define *adjacent games* as those that differ from  $\mathcal{G}$  on the utility function of a single player. Then, once the algorithm for equilibrium finding is differentially private, an adversary cannot distinguish, with high probability from his observations, the original game from any other games that are adjacent to it. Hence, when differential privacy holds, the adversary is unable to determine the utility function of any player with high confidence, ensuring the equilibrium computation process is privacy-preserving.

Using this concept of adjacency, we begin by showing that finding an approximate equilibrium with arbitrary small differential privacy guarantees can be impossible, depending on the desiderata imposed on the process. Specifically, we show that if the adversary can access an arbitrary number of communication channels, or the target accuracy metric of interest is the Euclidean distance from the equilibrium set, then distributed computation of high-quality equilibria while providing vanishing privacy guarantees is impossible. In other words, one will inevitably suffer either a low approximation of the equilibrium, or a low level of privacy (high privacy budget). While these negative results serve as guardrails to guide algorithm design, it is important to realize that they do not preclude all paths to meaningful results. Indeed, it is important to realize that Euclidean distance to the equilibrium set is only one of the set of possible quality metrics for measuring equilibrium approximation. Another metric is *Nash gap* or *exploitability*, meaning how close a strategy is to equilibrium in terms of *expected utility* rather than metric distance to the equilibrium set in strategy space.

Finally, we complement our negative result with a positive result regarding exploitability guarantees. In particular, we assume that the adversary has access to a constant number of communication channels and propose a new distributed algorithm for computing a Nash equilibrium. As typical in the differential privacy literature, our algorithm involves each player communicating a noisy version of their strategy to their neighbors and updating their strategy using a regularized proximal gradient step. A key novelty of the algorithm is to scale the regularizer proportional to the harmonic mean of the degrees divided by the degree of the player. The inverse proportionality with the degree ensures that more regularization is introduced for players with a lower degree (since a low-degree player’s gradient is more sensitive to variations in the utility matrices of its neighbors).

Specifically, we propose Algorithm 1, which simultaneously achieve low exploitability and low DP budget, with guarantees that improve as the number of players  $N$  increases. Moreover, the algorithm achieves that whenever the associated graph with the polymatrix game is sparse or dense, by imposing an adaptive regularizer on players’ utility functions. When the graph associated with the polymatrix game is sparse, we leverage the fact that for most nodes that are at least of distance  $t$  to the edge of the changed utility matrices, changes in the utility matrices between adjacent polymatrix games have minimal impact on them during the first  $t$  updates. In contrast, when the graph is dense, additional regularization, which is inversely proportional to the players’ degrees, stabilizes the updates of players with low degrees so that the adversary’s observation in two adjacent games will be similar. For players with high degrees, the aggregative nature of the utility function in polymatrix games mitigates the effect of changes in the utility matrices.

## 2 Related Work

In this section, we discuss prior work specifically targeted at distributed *equilibrium computation*. We omit a discussion of prior work on distributed (single-agent) optimization as it is hardly applicable to

our purposes, for several important reasons. For one, distributed optimization with differential privacy guarantee (Huang et al., 2015; Nozari et al., 2016; Wei et al., 2020) aims to find the minimum of the weighted sum of players' utility functions. However, distributed equilibrium finding is more complex since it aims to find the equilibrium point of those utility functions. Furthermore, the target output of the distributed optimization is of constant dimensionality with respect to the number of agents involved in the optimization process, the best parameters for a neural network (Wei et al., 2020). However, the output's dimension of equilibrium finding scales linearly with respect to the number of agents, since we need to specify the strategy of each individual player.

**Online Learning with Differential Privacy.** Prior work, including Jain et al. (2012); Guha Thakurta and Smith (2013); Agarwal and Singh (2017), has studied no-regret learning with DP guarantee, focusing on ensuring that an adversary cannot determine the utility (loss) vector at any single timestep. However, that condition does not imply differential privacy with respect to the payoff function itself. For instance, in a two-player normal-form game, the utility vector at timestep  $t$  for player 1 is  $\mathbf{U}\pi_2^{(t)}$ , where  $\mathbf{U}$  is the utility matrix and  $\pi_2^{(t)}$  is the strategy of player 2 at timestep  $t$ . In this case, an algorithm that reveals  $\mathbf{U}\pi_2^{(t)} + \mathbf{n}^{(t)}$ , where  $\mathbf{n}^{(t)}$  is sampled from a multivariate Gaussian, ensures differential privacy with respect to each individual utility vector. However, consider the case when  $\pi_2^{(t)}$  is sampled uniformly on each axis. By taking the average over all utility vectors, the Gaussian noise cancels out, and we will get the average of  $\mathbf{U}$ 's columns. Therefore,  $\mathbf{U}$  is not differentially private while each utility vector is still private.

**Differential Privacy in Games.** Other prior work, including Gade et al. (2020); Ye et al. (2021); Wang et al. (2022); Shakarami et al. (2022), has focused on finding equilibrium in aggregative games with differential privacy guarantees. However, they cannot achieve accuracy and differential privacy simultaneously. In other words, to achieve a fixed accuracy, the privacy budget will not continue to decrease while the number of players increases. The reason is that they did not exploit the game's structure to cancel out the noise imposed during the update. Therefore, accuracy will be harmed when achieving a low differential privacy budget by adding a large amount of noise. Furthermore, the privacy guarantee of Wang and Nedić (2024); Wang and Başar (2024) is weak because the adversary may determine the value of the utility function in a small region around the equilibrium point. Therefore, when the utility function is multilinear, such as polymatrix games, normal-form games, and extensive-form games, the adversary can fully determine the utility functions with certainty. Instead, in this paper, we exploit the structure of polymatrix games, whether the associated graph is dense or sparse, to mitigate the impact of the imposed noise, thereby achieving accuracy and differential privacy simultaneously.

### 3 Preliminaries

We now review several key concepts related to differential privacy and polymatrix games.

**Basics.** For any integer  $N > 0$ , let  $[N] := \{1, 2, 3, \dots, N - 1, N\}$ . For any vector  $\mathbf{x}$ , we use  $\|\mathbf{x}\|_p$  to denote its  $p$ -norm. By default,  $\|\mathbf{x}\|$  represents the 2-norm. Let  $\Delta^n := \{\mathbf{x} \in [0, 1]^n : \sum_{i=1}^n x_i = 1\}$  denotes the  $n - 1$  dimensional probability simplex. Moreover, for any discrete set  $\mathcal{S}$ , we can define  $\Delta^{\mathcal{S}}$  as the probability simplex over  $\mathcal{S}$ , with each index as the element in  $\mathcal{S}$ . For instance,  $\Delta^{[N]} = \Delta^N$ . For any convex set  $\mathcal{C}$ , we use  $\text{Proj}_{\mathcal{C}}(\mathbf{x}') := \arg\min_{\mathbf{x} \in \mathcal{C}} \|\mathbf{x} - \mathbf{x}'\|^2$  to denote the projection to  $\mathcal{C}$  from a vector  $\mathbf{x}'$  with respect to Euclidean distance.

**Polymatrix Games.** Polymatrix games can be written as a tuple

$$\mathcal{G} := \left( [N], E, \{\mathcal{A}_i\}_{i \in [N]}, \{\mathbf{U}_{i,j}\}_{(i,j) \in E} \right),$$

where

- $[N]$  is the set of players so there are  $N$  players.
- $E$  is the set of edges, where each edge  $(i, j) \in E$  indicates that players  $i, j$  interact with each other.
- $\{\mathcal{A}_i\}_{i \in [N]}$  is the set of players' action set, which means player  $i \in [N]$  chooses actions in  $\mathcal{A}_i$ .
- Let  $A := \max_{i \in [N]} |\mathcal{A}_i|$  be the size of the largest action set.
- $\mathbf{U}_{i,j} \in [-1, 1]^{\mathcal{A}_i \times \mathcal{A}_j}$  is the utility matrix between player  $(i, j) \in E$ .

For each Player  $i \in [N]$ , we denote with the symbol  $\mathcal{N}(i) := \{j \in [N] : (i, j) \in E\}$  the set of her neighbors. Then, when each player  $k \in [N]$  chooses her strategy as  $\pi_k \in \Delta^{\mathcal{A}_k}$ , the utility of player  $i$  is  $\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \pi_i^\top \mathbf{U}_{i,j} \pi_j$ <sup>1</sup>. Moreover, when the strategy profile of all players is  $\boldsymbol{\pi} := (\pi_1, \pi_2, \dots, \pi_N)$ , we can define the gradient of player  $i$ 's strategy with respect to her loss function (the negative of utility function) as  $\mathbf{g}_i^\pi := -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j$ .

**Differential Privacy.** We adopt the notion of adjacency in Huang et al. (2015). For any two polymatrix games, they are *adjacent* to each other if and only if they only differ from the utility matrices on a single edge. Formally,

**Definition 3.1** (Game adjacency). Given two polymatrix games  $\mathcal{G} = ([N], E, \{\mathcal{A}_i\}_{i \in [N]}, \{\mathbf{U}_{i,j}\}_{(i,j) \in E})$  and  $\mathcal{G}' = ([N'], E', \{\mathcal{A}'_i\}_{i \in [N']}, \{\mathbf{U}'_{i,j}\}_{(i,j) \in E'})$ , they are said to be *adjacent*, indicated as  $\mathcal{G} \sim \mathcal{G}'$ , if

1.  $N = N'$ ,  $E = E'$  and  $\mathcal{A}_i = \mathcal{A}'_i$  for any  $i \in [N]$ ; and
2. except for an edge  $(i, j) \in E$ ,  $\mathbf{U}_{i',j'} = \mathbf{U}'_{i',j'}$  and  $\mathbf{U}_{j',i'} = \mathbf{U}'_{j',i'}$  for any  $(i', j') \in E \setminus \{(i, j), (j, i)\}$ .

Consider now a generic iterative algorithm for finding equilibrium in polymatrix games, and denote with  $\Theta$  the space of all possible internal states of the algorithm at any time. Then, for a given polymatrix game  $\mathcal{G}$ , when the algorithm runs  $t > 0$  timesteps, it will generate intermediate variables  $\boldsymbol{\theta} := (\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}) \in \Theta^t$ , which are called *executions* of the algorithm.  $\Theta^t$  is the  $t$ -fold Cartesian product of  $\Theta$  with itself. Due to the randomness in the algorithm, this will result in a distribution  $\mathcal{P}_{\mathcal{G}}$  over  $\Theta^t$ . Moreover, there is an adversary with access to the communication channels between the players. Therefore, for any execution  $\boldsymbol{\theta} \in \Theta^t$ , the adversary can observe  $\mathcal{R}_{\mathcal{G}}(\boldsymbol{\theta}) = (\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}) =: \boldsymbol{o} \in \mathcal{O}^t$ , where  $\mathcal{O}$  is the space of all *observations* at a single timestep. Hence, we can define the  $(\epsilon, \delta)$ -differential privacy  $((\epsilon, \delta)$ -DP) as follows.

**Definition 3.2**  $((\epsilon, \delta)$ -Differential Privacy). For an  $\epsilon \geq 0$  and  $\delta \geq 0$ , an iterative distributed algorithm for finding equilibria is  $(\epsilon, \delta)$ -differentially private, if and only if for any two adjacent polymatrix game  $\mathcal{G}, \mathcal{G}'$ , any timestep  $t > 0$  and any set of observations  $\mathcal{S} \subseteq \mathcal{O}^t$ ,

$$\mathcal{P}_{\mathcal{G}}(\{\boldsymbol{\theta} : \mathcal{R}_{\mathcal{G}}(\boldsymbol{\theta}) \in \mathcal{S}\}) \leq e^\epsilon \mathcal{P}_{\mathcal{G}'}(\{\boldsymbol{\theta} : \mathcal{R}_{\mathcal{G}'}(\boldsymbol{\theta}) \in \mathcal{S}\}) + \delta. \quad (1)$$

Intuitively, Definition 3.2 guarantees that the algorithm will generate similar observations while deployed on adjacent games. Therefore, the adversary cannot differentiate  $\mathcal{G}, \mathcal{G}'$  from the observation so he cannot infer the utility matrices with certainty from observations.

<sup>1</sup>Typically the definition of player  $i$ 's utility is  $\sum_{j \in \mathcal{N}(i)} \pi_i^\top \mathbf{U}_{i,j} \pi_j$ . However, to ensure that the function and its gradient are bounded by constants regardless of the game size, which aligns with the DP literature (Huang et al., 2015; Wei et al., 2020; Ye et al., 2021), we divide it by  $|\mathcal{N}(i)|$ . Note that such modification will not change the equilibrium.

In this paper, we consider a generalization of Definition 3.2, the Rényi differential privacy (Rényi DP) (Mironov, 2017). The distribution of observations up to timestep  $t$  can be written as  $\mu_{\mathcal{G}}^{(t)}(\mathbf{o}) := \mathcal{P}_{\mathcal{G}}\left(\mathcal{R}_{\mathcal{G}}^{-1}(\mathbf{o})\right)$  for  $\mathbf{o} \in \mathcal{O}^t$  and thus we can define the  $(\alpha, \epsilon)$ -Rényi DP as follows.

**Definition 3.3** ( $(\alpha, \epsilon)$ -Rényi Differential Privacy). For  $\alpha > 1$  and  $\epsilon \geq 0$ , an iterative distributed algorithm for finding equilibria is  $(\alpha, \epsilon)$ -Rényi differentially private, if and only if for any two adjacent polymatrix game  $\mathcal{G}, \mathcal{G}'$  and timestep  $t > 0$ ,

$$D_\alpha\left(\mu_{\mathcal{G}}^{(t)}, \mu_{\mathcal{G}'}^{(t)}\right) := \frac{1}{\alpha - 1} \log \mathbb{E}_{\mathbf{o} \sim \mu_{\mathcal{G}'}^{(t)}} \left[ \left( \frac{\mu_{\mathcal{G}}^{(t)}(\mathbf{o})}{\mu_{\mathcal{G}'}^{(t)}(\mathbf{o})} \right)^\alpha \right] \leq \epsilon. \quad (2)$$

$\epsilon$  is also called the *privacy budget*.

Moreover, Definition 3.3 can be extended to  $\alpha = 1$  and  $\alpha = +\infty$  by defining  $D_1(\cdot, \cdot) = \lim_{\alpha \rightarrow 1^+} D_\alpha(\cdot, \cdot)$ , which is KL-divergence (Mironov, 2017), and  $D_\infty(\cdot, \cdot) := \lim_{\alpha \rightarrow +\infty} D_\alpha(\cdot, \cdot)$ . Moreover, Rényi DP can be converted to  $(\epsilon, \delta)$ -DP according to the following lemma from Mironov (2017).

**Lemma 3.4.** *If an algorithm satisfies  $(\alpha, \epsilon)$ -Rényi DP for  $\alpha > 1$ , then for any  $\delta \in (0, 1)$ , the algorithm also satisfies  $\left(\epsilon + \frac{\log(1/\delta)}{\alpha-1}, \delta\right)$ -DP. Moreover,  $(\infty, \epsilon)$ -Rényi DP is equivalent to  $(\epsilon, 0)$ -DP.*

## 4 Necessary Conditions for Achieving High Accuracy and Differential Privacy Simultaneously

In this section, we will show the necessary conditions to achieve differential privacy and high accuracy in approximating equilibria in polymatrix games. Specifically, we will show that: (i) the adversary should not have access to all communication channels; and (ii) the accuracy metric cannot be the Euclidean distance to the equilibrium. If even one of the conditions is violated, one can find a constant  $c_0 > 0$  so that when  $N$  is large enough, there exists a polymatrix game with  $N$  players so that the approximation error  $\zeta$  and differential privacy budget  $\epsilon$  cannot be guaranteed to be smaller than  $c_0$  simultaneously in that game.

All hardness results in this section are constructed with zero-sum polymatrix games (Cai et al., 2016), a special class of polymatrix games, that is, polymatrix games  $\mathcal{G} = ([N], E, \{\mathcal{A}_i\}_{i \in [N]}, \{\mathbf{U}_{i,j}\}_{(i,j) \in E})$  such that, for any strategy profile  $\pi$ ,

$$\sum_{i \in [N]} \sum_{j \in \mathcal{N}(i)} \pi_i^\top \mathbf{U}_{i,j} \pi_j \equiv 0. \quad (\text{Zero-sum condition})$$

One can see that when  $\mathbf{U}_{i,j} = -\mathbf{U}_{j,i}$  for any  $(i, j) \in E$ , the polymatrix game is zero-sum. Since the coarse correlated equilibria (CCE) in zero-sum polymatrix games collapse to NE (Cai et al., 2016), an algorithm that approximates CCE will approximate NE in zero-sum polymatrix games. Therefore, without loss of generality, in this section, we assume the algorithm will approximate the NE instead of CCE.

