Novel Privacy-Preserving Nash Equilibrium Computation with Pointwise Maximal Leakage Guarantees

Guanpu Chen, Zhaoyang Cheng, Tobias J. Oechtering, Mikael Skoglund

School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden.
E-mail: *guanpu, zhcheng, oech, skoglund@kth.se*

Abstract: This paper investigates privacy-preserving Nash equilibrium (NE) computation in non-cooperative games where players may have correlated payoff functions with prior knowledge. Although mechanisms with differential privacy (DP) guarantees are widely used to mitigate information leakage, their privacy guarantees are ineffective for correlated datasets. To address this limitation, we are inspired by pointwise maximal leakage (PML), a recently proposed privacy measure that exploits prior knowledge for assessing information leakage. We first revisit the traditional privacy-preserving mechanism and demonstrate that its PML guarantee averaged over players can be bounded by its DP guarantee. On this basis, we propose a novel NE-computing mechanism that integrates prior knowledge of players' payoff datasets into noise design, ensuring the adaptation of existing techniques for convergence guarantees. Furthermore, we show that the proposed mechanism offers a tighter bound with PML guarantees than the traditional mechanism with DP guarantees, which refines the over-conservative assessment of information leakage risks with correlated payoff datasets. Numerical experiments illustrate the effectiveness of our theoretical results.

Key Words: Nash equilibrium, differential privacy, pointwise maximal leakage, correlated dataset

1 Introduction

Non-cooperative game-theoretical models investigate strategic interactions among multiple players in competitive environments. A fundamental concept in such game problems is the Nash Equilibrium (NE), a strategy profile in which no player can improve their outcome by unilaterally altering their decision-making. The computation of NE has been a key focus in practice, driving the development of NE-seeking algorithms over the past decade [1–5].

Since the computational mechanisms of NE involve information exchange among players, privacy protection has become a crucial consideration in algorithm design to mitigate information leakage. Differential Privacy (DP), recognized for its explicit mathematical formulation, has emerged as the popular standard for assessing privacy leakage risks [6–8]. Recent studies have incorporated DP into game-theoretic problems to develop privacy-preserving NE-computing mechanisms. For example, an NE-seeking algorithm for aggregate games was designed to achieve both approximate convergence and bounded privacy leakage with DP guarantees [9], followed by the extended study on stochastic game settings [10]. Further research focused on general game models with convergence to the exact NE, as well as the investigations on the privacy budget under DP guarantees [11, 12].

Nevertheless, DP is not without limitations. A primary criticism is that DP focuses on individual entry differences and ignores prior distribution knowledge. Thus, DP privacy guarantees are ineffective for assessing privacy leakage risks in correlated datasets, as first pointed out in [13] and analytically shown in [14]. In fact, practical datasets are frequently accompanied by significant distributional patterns and interdependencies [15]. For instance, medical databases exhibit that some sensitive attributes of different individuals may be correlated across the population [13], while social network datasets reveal interdependencies in user preferences due to familial or social ties [16]. Such inherent correlations enhance the privacy leakage risk, as adversaries can exploit these relationships to infer sensitive information. This limitation highlights the need for privacy measures that account for the underlying distributional structures, including correlated datasets.

To address this concern, recently, a promising information leakage measure, known as pointwise maximal leakage (PML), was introduced that has quickly gained significant attention [17]. As the name suggests, PML generalizes the concept of maximal leakage [18] in a pointwise view. PML exploits information leakage in datasets with prior distribution by examining the worst-case ratio between two scenarios–whether the output sequence is observed or not by an adversary. Unlike DP, PML can effectively assess information leakage with underlying prior distributions, making it particularly advantageous for correlated datasets. The recent research in [19] further refined the PML verification, enhancing its practical applicability. As a result, PML becomes an operationally meaningful, robust, and flexible measure for designing privacy-preserving mechanisms.

In this paper, we focus on the privacy-preserving NE computation in multi-player non-cooperative games. We consider that the players have correlated payoff function sets with prior distributions in practice. Traditional mechanisms with DP guarantees are inadequate to assess information leakage in such scenarios. To address this limitation, we are inspired by PML to design a novel NE-computing mechanism that effectively accounts for correlated payoffs with prior knowledge.

This work was supported in part by the Digital Futures. Corresponding author: *Z. Cheng*.

Our main contributions are as follows.

- We revisit the traditional privacy-preserving NE-computing mechanism and assess its information leakage using DP and PML. Our finding shows that its PML guarantee averaged over players can be bounded by the same DP guarantee. This result highlights the broader applicability of PML in assessing information leakage with players' correlated payoffs.
- We propose a novel NE-computing mechanism that integrates the information of prior distributions of players' payoff datasets into the noise design. The design maintains bounded noise, allowing for the seamless adoption of developed convergence analysis in the existing privacy-preserving algorithms.
- We analyze the proposed NE-computing mechanism. On the one hand, the DP guarantee maintains the same bound as in the traditional privacy-preserving mechanism. On the other hand, given the prior distribution of correlated payoff datasets, the PML guarantee indicates a tighter bound, which refines the over-conservative assessment of information leakage risks.

The organization of the paper is as follows. Section II shows the problem formulation. Section III introduces the PML measure and shows its relation with DP. Section IV proposes a novel NE-computing mechanism with both algorithm convergence analysis and better privacy protection. Section V gives numerical experiments to evaluate our theoretical results, followed by the conclusions in Section VI.

2 Problem Formulation

In this section, we show the non-cooperative game-theoretical model, along with the NE concept, the traditional privacy-preserving mechanism, the DP guarantee, and the problem statement with the current challenge.

2.1 Nash Equilibrium and Computation

Consider a non-cooperative game G with multiple players indexed by $[N]=\{1,\ldots,N\}$. Player $i\in[N]$ takes a strategy x_i subject to a local constraint $\Omega_i\subseteq\mathbb{R}^n$. Let $x=\left(x_1^T,\ldots,x_N^T\right)^T$ be the stacked strategy profile for all players and $x_{-i}=\left(x_1^T,\ldots,x_{i-1}^T,x_{i+1}^T,\ldots,x_N^T\right)^T$ be the profile except for player i. Take $\Omega=\prod_{i=1}^N\Omega_i\subseteq\mathbb{R}^{nN}$ as all players' constraints. Player i has the continuously differentiable payoff function $f_i:\mathbb{R}^{nN}\to\mathbb{R}$. Take $f=\{f_1,\ldots,f_N\}$ as the set of all player's payoff functions, and similarly define the set $f_{-i}=\{f_1,\ldots,f_{i-1},f_{i+1},\ldots f_N\}$. Given x_{-i}, the objective of player i is to solve the following problem:

$$\min_{x_i\in\Omega_i} f_i\left(x_i,x_{-i}\right).$$

With these factors, the multi-player game can be represented by a triple $G=\{[N],\Omega,f\}$.

