IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 26 June 2024, accepted 19 July 2024, date of publication 25 July 2024, date of current version 5 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3433440

RESEARCH ARTICLE

Noise Variance Optimization in Differential Privacy: A Game-Theoretic Approach Through Per-Instance Differential Privacy

SEHYUN RYU, JONGGYU JANG, (Member, IEEE), AND HYUN JONG YANG, (Member, IEEE)
Department of Electrical Engineering, Pohang University of Science and Technology, Pohang, Gyeongsangbuk-do 37673, Republic of Korea

Corresponding author: Hyun Jong Yang (hyunyang@postech.ac.kr)

This research was supported in part by an IITP (Institute of Information & communications Technology Planning & Evaluation) grant funded by the Ministry of Science and ICT, Korea, as the project "6G MIMO System Research" (No. 2021-0-00161), in part by the Ministry of Oceans and Fisheries, Korea, as the research project "Development of automatic screening and hybrid detection system for hazardous material detecting in port container" (1525013872, No. 20200611), and in part by NRF (National Research Foundation of Korea) grant funded by the Ministry of Science and ICT, Korea (No. RS-2023-00250191).

ABSTRACT The concept of differential privacy (DP) can quantitatively measure privacy loss by observing the changes in the distribution caused by the inclusion of individuals in the target dataset. The DP, which is generally used as a constraint, has been prominent in safeguarding datasets in machine learning in industry giants like Apple and Google. A common methodology for guaranteeing DP is incorporating appropriate noise into query outputs, thereby establishing statistical defense systems against privacy attacks such as membership inference and linkage attacks. However, especially for small datasets, existing DP mechanisms occasionally add excessive amount of noise to query output, thereby discarding data utility. This is because the traditional DP computes privacy loss based on the worst-case scenario, i.e., statistical outliers. In this work, to tackle this challenge, we utilize per-instance DP (pDP) as a constraint, measuring privacy loss for each data instance and optimizing noise tailored to individual instances. In a nutshell, we propose a per-instance noise variance optimization (NVO) game, framed as a common interest sequential game, and show that the Nash equilibrium (NE) points of it inherently guarantee pDP for all data instances. Through extensive experiments, our proposed pDP algorithm demonstrated an average performance improvement of up to 99.53 % compared to the conventional DP algorithm in terms of KL divergence.

INDEX TERMS Differential privacy, game theory, per-instance differential privacy, security.

I. INTRODUCTION

Recently, the surge in machine learning has not only spotlighted the importance of statistical datasets but has also intensified the focus on privacy protections. Leading tech companies use the personal information we readily submit for training their machine learning models. If latent data is extracted from these models, our personal information could be exploited to facilitate financial crimes or terrorist activities. Meanwhile, the innovative concept of differential privacy (DP) has emerged as a key solution for quantitatively measuring privacy risks [1]. As depicted in Fig. 1, the DP, a method that balances the utility of data with individual privacy by injecting controlled randomness into datasets, ensures that the privacy of individual data points is preserved by making it *mathematically indiscernible* whether any specific individual data is included or excluded. In the figure, from a dataset $\mathcal{D}$, the data agent answers a query $Q(\mathcal{D})$. If there is no randomness in the query output, one potential attacker may access the private information via membership inference [2] or data linkage attacks [3]. In DP, the basic concept is forwarding randomized query output $Q(\mathcal{M}(\mathcal{D}))$ or $\mathcal{M}(Q(\mathcal{D}))$, where $\mathcal{M}$ is a randomized mechanism, thereby making the query output mathematically indiscernible.

The associate editor coordinating the review of this manuscript and approving it for publication was Sedat Akleylek.