Assume that there is a distributed algorithm for finding equilibrium with the following guarantees of accuracy and privacy. In any polymatrix game  $\mathcal{G}$  with  $N$  players, the algorithm outputs a strategy profile  $\pi$  with distribution  $\mu_{\mathcal{G}}$  on the observations. Here we omit the timestep superscript on  $\mu_{\mathcal{G}}$  to emphasize that it is the distribution over all observations till termination of the algorithm. Then, there exists a parameter  $\alpha \geq 1$  such that the algorithm satisfies the following accuracy and privacy guarantees in any

$N$ -player game  $\mathcal{G}$ :

$$\text{Accuracy: } \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \max_{\hat{\pi}_i \in \Delta^{\mathcal{A}_i}} \langle \pi_i - \hat{\pi}_i, g_i^\pi \rangle \right] \leq \zeta \quad (3)$$

$$\text{Privacy: } D_\alpha(\mu_{\mathcal{G}}, \mu_{\mathcal{G}'}) \leq \epsilon \quad \forall \mathcal{G} \sim \mathcal{G}'. \quad (4)$$

When  $\max_{\hat{\pi}_i \in \Delta^{\mathcal{A}_i}} \langle \pi_i - \hat{\pi}_i, g_i^\pi \rangle \leq \epsilon$  for every player  $i \in [N]$ , the strategy profile  $\pi$  is called an  $\epsilon$ -approximate NE. The accuracy metric  $\max_{\hat{\pi}_i \in \Delta^{\mathcal{A}_i}} \langle \pi_i - \hat{\pi}_i, g_i^\pi \rangle$  is also known as *exploitability*. It measures how much player  $i$  can benefit herself by unilaterally deviating from the current strategy profile  $\pi$ , which should be 0 when  $\pi$  is the equilibrium. It is a weaker accuracy guarantee than the Euclidean distance to the equilibrium. Because when  $\pi$  is close to the equilibrium in terms of Euclidean distance, then the exploitability must be small. But the reverse does not always hold.

In the following, we will show that if the adversary has access to all communication channels, it is impossible to achieve low exploitability  $\zeta$  and low privacy budget  $\epsilon$  simultaneously.

**Lemma 4.1.** *For any  $N \geq 12$ , there exists two zero-sum adjacent polymatrix games with  $N$  players so that for any algorithm guaranteeing (3) and (4), we have*

$$\zeta \geq \min \left\{ \frac{3 \exp(-2\epsilon)}{112}, \frac{1}{112} \right\}. \quad (5)$$

The proof is postponed to Appendix A.1, where we provide a lower bound for  $(\epsilon, \delta)$ -DP. Lemma 4.1 implies that even when  $N$  goes to infinity, once  $\zeta \leq \frac{1}{112}$ , we have  $\epsilon \geq \frac{\log 3}{2}$ .

Moreover, even when the adversary only has access to the observation of a single player, it is still impossible to find an accurate approximation of the equilibrium, in terms of Euclidean distance to the equilibrium, while guaranteeing privacy.

Let  $\Delta^{\mathcal{A}_i,*}$  be the set of Nash equilibrium (NE) of player  $i$  in a zero-sum polymatrix game. Moreover,  $\mu_{\mathcal{G},i}$  denotes marginal distribution on the observation of player  $i$  given the joint distribution  $\mu_{\mathcal{G}}$  on all observations. Then, consider an algorithm with the following guarantee in any  $N$ -player game  $\mathcal{G}$ . There exists a parameter  $\alpha \geq 1$  such that for any player  $i \in [N]$ , the algorithm satisfies that

$$\text{Accuracy: } \mathbb{E} \left[ \|\pi_i - \text{Proj}_{\Delta^{\mathcal{A}_i,*}}(\pi_i)\|^2 \right] \leq \zeta_i \quad (6)$$

$$\text{Privacy: } D_\alpha(\mu_{\mathcal{G},i}, \mu_{\mathcal{G}',i}) \leq \epsilon_i \quad \forall \mathcal{G} \sim \mathcal{G}'. \quad (7)$$

(6) guarantees that the output strategy of the algorithm is close to the set of NE, in terms of Euclidean distance. (7) states that for any two adjacent games  $\mathcal{G} \sim \mathcal{G}'$ , the Rényi divergence between the distribution of player  $i$ 's observations is small. This implies that by accessing the communication channel of player  $i$ , the adversary cannot distinguish  $\mathcal{G}$  and  $\mathcal{G}'$ , so he cannot speculate the utility matrices.

**Lemma 4.2.** *For any  $N \geq 8$ , there exists two zero-sum adjacent polymatrix games with  $N$  players so that for any algorithm guaranteeing (6) and (7), then*

$$\frac{1}{N} \sum_{i=1}^N \zeta_i \geq \min \left\{ \frac{3}{8} \exp \left( -\frac{4}{N} \sum_{i=1}^N \epsilon_i \right), \frac{1}{16} \right\}. \quad (8)$$

The proof is postponed to Appendix A.2. Therefore, Lemma 4.2 implies that once  $\frac{1}{N} \sum_{i=1}^N \zeta_i \leq \frac{1}{16}$ , the average differential privacy budget  $\frac{1}{N} \sum_{i=1}^N \epsilon_i \geq \frac{\log 6}{4}$ , even when  $N$  goes to infinity.

## 5 Methods

In this section, we will propose an iterative algorithm for finding coarse correlated equilibrium (CCE) in general-sum polymatrix games (Moulin and Vial, 1978). As we will show in Section 7, this algorithm guarantees differential privacy. We start by recalling a well-known connection between regret and convergence to CCE.

**Proposition 5.1.** *For an iterative algorithm that produces strategy profile  $\pi^{(t)}$  at timestep  $t \in [T]$ , its average strategy is the  $\zeta$ -approximate CCE of a polymatrix game  $\mathcal{G}$  with  $N$  players if and only if for any player  $i \in [N]$*

$$\max_{\hat{\pi}_i \in \Delta^{\mathcal{A}_i}} \frac{1}{T} \sum_{t=1}^T \left\langle g_i^{\pi^{(t)}}, \pi_i^{(t)} - \hat{\pi}_i \right\rangle \leq \zeta. \quad (9)$$

By picking a timestep  $t \in [T]$  uniformly at random and all players will play according to  $\pi^{(t)}$ , the expected increase of utility when player  $i$  deviates to  $\hat{\pi}_i$  is  $\frac{1}{T} \sum_{t=1}^T \left\langle g_i^{\pi^{(t)}}, \pi_i^{(t)} - \hat{\pi}_i \right\rangle$ .

Let  $g_i^{(t)} := g_i^{\bar{\pi}^{(t)}}, I^{\mathcal{A}_i}$  as the identity matrix with index  $\mathcal{A}_i \times \mathcal{A}_i$ , and  $\mathcal{N}(\mathbf{0}, \sigma^2 I^{\mathcal{A}_i})$  as the i.i.d. normal distribution over  $\mathcal{A}_i$  with zero mean and  $\sigma^2$  variance. Then, the algorithm is as follows: for any player  $i \in [N]$ , at each iteration  $t \in \{0, 1, 2, \dots, T-1, T\}$  for some  $T > 0$  ( $\pi_i^{(0)}$  is initialized as the uniform distribution over  $\mathcal{A}_i$ ), player  $i$  will sample a random noise  $n_i^{(t)} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I^{\mathcal{A}_i})$ . Then, she will broadcast her strategy  $\pi_i^{(t)}$  plus the noise to her neighbors. After receiving the noisy strategies from her neighbors, she will do one step of the projected gradient descent with the gradient  $g_i^{(t)}$ , learning rate  $\eta$ , and weight-decay parameter  $\tau_i$ . The complete algorithm is shown in Algorithm 1.

---

**Algorithm 1** Find CCE in polymatrix games with DP guarantee

---

```

Input: Player index  $i$ 
Initialize  $\pi_i^{(0)}$  as uniform distribution over  $\mathcal{A}_i$ 
Let  $\bar{N} := N \cdot (\sum_{i=1}^N \frac{1}{|\mathcal{N}(i)|})^{-1}$  be the harmonic mean of players' degrees
Initialize  $\tau_i \leftarrow \frac{(\bar{N})^{5/9}}{|\mathcal{N}(i)| \log N}$ 
for  $t = 0, 1, \dots, T$  do
    Sample  $n_i^{(t)} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I^{\mathcal{A}_i})$ 
    Broadcast  $\pi_i^{(t)} + n_i^{(t)}$  to neighbors  $j \in \mathcal{N}(i)$ 
    for  $j \in \mathcal{N}(i)$  do
        Receive  $\pi_j^{(t)} + n_j^{(t)}$ 
         $\bar{\pi}_j^{(t)} \leftarrow \text{Proj}_{\Delta^{\mathcal{A}_j}} \left( \pi_j^{(t)} + n_j^{(t)} \right)$ 
    end for
    Gradient  $g_i^{(t)} \leftarrow -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} U_{i,j} \bar{\pi}_j^{(t)}$ 
     $\pi_i^{(t+1)} \leftarrow \underset{\pi_i \in \Delta^{\mathcal{A}_i}}{\text{argmin}} \left\langle \pi_i, g_i^{(t)} \right\rangle + \tau_i \|\pi_i\|^2 + \frac{1}{\eta} \left\| \pi_i - \bar{\pi}_i^{(t)} \right\|^2. \quad (10)$ 
     $\{(10)\} \text{ is equivalent to } \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \frac{\bar{\pi}_i^{(t)} - \eta g_i^{(t)}}{1 + \eta \tau_i} \right).$ 
end for

```

---

The intuition of  $\tau_i \geq 0$  being inversely proportional to  $|\mathcal{N}(i)|$  is that for players with fewer neighbors, her

gradient is more sensitive to the variation of a single utility matrix, so more additional regularization (larger decay rate) is required to hide the variation of  $\pi_i^{(t)}$  due to the change of utility matrix from the adversary.

Moreover, to avoid observations from two adjacent games from increasingly diverging as updating more timesteps, the gradient descent of  $\pi_i^{(t+1)}$  starts from  $\bar{\pi}_i^{(t)}$  instead of  $\pi_i^{(t)}$ . However, we need to pay the price of suffering additional approximation error caused by the noise added to  $\bar{\pi}_i^{(t)}$ .

*Remark 5.2.* Let  $\mathbf{n}^{(t)} := (\mathbf{n}_1^{(t)}, \mathbf{n}_2^{(t)}, \dots, \mathbf{n}_N^{(t)})$ . In Algorithm 1, the execution of the algorithm at timestep  $t$  is  $(\boldsymbol{\pi}^{(t)}, \mathbf{n}^{(t)})$ . Then,  $\mathcal{R}_{\mathcal{G}} \left( \left( \boldsymbol{\pi}^{(s)}, \mathbf{n}^{(s)} \right)_{s=1}^t \right) = \left( \bigcup_{i \in [N]} \left\{ \pi_i^{(s)} + \mathbf{n}_i^{(s)} \right\} \right)_{s=1}^t$ .

## 6 Convergence Analysis

This section will show the convergence of Algorithm 1 in general-sum polymatrix games.

**Theorem 6.1.** Consider Algorithm 1. Let  $A = \max_{i \in [N]} |\mathcal{A}_i|$ . The update-rule can achieve the following guarantee in any polymatrix game. For any  $T > 0$ , player  $i \in [N]$  and strategy  $\pi_i \in \Delta^{\mathcal{A}_i}$ , we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left[ \langle \mathbf{g}_i^{\boldsymbol{\pi}^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle \right] \\ & \leq \frac{1}{\eta T} + A \frac{\sigma^2}{2\eta} + \left( \frac{2\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} + \frac{1}{2(\mathcal{N})^{4/9} \log N} + \frac{2\eta\sqrt{A}}{\sigma(\mathcal{N})^{4/9} \log N}. \end{aligned} \quad (11)$$

The proof is postponed in Appendix B. Theorem 6.1 implies that the average strategy converges to coarse correlated equilibrium (CCE). Besides, it is noteworthy that when applying more noise for better differential privacy guarantees, i.e. larger  $\sigma$ , the convergence rate will worsen since the gradients provide less information.

### 6.1 Proof Sketch

To show the convergence of Algorithm 1, by Lemma D.4. in Liu et al. (2024), we have

$$\begin{aligned} & \frac{\tau_i}{2} \|\pi_i^{(t+1)}\|^2 - \frac{\tau_i}{2} \|\pi_i\|^2 + \langle \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \\ & \leq \frac{1}{2\eta} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1 + \eta\tau_i}{2\eta} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2\eta} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2. \end{aligned}$$

Moreover,  $\left| \frac{\tau_i}{2} \|\pi_i^{(t+1)}\|^2 - \frac{\tau_i}{2} \|\pi_i\|^2 \right| \leq \frac{\tau_i}{2}$ . Then, by adding  $\langle \mathbf{g}_i^{\boldsymbol{\pi}^{(t+1)}} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle$  on both sides, we have

$$\begin{aligned} & \langle \mathbf{g}_i^{\boldsymbol{\pi}^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle - \frac{\tau_i}{2} \\ & \leq \frac{1}{2\eta} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1 + \eta\tau_i}{2\eta} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2\eta} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2 + \langle \mathbf{g}_i^{\boldsymbol{\pi}^{(t+1)}} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle. \end{aligned}$$

Moreover,  $\mathbb{E} \left[ \langle g_i^{\pi^{(t+1)}} - g_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \right] \leq \left( \frac{\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j$ , where the complete proof is postponed to Lemma B.1. Therefore, by taking expectation on both sides and telescoping, we have

$$\begin{aligned} & \sum_{t=1}^T \mathbb{E} \left[ \langle g_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle \right] \\ & \leq \frac{1}{2\eta} \mathbb{E} \left[ \|\pi_i - \bar{\pi}_i^{(1)}\|^2 \right] + \left( \frac{\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} T + \frac{\tau_i}{2} T \\ & \quad + \sum_{t=2}^T \mathbb{E} \left[ \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \|\pi_i - \bar{\pi}_i^{(t-1)}\|^2 \right] + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j. \end{aligned}$$

Further,  $\mathbb{E} \left[ \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \|\pi_i - \bar{\pi}_i^{(t-1)}\|^2 \right]$  is bounded by,

$$\begin{aligned} \mathbb{E} \left[ \|\pi_i - \text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi_i^{(t)} + \mathbf{n}_i^{(t)})\|^2 - \|\pi_i - \bar{\pi}_i^{(t-1)}\|^2 \right] & \leq \mathbb{E} \left[ \|\pi_i - \bar{\pi}_i^{(t-1)} - \mathbf{n}_i^{(t)}\|^2 - \|\pi_i - \bar{\pi}_i^{(t-1)}\|^2 \right] \\ & = \mathbb{E} \left[ 2 \langle \mathbf{n}_i^{(t)}, \pi_i - \bar{\pi}_i^{(t-1)} \rangle + \|\mathbf{n}_i^{(t)}\|^2 \right] \\ & \stackrel{(i)}{=} \mathbb{E} \left[ \|\mathbf{n}_i^{(t)}\|^2 \right] = A\sigma^2. \end{aligned}$$

(i) is because  $\mathbb{E} [\mathbf{n}_i^{(t)}] = \mathbf{0}$  and  $\mathbf{n}_i^{(t)}$  is independent of  $\pi_i - \bar{\pi}_i^{(t)}$ . By aggregating the results above and substitute  $\tau_i$  by its value in Algorithm 1, the proof is concluded.

## 7 Analysis of Differential Privacy

In this section, we show that Algorithm 1 can simultaneously guarantee differential privacy. Since Lemma 4.1 states that the adversary should not have access to all communication channels, we will bound the Rényi DP when the adversary only has access to the communication channel of a single player, i.e. Rényi divergence of  $\mu_{\mathcal{G},i}^{(t)}$  the marginal distribution of  $\mu_{\mathcal{G}}^{(t)}$  over  $(\pi_i^{(s)} + \mathbf{n}_i^{(s)})_{s=1}^t$ . According to the composition rule of Rényi differential privacy (Mironov, 2017), when the adversary has access to multiple communication channels, the privacy budget is the product of the number of communication channels he has access to, and the privacy budget when he only has access to a single communication channel.