Definition 1 (Nash Equilibrium) *An strategy profile* $\mathbf{x}^*=\left(x_i^*,x_{-i}^*\right)\in\Omega$ *is said to be an NE of game G if for all* $i\in[N]$ *and all* $x_i\in\Omega_i$, *we have*

$$f_i\left(x_i^*,x_{-i}^*\right)\le f_i\left(x_i,x_{-i}^*\right).$$

NE is a primary concept in non-cooperative games because when achieving an NE, no players can obtain a better payoff by unilaterally changing strategies.

Consider designing discrete-time algorithms to compute an NE. Regarding the protection of players' private information in the payoffs f, players are usually not willing to send their strategy x directly to others. Instead, at each time k in the computation iterations, player i adds noise ζ_i^k to its strategy x_i^k. Then, player i broadcasts its obscured strategy $o_i^k=x_i^k+\zeta_i^k$ to other players, which is also called observation. Take o_{-i}^k as player i's observations of other players.

Furthermore, let $\boldsymbol{x}=\{x^k\}_{k=0}^T$ with $x^k=\{x_i^k\}_{i=1}^N$ represent the strategy sequence from time 0 to T. Analogously, we can define the observation sequence by $\boldsymbol{o}=\{o^k\}_{k=0}^T$ with $o^k=\{o_i^k\}_{i=1}^N$. In the non-cooperative game G, player i only knows its payoff function f_i, its strategy x_i^k, and observation o^k at current time k. To achieve an NE strategy, player i may update its strategy variable x_i^k by $x_i^{k+1}=h(x_i^k,o^k,\lambda^k,f_i)$, where λ^k is the iteration stepsize and h is an operator constructed according to difference computational mechanisms. In this view, the potentially designed algorithm for each player $i\in[N]$ can be expressed as the following mechanism:

$$\begin{cases}o_i^k=x_i^k+\zeta_i^k, & x_i^0\in\Omega_i,\\ x_i^{k+1}=h(x_i^k,o^k,\lambda^k,f_i). & \end{cases}\qquad(\mathcal{A}_1)$$

Concretely, there exist various ways to design the operator $h(x_i^k,o^k,\lambda^k,f_i)$ in practice. For instance, a basic idea is to update the strategy through the best response mechanism [6], *i.e.*, $x_i^{k+1}=\arg\max_{x_i\in\Omega_i}f_i(x_i,o_{-i}^k)$. The first-order derivative of payoff functions can also be adopted [11], *i.e.*, $x_i^{k+1}=x_i^k+\lambda^k\nabla_{x_i}f_i(x_i^k,o_{-i}^k)$. No matter the detailed kind of computation mechanisms in $\mathcal{A}_1$, one main purpose is the guarantee of algorithm convergence. Besides convergence, such an NE-computing mechanism $\mathcal{A}_1$ with noise is employed for privacy protection. Thus, it is also important to consider the privacy guarantees along with iterations.

2.2 Revisiting Differential Privacy

As a popular standard to protect privacy, DP has been widely applied to algorithm design in optimization and game problems [9, 10, 12, 20]. DP assesses the information leakage risk in mechanisms from the perspective of an individual partaking. The expression of DP is based on adjacent datasets, namely, two datasets that only differ in one entry [7]. In the game model G, the players' payoff function set f is considered as the private dataset. Two function sets f and f' are accordingly adjacent if only one player's payoff function is different [9], *i.e.*, there exists $i\in[N]$ such that $f_i\neq f_i'$ but $f_j=f_j'$ for all $j\neq i$.

Let E_f be the set containing all possible payoff functions f, while $\mathcal{F}$ be the σ-algebra on E_f. Similarly, we can define $\mathcal{F}_i$ and $\mathcal{F}_{-i}$ generated by f_i and f_{-i}, respectively. Suppose that $(E_f,\mathcal{F})$ is a standard Borel space. In this context, take F as a random variable from $(E_f,\mathcal{F})$. Besides, regarding the observation sequence $\boldsymbol{o}$, we can similarly define the set E_o, the σ-algebra $\mathcal{O}$, the standard Borel space $(E_o,\mathcal{O})$, and the random variable O.

With these preparations, algorithm $\mathcal{A}_1$ could be viewed as a mapping $\mathcal{A}_1:E_f\to\mathcal{O}$. We can further define a mapping $P_{O|F}:E_f\times\mathcal{O}\to[0,1]$ such that $P_{O|F=f}(\cdot)$ is the conditional probability of the observation sequence $\boldsymbol{o}\in\mathcal{O}$ given the payoff function set $f\in E_f$. The definition of DP regarding algorithm $\mathcal{A}_1$ is accordingly as follows [8].

Definition 2 (ϵ-Differential Privacy) *Given $\epsilon \geq 0$, algorithm $\mathcal{A}_1$ is said to be ϵ-DP if for any possible $\boldsymbol{o} \in \mathcal{O}$ and any adjacent datasets f and f', we have*

$$\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \leq \exp(\epsilon).$$

The formulation of DP guarantees in Definition 2 is consistent with other common formulations. For example, according to the results in the field of control and optimization [6, 9], algorithm $\mathcal{A}_1$ is ϵ-DP if for any two adjacent datasets $f, f' \in E_f$ and any observation sequence $\boldsymbol{o}_s \subseteq E_o$, we have

$$\mathbb{P}\left[\mathcal{A}_1\left(f\right) \in \boldsymbol{o}_s\right] \leq e^{\epsilon}\mathbb{P}\left[\mathcal{A}_1\left(f'\right) \in \boldsymbol{o}_s\right].$$

Both the above formulation and Definition 2 illustrate the fact that, if the privacy guarantee ϵ is small enough, then the likelihoods of all possible observation sequences $\boldsymbol{o}$ and all adjacent datasets are close, meaning the difficulty of distinguishing the two adjacent datasets with high probability.

Given an observation sequence $\boldsymbol{o}$ and the payoff function set f, we can rewrite the strategy sequence $\boldsymbol{x}$ generated by an iterated algorithm (here it refers to algorithm $\mathcal{A}_1$) as $\boldsymbol{x} = A(\boldsymbol{o}, f)$. At each time k, based on any adjacent datasets f and f', define DP sensitivity [12, 20] of algorithm $\mathcal{A}_1$ by

$$\Delta_{DP}(k) = \sup_{\boldsymbol{o}\in\mathcal{O}} \left\{ \sup_{\substack{x_i^k \in A(\boldsymbol{o},f) \\ x_i'^k \in A(\boldsymbol{o},f')}} \|x^k - x'^k\|_1 \right\}. \quad (1)$$

The following lemma has established a bridge between privacy guarantee ϵ and DP sensitivity Δ_{DP}.