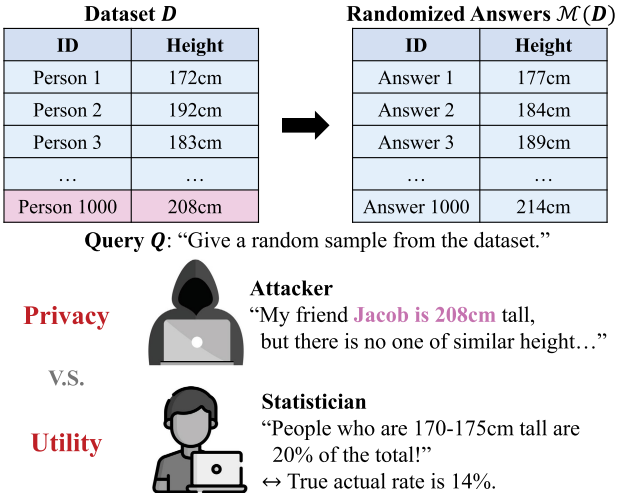

FIGURE 1. The essence of DP's data security lies in injecting noise into query outputs. The noise level has to be delicately balanced, considering the tradeoff between privacy and utility. From above, the attacker tries to identify if Jacob is in the dataset based on his known height of 208cm, but privacy is protected by noise, making it impossible. Meanwhile, the statistician struggles to analyze the feature of the dataset accurately due to noisy responses, resulting in utility loss.

In today's digital landscape, DP has been a promising technology as a strong safeguard against the looming threats of privacy violations, where the most common approach is simply adding noise to query outputs, i.e., *additive noise mechanism.* Specifically, compared to its rivals like homomorphic encryption and federated learning [4], [5], the additive noise DP mechanisms are computationally simple and generally applicable. By virtue of its simplicity and straightforwardness, DP has been widely employed by industry giants such as Apple [6] Google [7], and Microsoft [8] for data protection. Recently, many research endeavors have aimed to apply DP in distributed networks [9] and indoor localization scenarios [10], [11], adapting to evolving communication network environments.

A. LIMITATIONS OF DP

Although the DP mechanisms have been found in our daily lives, they sometimes add an excessive amount of noise to the query output, thereby making the datasets' utility almost statistically useless. To introduce this limitation, we bring the definition of the ϵ-DP, the basic concept of DP, where the ϵ is a privacy parameter.

Definition 1 (ϵ-DP): A randomized mechanism with domain $\mathbb{R}^d$, denoted by $\mathcal{M}$, has a range $\mathcal{R}(\mathcal{M})$. The mechanism $\mathcal{M}$ is ϵ-differential private, if for all dataset $\mathcal{D}, \mathcal{D}' \in \mathbb{R}^d$ such that $||\mathcal{D} - \mathcal{D}'||_1 \leq 1$:

$$\left| \ln \frac{\Pr[\mathcal{M}(\mathcal{D}) \in S]}{\Pr[\mathcal{M}(\mathcal{D}') \in S]} \right| \leq \epsilon, \ \forall S \subseteq \mathcal{R}(\mathcal{M}). \tag{1}$$

In Definition 1, the smaller ϵ, the stronger the privacy guarantees in DP. More importantly, the ϵ-DP is broadly defined across arbitrary datasets $\mathcal{D} \in \mathbb{R}^d$ and data instances. That is, the traditional DP focuses on designing mechanism $\mathcal{M}$ that satisfies the condition (1) for all statistical datasets $\mathcal{D}$, i.e., the mechanism $\mathcal{M}$ is designed for *the worst-case scenario* of all the statistical datasets. In general, the worst-case scenario indicates that $n-1$ data instances are the same, while one outlier has totally different information. Because this outlier can be placed anywhere, the conventional DP mechanisms, such as Logistic and Gaussian mechanisms [12], reasonably add *excessive* amount of *identical* noise to the query outputs.

B. PER-INSTANCE DP

Although adding *identical* noise in the conventional mechanisms is a reasonable choice in DP; however, if we only consider a specific dataset $\mathcal{D}$, we can further enhance the utility of the dataset by *adopting non-identical noise* while preserving privacy. The recently proposed concept, per-instance DP (pDP), gives us a motivation that the privacy loss of an instance in a fixed dataset can be measured.

Definition 2 (ϵ-pDP [13]): A randomized mechanism, denoted by $\mathcal{M}$, has a range $\mathcal{R}(\mathcal{M})$. For a fixed dataset $\mathcal{Z}$ and a fixed data instance $z \in \mathcal{Z}$, the mechanism $\mathcal{M}$ meets ϵ-pDP, if the following condition holds:

$$\left| \ln \frac{\Pr[\mathcal{M}(\mathcal{Z}) \in S]}{\Pr[\mathcal{M}(\mathcal{Z} \backslash \{z\}) \in S]} \right| \leq \epsilon, \ \forall S \subseteq \mathcal{R}(\mathcal{M}), \tag{2}$$

where red-colored contents are different from Definition 1. We note that the pDP is defined for fixed dataset $\mathcal{Z}$ and data instance z; thus, the DP in (1) is the supremum of pDP for dataset $\mathcal{Z}$ and instance z.