We break the analysis into two cases: dense graphs and sparse graphs. In the following, we will show that the privacy budget of Algorithm 1 is proportional to  $\eta^2/T$  but inversely proportional to  $\sigma^2$ . Intuitively, when the learning rate  $\eta$  is larger or the algorithm runs more timesteps, the observation will diverge more due to the different utility matrices. But when we add more noise, i.e. larger  $\sigma$ , it will be harder to determine which distribution a sample  $\pi_i^{(t)} + \mathbf{n}_i^{(t)}$  is from.

**Theorem 7.1.** Consider Algorithm 1 and any two adjacent polymatrix games  $\mathcal{G}, \mathcal{G}'$  that differs at  $(v_1, v_2)$ . Let  $\text{dist}(i, j)$  be the length of the shortest path between players  $i, j$ , which is  $\infty$  when  $i, j$  are not connected. The update-rule guarantees the following for any  $T > 0$  and player  $i \in [N]$ ,

$$\frac{1}{N} \sum_{i=1}^N D_\alpha \left( \mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)} \right) \leq \frac{\alpha\eta^2}{\sigma^2} \min \{ \clubsuit, \spadesuit \} T, \quad (12)$$

where

$$\clubsuit := \frac{16A^3 (\log N)^2}{(\bar{\mathcal{N}})^{4/9}} + \frac{4A}{N} \quad (13)$$

$$\spadesuit := \frac{2A}{N} \sum_{i=1}^N \mathbb{1}(T > \min\{dist(i, v_1), dist(i, v_2)\}). \quad (14)$$

Theorem 7.1 implies that when the graph is dense, i.e.  $\bar{\mathcal{N}} \geq N^p$  for some  $p \in (0, 1)$ ,  $\clubsuit$  will be sublinear in  $N$ . When the graph is sparse, i.e.  $\mathcal{N}^{\max} := \max_{i \in [N]} |\mathcal{N}(i)|$  is a constant,  $\spadesuit$  will be small since the most players will be distant from  $v_1, v_2$  (approximately  $\Omega(\log N)$ ).

In the following, we sketch the important steps in the proof of Theorem 7.1. A complete proof can be found in Appendix C of the Appendix.

## 7.1 Proof Sketch: the Case of Dense Graphs ( $\clubsuit$ )

Firstly, we will introduce the chain rule of Rényi divergence.

**Lemma 7.2** (Chain Rule of Rényi divergence). *For any distribution  $p, q$  over random variables  $X^1, X^2$ , for any  $\alpha \geq 1$ , we have*

$$D_\alpha(p(X^1, X^2), q(X^1, X^2)) \leq D_\alpha(p(X^1), q(X^1)) + \sup_{\hat{x}^1} D_\alpha(p(X^2 | X^1 = \hat{x}^1), q(X^2 | X^1 = \hat{x}^1)).$$

A proof of the above lemma is provided in Appendix C.3 for completeness. By Lemma 7.2, the divergence between the distribution over all observations can be decomposed into the divergence between the distribution over the observation at a single timestep. Formally, for any  $\alpha \geq 1$ , adjacent games  $\mathcal{G} \sim \mathcal{G}'$ , and player  $i \in [N]$ ,

$$D_\alpha(\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)}) \leq \sum_{t=1}^T \sup_{(\hat{o}^{(1)}, \dots, \hat{o}^{(t-1)})} D_\alpha(\mu_{\mathcal{G},i}^{(t)}(\cdot | \hat{o}^{(1)}, \dots, \hat{o}^{(t-1)}), \mu_{\mathcal{G}',i}^{(t)}(\cdot | \hat{o}^{(1)}, \dots, \hat{o}^{(t-1)})).$$

Note that the observation here  $\hat{o}^{(s)} = \pi_i^{(s)} + \mathbf{n}_i^{(s)}$ . When the graph is dense, i.e.  $\bar{\mathcal{N}} = N^p$  for some  $p > 0$ , the degree  $|\mathcal{N}(i)|$  of most players will close to  $\bar{\mathcal{N}}$ .

Then, since post-processing will not increase the privacy budget, we can augment the set of observations to bound  $D_\alpha(\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)})$ . Specifically, for two adjacent games  $\mathcal{G} \sim \mathcal{G}'$  that differs at the utility matrices on edge  $(v_1, v_2)$ , the augmented observation will include  $\{\pi_j^{(t)} + \mathbf{n}_j^{(t)}\}_{j \in \mathcal{S}'}$  where  $\mathcal{S} \supseteq [N] \setminus \{v_1, v_2\}$  and  $v_1, v_2$  will be included in  $\mathcal{S}$  if their degrees are no less than  $\bar{\mathcal{N}}^{2/9}$ . Let  $\mu_{\mathcal{G},\mathcal{S}}^{(t)}$  denotes the marginal distribution of  $\mu_{\mathcal{G}}^{(t)}$  on  $\{\pi_j^{(t)} + \mathbf{n}_j^{(t)}\}_{j \in \mathcal{S}}$ . We can see that for any  $i \in \mathcal{S}$ ,

$$D_\alpha(\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)}) \leq D_\alpha(\mu_{\mathcal{G},\mathcal{S}}^{(T)}, \mu_{\mathcal{G}',\mathcal{S}}^{(T)}).$$

For  $i \notin \mathcal{S}$ , we can show that  $D_\alpha(\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)})$  is bounded by some constant. Therefore, since  $|\mathcal{S}| \geq N - 2$ ,  $\frac{1}{N} \sum_{i=1}^N D_\alpha(\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)})$  is bounded when  $D_\alpha(\mu_{\mathcal{G},\mathcal{S}}^{(T)}, \mu_{\mathcal{G}',\mathcal{S}}^{(T)})$  is sublinear in  $N$ .

Note that  $\mu_{\mathcal{G}, \mathcal{S}}^{(t)}(\cdot | \hat{o}^{(1)}, \dots, \hat{o}^{(t-1)})$  is a multivariate Gaussian distribution with mean  $(\pi_i^{(t)})_{i \in \mathcal{S}}$  and variance  $\sigma^2 \mathbf{I}^{\times_{i \in \mathcal{S}} \mathcal{A}_i}$ . Therefore, by the Rényi divergence of multivariate Gaussian (Gil et al., 2013), the divergence is bounded by  $\frac{\alpha}{2\sigma^2} \sum_{i \in \mathcal{S}} \left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\|^2$ .

If  $\mathcal{S} = [N]$ , then for any player  $i \in [N] \setminus \{v_1, v_2\}$ , given all past observations  $\left( \left\{ \pi_j^{(s)} + \mathbf{n}_j^{(s)} \right\}_{j \in \mathcal{S}} \right)_{s=1}^{t-1}$  are identical in  $\mathcal{G}, \mathcal{G}'$ ,  $\pi_i^{(t)} = \pi'^{(t)}_i$ . For  $i \in \{v_1, v_2\}$ , given  $\bar{\pi}_i^{(t-1)} = \bar{\pi}'_i^{(t-1)}$  and only  $\mathbf{U}_{v_1, v_2} \neq \mathbf{U}'_{v_1, v_2}$ ,  $\left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\| \leq \mathcal{O} \left( \frac{1}{|\mathcal{N}(i)|} \right) \leq \mathcal{O} \left( \frac{1}{(\bar{N})^{2/9}} \right)$  by definition of  $\mathcal{S}$ .  $\mathcal{O}$  hides constants invariant to the number of players  $N$ , e.g.  $A$ , the size of the largest action set.

If  $v_1, v_2 \notin \mathcal{S}$ , then for any  $i \in \mathcal{S} \setminus (\mathcal{N}(v_1) \cup \mathcal{N}(v_2))$ ,  $\pi_i^{(t)} = \pi'^{(t)}_i$ . For  $i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)$ , we will further augment the space of observations so that the adversary can observe  $\left( \left\{ \mathbf{n}_{v_1}^{(s)}, \mathbf{n}_{v_2}^{(s)} \right\} \right)_{s=1}^{t-1}$ . Assume  $i \in \mathcal{N}(v_1) \setminus \mathcal{N}(v_2)$  for ease of exposition. For any  $j \in \mathcal{N}(i) \setminus \{v_1, v_2\}$ , we have  $\bar{\pi}_j^{(t-1)} = \bar{\pi}'_j^{(t-1)}$ . Therefore,  $\left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\| \leq \mathcal{O} \left( \left\| \bar{\pi}_{v_1}^{(t-1)} - \bar{\pi}'_{v_1}^{(t-1)} \right\| \right) \leq \mathcal{O} \left( \left\| \pi_{v_1}^{(t-1)} - \pi'_{v_1}^{(t-1)} \right\| \right)$ . Moreover, due to the additional regularization imposed on each player, we have the following lemma.

**Lemma 7.3.** Consider the Algorithm 1. For any player  $i \in [N]$  and timestep  $t > 0$ , by updating the strategy  $\left\{ \pi_i^{(s)} \right\}_{s=0}^t, \left\{ \pi'^{(s)}_i \right\}_{s=0}^t$  with the same noise  $\left\{ \mathbf{n}_i^{(s)} \right\}_{s=0}^t$  but different gradients  $\left\{ \mathbf{g}_i^{(s)} \right\}_{s=0}^t, \left\{ \mathbf{g}'_i^{(s)} \right\}_{s=0}^t$  individually, we have

$$\left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\| \leq \frac{2\sqrt{\mathcal{A}_i}}{\tau_i}. \quad (15)$$

The proof is postponed to Appendix C.3. Therefore, finally,  $\left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\| \leq \mathcal{O} \left( \frac{1}{\tau_{v_1}} \right) \leq \mathcal{O} \left( \frac{\log N}{(\bar{N})^{1/3}} \right)$  because  $|\mathcal{N}(v_1)| \leq (\bar{N})^{2/9}$ . Moreover, since  $|\mathcal{N}(v_1)| \leq (\bar{N})^{2/9}$ ,

$$\sum_{i \in \mathcal{S}} \left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\|^2 = \sum_{i \in \mathcal{S} \cap (\mathcal{N}(v_1) \cup \mathcal{N}(v_2))} \left\| \pi_i^{(t)} - \pi'^{(t)}_i \right\|^2 \leq \mathcal{O} \left( \frac{(\log N)^2}{(\bar{N})^{4/9}} \right).$$

For cases of  $v_1 \in \mathcal{S}, v_2 \notin \mathcal{S}$  and  $v_1 \notin \mathcal{S}, v_2 \in \mathcal{S}$ , the proof is similar.

## 7.2 Proof Sketch: the Case of Sparse Graphs (♠)

For the sparse graph, the degree of all players is bounded by a constant  $\mathcal{N}^{\max}$ . For simplicity, let's consider two adjacent games  $\mathcal{G} \sim \mathcal{G}'$  which differs at the utility matrices on edge  $(v_1, v_2) \in E$ , and  $\sigma = 0$  in Algorithm 1, i.e. the noise-free case.

At timestep  $t$ ,  $\pi_i^{(t)}$  in  $\mathcal{G}$  and its counterpart  $\pi'^{(t)}_i$  in  $\mathcal{G}'$  are identical if  $t \leq \min \{ \text{dist}(i, v_1), \text{dist}(i, v_2) \}$ . This can be proved by mathematical induction.

Therefore, Rényi divergence between  $\mu_{\mathcal{G}, i}^{(t)}(\cdot | \hat{o}^{(1)}, \dots, \hat{o}^{(t-1)})$  and  $\mu_{\mathcal{G}', i}^{(t)}(\cdot | \hat{o}^{(1)}, \dots, \hat{o}^{(t-1)})$  is 0, since they are both normal distribution with mean  $\pi_i^{(t)}$  and variance 0.

Recall that the graph is sparse and the degree of each player is bounded by a constant  $\mathcal{N}^{\max}$ . In such cases, most players are at least of distance  $\mathcal{O}(\log_{\mathcal{N}^{\max}} N)$  to the edge  $(v_1, v_2)$ . Therefore, if  $T$  in Algorithm 1

is no larger than  $\mathcal{O}(\log_{\mathcal{N}^{\max}} N)$ , the Rényi divergence of most players at timestep  $t$  is 0. Therefore, by averaging across all players, the privacy budget will be 0 as the number of players goes to infinity.

The general proof for the cases when  $\sigma > 0$  is shown in Appendix C.2.

## 8 Trade-off Between Accuracy and Differential Privacy

According to Theorem 6.1 and Theorem 7.1, by picking the hyper-parameters of Algorithm 1 carefully, when the number of players grows to infinity, we may obtain a zero privacy budget while converging to CCE.

**Theorem 8.1.** *Consider Algorithm 1 with  $\sigma = \frac{1}{\sqrt{T}}$  and any two adjacent polymatrix games  $\mathcal{G}, \mathcal{G}'$  that differs at  $(v_1, v_2)$ . The update-rule guarantees the following,*

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \max \left\{ \frac{1}{T} \sum_{t=1}^T \mathbb{E} [\langle g_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle], 0 \right\} \\ & \leq \frac{A+1}{\eta T} + 2A^{3/2}\eta^2\sqrt{T} + \frac{7A^{3/2}}{2\sqrt{T}} + \frac{1}{2(\bar{\mathcal{N}})^{4/9} \log N} + \frac{2\eta\sqrt{AT}}{(\bar{\mathcal{N}})^{4/9} \log N} \\ & \quad \frac{1}{N} \sum_{i=1}^N D_\alpha (\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)}) \leq \alpha\eta^2 \min \{\clubsuit, \spadesuit\}. \end{aligned} \tag{16}$$

The proof is simply substituting the values of  $\sigma$  into Theorem 6.1 and Theorem 7.1. When the graph is dense, i.e.  $\bar{\mathcal{N}} \geq N^p$  for some  $p > 0$ , take  $\eta = \frac{1}{T\clubsuit^{1/3}}$ ,  $T = \frac{N^{8p/9}}{(\log N)^4}$ , then

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \mathbb{E} [\langle g_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle] \leq \mathcal{O} \left( \frac{(\log N)^{2/3}}{N^{4p/27}} \right),$$

where the  $\mathcal{O}$  notation takes  $A$  as constant. Simultaneously, the differential budget guarantees that

$$\frac{1}{N} \sum_{i=1}^N D_\alpha (\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)}) \leq \mathcal{O} \left( \frac{\alpha (\log N)^{2/3}}{N^{4p/27}} \right).$$

Therefore, when the graph is dense, the accuracy and differential privacy budget of Algorithm 1 would both be better as the number of players increases.

Consider when the graph is sparse, i.e.  $\mathcal{N}^{\max}$  is a constant. If  $\mathcal{N}^{\max} \geq 2$ , let  $\eta = \frac{1}{T\spadesuit^{1/3}}$ ,  $T = (1 - \log_N \log N) \log_{\mathcal{N}^{\max}} N$ , then  $\spadesuit \leq \frac{4A}{\log N}$ , since at most  $\frac{2(\mathcal{N}^{\max})^T}{\mathcal{N}^{\max}-1} \leq \frac{2N}{\log N}$  players satisfy  $\min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\} < T$ . If  $\mathcal{N}^{\max} = 1$ ,  $\spadesuit \leq \frac{2}{N}$  for any  $T > 0$ . Therefore,

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \mathbb{E} [\langle g_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle] \leq \mathcal{O} \left( \frac{1}{(\log N)^{1/3}} \right),$$

and

$$\frac{1}{N} \sum_{i=1}^N D_\alpha (\mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)}) \leq \mathcal{O} \left( \frac{\alpha}{(\log N)^{1/3}} \right).$$

## 9 Experiments

The experiment results are depicted in Figures 1 and 2. The implementation is based on Pytorch (Paszke et al., 2019) for efficient parallel update of all players' strategies. In dense graphs, we can see that as the number of players increases, both exploitability and the privacy budget decrease. However, in sparse graphs, only the privacy budget decreases. The reason might be the convergence rate of sparse graphs is  $\mathcal{O}\left(\frac{1}{(\log N)^{1/3}}\right)$ , which is not prominent compared to the constants/noise when  $N \leq 2^{15}$ . Parameters are selected according to Theorem 8.1.