Lemma 1 ([20]) *Suppose that the noise in algorithm $\mathcal{A}_1$ follows a Laplace distribution $\zeta_i^k \sim Lap(0, M^k)$ for $i \in [N]$. For any finite iteration T, algorithm $\mathcal{A}_1$ is ϵ-DP if*

$$\sum_{k=0}^{T} \frac{\Delta_{DP}(k)}{M^k} \leq \epsilon.$$

2.3 Problem Statement

In fact, a notable concern is that DP has been subject to scrutiny, with some studies suggesting that it may be vulnerable due to its omission of the intrinsic distributional characteristics of correlated datasets [13]. As outlined in Definition 2, DP does not incorporate any underlying distribution information within the payoff function datasets f, thereby failing to capture how these inherent distributions influence the information leakage risk. While extant research often assumes that datasets are comprised exclusively of independent records, the reality is that datasets frequently exhibit significant distributional patterns and interdependencies [15], which in turn affect the privacy-preserving mechanism. For example, medical databases exhibit that some sensitive attributes of different individuals may be correlated across the population [13]. In the social network [16], agents' preferences may be strongly correlated due to familial factors or friendship ties. Datasets characterized by underlying distribution patterns pose elevated privacy risks, especially when the distributions have inherent correlations and the population base is large. Traditional DP guarantees are inadequate to assess information leakage in such scenarios. We will elucidate more in the following example.

Example 1 *Suppose a company with 10 members, where each member works on either Task 1 or Task 2. A business rival aims to figure out "How many members work on Task 1?" Suppose that $Lap(0, \frac{1}{\beta})$ noise has been added to the truth with DP guarantees [13]. Imagine an extreme case where members' choices are 100% correlated due to, for instance, the limited resources. Then, the output of the query answer might be 12, which could be proved to be $\exp(10\beta)$ times more likely when the truth is 10 (refers to Task 1) than when the truth is 0 (refers to Task 2). Hence, DP falls short in preventing information leakage under these circumstances.*

Generally, for datasets with an underlying distribution, the risk of information leakage is a pressing concern. The traditional DP framework, focusing on the privacy of individual entry differences, is ineffective for assessing the amount of information leakage if the prior distribution of datasets is not independent, especially for large-scale datasets with high correlation. This limitation underscores the importance of an innovative privacy measure and the development of a privacy-preserving NE-computing mechanism.

Problem 1 *How can we design a privacy-preserving NE-computing mechanism with privacy guarantees when players have correlated payoff datasets?*

3 Pointwise Maximal Leakage

In this section, we will introduce PML and its relation with DP concerning discrete-time algorithm iterations.

PML has ascended as a promising privacy measure recently and garnered significant attention [17]. As the name suggests, PML generalizes the concept of maximal leakage [18] in a pointwise view and integrates robustness characteristics by quantifying the amount of worst-case information leakage of a mechanism. In other words, PML can assess any process from a database with arbitrary distribution to an outcome sequence and thus is operationally meaningful, robust, and flexible.

Recall the random variables F and O taking values from $(E_f, \mathcal{F})$ and $(E_o, \mathcal{O})$, respectively. To describe the prior information behind payoff functions f, consider the distributions P_F of F and P_O of O. Denote $\operatorname{supp}(P_F) = \{f \in \mathcal{F} : P_F(f) > 0\}$ and $\operatorname{supp}(P_O) = \{\boldsymbol{o} \in \mathcal{O} : P_O(\boldsymbol{o}) > 0\}$ as the support sets of F and O, respectively. When an adversary observes $\boldsymbol{o}$, it usually wants to guess the true F. Thus, take $\hat{F}$ as the random variable of the adversary's guess and $(E_{\hat{f}}, \hat{\mathcal{F}})$ as all its possible measurable spaces. Analogously, the guessing process can be denoted by a mapping $P_{\hat{F}|O} : E_o \times \hat{\mathcal{F}} \to [0, 1]$.

In the following, we introduce PML from the perspective of gain functions [19], and more comprehensive discussions can be found in [17]. Based on the guess, the objective of an adversary is to maximize its expected gain g to assess the performance of the guess, where the form of the gain function should be $g : \mathcal{F} \times \hat{\mathcal{F}} \to \mathbb{R}^+$. Noticing the adversary's preference is unknown, the adversary may construct gain functions in different ways, for example, by the identity function or by other metrics and distances. To address these scenarios, we consider that the adversary could pick any gain function g from $\Gamma \triangleq \left\{ g : \sup_{\hat{f} \in \hat{\mathcal{F}}} \mathbb{E}\left[g(F, \hat{f}) \right] < \infty \right\}$. Given a specific realization $\hat{f}$ of the guess, $\mathbb{E}[g(F, \hat{f})]$ should be the adversary's expected gain regarding a random variable F. Given the observation sequence $\boldsymbol{o}$, $\mathbb{E}[g(F, \hat{F})|O = \boldsymbol{o}]$ should be the adversary's expected gain between a random variable F and a random variable $\hat{F}$.

Based on the concept of the gain function g, when the adversary has no observation, $\sup_{\hat{f} \in \hat{\mathcal{F}}} \mathbb{E}[g(F, \hat{f})]$ denotes the worst-case value for protecting privacy since the adversary maximizes its expected gain. On the other hand, when the adversary observes $\boldsymbol{o}$, $\sup_{P_{\hat{F}|O}} \mathbb{E}[g(F, \hat{F})|O = \boldsymbol{o}]$ is the worst-case value for protecting privacy since the adversary maximizes its expected gain among all guesses $\hat{F}$ according to possible mappings $P_{\hat{F}|O}$.

Then, these two expected adversarial gains in the worst case over all possible measurable spaces and for all possible gain functions could be utilized to describe the PML of any mechanism [19]. Specifically, defined the PML from F to $\boldsymbol{o}$ as

$$\ell(F \to \boldsymbol{o}) := \log \sup_{(E_{\hat{F}}, \hat{\mathcal{F}}), g \in \Gamma} \frac{\sup_{P_{\hat{F}|O}} \mathbb{E}[g(F, \hat{F})|O = \boldsymbol{o}]}{\sup_{\hat{f} \in \hat{\mathcal{F}}} \mathbb{E}[g(F, \hat{f})]}. \quad (2)$$

Remark 1 *Upon examining the above formulation, PML is defined with the combination of the suprema over all possible guessing and the ratio*[1] *of worst-case values. The definition ensures that PML captures the worst-case scenario and quantifies the largest amount of potential information leakage. In this context, PML serves as a robust measure of privacy protection.*

Although PML possesses remarkable advantages, its formulation in (2) is complicated, thereby posing considerable challenges for its practical implementation. Fortunately, it has been shown in the following theorem that PML has a simple equivalent formulation beneficial for analysis.

Theorem 1 ([19, Theorem 3]) The PML from F to $\boldsymbol{o} \in \operatorname{supp}(P_O)$ can be expressed as

$$\ell(F \to \boldsymbol{o}) = \log \sup_{f \in \operatorname{supp}(P_F)} \frac{P_{O|F=f}(\boldsymbol{o})}{P_O(\boldsymbol{o})}.$$

More details on the fundamental result of PML can be found in [19]. With this, we can now define a privacy measure suitable for our NE computation.

Definition 3 (ϵ-PML averaged over players) *Given $\epsilon \geq 0$ and a prior distribution P_F of players' payoff function f, algorithm $\mathcal{A}_1$ is said to be ϵ-PML averaged over players if for all $\boldsymbol{o} \in \operatorname{supp}(P_O)$, we have*

$$\frac{1}{N} \ell(F \to \boldsymbol{o}) \leq \epsilon.$$

Note that PML assesses the information leakage across the entire dataset, while DP merely assesses the leakage of all possible adjacent datasets with one different entry. Hence, it is reasonable to introduce the factor $\frac{1}{N}$ in Definition 3, which aligns the privacy guarantee with the same level as DP regarding player numbers.