Although no specific solution to design noise distribution tightly guaranteeing pDP, the authors of [13] provide measuring different security levels for each data instance, defining privacy loss within the actual dataset and analyzing the variations. If a specific dataset is given, we no longer need to consider the worst-case for all arbitrary datasets. As depicted in Fig. 2, applying identical noise to all query outputs for DP satisfaction is undoubtedly not the optimal solution.

In this paper, we focus on designing appropriate noise distribution for each data instance for simultaneously satisfying pDP for all data instances. Intuitively, according to the definition of the pDP, the rarer the data instance, the harder it is to guarantee statistical indistinguishability. Let us consider the example depicted in Fig. 1. The person who is extremely taller than the others is easily specified; however, the person who is in the range (175 to 180 cm tall) is relatively hard to specify. In other words, the statistical outlier requires larger noise to guarantee indistinguishability. In the upcoming sections, we aim to answer the following question:

How can we optimize noise on a per-instance basis to satisfy pDP for a dataset while preserving the statistical features of the original data as much as possible?

In response to this question, our objective is to introduce a per-instance additive noise mechanism grounded in the principles of pDP. We propose solving the problem as a noise variance optimization (NVO) game and establish, through

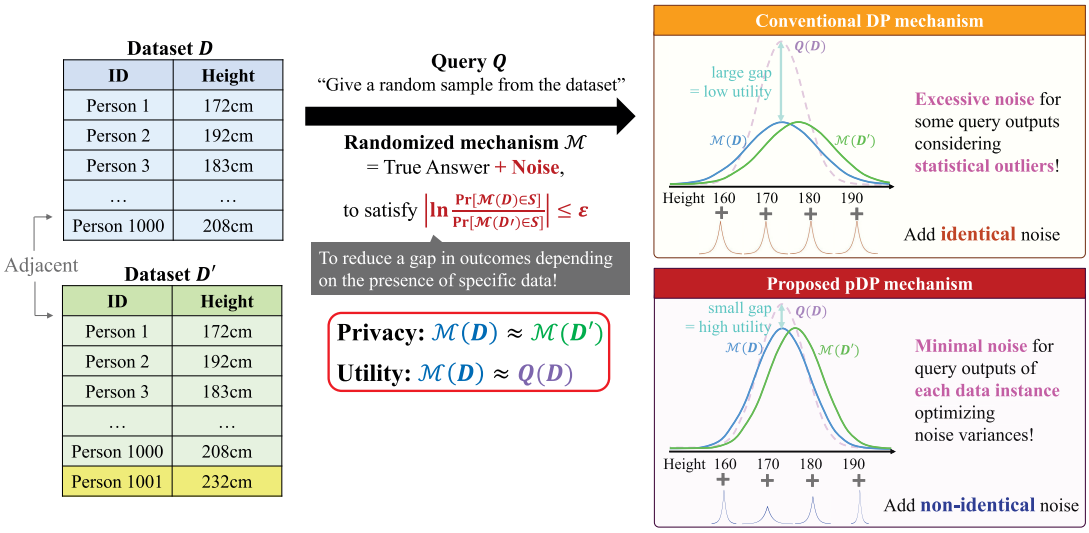

FIGURE 2. This figure illustrates the concept of DP guarantees. Conventional DP mechanisms sample noise values from the same probability distribution for all query outputs. Our proposed pDP mechanism aims to maintain the same level of privacy loss while mimicking the probability distribution of original query outputs by adding noise with varying variances to each instance's query output.

Theorem 1, that the Nash equilibrium (NE) of this game necessarily ensures pDP for all data instances.

C. CHALLENGES

Here, we introduce challenges on the horizon in optimizing the noise distribution to guarantee pDP. The main challenge is that ensuring pDP for a particular data instance is inherently dependent on the noise distribution of other data instances. Thus, altering the noise distribution for a data instance presents a tangible risk: certain instances might become non-compliant for pDP as a consequence of such modifications. Conventional additive noise mechanisms can guarantee mathematically well-proven assurance; however, when introducing non-identical noises, establishing these guarantees becomes more difficult. In summary, finding a balance between preserving the dataset's original statistical utility and ensuring ϵ-pDP requires intricate adjustments to the noise distribution, a challenge amplified by the curse of ***interdependency***.