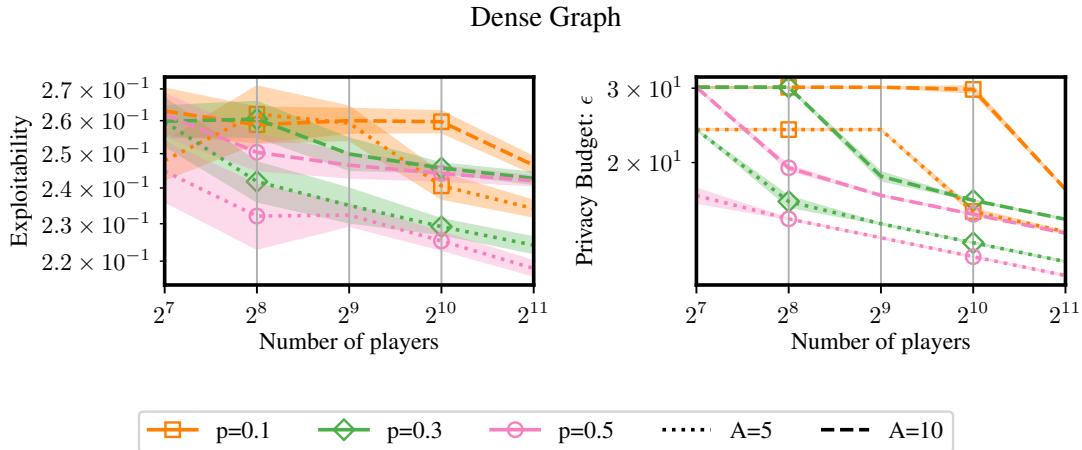


Figure 1: Experiment results of dense graphs. Each node (player) in the graph will be connected to another node with probability  $p$ , the connection is sampled i.i.d. for each node. Then, the duplicate edges will be removed. The action set sizes of all players are set to  $A$ . We can see from the result that the exploitability and the privacy budget both decrease as the number of players increases.

## 10 Conclusion

This paper shows the necessary conditions to find the equilibria of polymatrix games with DP guarantees. Moreover, to justify the conditions, we propose an algorithm to find the equilibria with DP guarantees under the conditions. Moreover, as the number of players in the game increases, the algorithm will gradually achieve perfect accuracy and DP guarantees.

## References

- Naman Agarwal and Karan Singh. The price of differential privacy for online learning. In *International Conference on Machine Learning (ICML)*, 2017.
- Yang Cai, Ozan Candogan, Constantinos Daskalakis, and Christos Papadimitriou. Zero-sum polymatrix games: A generalization of minmax. *Mathematics of Operations Research*, 41(2):648–655, 2016.
- Rachel Cummings, Michael Kearns, Aaron Roth, and Zhiwei Steven Wu. Privacy and truthful equilibrium

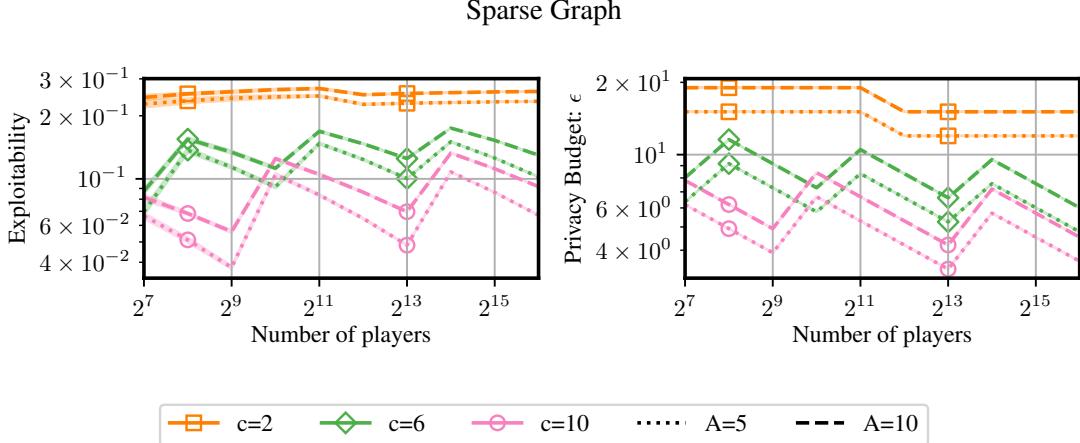


Figure 2: Experiment results of sparse graphs. We will randomly generate  $cN$  edges in total, with each node appearing  $c$  times. Then, duplicate edges and self-loops will be removed. The action set sizes of all players are set to  $A$ . The result shows that while the exploitability remains unchanged, the exploitability decreases as the number of players increases.

- selection for aggregative games. In *Web and Internet Economics: 11th International Conference, WINE 2015, Amsterdam, The Netherlands, December 9-12, 2015, Proceedings 11*, pages 286–299. Springer, 2015.
- Giuseppe De Nittis, Alberto Marchesi, and Nicola Gatti. Computing the strategy to commit to in polymatrix games. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2018.
- Argyrios Deligkas, John Fearnley, Rahul Savani, and Paul Spirakis. Computing approximate nash equilibria in polymatrix games. *Algorithmica*, 77(2):487–514, 2017.
- Ahmed Said Donmez, Yuksel Arslantas, and Muhammed O Sayin. Team collaboration vs competition: New fictitious play dynamics for multi-team zero-sum games. *arXiv preprint arXiv:2402.02147*, 2024.
- Cynthia Dwork. Differential privacy. automata, languages and programming. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2006.
- Shripad Gade, Anna Winnicki, and Subhonmesh Bose. On privatizing equilibrium computation in aggregate games over networks. *IFAC-PapersOnLine*, 53(2):3272–3277, 2020.
- Manuel Gil, Fady Alajaji, and Tamas Linder. Rényi divergence measures for commonly used univariate continuous distributions. *Information Sciences*, 249:124–131, 2013.
- Abhradeep Guha Thakurta and Adam Smith. (nearly) optimal algorithms for private online learning in full-information and bandit settings. *Neural Information Processing Systems (NeurIPS)*, 2013.
- Zhenqi Huang, Sayan Mitra, and Nitin Vaidya. Differentially private distributed optimization. In *International Conference on Distributed Computing and Networking (ICDCN)*, 2015.
- Prateek Jain, Pravesh Kothari, and Abhradeep Thakurta. Differentially private online learning. In *Conference on Learning Theory (COLT)*, 2012.
- Elena Janovskaja. Equilibrium points in polymatrix games. *Lithuanian Mathematical Journal*, 8(2):381–384, 1968.

- Michael Kearns, Mallesh Pai, Aaron Roth, and Jonathan Ullman. Mechanism design in large games: Incentives and privacy. In *Innovations in Theoretical Computer Science (ITCS)*, 2014.
- Mingyang Liu, Gabriele Farina, and Asuman Ozdaglar. A policy-gradient approach to solving imperfect-information games with iterate convergence. *arXiv preprint arXiv:2408.00751*, 2024.
- Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pages 263–275. IEEE, 2017.
- H. Moulin and J.-P. Vial. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *Int. J. Game Theory*, 7(3):201–221, September 1978. ISSN 1432-1270. doi: 10.1007/BF01769190.
- Erfan Nozari, Pavankumar Tallapragada, and Jorge Cortés. Differentially private distributed convex optimization via functional perturbation. *IEEE Transactions on Control of Network Systems*, 5(1):395–408, 2016.
- Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *Neural Information Processing Systems (NeurIPS)*, 2019.
- Ryan M Rogers and Aaron Roth. Asymptotically truthful equilibrium selection in large congestion games. In *ACM Conference on Economics and Computation (EC)*, 2014.
- Mehran Shakarami, Claudio De Persis, and Nima Monshizadeh. Distributed dynamics for aggregative games: Robustness and privacy guarantees. *International Journal of Robust and Nonlinear Control*, 32(9): 5048–5069, 2022.
- Jimin Wang, Ji-Feng Zhang, and Xingkang He. Differentially private distributed algorithms for stochastic aggregative games. *Automatica*, 142:110440, 2022.
- Yongqiang Wang and Tamer Başar. Ensuring both almost sure convergence and differential privacy in nash equilibrium seeking on directed graphs. *IEEE Transactions on Automatic Control*, 2024.
- Yongqiang Wang and Angelia Nedić. Differentially-private distributed algorithms for aggregative games with guaranteed convergence. *IEEE Transactions on Automatic Control*, 2024.
- Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard H Yang, Farhad Farokhi, Shi Jin, Tony QS Quek, and H Vincent Poor. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE transactions on information forensics and security*, 15:3454–3469, 2020.
- Maojiao Ye, Guoqiang Hu, Lihua Xie, and Shengyuan Xu. Differentially private distributed nash equilibrium seeking for aggregative games. *IEEE Transactions on Automatic Control*, 67(5):2451–2458, 2021.

## A Proof of Section 4

The section is arranged as follows. For both Lemma 4.1 and Lemma 4.2, we will show that there exists a polymatrix game, so that achieving accuracy and  $(\epsilon, \delta)$ -DP simultaneously is impossible. Then, we further show that achieving accuracy and Rényi DP is impossible by showing that for  $(1, \epsilon)$ -Rényi DP. Because the  $\alpha$ -Rényi divergence grows monotonically with respect to  $\alpha$  (Mironov, 2017).

### A.1 Proof of Lemma 4.1

**Lemma 4.1.** *For any  $N \geq 12$ , there exists two zero-sum adjacent polymatrix games with  $N$  players so that for any algorithm guaranteeing (3) and (4), we have*

$$\zeta \geq \min \left\{ \frac{3 \exp(-2\epsilon)}{112}, \frac{1}{112} \right\}. \quad (5)$$

Table 1:  $i$  is the row player and  $j$  is the column player, the utility in the table is  $U_{i,j}(a_i, a_j) = -U_{j,i}(a_j, a_i)$ .

$a_1$	$a_1$	$a_1$	$a_1$	$a_1$	$a_1$
$a_1$	1	0	0.5	0	0.5
$a_2$	0	1	0	0.5	0.5

*Proof.* Consider a zero-sum polymatrix game with  $3N$  players, where its corresponding graph is a chain, i.e. the edge set  $E = \{(i, i+1)\}_{i \in [3N-1]}$ . For  $i \in [N]$ ,  $U_{3i-1, 3i-2}$  is Table 1 (left) and  $U_{3i, 3i-1}$  is Table 1 (right). All other utility matrices are zero matrices. Therefore, the only equilibrium is all players playing  $a_1$ . Let denote the game as  $\mathcal{G}^\emptyset$ .

In this case, for any  $i \in [N]$ , when changing  $U_{3i-1, 3i-2}$  to Table 1 (middle), the equilibrium of both  $3i-1$  and  $3i$  will be choosing  $a_2$  deterministically, and the equilibrium of other players keep unchanged. Let denote this new game as  $\mathcal{G}^{\{i\}}$ .

#### A.1.1 Lowerbounds on $(\epsilon, \delta)$ -DP

This section will show that if the algorithm satisfies  $(\epsilon, \delta)$ -DP, then

$$\zeta \geq \frac{1-\delta}{28(1+e^\epsilon)}. \quad (17)$$

Suppose the distributed algorithm satisfies the following guarantee of accuracy and privacy. Let  $\mathbf{g}_i^\pi(\mathcal{G})$  be the gradient of player  $i$  in  $\mathcal{G}$  and  $\boldsymbol{\pi}^{\{i\}}$  be the output of the algorithm in  $\mathcal{G}^{\{i\}}$ . Then,

$$\forall i \in [N], \sum_{j=1}^{3N} \mathbb{E} \left[ \max_{\hat{\pi}_j \in \Delta^{\mathcal{A}_j}} \langle \pi_j^{\{i\}} - \hat{\pi}_j, \mathbf{g}_j^{\boldsymbol{\pi}^{\{i\}}}(\mathcal{G}^{\{i\}}) \rangle \right] \leq N\zeta \quad (\text{Accuracy})$$

$$\forall i \in [N], \text{ set } \mathcal{C}, \quad \Pr(\boldsymbol{\pi}^\emptyset \in \mathcal{C}) \leq e^\epsilon \Pr(\boldsymbol{\pi}^{\{i\}} \in \mathcal{C}) + \delta, \quad (\text{Privacy})$$

Moreover, it is easy to verify that for any strategy profile  $\boldsymbol{\pi} \in \times_{i=1}^{3N} \Delta^{\mathcal{A}_i}$  and  $i \in [N], k \in \{0\} \cup [N]$  (let  $\mathcal{G}^{\{0\}} = \mathcal{G}^\emptyset$  for simplicity), we have

$$\max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1} - \hat{\pi}_{3i-1}, \mathbf{g}_{3i-1}^{\boldsymbol{\pi}}(\mathcal{G}^{\{k\}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i} - \hat{\pi}_{3i}, \mathbf{g}_{3i}^{\boldsymbol{\pi}}(\mathcal{G}^{\{k\}}) \rangle \geq \begin{cases} 0.5\pi_{3i}(a_2) & i \neq k \\ 0.5\pi_{3i}(a_1) & i = k. \end{cases}$$

Since the algorithm is distributed, the findr for each player can be written as  $f_i: \mathcal{O}^T \times \mathbb{R}^{\mathcal{A}_i \times \mathcal{A}_{i-1}} \times \mathbb{R}^{\mathcal{A}_i \times \mathcal{A}_{i+1}} \times \Delta^{\mathcal{A}_i} \rightarrow \mathbb{R}$  (let  $\mathcal{A}_0 = \mathcal{A}_{N+1} = \emptyset$  for notational simplicity), which is a conditional distribution over all possible strategies in  $\Delta^{\mathcal{A}_i}$ , conditioned on all past observations and player's utility matrices.

Let  $\text{adv}(\mathbf{o}, \pi_{3i}) := f_{3i}(\mathbf{o}_{3i}, \mathbf{U}_{3i,3i-1}, \pi_{3i})$  for  $\mathbf{o} \in \mathcal{O}^T$  and  $\pi_{3i} \in \Delta^{\mathcal{A}_{3i}}$ , where  $\mathbf{U}_{3i,3i-1}$  is the utility matrix between player  $3i, 3i-1$  in  $\mathcal{G}^\emptyset$ . Note that since the adversary cannot access utility matrices,  $\text{adv}(\mathbf{o}, \pi_{3i})$  is fixed even though the observation now comes from games  $\mathcal{G}^{\{k\}}$  for some  $k \in [N]$ .

Then, the adversary may sample  $\pi_{3i}^{\text{adv}, k}$  according to  $\text{adv}(\mathbf{o}^k, \pi_{3i})$ , where  $\mathbf{o}^k \in \mathcal{O}^T$  is the observation in game  $\mathcal{G}^{\{k\}}$ . Because the utility matrices of player  $3i$  in  $\mathcal{G}^\emptyset$  and  $\mathcal{G}^{\{k\}}$  are identical, we have

$$\begin{aligned} 0.5\mathbb{E} [\pi_{3i}^{\text{adv}, k}(a_2)] &= 0.5\mathbb{E} [\pi_{3i}^{\{k\}}(a_2)] \\ &\leq \mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^{\{k\}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\{k\}}}(\mathcal{G}^{\{k\}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^{\{k\}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\{k\}}}(\mathcal{G}^{\{k\}}) \rangle \right]. \end{aligned}$$

Similar bounds also holds for  $\mathbb{E} [\pi_{3i}^{\text{adv}}(a_1)]$  when  $i = k$ .

Then, by Markov inequality and the definition of  $(\epsilon, \delta)$ -DP, we have

$$\begin{aligned} &4\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^\emptyset - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^\emptyset - \hat{\pi}_{3i}, g_{3i}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle \right] \\ &\geq 2\mathbb{E} [\pi_{3i}^{\text{adv}, \emptyset}(a_2)] \\ &\geq \Pr (\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5) \\ &= \Pr (\pi_{3i}^{\text{adv}, \emptyset}(a_1) \leq 0.5) \\ &\geq e^{-\epsilon} \Pr (\pi_{3i}^{\text{adv}, \{i\}}(a_1) \leq 0.5) - \delta e^{-\epsilon} \\ &\geq e^{-\epsilon} (1 - 2\mathbb{E} [\pi_{3i}^{\text{adv}, \{i\}}(a_1)]) - \delta e^{-\epsilon} \\ &\geq e^{-\epsilon} \left( 1 - 4\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^{\{i\}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\{i\}}}(\mathcal{G}^{\{i\}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^{\{i\}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\{i\}}}(\mathcal{G}^{\{i\}}) \rangle \right] \right) - \delta e^{-\epsilon}. \end{aligned} \tag{18}$$

Moreover, for each  $i \in [N]$  and  $\mathcal{G}^{\{i\}}$ , we can further define  $\mathcal{G}^{\{i,j\}}$  for  $i \neq j \in [N]$  as the adjacent game to  $\mathcal{G}^{\{i\}}$  that differs from  $\mathcal{G}^{\{i\}}$  only at  $\mathbf{U}_{3j-1,3j-2}$ , by changing it from Table 1 (left) to Table 1 (middle). Also note that  $\mathcal{G}^{\{i,j\}}$  is also adjacent to  $\mathcal{G}^{\{j\}}$ . Such a process can be recursively applied to  $\mathcal{G}^{\{i,j\}}$  and finally we will get  $2^N$  games  $\{\mathcal{G}^{\mathcal{S}}\}_{\mathcal{S} \subseteq [N]}$ .