A general equivalence between the formulations of DP and PML of the worst case over individual data entries has been provided in [21, Theorem 4.2] by considering finite alphabets. Here, we consider continuous payoff functions in the game G setting. In particular, we investigate Laplace noise in algorithm $\mathcal{A}_1$ and reveal the privacy guarantees of DP and PML in the following result.

Theorem 2 *Suppose that the noise in algorithm $\mathcal{A}_1$ follows a Laplace distribution $\zeta_i^k \sim Lap(0, M^k)$ for $i \in [N]$. If there exists $\epsilon > 0$ such that, for any finite iteration T,*

$$\sum_{k=0}^{T} \frac{\Delta_{DP}(k)}{M^k} \leq \epsilon,$$

then algorithm $\mathcal{A}_1$ is both ϵ-DP and ϵ-PML averaged over players (for any prior distribution P_F).

Proof *If $\sum_{k=0}^{T} \frac{\Delta_{DP}(k)}{M^k} \leq \epsilon$, then algorithm $\mathcal{A}_1$ is ϵ-DP according to [20]. By Definition 2, for any possible $\boldsymbol{o} \in \mathcal{O}$ and two adjacent payoff function sets f and f', we have $\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \leq \exp(\epsilon)$. Recall P_{F_i} and $P_{F_{-i}}$ as the distributions of F_i and F_{-i}, respectively. The ϵ-DP result holds over the support sets, i.e., for any $\boldsymbol{o} \in \operatorname{supp}(P_O)$, $i \in [N]$, $f_{-i} \in \operatorname{supp}(P_{F_{-i}})$, $f_i, f_i' \in \operatorname{supp}(P_{F_i})$, we can obtain $\frac{P_{O|F=\{f_i, f_{-i}\}}(\boldsymbol{o})}{P_{O|F=\{f_i', f_{-i}\}}(\boldsymbol{o})} \leq \exp(\epsilon)$.*

For $f = \{f_1, \ldots, f_N\}$ and $f' = \{f_1', \ldots, f_N'\}$, take the sets $f_{[i:j]} = \{f_i, \ldots, f_j\}$ and $f'_{[i:j]} = \{f_i', \ldots, f_j'\}$ if $i \leq j$, while $f_{[i:j]} = f'_{[i:j]} = \varnothing$, if $i > j$. In this view, for any two datasets f and f' which are not restricted as adjacent payoffs, we have

$$\begin{aligned}
&\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \\
=&\frac{P_{O|F=\{f_{[1:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f'_{[1:N]}\}}(\boldsymbol{o})} \\
=&\frac{P_{O|F=\{f_1, f_{[2:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f_1', f_{[2:N]}\}}(\boldsymbol{o})} \cdot \frac{P_{O|F=\{f_1', f_{[2:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f_1', f'_{[2:N]}\}}(\boldsymbol{o})} \\
=&\frac{P_{O|F=\{f_1, f_{[2:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f_1', f_{[2:N]}\}}(\boldsymbol{o})} \cdot \frac{P_{O|F=\{f'_{[1:1]}, f_2, f_{[3:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f'_{[1:1]}, f_2', f_{[3:N]}\}}(\boldsymbol{o})} \\
&\cdot \frac{P_{O|F=\{f'_{[1:2]}, f_{[3:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f'_{[1:2]}, f'_{[3:N]}\}}(\boldsymbol{o})} \\
=&\cdots \\
=&\Pi_{i=1}^{N} \frac{P_{O|F=\{f'_{[1:i-1]}, f_i, f_{[i+1:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f'_{[1:i-1]}, f_i', f_{[i+1:N]}\}}(\boldsymbol{o})}.
\end{aligned}$$

Notice that $\frac{P_{O|F=\{f'_{[1:i-1]}, f_i, f_{[i+1:N]}\}}(\boldsymbol{o})}{P_{O|F=\{f'_{[1:i-1]}, f_i', f_{[i+1:N]}\}}(\boldsymbol{o})} \leq \exp(\epsilon)$ since $\{f'_{[1:i-1]}, f_i, f_{[i+1:N]}\}$ and $\{f'_{[1:i-1]}, f_i', f_{[i+1:N]}\}$ are adja-

[1] We use the conventions that $\frac{0}{0} = 1$ and $\frac{x}{0} = \infty$ if $x > 0$.

cent. Thus, we have the following inequality

$$\sup_{o\in \operatorname{supp}(P_O)} \sup_{f,f'\in \operatorname{supp}(P_F)} \frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \leq \exp(N\epsilon).$$

By taking the reciprocal of both sides and swapping f and f' in the above inequality,

$$\inf_{o\in \operatorname{supp}(P_O)} \inf_{f,f'\in \operatorname{supp}(P_F)} \frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \geq \exp(-N\epsilon).$$

To analyze the bounds of PML, we first introduce the following inequalities with integrals

$$\begin{aligned}
&\int_{f\in \operatorname{supp}(P_F)} \inf_{\boldsymbol{o}\in \operatorname{supp}(P_O)} \inf_{f'\in \operatorname{supp}(P_F)} \frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} P_F(f)df \\
\geq & \inf_{\boldsymbol{o}\in \operatorname{supp}(P_O)} \inf_{\hat{f},f'\in \operatorname{supp}(P_F)} \frac{P_{O|F=\hat{f}}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \int_{f\in \operatorname{supp}(P_F)} P_F(f)df \\
\geq & \exp(-N\epsilon).
\end{aligned}$$

Following the above result, we can derive

$$\begin{aligned}
&\sup_{\boldsymbol{o}\in \operatorname{supp}(P_O)} \exp \ell(F \to \boldsymbol{o}) \\
=& \sup_{\boldsymbol{o}\in \operatorname{supp}(P_O)} \frac{\sup_{f'\in \operatorname{supp}(P_F)} P_{O|F=f'}(\boldsymbol{o})}{\int_{f\in \operatorname{supp}(P_F)} P_{O|F=f}(\boldsymbol{o})P(f)df} \\
\leq& \frac{1}{\int_{f\in \operatorname{supp}(P_F)} \inf_{\boldsymbol{o}\in \operatorname{supp}(P_O)} \inf_{f'\in \operatorname{supp}(P_F)} \frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} P_F(f)df} \\
\leq& \exp(N\epsilon),
\end{aligned} \tag{3}$$

where the first inequality holds because of Fatou's Lemma. Therefore, the above conclusion yields that algorithm $\mathcal{A}_1$ is ϵ-PML averaged over players, which completes the proof. □

The result in Theorem 2 shows that the applicability area of PML is wider than that of DP. If an algorithm adopts the noise with a Laplace distribution, then the bound of DP guarantees is the same as PML guarantees. On the other side, the conclusion in Theorem 2 is true for any prior distribution behind the payoff function dataset f, meaning the capability of PML to describe the information leakage over the correlated datasets.