**Lemma A.1.** Consider the  $2^N$  games  $\{\mathcal{G}^{\mathcal{S}}\}_{\mathcal{S} \subseteq [N]}$  constructed above. There must exist two adjacent games  $\mathcal{G}^{\mathcal{S}}$  and  $\mathcal{G}^{\mathcal{S} \cup \{i\}}$  with  $i \in [N] \setminus \mathcal{S}$ , so that

$$\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^{\mathcal{S}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\mathcal{S}}}(\mathcal{G}^{\mathcal{S}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^{\mathcal{S}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\mathcal{S}}}(\mathcal{G}^{\mathcal{S}}) \rangle \right] \leq 7\zeta \tag{19}$$

$$\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^{\mathcal{S} \cup \{i\}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\mathcal{S} \cup \{i\}}}(\mathcal{G}^{\mathcal{S} \cup \{i\}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^{\mathcal{S} \cup \{i\}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\mathcal{S} \cup \{i\}}}(\mathcal{G}^{\mathcal{S} \cup \{i\}}) \rangle \right] \leq 7\zeta. \tag{20}$$

The proof is postponed to Appendix A.3. Since (18) also holds for the  $\mathcal{G}^{\mathcal{S}}$  and  $\mathcal{G}^{\mathcal{S} \cup \{i\}}$  in Lemma A.1, we have

$$28(e^{-\epsilon} + 1)\zeta \geq e^{-\epsilon} - \delta e^{-\epsilon}.$$

Therefore,  $\zeta \geq \frac{e^{-\epsilon} - \delta e^{-\epsilon}}{28(e^{-\epsilon} + 1)} = \frac{1-\delta}{28(1+e^\epsilon)}$ .

### A.1.2 Lowerbounds on $(\alpha, \epsilon)$ -Rényi DP

This section will show that if the algorithm satisfies  $(\epsilon, \delta)$ -DP, then

$$\frac{1}{N} \sum_{i=1}^N \zeta_i \geq \frac{1 - \max_{i \in [N]} \delta_i}{1 + \exp\left(\frac{1}{N} \sum_{i=1}^N \epsilon_i\right)} - \frac{2}{N}. \quad (21)$$

By monotonicity of Rényi Divergence (Mironov, 2017), we only need to show the lowerbound of KL-divergence. By further post-processing on  $\pi_{3i}^{\text{adv}, \mathcal{S}}$  for  $i \in [N]$  and  $\mathcal{S} \subseteq [N]$ , the adversary can output whether  $\pi_{3i}^{\text{adv}, \mathcal{S}}(a_2) \geq 0.5$ . Let the distribution of the output be  $\mu_{3i}^{\mathcal{S}}$ . Then, for the adjacent games  $\mathcal{G}^\emptyset, \mathcal{G}^{\{i\}}$ , where  $i \in [N]$ , we have

$$\begin{aligned} \epsilon &\geq \text{KL}\left(\mu_{3i}^\emptyset, \mu_{3i}^{\{i\}}\right) = \Pr\left(\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5\right) \log \frac{\Pr\left(\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5\right)}{\Pr\left(\pi_{3i}^{\text{adv}, \{i\}}(a_2) \geq 0.5\right)} \\ &\quad + \left(1 - \Pr\left(\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5\right)\right) \log \frac{1 - \Pr\left(\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5\right)}{1 - \Pr\left(\pi_{3i}^{\text{adv}, \{i\}}(a_2) \geq 0.5\right)} \\ &\stackrel{(i)}{\geq} (1 - \text{expl}_i^\emptyset) \log \frac{1 - \text{expl}_i^\emptyset}{\text{expl}_i^{\{i\}}} + \text{expl}_i^\emptyset \log \frac{\text{expl}_i^\emptyset}{1 - \text{expl}_i^{\{i\}}}, \end{aligned}$$

where

$$\text{expl}_i^{\mathcal{S}} := 4\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \left\langle \pi_{3i-1}^{\mathcal{S}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\mathcal{S}}_{3i-1}}(\mathcal{G}^{\mathcal{S}}) \right\rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \left\langle \pi_{3i}^{\mathcal{S}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\mathcal{S}}_{3i}}(\mathcal{G}^{\mathcal{S}}) \right\rangle \right].$$

(i) is because  $\Pr\left(\pi_{3i}^{\text{adv}, \emptyset}(a_2) \geq 0.5\right) \leq \text{expl}_i^\emptyset$  and  $\Pr\left(\pi_{3i}^{\text{adv}, \{i\}}(a_2) \geq 0.5\right) \geq 1 - \text{expl}_i^{\{i\}}$ .

By Lemma A.1, there exists  $\mathcal{S} \subseteq [N]$  and  $i \in [N] \setminus s$  so that  $\text{expl}_i^{\mathcal{S}}, \text{expl}_i^{\mathcal{S} \cup \{i\}} \leq 28\zeta$ . Therefore,

$$\epsilon \geq (1 - 28\zeta) \log \frac{1 - 28\zeta}{28\zeta} + 28\zeta \log \frac{28\zeta}{1 - 28\zeta} = (1 - 56\zeta) \log \frac{1 - 28\zeta}{28\zeta}.$$

It implies that when  $\zeta \leq \frac{1}{112}$ ,  $\epsilon \geq \frac{1}{2} \log \frac{3}{112\zeta}$  so that  $\zeta \geq \frac{3}{112\exp(2\epsilon)}$ .  $\square$

## A.2 Proof of Lemma 4.2

**Lemma 4.2.** *For any  $N \geq 8$ , there exists two zero-sum adjacent polymatrix games with  $N$  players so that for any algorithm guaranteeing (6) and (7), then*

$$\frac{1}{N} \sum_{i=1}^N \zeta_i \geq \min \left\{ \frac{3}{8} \exp\left(-\frac{4}{N} \sum_{i=1}^N \epsilon_i\right), \frac{1}{16} \right\}. \quad (8)$$

*Proof.* Consider a zero-sum polymatrix game, where its corresponding graph is a chain. Specifically, there is an edge  $(i, j) \in E$ , if and only if  $i = j - 1$ . Moreover, each player has two actions and  $\mathbf{U}_{i,i+1}$  is shown in Table 2 (left).

Table 2:  $i$  is the row player and  $j$  is the column player, the utility in the table is  $\mathbf{U}_i(a_i, a_j)$  and  $\mathbf{U}_j(a_j, a_i) = -\mathbf{U}_i(a_i, a_j)$ .

	$a_1$	$a_2$		$a_1$	$a_2$
$a_1$	0.5	$0.5 - \epsilon_i$	$a_1$	$0.5 - 2\epsilon_i$	$0.5 - 3\epsilon_i$
$a_2$	$0.5 - 3\epsilon_i$	$0.5 - 2\epsilon_i$	$a_2$	$0.5 - \epsilon_i$	0.5

Moreover,  $0.1 = \epsilon_1 = 10\epsilon_2 \geq 10^2\epsilon_3 = \dots = 10^{N-1}\epsilon_N > 0$ . Then, the Nash equilibrium (NE) is  $(a_1, a_2, a_1, a_2, \dots, a_{\frac{(-1)^N+3}{2}})$ . However, when the utility matrix between player (1,2) changes to Table 2 (right), the NE becomes  $(a_2, a_1, a_2, a_1, \dots, a_{\frac{(-1)^{N+1}+3}{2}})$ . Let the original game be denoted as  $\mathcal{G}^0$  and the new game as  $\mathcal{G}^1$ .

With the observation  $\mathbf{o}_i \in \mathcal{O}^T$  of any player  $i > 2$ , the adversary may sample strategy  $\pi_i^{\text{adv}}$  from the distribution  $\text{adv}(\mathbf{o}_i, \pi_i) := f_i(\mathbf{o}_i, \mathbf{U}_{i,i-1}, \mathbf{U}_{i,i+1}, \pi_i)$  to obtain the equilibrium of player  $i$ , where  $\mathbf{U}_{i,i-1}, \mathbf{U}_{i,i+1}$  are both utility matrices in  $\mathcal{G}^0$ .

### A.2.1 Lowerbounds on $(\epsilon, \delta)$ -DP

For each player  $i \in [N] \setminus \{1, 2\}$ , assume the distributed algorithm may achieve the following guarantee.

$$\forall k \in \{0, 1\}, \quad \mathbb{E} \left[ \left\| \pi_i^k - \pi_i^{*,k} \right\|^2 \right] \leq \zeta_i \quad (\text{Accuracy})$$

$$\forall \mathcal{C}_i \subseteq \Delta^{\mathcal{A}_i}, \quad \Pr(\pi_i^0 \in \mathcal{C}_i) \leq e^{\epsilon_i} \Pr(\pi_i^1 \in \mathcal{C}_i) + \delta_i, \quad (\text{Privacy})$$

where  $\pi_i^0, \pi_i^1$  is the output of the algorithm when deploying on player  $i$  for  $\mathcal{G}^0, \mathcal{G}^1$ .

Let  $\pi_i^{\text{adv},k}$  be the random variable sampled from  $\text{adv}(\mathbf{o}_i^k, \cdot)$ , where  $\mathbf{o}_i^k$  is the observation in  $\mathcal{G}^k$  with  $k \in \{0, 1\}$ . Since post-processing will not decrease the strength of DP, observing  $\pi_i^{\text{adv},k}$  instead of  $\mathbf{o}_i^k$  will leak less information. Moreover, since the utility matrices of players  $i > 2$  do not change, for any player  $i > 2$  and  $k \in \{0, 1\}$ , we have

$$\mathbb{E} \left[ \left\| \pi_i^k - \pi_i^{*,k} \right\|^2 \right] = \mathbb{E} \left[ \left\| \pi_i^{\text{adv},k} - \pi_i^{*,k} \right\|^2 \right].$$

Let the equilibrium of  $\mathcal{G}^0$  be  $\boldsymbol{\pi}^{*,0} := ((1, 0), (0, 1), \dots)$  and the equilibrium of  $\mathcal{G}^1$  be  $\boldsymbol{\pi}^{*,1} := ((0, 1), (1, 0), \dots)$ . Then, by Markov inequality, for any  $k \in \{0, 1\}$  and  $i > 2$ , we have

$$\Pr \left( \left\| \pi_i^{\text{adv},k} - \pi_i^{*,k} \right\|^2 \geq 1 \right) \leq \frac{\mathbb{E} \left[ \left\| \pi_i^{\text{adv},k} - \pi_i^{*,k} \right\|^2 \right]}{1} = \frac{\mathbb{E} \left[ \left\| \pi_i^k - \pi_i^{*,k} \right\|^2 \right]}{1} \leq \zeta_i.$$

Therefore, according to the privacy guarantee, for any  $i > 2$ , we have

$$\begin{aligned} 1 - \zeta_i &\leq \Pr \left( \left\| \pi_i^{\text{adv},0} - \pi_i^{*,0} \right\|^2 \leq 1 \right) \stackrel{(i)}{\leq} \Pr \left( \left\| \pi_i^{\text{adv},0} - \pi_i^{*,1} \right\|^2 \geq 1 \right) \\ &\stackrel{(ii)}{\leq} e^{\epsilon_i} \Pr \left( \left\| \pi_i^{\text{adv},1} - \pi_i^{*,1} \right\|^2 \geq 1 \right) + \delta_i \leq e^{\epsilon_i} \zeta_i + \delta_i. \end{aligned}$$

(i) is because  $\|\pi_i^{*,0} - \pi_i^{*,1}\|^2 = 2$ . (ii) is because post-processing will not decrease the strength of DP.

Therefore, we have  $\zeta_i \geq \frac{1-\delta_i}{1+e^{\epsilon_i}}$  for  $i > 2$ . For  $i \in \{1, 2\}$ , we have  $1 \geq \frac{1-\delta_i}{1+e^{\epsilon_i}}$  since  $1 - \delta_i \leq 1$ .

Moreover, we have

$$\frac{1}{N} \left( 2 + \sum_{i=3}^N \zeta_i \right) \geq \frac{1}{N} \sum_{i=1}^N \frac{1-\delta_i}{1+e^{\epsilon_i}} \geq \frac{1}{N} \sum_{i=1}^N \frac{1 - \max_{i \in [N]} \delta_i}{1+e^{\epsilon_i}} \stackrel{(i)}{\geq} \frac{1 - \max_{i \in [N]} \delta_i}{1 + \exp\left(\frac{1}{N} \sum_{i=1}^N \epsilon_i\right)}.$$

(i) uses Jensen's inequality.

### A.2.2 Lowerbounds on $(\alpha, \epsilon)$ -Rényi DP

By further post-processing on  $\pi_i^{\text{adv},k}$ , we can get a random variable sampled from  $\mu_i^k$ , which is the distribution over  $\{0, 1\}$  indicating whether  $\|\pi_i^{\text{adv},k} - \pi_i^{*,0}\| \leq 1$ . Then, for any  $i > 2$ , we have

$$\epsilon_i \stackrel{(i)}{\geq} \text{KL}(\mu_i^0, \mu_i^1) \stackrel{(ii)}{\geq} (1 - \zeta_i) \log \frac{1 - \zeta_i}{\zeta_i} + \zeta_i \log \frac{\zeta_i}{1 - \zeta_i} = (1 - 2\zeta_i) \log \frac{1 - \zeta_i}{\zeta_i}.$$

(i) is because post-processing will not weaken Rényi DP (Mironov, 2017). (ii) is because

$$\begin{aligned} \mu_i^0(0) &= \Pr\left(\|\pi_i^{\text{adv},0} - \pi_i^{*,0}\| > 1\right) \leq \zeta_i \\ \mu_i^1(1) &= \Pr\left(\|\pi_i^{\text{adv},1} - \pi_i^{*,0}\| \leq 1\right) \leq \Pr\left(\|\pi_i^{\text{adv},1} - \pi_i^{*,1}\| \geq 1\right) \leq \zeta_i. \end{aligned}$$

Therefore, the KL-divergence is lowerbounded by  $(1 - \zeta_i) \log \frac{1 - \zeta_i}{\zeta_i} + \zeta_i \log \frac{\zeta_i}{1 - \zeta_i}$ .

When  $\zeta_i \leq \frac{1}{4}$ , we have  $\epsilon_i \geq \frac{1}{2} \log \frac{3}{4\zeta_i}$ . Therefore,  $\zeta_i \geq \frac{3}{4e^{2\epsilon_i}}$ . Moreover, if  $\frac{1}{N} \sum_{i=1}^N \zeta_i \leq \frac{1}{16}$ , by pigeon-hole principle, there exists a set of players  $\mathcal{I} \subseteq [N] \setminus \{1, 2\}$  with  $|\mathcal{I}| \geq \frac{N}{2}$ , so that for any  $i \in \mathcal{I}$ ,  $\zeta_i \leq \frac{1}{4}$ . This further implies that

$$\frac{1}{N} \sum_{i=1}^N \zeta_i \geq \frac{1}{2|\mathcal{I}|} \sum_{i \in \mathcal{I}} \zeta_i \quad \frac{1}{N} \sum_{i=1}^N \epsilon_i \geq \frac{1}{2|\mathcal{I}|} \sum_{i \in \mathcal{I}} \epsilon_i,$$

since  $\zeta_i, \epsilon_i \geq 0$  for any  $i \in [N]$ .

Therefore, finally,

$$\frac{1}{N} \sum_{i=1}^N \zeta_i \geq \frac{1}{2|\mathcal{I}|} \sum_{i \in \mathcal{I}} \zeta_i \geq \frac{3}{8|\mathcal{I}|} \sum_{i \in \mathcal{I}} e^{-2\epsilon_i} \stackrel{(i)}{\geq} \frac{3}{8} \exp\left(-\frac{2}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \epsilon_i\right) \geq \frac{3}{8} \exp\left(-\frac{4}{N} \sum_{i=1}^N \epsilon_i\right).$$

(i) uses Jensen's inequality.  $\square$

## A.3 Omitted proof of lemmas

**Lemma A.1.** Consider the  $2^N$  games  $\{\mathcal{G}^S\}_{S \subseteq [N]}$  constructed above. There must exist two adjacent games  $\mathcal{G}^S$  and  $\mathcal{G}^{S \cup \{i\}}$  with  $i \in [N] \setminus S$ , so that

$$\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \left\langle \pi_{3i-1}^S - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^S}(\mathcal{G}^S) \right\rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \left\langle \pi_{3i}^S - \hat{\pi}_{3i}, g_{3i}^{\pi^S}(\mathcal{G}^S) \right\rangle \right] \leq 7\zeta \quad (19)$$

$$\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \left\langle \pi_{3i-1}^{S \cup \{i\}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{S \cup \{i\}}}(\mathcal{G}^{S \cup \{i\}}) \right\rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \left\langle \pi_{3i}^{S \cup \{i\}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{S \cup \{i\}}}(\mathcal{G}^{S \cup \{i\}}) \right\rangle \right] \leq 7\zeta. \quad (20)$$

*Proof.* By the accuracy condition and non-negativity of  $\max_{\hat{\pi}_j \in \Delta^{\mathcal{A}_j}} \langle \pi_j^\emptyset - \hat{\pi}_j, g_j^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle$ , we have

$$\sum_{i \in [N]} \mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^\emptyset - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^\emptyset - \hat{\pi}_{3i}, g_{3i}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle \right] \leq 3N\zeta.$$

Therefore, by the pigeon-hole principle, there must exist a subset  $\mathcal{I} \subseteq [N]$  with  $|\mathcal{I}| \geq \left\lceil \frac{N}{2} \right\rceil + 1$ , so that for any  $i \in \mathcal{I}$ ,  $\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^\emptyset - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^\emptyset - \hat{\pi}_{3i}, g_{3i}^{\pi^\emptyset}(\mathcal{G}^\emptyset) \rangle \right] \leq 7\zeta$ . The same guarantee holds for any  $\mathcal{G}^{\{i\}}$  with  $i \in [N]$ .

Then, we will form a meta graph, with each node as a subset of  $[N]$  so that there are  $2^N$  nodes in total. For each two nodes  $\mathcal{S}^1, \mathcal{S}^2$  with  $|\mathcal{S}^1| \leq |\mathcal{S}^2|$ , they are connected in the meta graph if and only if  $\mathcal{G}^{\mathcal{S}^1}$  and  $\mathcal{G}^{\mathcal{S}^2}$  are adjacent. In other words,  $\mathcal{S}^2 = \mathcal{S}^1 \cup \{i\}$  and  $i \in [N] \setminus \mathcal{S}^1$ . The edge between  $(\mathcal{S}^1, \mathcal{S}^2)$  is labeled as  $i \in [N]$  if  $\mathcal{S}^2 \setminus \mathcal{S}^1 = \{i\}$ . Therefore, each node has  $N$  edges and their labels are different from each other. Then, there are  $2^N$  nodes in the meta graph and  $N2^{N-1}$  edges.

For each node  $\mathcal{S} \subseteq [N]$ , the set  $\mathcal{I}$  constructed above can be viewed as selecting edges with labels in  $\mathcal{I}$ . Then, for each node  $\mathcal{S}$  and its selected edge with label  $i$ , it is guaranteed that

$$\mathbb{E} \left[ \max_{\hat{\pi}_{3i-1} \in \Delta^{\mathcal{A}_{3i-1}}} \langle \pi_{3i-1}^{\mathcal{S}} - \hat{\pi}_{3i-1}, g_{3i-1}^{\pi^{\mathcal{S}}}(\mathcal{G}^{\mathcal{S}}) \rangle + \max_{\hat{\pi}_{3i} \in \Delta^{\mathcal{A}_{3i}}} \langle \pi_{3i}^{\mathcal{S}} - \hat{\pi}_{3i}, g_{3i}^{\pi^{\mathcal{S}}}(\mathcal{G}^{\mathcal{S}}) \rangle \right] \leq 7\zeta. \quad (22)$$

Because  $|\mathcal{I}| \geq \left\lceil \frac{N}{2} \right\rceil + 1$  and there are  $N2^{N-1}$  edges in total, by pigeon-hole principle, there must exist an edge being selected twice. In other words, there exist two adjacent nodes  $\mathcal{S}^1, \mathcal{S}^2$  that both select edge with label  $i$ . By definition, assume  $|\mathcal{S}^1| \leq |\mathcal{S}^2|$  without loss of generality, then  $\mathcal{S}^2 = \mathcal{S}^1 \cup \{i\}$ . Therefore, (22) is satisfied for both  $\mathcal{S}^1, \mathcal{S}^2$  on  $i$ .  $\square$

## B Proof of Theorem 6.1

**Theorem 6.1.** Consider Algorithm 1. Let  $A = \max_{i \in [N]} |\mathcal{A}_i|$ . The update-rule can achieve the following guarantee in any polymatrix game. For any  $T > 0$ , player  $i \in [N]$  and strategy  $\pi_i \in \Delta^{\mathcal{A}_i}$ , we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left[ \langle g_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle \right] \\ & \leq \frac{1}{\eta T} + A \frac{\sigma^2}{2\eta} + \left( \frac{2\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} + \frac{1}{2(\bar{N})^{4/9} \log N} + \frac{2\eta\sqrt{A}}{\sigma(\bar{N})^{4/9} \log N}. \end{aligned} \quad (11)$$

*Proof.* By Lemma D.4. in Liu et al. (2024), since Algorithm 1 is equivalent to  $\pi_i^{(t+1)} = \operatorname{argmin}_{\pi_i \in \Delta^{\mathcal{A}_i}} \langle \pi_i, g_i^{(t)} \rangle + \frac{\tau_i}{2} \|\pi_i\|^2 + \frac{1}{2\eta} \|\pi_i - \bar{\pi}_i^{(t)}\|^2$ , for any  $\pi_i \in \Delta^{\mathcal{A}_i}$ , we have

$$\begin{aligned} & \frac{\eta\tau_i}{2} \|\pi_i^{(t+1)}\|^2 - \frac{\eta\tau_i}{2} \|\pi_i\|^2 + \eta \langle g_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \\ & \leq \frac{1}{2} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1+\eta\tau_i}{2} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2. \end{aligned}$$

Then, by adding  $\eta \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle$  on both sides, we have

$$\begin{aligned} & \frac{\eta\tau_i}{2} \|\pi_i^{(t+1)}\|^2 - \frac{\eta\tau_i}{2} \|\pi_i\|^2 + \eta \langle \mathbf{g}_i^{(t+1)}, \pi_i^{(t+1)} - \pi_i \rangle \\ & \leq \frac{1}{2} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1+\eta\tau_i}{2} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2 + \eta \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle. \end{aligned} \quad (23)$$

Moreover,

$$\begin{aligned} \langle \mathbf{g}_i^{(t+1)}, \pi_i^{(t+1)} - \pi_i \rangle &= \langle \mathbf{g}_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle + \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle \\ &\geq \langle \mathbf{g}_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle - \|\mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{\pi^{(t+1)}}\|_{\infty} \cdot \|\pi_i^{(t+1)} - \pi_i\|_1 \\ &\geq \langle \mathbf{g}_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle - 2 \left\| \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} (\pi_j^{(t+1)} - \bar{\pi}_j^{(t+1)}) \right\|_{\infty} \\ &\stackrel{(i)}{\geq} \langle \mathbf{g}_i^{\pi^{(t+1)}}, \pi_i^{(t+1)} - \pi_i \rangle - \frac{2}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \|\pi_j^{(t+1)} - \bar{\pi}_j^{(t+1)}\|_1. \end{aligned}$$

(i) is because  $\mathbf{U}_{i,j} \in [-1, 1]^{\mathcal{A}_i \times \mathcal{A}_j}$ . Recall that  $A := \max_{i \in [N]} |\mathcal{A}_i|$ , we have

$$\begin{aligned} \|\pi_j^{(t+1)} - \bar{\pi}_j^{(t+1)}\|_1 &\leq \sqrt{A} \|\pi_j^{(t+1)} - \bar{\pi}_j^{(t+1)}\| = \sqrt{A} \|\pi_j^{(t+1)} - \text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)})\| \\ &\leq \sqrt{A} \|\mathbf{n}_j^{(t+1)}\|. \end{aligned}$$

By taking the expectation on both sides of (23), we have,

$$\begin{aligned} & \mathbb{E} \left[ \left( \frac{\eta\tau_i}{2} \|\pi_i^{(t+1)}\|^2 - \frac{\eta\tau_i}{2} \|\pi_i\|^2 + \eta \left\langle -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j^{(t+1)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right) \right] \\ & \leq \mathbb{E} \left[ \frac{1}{2} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1+\eta\tau_i}{2} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2 \right] \\ & \quad + \eta \mathbb{E} \left[ \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle + \frac{2\sqrt{A}}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \|\mathbf{n}_j^{(t+1)}\| \right]. \end{aligned} \quad (24)$$

By Jensen's Inequality,  $\mathbb{E} [\|\mathbf{n}_j^{(t+1)}\|] \leq \sqrt{\mathbb{E} [\|\mathbf{n}_j^{(t+1)}\|^2]} = \sqrt{|\mathcal{A}_j|} \sigma$ . Therefore,

$$\frac{2\sqrt{A}}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbb{E} [\|\mathbf{n}_j^{(t+1)}\|] \leq 2A\sigma.$$

Moreover, since  $\|\pi_i\|^2 \in [\frac{1}{|\mathcal{A}_i|}, 1]$ , we have

$$\begin{aligned} & \mathbb{E} \left[ \eta \left\langle -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j^{(t+1)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \\ & \leq \mathbb{E} \left[ \frac{1}{2} \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \frac{1+\eta\tau_i}{2} \|\pi_i - \pi_i^{(t+1)}\|^2 - \frac{1}{2} \|\pi_i^{(t+1)} - \bar{\pi}_i^{(t)}\|^2 + \eta \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \right] \\ & \quad + 2A\eta\sigma + \frac{\eta}{2}\tau_i. \end{aligned}$$

To bound  $\mathbb{E} \left[ \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \right]$ , we have the following lemma.

**Lemma B.1.** Consider Algorithm 1. For any player  $i \in [N]$ , we have

$$\mathbb{E} \left[ \langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \rangle \right] \leq \left( \frac{2\eta^2}{\sigma} + \frac{3\sigma}{2} \right) A^{\frac{3}{2}} + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j. \quad (25)$$

The proof is postponed to the end of this section.

We can further bound  $\mathbb{E} \left[ \|\pi_i - \bar{\pi}_i^{(t)}\|^2 - \|\pi_i - \pi_i^{(t)}\|^2 \right]$  by

$$\begin{aligned} \mathbb{E} \left[ \left\| \pi_i - \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \pi_i^{(t)} + \mathbf{n}_i^{(t)} \right) \right\|^2 - \left\| \pi_i - \pi_i^{(t)} \right\|^2 \right] &\leq \mathbb{E} \left[ \left\| \pi_i - \pi_i^{(t)} - \mathbf{n}_i^{(t)} \right\|^2 - \left\| \pi_i - \pi_i^{(t)} \right\|^2 \right] \\ &= \mathbb{E} \left[ 2 \langle \mathbf{n}_i^{(t)}, \pi_i^{(t)} - \pi_i \rangle + \|\mathbf{n}_i^{(t)}\|^2 \right] \\ &\stackrel{(i)}{=} \mathbb{E} \left[ \|\mathbf{n}_i^{(t)}\|^2 \right] = A\sigma^2. \end{aligned}$$

(i) is because  $\mathbb{E} [\mathbf{n}_i^{(t)}] = \mathbf{0}$  and  $\mathbf{n}_i^{(t)}$  is independent of  $\pi_i - \pi_i^{(t)}$ .

Finally, by telescoping,

$$\begin{aligned} &\sum_{t=1}^T \mathbb{E} \left[ \left\langle -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j^{(t+1)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \\ &\leq \frac{1}{2\eta} \left\| \pi_1 - \bar{\pi}_1^{(1)} \right\|^2 + A \frac{\sigma^2}{2\eta} T + \left( \frac{2\eta^2}{\sigma} + \frac{3\sigma}{2} \right) A^{\frac{3}{2}} T + 2A\sigma T + \frac{1}{2} \tau_i T + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j T \\ &\leq \frac{1}{\eta} + A \frac{\sigma^2}{2\eta} T + \left( \frac{2\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} T + \frac{1}{2} \tau_i T + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j T. \end{aligned} \quad \square$$

Further, by taking summation over all player  $i \in [N]$  and averaging, we have

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^T \mathbb{E} \left[ \left\langle -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j^{(t+1)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \\ &\leq \frac{1}{\eta} + A \frac{\sigma^2}{2\eta} T + \left( \frac{2\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} T + \frac{T}{2N} \sum_{i=1}^N \tau_i + \frac{1}{N} \sum_{i=1}^N \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j T. \end{aligned}$$

Then, we can further bound the summation over  $\tau_i$  by Lemma B.2.

**Lemma B.2.** When  $\tau_i = \frac{c}{|\mathcal{N}(i)|}$  for any  $i \in [N]$ , where  $c > 0$  is a game-dependent constant, we have

$$\frac{1}{N} \sum_{i=1}^N \tau_i = \frac{c}{\bar{N}} \quad \frac{1}{N} \sum_{i=1}^N \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \tau_j \leq \frac{c}{\bar{N}}. \quad (26)$$

The proof is postponed to the end of this section. Therefore, since  $c = \frac{(\bar{\mathcal{N}})^{5/9}}{\log N}$  in Algorithm 1, we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left[ \left\langle -\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \pi_j^{(t+1)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \\ & \leq \frac{1}{\eta T} + A \frac{\sigma^2}{2\eta} + \left( \frac{2\eta^2}{\sigma} + \frac{7\sigma}{2} \right) A^{\frac{3}{2}} + \frac{1}{2(\bar{\mathcal{N}})^{4/9} \log N} + \frac{2\eta\sqrt{A}}{\sigma (\bar{\mathcal{N}})^{4/9} \log N}. \end{aligned} \quad \square$$

## B.1 Omitted Proof of Appendix B

**Lemma B.1.** Consider Algorithm 1. For any player  $i \in [N]$ , we have

$$\mathbb{E} \left[ \left\langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \leq \left( \frac{2\eta^2}{\sigma} + \frac{3\sigma}{2} \right) A^{\frac{3}{2}} + \frac{2\eta\sqrt{A}}{|\mathcal{N}(i)|\sigma} \sum_{j \in \mathcal{N}(i)} \tau_j. \quad (25)$$

*Proof.* By Hölder's Inequality, we have

$$\mathbb{E} \left[ \left\langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \leq \mathbb{E} \left[ \left\| \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)} \right\| \cdot \left\| \pi_i^{(t+1)} - \pi_i \right\| \right].$$

Moreover,

$$\begin{aligned} \left\| \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)} \right\| &= \frac{1}{|\mathcal{N}(i)|} \left\| \sum_{j \in \mathcal{N}(i)} \mathbf{U}_{i,j} \left( \text{Proj}_{\Delta^{\mathcal{A}_j}} \left( \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} \right) - \bar{\pi}_j^{(t)} \right) \right\| \\ &\leq \frac{\sqrt{|\mathcal{A}_i|}}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \left\| \text{Proj}_{\Delta^{\mathcal{A}_j}} \left( \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} \right) - \bar{\pi}_j^{(t)} \right\|_1 \\ &\leq \frac{\sqrt{|\mathcal{A}_i|}}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \sqrt{|\mathcal{A}_j|} \left\| \text{Proj}_{\Delta^{\mathcal{A}_j}} \left( \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} \right) - \bar{\pi}_j^{(t)} \right\| \\ &\leq \frac{\sqrt{|\mathcal{A}_i|}}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \sqrt{|\mathcal{A}_j|} \left\| \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} - \bar{\pi}_j^{(t)} \right\|. \end{aligned}$$

Therefore,

$$\begin{aligned} & \mathbb{E} \left[ \left\langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \\ & \leq \frac{A}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbb{E} \left[ \left\| \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} - \bar{\pi}_j^{(t)} \right\| \cdot \left\| \pi_i^{(t+1)} - \pi_i \right\| \right] \\ & \leq \frac{A}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \left( \frac{1}{\sigma\sqrt{A}} \mathbb{E} \left[ \left\| \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} - \bar{\pi}_j^{(t)} \right\|^2 \right] + \frac{\sigma\sqrt{A}}{4} \mathbb{E} \left[ \left\| \pi_i^{(t+1)} - \pi_i \right\|^2 \right] \right). \end{aligned}$$