4 Privacy-preserving NE-computing Mechanism

In this section, we aim to design a novel NE-computing mechanism and show both the convergence guarantee and better privacy guarantees than algorithm $\mathcal{A}_1$.

Following the game setting for introducing algorithm $\mathcal{A}_1$, we further consider a prior distribution P_{F_i} behind the payoff function f_i of each player i. We redesign the noise in algorithm $\mathcal{A}_1$ by utilizing each player's prior distribution P_{F_i}. Here, denote the entropy $H_i = -\int_{f_i\in \operatorname{supp}(P_{F_i})} P_{F_i}(f_i)\log(P_{F_i}(f_i))df_i$ and the function $\sigma(s) = 1 + \frac{1}{1+\exp(-s)}$ generated by a sigmoid function. The following mechanism shows the novel discrete-time NE computation. For $i \in [N]$,

$$\begin{cases} o_i^k = x_i^k + \zeta_i^k \cdot \sigma(H_i), \qquad x_i^0 \in \Omega_i, \\ x_i^{k+1} = h(x_i^k, o^k, \lambda^k, f_i). \end{cases} \tag{$\mathcal{A}_2$}$$

Similarly to algorithm $\mathcal{A}_1$, we continue using the scalar λ^k and the operator h in algorithm $\mathcal{A}_2$ to represent an iteration stepsize and a general updating process regarding player i's strategy variable x_i^k, respectively.

4.1 Convergence Analysis

As a newly proposed algorithm, its convergence analysis is of primary concern. Similarly to most NE-computing algorithms in non-cooperative games [1–5, 9, 11], suppose the payoff function set f equipped with good properties leading to a strongly monotone and Lipchitz continuous pseudo-gradient, as well as the constraint set Ω with convexity and compactness. We give the convergence result of algorithm $\mathcal{A}_2$ based on the noise following a Laplace distribution and the payoff prior distribution P_{F_i}.

Theorem 3 *Suppose that the noise following Laplace distribution $\zeta_i^k \sim Lap(0, M^k)$ satisfies $\sum_{k=0}^{\infty}(M^k)^2 < \infty$. If algorithm $\mathcal{A}_1$ is convergent to an NE, then with the same updating operator h and stepsize setting λ^k, algorithm $\mathcal{A}_2$ is also convergent to the same NE.*

Proof *It is not hard to find that the main distinction between algorithms $\mathcal{A}_1$ and $\mathcal{A}_2$ lies in the scaling of the noise variance from ζ_i^k to $\zeta_i^k\sigma(H_i)$. Generally, when proving the convergence of algorithm $\mathcal{A}_1$, all concerns about the noise are to examine whether some constants can bound them. For each time k, take $\mathcal{X}^k = \{x_i^l, i \in [N], 0 \leq l \leq k\}$. Specifically, after some process like [11, Lemma 5], the following two properties of noise are necessary to verify for the convergence of algorithm $\mathcal{A}_1$: $\mathbb{E}[\langle C_1, C_2\zeta_i^k\rangle|\mathcal{X}^k] = 0$ and $\mathbb{E}[\langle C_3\zeta_i^k, C_4\zeta_i^k\rangle|\mathcal{X}^k] \leq c$, where the inner product is well-defined over vectors in $\mathbb{R}^n$, and constants satisfy $C_1 \in \mathbb{R}^n$, $C_2, C_3, C_4 \in \mathbb{R}^{n\times n}$, and $c \in \mathbb{R}$.*

Regarding our noise setting in algorithm $\mathcal{A}_2$, we need to examine $\zeta_i^k\sigma(H_i)$. Recalling $\zeta_i^k \sim Lap(0, M^k)$, we obtain $\zeta_i^k\sigma(H_i) \sim Lap(0, \sigma(H_i)M^k)$. Hence, we can further derive $\mathbb{E}[\zeta_i^k\sigma(H_i)|\mathcal{X}^k] = 0$ and $\mathbb{E}[\|\zeta_i^k\sigma(H_i)\|_2^2|\mathcal{X}^k] = 2\sigma(H_i)^2(M^k)^2$. In this context, we have

$$\mathbb{E}[\langle C_1, C_2\zeta_i^j\sigma(H_i)\rangle|\mathcal{X}^k] = C_1^T C_2\mathbb{E}[\zeta_i^j\sigma(H_i)|\mathcal{X}^k] = 0,$$

$$\mathbb{E}[\langle C_3\zeta_i^k\sigma(H_i), C_4\zeta_i^k\sigma(H_i)\rangle|\mathcal{X}^k] \leq \|C_3^T C_4\|_2\sigma(H_i)^2(M^k)^2.$$

Thus, the newly introduced noise terms in algorithm $\mathcal{A}_2$ will not impact its convergence analysis. As a result, provided that algorithm $\mathcal{A}_1$ converges to an NE, algorithm $\mathcal{A}_2$ is guaranteed to converge to the same NE. □

4.2 Information Leakage Analysis

We can follow the definitions of ϵ-DP and ϵ-PML averaged over players of algorithm $\mathcal{A}_1$ to analogously define these two guarantees of algorithm $\mathcal{A}_2$. The following theorem assesses the PML guarantee of algorithm $\mathcal{A}_2$.

Theorem 4 *Suppose that the noise ζ_i^k follows a Laplace distribution with $\zeta_i^k \sim Lap(0, M^k)$. If there exists $\epsilon_1 > 0$ such that, for any finite iteration T,*

$$\sum_{k=0}^{T} \frac{\Delta_{DP}(k)}{M^k} \leq \epsilon_1,$$

then algorithm $\mathcal{A}_2$ is ϵ_1-DP and there exists $\epsilon_2 < \epsilon_1$ such that algorithm $\mathcal{A}_2$ is ϵ_2-PML averaged over players.

To give a rigorous proof of Theorem 4, we need to introduce a new sensitivity term to assess the bound of PML guarantees in algorithm $\mathcal{A}_2$. Recalling the DP sensitivity in (1), we denote the PML sensitivity of a fixed f at time k by

$$\Delta_{PML}(k,f) = \sup_{\substack{\boldsymbol{o}\in\mathrm{supp}(P_O)\\ f'\in\mathrm{supp}(P_F)}} \left\{ \sup_{\substack{x^k\in A(\boldsymbol{o},f)\\ x'^k\in A(\boldsymbol{o},f')}} \|x^k - x'^k\|_1 \right\}. \tag{4}$$

The distinction between $\Delta_{DP}(k)$ and $\Delta_{PML}(k,f)$ emerges in their functional domains: Δ_{PML} considers all possible payoff function sets over $\mathcal{F}$ while Δ_{DP} just operates with the adjacent sets of f. Following the theoretical results in DP [9, 12], if f and f' are adjacent where $f_i \neq f_i'$ but $f_j = f_j'$ for $j \neq i$, then $\|x^k - x'^k\|_1 = \|x_i^k - x_i'^k\|_1$. In this special case, both Δ_{PML} and Δ_{DP} yield equivalent results with the almost-sure guarantee. Consequently, Δ_{PML} demonstrates superior generality when considering non-adjacent functions, since $\|x^k - x'^k\|_1$ may not equal to $\|x_i^k - x_i'^k\|_1$ in most cases and Δ_{DP} could no longer characterize this scenario. In other words, Δ_{PML} expands the descriptive domain beyond the limitations of Δ_{DP} and offers flexibility for privacy analysis across players' general payoff function spaces.