Furthermore, since  $\mathbb{E} [\mathbf{n}_j^{(t+1)}] = 0$  and  $\mathbf{n}_j^{(t+1)}$  is independent of  $\pi_j^{(t+1)}, \bar{\pi}_j^{(t)}$ , we have

$$\begin{aligned}
\mathbb{E} \left[ \left\| \pi_j^{(t+1)} + \mathbf{n}_j^{(t+1)} - \bar{\pi}_j^{(t)} \right\|^2 \right] &= \mathbb{E} \left[ \left\| \pi_j^{(t+1)} - \bar{\pi}_j^{(t)} \right\|^2 \right] + \mathbb{E} \left[ \left\| \mathbf{n}_j^{(t+1)} \right\|^2 \right] \\
&\leq \mathbb{E} \left[ \left\| \text{Proj}_{\Delta^{\mathcal{A}_j}} \left( \frac{\bar{\pi}_j^{(t)} - \eta \mathbf{g}_j^{(t)}}{1 + \eta \tau_j} \right) - \bar{\pi}_j^{(t)} \right\|^2 \right] + A\sigma^2 \\
&\leq \mathbb{E} \left[ \left\| \frac{\bar{\pi}_j^{(t)} - \eta \mathbf{g}_j^{(t)}}{1 + \eta \tau_j} - \bar{\pi}_j^{(t)} \right\|^2 \right] + A\sigma^2 \\
&= \mathbb{E} \left[ \left\| \frac{\eta \tau_j}{1 + \eta \tau_j} \bar{\pi}_j^{(t)} - \frac{\eta}{1 + \eta \tau_j} \mathbf{g}_j^{(t)} \right\|^2 \right] + A\sigma^2 \\
&\stackrel{(i)}{\leq} 2\eta \tau_j \mathbb{E} \left[ \left\| \bar{\pi}_j^{(t)} \right\|^2 \right] + 2\eta^2 \mathbb{E} \left[ \left\| \mathbf{g}_j^{(t)} \right\|^2 \right] + A\sigma^2 \\
&\stackrel{(ii)}{\leq} 2\eta \tau_j + 2\eta^2 A + A\sigma^2.
\end{aligned}$$

(i) is because  $(a+b)^2 \leq 2a^2 + 2b^2$  for any  $a, b \in \mathbb{R}$  and  $1 + \eta \tau_j \geq \max \{1, \eta \tau_j\}$ . (ii) is because  $\mathbf{U}_{i,j} \in [-1, 1]^{\mathcal{A}_i \times \mathcal{A}_j}$ . Also,  $\frac{\sigma \sqrt{A}}{4} \mathbb{E} \left[ \left\| \pi_i^{(t+1)} - \pi_i \right\|^2 \right] \leq \frac{\sigma \sqrt{A}}{2}$ . Therefore,

$$\mathbb{E} \left[ \left\langle \mathbf{g}_i^{(t+1)} - \mathbf{g}_i^{(t)}, \pi_i^{(t+1)} - \pi_i \right\rangle \right] \leq \left( \frac{2\eta^2}{\sigma} + \frac{3\sigma}{2} \right) A^{\frac{3}{2}} + \frac{2\eta \sqrt{A}}{|\mathcal{N}(i)| \sigma} \sum_{j \in \mathcal{N}(i)} \tau_j. \quad \square$$

**Lemma B.2.** When  $\tau_i = \frac{c}{|\mathcal{N}(i)|}$  for any  $i \in [N]$ , where  $c > 0$  is a game-dependent constant, we have

$$\frac{1}{N} \sum_{i=1}^N \tau_i = \frac{c}{\bar{N}} \quad \frac{1}{N} \sum_{i=1}^N \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \tau_j \leq \frac{c}{\bar{N}}. \quad (26)$$

*Proof.* By definition,

$$\frac{1}{N} \sum_{i=1}^N \tau_i = \frac{c}{N} \sum_{i=1}^N \frac{1}{|\mathcal{N}(i)|} = \frac{c}{\bar{N}}.$$

Moreover,

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^N \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \tau_j &= \frac{c}{N} \sum_{i=1}^N \sum_{j \in \mathcal{N}(i)} \frac{1}{|\mathcal{N}(i)| \cdot |\mathcal{N}(j)|} \\
&\leq \frac{c}{N} \sum_{i=1}^N \sum_{j \in \mathcal{N}(i)} \left( \frac{1}{2|\mathcal{N}(i)|^2} + \frac{1}{2|\mathcal{N}(j)|^2} \right) \\
&\stackrel{(i)}{=} \frac{c}{N} \sum_{i=1}^N \frac{2|\mathcal{N}(i)|}{2|\mathcal{N}(i)|^2} \\
&= \frac{c}{\bar{N}}.
\end{aligned}$$

(i) is because each player contributes  $\frac{1}{2|\mathcal{N}(i)|^2}$  to the summation 2 times on each edge, i.e.  $2|\mathcal{N}(i)|$  times in total.  $\square$

## C Proof of Section 7

**Theorem 7.1.** Consider Algorithm 1 and any two adjacent polymatrix games  $\mathcal{G}, \mathcal{G}'$  that differs at  $(v_1, v_2)$ . Let  $dist(i, j)$  be the length of the shortest path between players  $i, j$ , which is  $\infty$  when  $i, j$  are not connected. The update-rule guarantees the following for any  $T > 0$  and player  $i \in [N]$ ,

$$\frac{1}{N} \sum_{i=1}^N D_\alpha \left( \mu_{\mathcal{G},i}^{(T)}, \mu_{\mathcal{G}',i}^{(T)} \right) \leq \frac{\alpha \eta^2}{\sigma^2} \min \{ \clubsuit, \spadesuit \} T, \quad (12)$$

where

$$\clubsuit := \frac{16A^3 (\log N)^2}{(\bar{N})^{4/9}} + \frac{4A}{N} \quad (13)$$

$$\spadesuit := \frac{2A}{N} \sum_{i=1}^N \mathbb{1}(T > \min \{ dist(i, v_1), dist(i, v_2) \}). \quad (14)$$

*Proof.* In the following, we will introduce the proof for dense graphs ( $\clubsuit$ ) first and that for the sparse graphs ( $\spadesuit$ ) later.

### C.1 Proof for Dense Graphs ( $\clubsuit$ )

Let  $\bar{N} := \left( \frac{1}{N} \sum_{i \in [N]} \frac{1}{|\mathcal{N}(i)|} \right)^{-1}$  be the harmonic mean of players' degrees. When the graph is dense,  $\bar{N}$  will be relatively large. Now, we will bound the Rényi DP. Let  $(v_1, v_2)$  be the edge on which the utility matrix differs in  $\mathcal{G}$  and  $\mathcal{G}'$ . Then, we define the set  $\mathcal{S}$  as

$$\mathcal{S} := \begin{cases} [N] & |\mathcal{N}(v_1)|, |\mathcal{N}(v_2)| \geq (\bar{N})^{2/9} \\ [N] \setminus \{v_1\} & |\mathcal{N}(v_1)| < (\bar{N})^{2/9}, |\mathcal{N}(v_2)| \geq (\bar{N})^{2/9} \\ [N] \setminus \{v_2\} & |\mathcal{N}(v_2)| < (\bar{N})^{2/9}, |\mathcal{N}(v_1)| \geq (\bar{N})^{2/9} \\ [N] \setminus \{v_1, v_2\} & |\mathcal{N}(v_1)|, |\mathcal{N}(v_2)| < (\bar{N})^{2/9}. \end{cases}$$

Consider the marginal distribution  $\mu_{\mathcal{G},\mathcal{S}}^{(t)}$  of  $\mu_{\mathcal{G}}^{(t)}$  over  $\left( \left\{ \pi_i^{(s)} + n_i^{(s)} \right\}_{i \in \mathcal{S}} \right)_{s=1}^t$ . Then,  $D_\alpha \left( \mu_{\mathcal{G},\mathcal{S}}^{(t)}, \mu_{\mathcal{G}',\mathcal{S}}^{(t)} \right) \geq D_\alpha \left( \mu_{\mathcal{G},i}^{(t)}, \mu_{\mathcal{G}',i}^{(t)} \right)$  for any  $i \in [N] \setminus \{v_1, v_2\}$  and  $t > 0$  since  $i \in \mathcal{S}$  and enlarging the set of observations will not decrease the Rényi divergence. This is because post-processing (delete observations) will not increase Rényi divergence (Mironov, 2017).

#### C.1.1 $v_1, v_2 \in \mathcal{S}$

By Lemma 7.2, let  $\pi'_i^{(t)}$  be the counterpart of  $\pi_i^{(t)}$  for any  $i \in [N]$ , we have

$$\begin{aligned} D_\alpha \left( \mu_{\mathcal{G},\mathcal{S}}^{(T)}, \mu_{\mathcal{G}',\mathcal{S}}^{(T)} \right) &\leq \sum_{t=1}^T \sup_{(\hat{o}^{(1)}, \hat{o}^{(2)}, \dots, \hat{o}^{(t-1)})} D_\alpha \left( \mu_{\mathcal{G},\mathcal{S}}(\cdot | \hat{o}^{(1)}, \hat{o}^{(2)}, \dots, \hat{o}^{(t-1)}), \mu_{\mathcal{G}',\mathcal{S}}(\cdot | \hat{o}^{(1)}, \hat{o}^{(2)}, \dots, \hat{o}^{(t-1)}) \right) \\ &\stackrel{(i)}{=} \frac{\alpha}{2\sigma^2} \sum_{t=1}^T \sum_{i \in [N]} \sup_{(\hat{o}^{(1)}, \hat{o}^{(2)}, \dots, \hat{o}^{(t-1)})} \left\| \pi_i^{(t)} - \pi'_i^{(t)} \right\|^2. \end{aligned}$$

(i) is by the Rényi divergence of multi-variate Gaussian distribution (Gil et al., 2013), since the distribution of  $\pi_i^{(t)} + \mathbf{n}_i^{(t)}$  is a multivariate gaussian with mean  $\pi_i^{(t)}$  and variance  $\sigma^2 I_{\mathcal{A}_i}$  ( $I_{\mathcal{A}_i}$  is the identity matrix indexed by  $\mathcal{A}_i \times \mathcal{A}_i$ ). The inequality above implies that the Rényi divergence can be bounded by the squared 2-norm of  $\pi_i^{(t)} - \pi'_i^{(t)}$ , given all past observations  $(\hat{o}^{(1)}, \hat{o}^{(2)}, \dots, \hat{o}^{(t-1)})$  are identical in  $\mathcal{G}, \mathcal{G}'$ .

Given the observations are identical in  $\mathcal{G}, \mathcal{G}'$ , the gradient  $\mathbf{g}_i^{(t-1)}$  of all players except  $v_1, v_2$  are identical in  $\mathcal{G}, \mathcal{G}'$  since the gradient only depends on  $\pi_i^{(t-1)} + \mathbf{n}_i^{(t-1)}$  and the utility matrices, which are identical for  $i \in [N] \setminus \{v_1, v_2\}$ . Therefore,  $\pi_i^{(t)} = \pi'_i^{(t)}$  for any  $i \in [N] \setminus \{v_1, v_2\}$ . For  $i \in \{v_1, v_2\}$ , we have

$$\begin{aligned} \|\pi_i^{(t)} - \pi'_i^{(t)}\| &= \left\| \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \frac{\bar{\pi}_i^{(t-1)} - \eta \mathbf{g}_i^{(t-1)}}{1 + \eta \tau_i} \right) - \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \frac{\bar{\pi}_i^{(t-1)} - \eta \mathbf{g}'_i^{(t-1)}}{1 + \eta \tau_i} \right) \right\| \\ &\leq \left\| \frac{\bar{\pi}_i^{(t-1)} - \eta \mathbf{g}_i^{(t-1)}}{1 + \eta \tau_i} - \frac{\bar{\pi}_i^{(t-1)} - \eta \mathbf{g}'_i^{(t-1)}}{1 + \eta \tau_i} \right\| \\ &\leq \eta \left\| \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} (\mathbf{U}_{i,j} - \mathbf{U}'_{i,j}) \bar{\pi}_j^{(t-1)} \right\| \\ &\stackrel{(i)}{\leq} \frac{2\eta}{|\mathcal{N}(i)|} \sqrt{|\mathcal{A}_i|}. \end{aligned}$$

(i) is because  $\mathbf{U}_{i,j} \in [-1, 1]^{\mathcal{A}_i \times \mathcal{A}_j}$  and  $\mathbf{U}_{i,j} \neq \mathbf{U}'_{i,j}$  only if  $(i, j) = (v_1, v_2)$  or  $(i, j) = (v_2, v_1)$ . Therefore, since the distribution of  $\pi_i^{(t)} + \mathbf{n}_i^{(t)}$  is actually a multivariate gaussian with mean  $\pi_i^{(t)}$  and variance  $\sigma^2 I_{\mathcal{A}_i}$ , so that the Rényi divergence is bounded by  $\left( \frac{2\alpha\eta^2}{\sigma^2 |\mathcal{N}(v_1)|^2} |\mathcal{A}_{v_1}| + \frac{2\alpha\eta^2}{\sigma^2 |\mathcal{N}(v_2)|^2} |\mathcal{A}_{v_2}| \right) T$ . By the definition of  $\mathcal{S}$ , it is further bounded by  $\frac{4\alpha\eta^2}{\sigma^2 (\mathcal{N})^{4/9}} AT$ .

### C.1.2 $v_1, v_2 \notin \mathcal{S}$

When the adversary may only observe  $(\pi_k^{(s)} + \mathbf{n}_k^{(s)})_{s=1}^t$  for some  $k \in \{v_1, v_2\}$ , by Lemma 7.2, we only need to bound  $\|\pi_k^{(t)} - \pi'_k^{(t)}\|^2$  when all past observations  $(\pi_k^{(s)} + \mathbf{n}_k^{(s)})_{s=1}^{t-1}$  are the same in  $\mathcal{G}, \mathcal{G}'$ . Then,

$$\begin{aligned} \|\pi_k^{(t)} - \pi'_k^{(t)}\| &= \left\| \text{Proj}_{\Delta^{\mathcal{A}_k}} \left( \frac{\bar{\pi}_k^{(t-1)} - \eta \mathbf{g}_k^{(t-1)}}{1 + \eta \tau_k} \right) - \text{Proj}_{\Delta^{\mathcal{A}_k}} \left( \frac{\bar{\pi}_k^{(t-1)} - \eta \mathbf{g}'_k^{(t-1)}}{1 + \eta \tau_k} \right) \right\| \\ &\leq \left\| \frac{\bar{\pi}_k^{(t-1)} - \eta \mathbf{g}_k^{(t-1)}}{1 + \eta \tau_k} - \frac{\bar{\pi}_k^{(t-1)} - \eta \mathbf{g}'_k^{(t-1)}}{1 + \eta \tau_k} \right\| \\ &\leq \eta \|\mathbf{g}_k^{(t-1)} - \mathbf{g}'_k^{(t-1)}\| \\ &\leq 2\eta \sqrt{|\mathcal{A}_k|}. \end{aligned}$$

For  $k \in [N] \setminus \{v_1, v_2\}$ , similar to the proof in the previous section, we will augment the observations of the adversary first. The adversary may observe  $\left( \{\pi_i^{(s)} + \mathbf{n}_i^{(s)}\}_{i \in [N] \setminus \{v_1, v_2\}} \right)_{s=1}^t$  and  $\left( \{\mathbf{n}_{v_1}^{(s)}, \mathbf{n}_{v_2}^{(s)}\} \right)_{s=1}^t$  at timestep  $t$ . Note that the actual observation of the adversary is still  $(\pi_k^{(s)} + \mathbf{n}_k^{(s)})_{s=1}^t$ . We augment

the observations to simplify the proof, since the Rényi divergence of the augmented observation's distribution upperbounds that of the actual observation.

Still, by the chain rule of Rényi divergence, we need to bound  $\sum_{i \in [N] \setminus \{v_1, v_2\}} \|\pi_i^{(t)} - \pi'^{(t)}_i\|$  for any  $t \in [T]$ . For any  $i \notin \mathcal{N}(v_1) \cup \mathcal{N}(v_2)$ , similar to the discussion in previous section, the gradients  $\mathbf{g}_i^{(t)}$  and  $\bar{\pi}_i^{(t-1)}$  are identical in  $\mathcal{G}, \mathcal{G}'$ , so that the  $\pi_i^{(t)} = \pi'^{(t)}_i$ .