Next, we use the concept of Δ_{PML} to prove Theorem 4.

Proof *Take O^k and X^k as the random variables to represent the observation o^k and the strategy x^k at time k, respectively. Also, take $O^{[k]}$ and $X^{[k]}$ as the random variables to represent the observation $\{o^l\}_{l=0}^k$, the strategy $\{x^l\}_{l=0}^k$ from time $l = 0$ to k, respectively. To avoid ambiguity, if $l = 0$, take $O^{[l-1]} = \varnothing$ and $\{o^{l-1}, x^l\} = \{x^l\}$. According to [20], for any fixed $\boldsymbol{o}$, f and T,*

$$\begin{aligned}
&P_{O|F=f}(\boldsymbol{o})\\
=&\int P_{O^{[T]},X^{[T]}|F=f}(\{o^k, x^k\}_{k=0}^T) d\boldsymbol{x}\\
=&\int \Pi_{k=0}^T P_{O^k,X^k|F=f,\{O^{[k-1]},X^{[k-1]}\}=\{o^l,x^l\}_{l=0}^{k-1}}(\{o^k, x^k\}) d\boldsymbol{x}\\
=&\int \Pi_{k=0}^T \{P_{O^k|F=f,\{O^{[k-1]},X^{[k]}\}=\{o^{l-1},x^l\}_{l=0}^k}(o^k)\\
&\qquad \cdot P_{X^k|F=f,\{O^{[k-1]},X^{[k-1]}\}=\{o^l,x^l\}_{l=0}^{k-1}}(x^k)\} d\boldsymbol{x}\\
=&\int_{\boldsymbol{x}\in A(\boldsymbol{o},f)} \Pi_{k=0}^T \Pi_{i=1}^N P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k) d\boldsymbol{x}.
\end{aligned}$$

Due to $\zeta_i^k \sim Lap(0, M^k)$, we have $x_i^k + \zeta_i^k \cdot \sigma(H_i) \sim Lap(x_i^k, \sigma(H_i)M^k)$, and

$$P_{O_i^k|F=f,X_i^k=x_i^k}(x_i^k + \zeta_i^k \cdot \sigma(H_i)) = \frac{\exp\left(-\frac{\|o_i^k - x_i^k\|_1}{\sigma(H_i)M^k}\right)}{2\sigma(H_i)M^k}.$$

Thus, for any f and f', the following inequality holds

$$\begin{aligned}
\frac{P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k)}{P_{O_i^k|F=f',X_i^k=x_i'^k}(o_i^k)} &= \frac{\frac{1}{2\sigma(H_i)M^k}\exp\left(-\frac{\|o_i^k - x_i^k\|_1}{\sigma(H_i)M^k}\right)}{\frac{1}{2\sigma(H_i)M^k}\exp\left(-\frac{\|o_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right)}\\
&= \exp\left(\frac{\|o_i^k - x_i'^k\|_1 - \|o_i^k - x_i^k\|_1}{\sigma(H_i)M^k}\right).
\end{aligned} \tag{5}$$

Now consider $\|o_i^k - x_i'^k\|_1 - \|o_i^k - x_i^k\|_1 \leq \|x_i^k - x_i'^k\|_1$. By substituting this in (5), we have

$$\begin{aligned}
&\Pi_{k=0}^T \Pi_{i=1}^N \frac{P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k)}{P_{O_i^k|F=f',X_i^k=x_i'^k}(o_i^k)}\\
\leq& \exp\left(\sum_{k=0}^T \sum_{i=1}^N \frac{\|x_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right)\\
\leq& \exp\left(\sum_{k=0}^T \sum_{i=1}^N \frac{\|x_i^k - x_i'^k\|_1}{M^k}\right).
\end{aligned}$$

According to [20], if f and f' are adjacent, then we have

$$\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})} \leq \exp\left(-\sum_{k=0}^T \frac{\Delta_{DP}(k)}{M^k}\right).$$

Thus, algorithm $\mathcal{A}_2$ is ϵ_1-DP due to $\sum_{k=0}^T \frac{\Delta_{DP}(k)}{M^k} \leq \epsilon_1$.

Similarly, we can find $\|o_i^k - x_i'^k\|_1 - \|o_i^k - x_i^k\|_1 \geq -\|x_i^k - x_i'^k\|_1$. Again, by substituting the above condition in (5),

$$\begin{aligned}
&\Pi_{k=0}^T \Pi_{i=1}^N \frac{P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k)}{P_{O_i^k|F=f',X_i^k=x_i'^k}(o_i^k)}\\
=&\Pi_{k=0}^T \Pi_{i=1}^N \exp\left(\frac{\|o_i^k - x_i'^k\|_1 - \|o_i^k - x_i^k\|_1}{\sigma(H_i)M^k}\right)\\
\geq&\Pi_{k=0}^T \Pi_{i=1}^N \exp\left(-\frac{\|x_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right)\\
=&\exp\left(-\sum_{k=0}^T \sum_{i=1}^N \frac{\|x_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right).
\end{aligned} \tag{6}$$

By investigating a relation

$$x_i^{k+1} = x_i'^{k+1} + h(x_i^k, o^k, \lambda^k, f_i) - h(x_i'^k, o^k, \lambda^k, f_i'),$$

let us define a mapping $\mathcal{B} : A(\boldsymbol{o}, f') \to A(\boldsymbol{o}, f)$ such that $\boldsymbol{x} = \mathcal{B}(\boldsymbol{x}')$. Obviously, $\mathcal{B}$ is a bijection from set $A(\boldsymbol{o}, f')$ to set $A(\boldsymbol{o}, f)$. Then,

$$\begin{aligned}
&\int_{\boldsymbol{x}\in A(\boldsymbol{o},f)} \Pi_{k=0}^T \Pi_{i=1}^N P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k) d\boldsymbol{x}\\
=&\int_{\boldsymbol{x}\in A(\boldsymbol{o},f)} \frac{\Pi_{k=0}^T \Pi_{i=1}^N P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k)}{\Pi_{k=0}^T \Pi_{i=1}^N P_{O_i^k|F=f',X_i^k=x_i'^k}(o_i^k)}\\
&\qquad \cdot \Pi_{k=0}^T \Pi_{i=1}^N P_{O^k|F=f',X_i^k=x_i'^k}(o_i^k) d\boldsymbol{x}\\
\geq&\int_{\boldsymbol{x}'\in A(\boldsymbol{o},f')} \exp\left(-\sum_{k=0}^T \sum_{i=1}^N \frac{\|x_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right)\\
&\qquad \cdot \Pi_{k=0}^T \Pi_{i=1}^N P_{O^k|F=f',X_i^k=x_i'^k}(o_i^k) d\boldsymbol{x}'\\
\geq&\inf_{\substack{\boldsymbol{o}\in\mathrm{supp}(P_O)\\ f'\in\mathrm{supp}(P_F)}} \left\{ \inf_{\substack{x^k\in A(\boldsymbol{o},f)\\ x'^k\in A(\boldsymbol{o},f')}} \exp\left(-\sum_{k=0}^T \sum_{i=1}^N \frac{\|x_i^k - x_i'^k\|_1}{\sigma(H_i)M^k}\right)\right\}\\
&\cdot \int_{\boldsymbol{x}'\in A(\boldsymbol{o},f')} \Pi_{k=0}^T \Pi_{i=1}^N P_{O^k|F=f',X_i^k=x_i'^k}(o_i^k) d\boldsymbol{x}'.
\end{aligned} \tag{7}$$

By considering both (6) and (7), we have

$$\begin{aligned}
&\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})}\\
=&\frac{\int_{\boldsymbol{x}\in A(\boldsymbol{o},f)}\Pi_{k=0}^{T}\Pi_{i=1}^{N}P_{O_i^k|F=f,X_i^k=x_i^k}(o_i^k)d\boldsymbol{x}.}{\int_{\boldsymbol{x}'\in A(\boldsymbol{o},f')}\Pi_{k=0}^{T}\Pi_{i=1}^{N}P_{O_i^k|F=f',X_i^k=x_i'^k}(o_i^k)d\boldsymbol{x}'.}\\
\geq&\exp\left(-\sum_{k=0}^{T}\frac{1}{M^k}\sup_{\substack{\boldsymbol{o}\in\mathrm{supp}(P_O)\\ f'\in\mathrm{supp}(P_F)}}\left\{\sup_{\substack{x^k\in A(\boldsymbol{o},f)\\ x'^k\in A(\boldsymbol{o},f')}}\sum_{i=1}^{N}\frac{\|x_i^k-x_i'^k\|_1}{\sigma(H_i)}\right\}\right)\\
>&\exp\left(-\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f)}{M^k}\right),
\end{aligned}$$

where the last inequality holds due to $\sigma(H_i)>1$ and the definition of Δ_{PML} in (4)*. Recalling the sensitives of Δ_{PML} in* (4) *and Δ_{DP} in* (1)*, we can obtain the following inequality:*

$$\begin{aligned}
&\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f)}{M^k}\\
=&\sum_{k=0}^{T}\frac{1}{M^k}\sup_{\substack{\boldsymbol{o}\in\mathrm{supp}(P_O)\\ f'\in\mathrm{supp}(P_F)}}\left\{\sup_{\substack{x^k\in A(\boldsymbol{o},f)\\ x'^k\in A(\boldsymbol{o},f')}}\sum_{i=1}^{N}\|x_i^k-x_i'^k\|_1\right\}\\
\leq&\sum_{k=0}^{T}\frac{1}{M^k}\sup_{\substack{\boldsymbol{o}\in\mathrm{supp}(P_O)\\ f'\in\mathrm{supp}(P_F)}}\left\{\sup_{\substack{x^k\in A(\boldsymbol{o},f)\\ x'^k\in A(\boldsymbol{o},f')}}\max_{i\in[N]}N\|x_i^k-x_i'^k\|_1\right\}\\
\leq&\sum_{k=0}^{T}N\frac{\Delta_{DP}(k)}{M^k}\\
\leq&N\epsilon_1.
\end{aligned}$$

Suppose there exists T and $f,f'\in\mathrm{supp}\,(P_F)$ such that $\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f)}{M^k}\neq\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f')}{M^k}$; Otherwise, f and f' are identical almost surely in $\boldsymbol{x}$. On this basis, we have

$$\begin{aligned}
&\int_{f\in\mathrm{supp}(P_F)}\inf_{\boldsymbol{o}\in\mathrm{supp}(P_O)}\inf_{f'\in\mathrm{supp}(P_F)}\frac{P_{O|F=f}(\boldsymbol{o})}{P_{O|F=f'}(\boldsymbol{o})}P_F(f)df\\
\geq&\int_{f\in\mathrm{supp}(P_F)}\exp\left(-\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f)}{M^k}\right)P_F(f)df\\
>&\exp(-N\epsilon_1).
\end{aligned}$$

Thus, there exists $\epsilon_2<\epsilon_1$ such that

$$\exp(-N\epsilon_2)\leq\int_{f\in\mathrm{supp}(P_F)}\exp\left(-\sum_{k=0}^{T}\frac{\Delta_{PML}(k,f)}{M^k}\right)P_F(f)df.$$

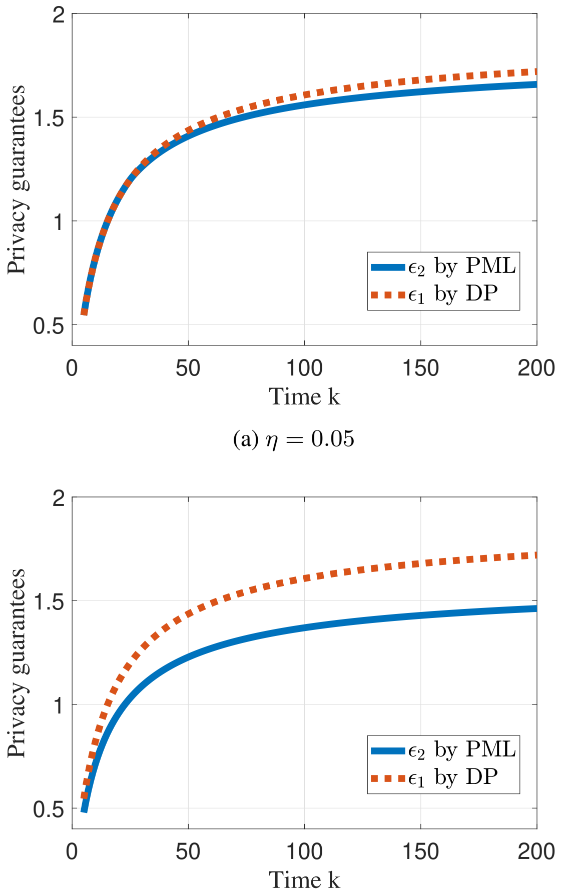
Recalling (3) in the Proof of Theorem 2, we have $\frac{1}{N}\sup_{\boldsymbol{o}\in\mathrm{supp}(P_O)}\ell(F\to\boldsymbol{o})\leq\epsilon_2$. Algorithm $\mathcal{A}_2$ is then ϵ_2-PML averaged over players with $\epsilon_2<\epsilon_1$, and the conclusion is proved. □

As shown in Theorem 4, compared to algorithm $\mathcal{A}_1$, our proposed algorithm $\mathcal{A}_2$ demonstrates better privacy protection. On one hand, from the perspective of DP, algorithm $\mathcal{A}_2$ maintains the same privacy guarantees as algorithm $\mathcal{A}_1$. This is because DP focuses on adjacent datasets and is insensitive to the distribution of correlated datasets, making it unable to distinguish variations in privacy leakage caused by different distributions within the algorithms. On the other hand, by integrating the prior knowledge of players' payoff function datasets into the noise, algorithm $\mathcal{A}_2$ achieves a tighter bound ϵ_2 of PML guarantees compared to ϵ_1 in the traditional algorithm $\mathcal{A}_1$.