For  $i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2) \setminus \{v_1, v_2\}$ , we have

$$\begin{aligned} \|\pi_i^{(t)} - \pi'^{(t)}_i\| &\leq \eta \|\mathbf{g}_i^{(t-1)} - \mathbf{g}'_i^{(t-1)}\| = \frac{\eta}{|\mathcal{N}(i)|} \left\| \sum_{j \in \mathcal{N}(i)} \mathbf{u}_{i,j} (\bar{\pi}_j^{(t-1)} - \bar{\pi}'_j^{(t-1)}) \right\| \\ &\stackrel{(i)}{=} \frac{\eta}{|\mathcal{N}(i)|} \left\| \sum_{j \in \{v_1, v_2\} \cap \mathcal{N}(i)} \mathbf{u}_{i,j} (\bar{\pi}_j^{(t-1)} - \bar{\pi}'_j^{(t-1)}) \right\| \\ &\leq \frac{\eta \sqrt{|\mathcal{A}_i|}}{|\mathcal{N}(i)|} \sum_{j \in \{v_1, v_2\} \cap \mathcal{N}(i)} \|\bar{\pi}_j^{(t-1)} - \bar{\pi}'_j^{(t-1)}\|_1 \\ &\leq \frac{\eta A}{|\mathcal{N}(i)|} \sum_{j \in \{v_1, v_2\} \cap \mathcal{N}(i)} \|\bar{\pi}_j^{(t-1)} - \bar{\pi}'_j^{(t-1)}\|. \end{aligned}$$

(i) uses the fact that  $\bar{\pi}_i^{(t-1)} = \bar{\pi}'_i^{(t-1)}$  for  $i \notin \{v_1, v_2\}$ . Since,  $\mathbf{n}_j^{(t-1)} = \mathbf{n}'_j^{(t-1)}$ , for any  $j \in \{v_1, v_2\}$ , we have

$$\begin{aligned} \|\bar{\pi}_j^{(t-1)} - \bar{\pi}'_j^{(t-1)}\| &= \left\| \text{Proj}_{\Delta^{\mathcal{A}_j}} (\pi_j^{(t-1)} + \mathbf{n}_j^{(t-1)}) - \text{Proj}_{\Delta^{\mathcal{A}_j}} (\pi'_j^{(t-1)} + \mathbf{n}'_j^{(t-1)}) \right\| \\ &\leq \|\pi_j^{(t-1)} + \mathbf{n}_j^{(t-1)} - \pi'_j^{(t-1)} - \mathbf{n}'_j^{(t-1)}\| \\ &= \|\pi_j^{(t-1)} - \pi'_j^{(t-1)}\|. \end{aligned}$$

Moreover, due to the additional regularization, for  $j \in \{v_1, v_2\}$ ,  $\|\pi_j^{(t-1)} - \pi'_j^{(t-1)}\|$  will be bounded as follows according to Lemma 7.3.

$$\|\pi_j^{(t-1)} - \pi'_j^{(t-1)}\| \leq \frac{2\sqrt{\mathcal{A}_j}}{\tau_j} \stackrel{(i)}{\leq} 2\sqrt{\mathcal{A}_j} \frac{(\bar{\mathcal{N}})^{2/9} \log N}{(\bar{\mathcal{N}})^{5/9}} = \frac{2\sqrt{\mathcal{A}_j} \log N}{(\bar{\mathcal{N}})^{1/3}}.$$

(i) is because when  $v_1 \notin \mathcal{S}$ ,  $|\mathcal{N}(v_1)| < (\bar{\mathcal{N}})^{2/9}$  so that  $\tau_{v_1} \geq \frac{(\bar{\mathcal{N}})^{5/9}}{(\bar{\mathcal{N}})^{2/9} \log N}$  by definition. The same argument also holds for  $v_2$ . Therefore,  $\|\pi_i^{(t)} - \pi'^{(t)}_i\| \leq \frac{4\eta A^{3/2} \log N}{(\bar{\mathcal{N}})^{1/3} |\mathcal{N}(i)|}$ , where  $A = \max_{i \in [N]} |\mathcal{A}_i|$ . Furthermore, for  $i \in [N] \setminus \{v_1, v_2\}$

$$\begin{aligned} D_\alpha \left( \mu_{\mathcal{G}, i}^{(T)}, \mu_{\mathcal{G}', i}^{(T)} \right) &\leq D_\alpha \left( \mu_{\mathcal{G}, \mathcal{S}}^{(T)}, \mu_{\mathcal{G}', \mathcal{S}}^{(T)} \right) \leq \frac{\alpha}{2\sigma^2} |\mathcal{N}(v_1) \cup \mathcal{N}(v_2)| \cdot \left( \frac{16\eta^2 A^3 (\log N)^2}{(\bar{\mathcal{N}})^{2/3} \min_{i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)} |\mathcal{N}(i)|^2} \right) T \\ &\leq \frac{\alpha}{2\sigma^2} \left( 2(\bar{\mathcal{N}})^{2/9} \right) \cdot \left( \frac{16\eta^2 A^3 (\log N)^2}{(\bar{\mathcal{N}})^{2/3} \min_{i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)} |\mathcal{N}(i)|^2} \right) T \\ &= \frac{16\alpha\eta^2 A^3 (\log N)^2}{\sigma^2 (\bar{\mathcal{N}})^{4/9} \min_{i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)} |\mathcal{N}(i)|^2} T. \end{aligned}$$

Finally, since  $\min_{i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)} |\mathcal{N}(i)|^2 \geq 1$ , we can conclude the proof. Moreover,  $\min_{i \in \mathcal{N}(v_1) \cup \mathcal{N}(v_2)} |\mathcal{N}(i)|^2$  can be much larger than a constant in practice, e.g., Erdős–Rényi graphs.

The proofs for the rest possibilities ( $v_1 \in \mathcal{S}, v_2 \notin \mathcal{S}$  and  $v_1 \notin \mathcal{S}, v_2 \in \mathcal{S}$ ) are similar.  $\square$

## C.2 Proof for Sparse Graphs (♠)

In this section, we show the DP guarantee when the associated graph of the polymatrix game is sparse. Formally, there exists a constant  $\mathcal{N}^{\max} > 0$  so that  $|\mathcal{N}(i)| \leq \mathcal{N}^{\max}$  for any  $i \in [N]$ .

We will augment the observations from  $\left(\pi_i^{(s)} + \mathbf{n}_i^{(s)}\right)_{s=1}^{t-1}$  to  $\left(\left\{\pi_j^{(s)} + \mathbf{n}_j^{(s)}\right\}_{j \in \mathcal{S}_i^{(s)}}\right)_{s=1}^t$ , where

$$\mathcal{S}_i^{(s)} := \begin{cases} \{j: \min \{\text{dist}(j, v_1), \text{dist}(j, v_2)\} \geq s\} & \min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\} \geq s \\ \{i\} & \text{Otherwise.} \end{cases}$$

When  $s \leq \min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\}$ , we will show in the following that with all past observations identical in  $\mathcal{G}, \mathcal{G}'$ , then  $\pi_j^{(s)}$  and its counterpart  $\pi'^{(s)}_j$  in  $\mathcal{G}'$  are identical for any  $j \in \mathcal{S}_i^{(s)}$ . When  $s = 0$ ,  $\mathcal{S}_i^{(0)} = [N]$  and  $\pi_j^{(0)} = \pi'^{(0)}_j$  for any  $j \in [N]$  since they are both initialized to uniform distribution over  $\mathcal{A}_j$ . When  $s > 0$ , for any  $j \in \mathcal{S}_i^{(s)}$ , we must have  $\mathcal{N}(j) \subseteq \mathcal{S}_i^{(s-1)}$ , otherwise  $\min \{\text{dist}(j, v_1), \text{dist}(j, v_2)\} < s$ . Therefore, since all past observations are identical, we have  $\bar{\pi}_k^{(s-1)}$  are identical in  $\mathcal{G}, \mathcal{G}'$  for any  $k \in \mathcal{N}(j)$  so that the gradient  $\mathbf{g}_j^{(s-1)} = \mathbf{g}'_j^{(s-1)}$ . Moreover, given  $\bar{\pi}_j^{(s-1)}$  are identical,  $\pi_j^{(s)}$  are identical in  $\mathcal{G}, \mathcal{G}'$ .

When  $s > \min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\}$ ,  $\|\pi_i^{(s)} - \pi'^{(s)}_i\| \leq 2\eta\sqrt{|\mathcal{A}_i|}$ , given the past observation  $\pi_i^{(s-1)} + \mathbf{n}_i^{(s-1)}$  is identical.

By Lemma 7.2, for any  $T > 0$ ,

$$\begin{aligned} D_\alpha \left( \mu_{\mathcal{G}, i}^{(T)}, \mu_{\mathcal{G}', i}^{(T)} \right) &\leq \max \{0, T - \min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\}\} \frac{2\alpha\eta^2}{\sigma^2} |\mathcal{A}_i| \\ &\leq \mathbb{1}(T > \min \{\text{dist}(i, v_1), \text{dist}(i, v_2)\}) \frac{2\alpha\eta^2}{\sigma^2} |\mathcal{A}_i| T. \end{aligned} \quad \square$$

## C.3 Omitted Proof of Lemmas

**Lemma 7.3.** Consider the Algorithm 1. For any player  $i \in [N]$  and timestep  $t > 0$ , by updating the strategy  $\left\{\pi_i^{(s)}\right\}_{s=0}^t, \left\{\pi'^{(s)}_i\right\}_{s=0}^t$  with the same noise  $\left\{\mathbf{n}_i^{(s)}\right\}_{s=0}^t$  but different gradients  $\left\{\mathbf{g}_i^{(s)}\right\}_{s=0}^t, \left\{\mathbf{g}'_i^{(s)}\right\}_{s=0}^t$  individually, we have

$$\left\|\pi_i^{(t)} - \pi'^{(t)}_i\right\| \leq \frac{2\sqrt{|\mathcal{A}_i|}}{\tau_i}. \quad (15)$$

*Proof.*

$$\begin{aligned}
\|\pi_i^{(t+1)} - \pi'^{(t+1)}_i\| &= \left\| \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \frac{\text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi_i^{(t)} + \mathbf{n}_i^{(t)}) - \eta \mathbf{g}_i^{(t)}}{1 + \eta \tau_i} \right) - \text{Proj}_{\Delta^{\mathcal{A}_i}} \left( \frac{\text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi'^{(t)}_i + \mathbf{n}_i^{(t)}) - \eta \mathbf{g}'_i^{(t)}}{1 + \eta \tau_i} \right) \right\| \\
&\leq \left\| \frac{\text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi_i^{(t)} + \mathbf{n}_i^{(t)}) - \eta \mathbf{g}_i^{(t)}}{1 + \eta \tau_i} - \frac{\text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi'^{(t)}_i + \mathbf{n}_i^{(t)}) - \eta \mathbf{g}'_i^{(t)}}{1 + \eta \tau_i} \right\| \\
&\leq \frac{1}{1 + \eta \tau_i} \left\| \text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi_i^{(t)} + \mathbf{n}_i^{(t)}) - \text{Proj}_{\Delta^{\mathcal{A}_i}} (\pi'^{(t)}_i + \mathbf{n}_i^{(t)}) \right\| + \frac{\eta}{1 + \eta \tau_i} \|\mathbf{g}_i^{(t)} - \mathbf{g}'_i^{(t)}\| \\
&\leq \frac{1}{1 + \eta \tau_i} \|\pi_i^{(t)} - \pi'^{(t)}_i\| + \frac{\eta}{1 + \eta \tau_i} \|\mathbf{g}_i^{(t)} - \mathbf{g}'_i^{(t)}\|.
\end{aligned}$$

By recursively applying the inequality above, we have

$$\begin{aligned}
\|\pi_i^{(t+1)} - \pi'^{(t+1)}_i\| &\leq \left( \frac{1}{1 + \eta \tau_i} \right)^{t+1} \|\pi_i^{(0)} - \pi'^{(0)}_i\| + \eta \sum_{s=0}^t \left( \frac{1}{1 + \eta \tau_i} \right)^{t-s+1} \|\mathbf{g}_i^{(s)} - \mathbf{g}'_i^{(s)}\| \\
&\stackrel{(i)}{=} \eta \sum_{s=0}^t \left( \frac{1}{1 + \eta \tau_i} \right)^{t-s+1} \|\mathbf{g}_i^{(s)} - \mathbf{g}'_i^{(s)}\|.
\end{aligned}$$

(i) is because  $\pi_i^{(0)}, \pi'^{(0)}_i$  are both initialized as uniform distribution over  $\mathcal{A}_i$ . Therefore, since each element of the gradient is bounded by  $[-1, 1]$  by definition, we have

$$\|\pi_i^{(t+1)} - \pi'^{(t+1)}_i\| \leq 2\eta \sum_{s=0}^t \left( \frac{1}{1 + \eta \tau_i} \right)^{t-s+1} \sqrt{\mathcal{A}_i} \leq 2\eta \sqrt{\mathcal{A}_i} \frac{1}{(1 + \eta \tau_i) \left( 1 - \frac{1}{1 + \eta \tau_i} \right)} = \frac{2\sqrt{\mathcal{A}_i}}{\tau_i}. \quad \square$$

**Lemma 7.2** (Chain Rule of Rényi divergence). *For any distribution  $p, q$  over random variables  $X^1, X^2$ , for any  $\alpha \geq 1$ , we have*

$$D_\alpha \left( p(X^1, X^2), q(X^1, X^2) \right) \leq D_\alpha \left( p(X^1), q(X^1) \right) + \sup_{\hat{x}^1} D_\alpha \left( p(X^2 | X^1 = \hat{x}^1), q(X^2 | X^1 = \hat{x}^1) \right).$$

*Proof.* By definition, when  $\alpha > 1$ , we have

$$\begin{aligned}
&D_\alpha \left( p(X^1, X^2), q(X^1, X^2) \right) \\
&= \frac{1}{\alpha - 1} \log \int_{x^1, x^2} \left( p(x^1, x^2) \right)^\alpha \left( q(x^1, x^2) \right)^{1-\alpha} dx^1 dx^2 \\
&\stackrel{(i)}{=} \frac{1}{\alpha - 1} \log \int_{x^1} \left( \int_{x^2} \left( p(x^2 | x^1) \right)^\alpha \left( q(x^2 | x^1) \right)^{1-\alpha} dx^2 \right) \left( p(x^1) \right)^\alpha \left( q(x^1) \right)^{1-\alpha} dx^1 \\
&\stackrel{(ii)}{\leq} \frac{1}{\alpha - 1} \log \sup_{\hat{x}^1} \left( \int_{x^2} \left( p(x^2 | \hat{x}^1) \right)^\alpha \left( q(x^2 | \hat{x}^1) \right)^{1-\alpha} dx^2 \right) \left( \int_{x^1} \left( p(x^1) \right)^\alpha \left( q(x^1) \right)^{1-\alpha} dx^1 \right) \\
&= \frac{1}{\alpha - 1} \sup_{\hat{x}^1} \log \int_{x^2} \left( p(x^2 | \hat{x}^1) \right)^\alpha \left( q(x^2 | \hat{x}^1) \right)^{1-\alpha} dx^2 + \frac{1}{\alpha - 1} \log \int_{x^1} \left( p(x^1) \right)^\alpha \left( q(x^1) \right)^{1-\alpha} dx^1 \\
&= D_\alpha \left( p(X^1), q(X^1) \right) + \sup_{\hat{x}^1} D_\alpha \left( p(X^2 | X^1 = \hat{x}^1), q(X^2 | X^1 = \hat{x}^1) \right).
\end{aligned}$$

(i) is by Tonelli's theorem. (ii) is by Hölder's Inequality.

When  $\alpha = 1$ , the corresponding Rényi divergence is KL-divergence. Therefore,

$$\begin{aligned}
& D_1 \left( p(X^1, X^2), q(X^1, X^2) \right) \\
&= \int_{x^1, x^2} p(x^1, x^2) \log \frac{p(x^1, x^2)}{q(x^1, x^2)} dx^1 dx^2 \\
&= \int_{x^1, x^2} p(x^2 | x^1) p(x^1) \log \frac{p(x^2 | x^1) p(x^1)}{q(x^2 | x^1) q(x^1)} dx^1 dx^2 \\
&= \int_{x^1, x^2} p(x^2 | x^1) p(x^1) \log \frac{p(x^1)}{q(x^1)} dx^1 dx^2 + \int_{x^1, x^2} p(x^2 | x^1) p(x^1) \log \frac{p(x^2 | x^1)}{q(x^2 | x^1)} dx^1 dx^2 \\
&= D_1 \left( p(X^1), q(X^1) \right) + \mathbb{E}_{\hat{x}^1 \sim p(X^1)} \left[ D_1 \left( p(X^2 | X^1 = \hat{x}^1), q(X^2 | X^1 = \hat{x}^1) \right) \right] \\
&\leq D_1 \left( p(X^1), q(X^1) \right) + \sup_{\hat{x}^1} D_1 \left( p(X^2 | X^1 = \hat{x}^1), q(X^2 | X^1 = \hat{x}^1) \right).
\end{aligned}$$

□