At this point, by adopting PML as the measure to assess information leakage, we have successfully designed a privacy-preserving mechanism for computing NE in the non-cooperative game G. When players have correlated payoff datasets, algorithm $\mathcal{A}_2$ leads to better privacy protection with PML guarantees than the traditional algorithm $\mathcal{A}_1$ with DP guarantees. We have addressed Problem 1 so far.

5 Numerical Experiment

Consider a non-cooperative game model with N players exposed to a contagious disease [13]. Let α represent the probability of infection for player 1. Given the infected player 1, let η denote the correlation coefficient for the probability of other players being infected. With the availability of vaccines, each player must decide whether to vaccinate. The strategy of player i is represented by $x_i\in[0,1]$, which denotes the probability that this player vaccinates [22]. Each player's payoff function is defined by



(a) $\eta=0.05$

(b) $\eta=0.95$

Fig. 1: Privacy guarantee trends by PML and DP with different correlation coefficients: (a) $\eta=0.05$ indicates players' very low infection correlation; (b) $\eta=0.95$ reflects players' remarkably high infection correlation.

$f_i(x_i, x_{-i}) = x_i(-r_{va}) + (1 - x_i)(-r_{in}\delta(x))$, where r_{va} and r_{in} represent the morbidity risks for vaccination and infection, respectively, while $\delta(x)$ is the probability that an unvaccinated individual will be infected given the vaccination coverage status, that is, players' strategies x.

To compare with the existing privacy-preserving benchmarks [9, 11], we take the similar parameter settings, including $M^k = 1 + 0.2 * k^{0.8}$ in the Laplace distribution and $\lambda^k = \frac{0.1}{1+0.1k}$ in the stepsize, to investigate both the DP and PML guarantees in algorithm $\mathcal{A}_2$. An adversary aims to extract information related to the payoff function f, such as inquiring about morbidity risks or the number of infected individuals. Note that the privacy guarantees in both the DP and PML reflect the upper bounds of all possible information leakage. As can be seen from Fig. 1(a), when the infection correlation among players is very low, the upper bound of the privacy guarantee ϵ_2 of PML does not show a significant advantage over ϵ_1 of DP. However, in Fig. 1(b), when the infection correlation among players is remarkably high, the upper bound of the privacy guarantee ϵ_2 of PML is notably tighter than ϵ_1 of DP. These numerical illustrations support our main theoretical results in Theorem 4.

6 Concluding Remarks

This paper explored privacy-preserving NE computation in non-cooperative games with correlated payoffs. It is important to note that PML provides a wider information leakage assessment than DP, particularly for correlated datasets. By integrating prior knowledge into noise design, we proposed a novel mechanism ensuring seamless adaptation of existing convergence guarantees. Moreover, we realized a tighter bound with PML guarantees in the proposed mechanism, which refines the over-conservative assessment of information leakage risks with correlated payoff datasets.

To our knowledge, this is the first work to successfully employ PML in NE computation. Our work establishes a new connection between privacy protection and NE computation, which may pave the way for future advancements to enhance this research topic, for example, developing a necessary and sufficient bound for privacy guarantees, extending the mechanism to distributed networks, and refining noise design to accommodate specific models. These perspectives offer promising avenues for privacy-preserving mechanisms to prosper with multi-agent interactions. We are conducting ongoing work on extending our results to network game models and corresponding distributed algorithm design.

References

[1] J. Koshal, A. Nedić, and U. V. Shanbhag, "A gossip algorithm for aggregative games on graphs," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 4840–4845.

[2] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained deterministic mean field control," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2015.

[3] D. Gadjov and L. Pavel, "A passivity-based approach to Nash equilibrium seeking over networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1077–1092, 2018.

[4] G. Chen, Y. Ming, Y. Hong, and P. Yi, "Distributed algorithm for ε-generalized Nash equilibria with uncertain coupled constraints," *Automatica*, vol. 123, p. 109313, 2021.

[5] G. Chen, G. Xu, F. He, Y. Hong, L. Rutkowski, and D. Tao, "Approaching the global Nash equilibrium of non-convex multi-player games," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 10 797–10 813, 2024.

[6] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[7] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, "Differential privacy under continual observation," in *Proceedings of 42nd ACM Symposium on Theory of Computing*, 2010, pp. 715–724.

[8] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy, and mutual-information privacy," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5018–5029, 2016.

[9] M. Ye, G. Hu, L. Xie, and S. Xu, "Differentially private distributed Nash equilibrium seeking for aggregative games," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2451–2458, 2021.

[10] J. Wang, J.-F. Zhang, and X. He, "Differentially private distributed algorithms for stochastic aggregative games," *Automatica*, vol. 142, p. 110440, 2022.

[11] Y. Wang and T. Başar, "Ensuring both almost sure convergence and differential privacy in Nash equilibrium seeking on directed graphs," *IEEE Transactions on Automatic Control*, 2024.

[12] Y. Wang and A. Nedić, "Differentially-private distributed algorithms for aggregative games with guaranteed convergence," *IEEE Transactions on Automatic Control*, 2024.

[13] D. Kifer and A. Machanavajjhala, "No free lunch in data privacy," in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, 2011, pp. 193–204.

[14] S. Saeidian, T. J. Oechtering, and M. Skoglund, "Evaluating differential privacy on correlated datasets using pointwise maximal leakage," in *Annual Privacy Forum*. Springer, 2024, pp. 73–86.

[15] T. Zhu, P. Xiong, G. Li, and W. Zhou, "Correlated differential privacy: Hiding information in non-IID data set," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 229–242, 2014.

[16] I. Kayes and A. Iamnitchi, "Privacy and security in online social networks: A survey," *Online Social Networks and Media*, vol. 3, pp. 1–21, 2017.

[17] S. Saeidian, G. Cervia, T. J. Oechtering, and M. Skoglund, "Pointwise maximal leakage," *IEEE Transactions on Information Theory*, vol. 69, no. 12, pp. 8054–8080, 2023.

[18] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1625–1657, 2019.

[19] S. Saeidian, G. Cervia, T. J. Oechtering, and M. Skoglund, "Pointwise maximal leakage on general alphabets," in *2023 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2023, pp. 388–393.

[20] Z. Huang, S. Mitra, and N. Vaidya, "Differentially private distributed optimization," in *Proceedings of the 16th International Conference on Distributed Computing and Networking*, 2015, pp. 1–10.

[21] S. Saeidian, G. Cervia, T. J. Oechtering, and M. Skoglund, "Rethinking disclosure prevention with pointwise maximal leakage," *Journal of Privacy and Confidentiality*, vol. 15, no. 1, Mar. 2025.

[22] C. T. Bauch and D. J. Earn, "Vaccination and the theory of games," *Proceedings of the National Academy of Sciences*, vol. 101, no. 36, pp. 13 391–13 394, 2004.