D. CONTRIBUTIONS

We introduce an innovative approach to optimize non-identical noise distribution tailored to specific data instances. Our main contributions are three-fold:

- We propose the NVO game designed to find suitable non-identical per-instance additive Laplace noises within a dataset. Within this game, every player (representing data instances) collaboratively/sequentially acts to guarantee ϵ-pDP, all the while optimizing the utility of data statistics.
- We prove that an NE strategy in the NVO game ensures ϵ-pDP across all data instances.
- We simulate the best response dynamics (BRD) algorithm and the approximated enumeration (AE) algorithm as examples to obtain an NE strategy for the proposed NVO game. The proposed NVO game not only assures the same ϵ-pDP as the commonly adopted Laplace mechanism but also demonstrates superiority in preserving statistical utility.

II. RELATED WORKS

A. ADDITIVE MECHANISMS FOR DP

Traditional additive noise mechanisms offer straightforward and mathematically well-proven methods to ensure DP. Many attempts have been made to improve these traditional methods: [14], [15], [16] have proposed additive staircase-like noise as a substitute for the Laplace distribution, aiming to optimize a given statistical utility function while guaranteeing ϵ-DP. In addition, the IBM DP library demonstrates efforts to enhance additive noise mechanisms by constraining the randomized output within a predefined range [17]. In the context of optimizing the noise distribution, [18] has attempted to regulate additive noise to meet Rényi DP criteria. Similar concepts have been explored in parallel studies: sampling scenarios [19], [20] or deep learning [21], [22]. While there are efforts to alter noise distributions, prior studies have primarily employed uniform noise across all data. Such a method is not appropriate for guaranteeing tight pDP while preserving statistical utility through per-instance non-identical noise.

B. RELATION TO (ϵ, δ)-DP

The pioneer of DP [23], defined (ϵ, δ)-DP, demonstrating that the Gaussian mechanism can achieve this definition. Since the advent of DPSGD [24], there has been a surge

of applications in machine learning [25], [26], [27]. It is worth noting that our method can be easily adapted to the widely-known relaxation of DP, – namely (ϵ, δ)-pDP, with the per-instance Gaussian mechanism. Furthermore, when the dataset displays ample diversity among its values and there are constraints on accessing external data, (ϵ, δ)-pDP becomes equivalent to (ϵ, δ)-DP. Thus, our research on the pDP mechanism can be interpreted as striving to ensure greater effectiveness of DP in specific scenarios.

C. GAME THEORY

We introduce the NVO game, designed to simplify identifying the best variance for additive noises, inspired by the familiar game-theoretic approach introduced by [28]. Within the context of the NVO game, we show that the NE points of the game ensure ϵ-pDP. However, identifying this NE point brings its own set of complexities. To address this, we devise methods utilizing established algorithms to reach the NE point [29], [30].

III. PRELIMINARY

In this section, we introduce the preliminary concepts underpinning this paper. For brevity, we use scalar-form data instances in the remainder of this paper. We note that the concepts presented here can be seamlessly applied to vector-form data instances as well.

A. LAPLACE MECHANISM

In addition to the ϵ-DP and ϵ-pDP in Definitions 1 and 2, we begin with the definition of the Laplace mechanism, the most recognized method to guarantee ϵ-DP.

Definition 3 (Laplace Mechanism): Given any query function $f : \mathcal{X} \to \mathbb{R}$ with ℓ_1 sensitivity of $\Delta f \in \mathbb{R}$, the Laplace mechanism is defined as:

$$\mathcal{M}_{\mathrm{L}}(x, f(\cdot), \epsilon) = f(x) + y, \tag{3}$$

where y is a random number drawn from $\mathrm{Lap}(\Delta f/\epsilon)$ and $\mathcal{X}$ denotes the domain of variable x.

The Laplace mechanism ensures ϵ-DP by considering the worst-case scenario and adding identical noise to all data instances. However, we see an opportunity to enhance the preservation of statistical features in the data by employing a customized per-instance noise variance within the Laplace distribution.

B. RANDOM SAMPLING QUERY

In this study, we focus on the random sampling query, as detailed subsequently.

Definition 4 (Random sampling query): Given numeric dataset $\mathcal{D}$, the output of the random sampling query q is a random variable following the probability distribution of a dataset, *i.e.*,

$$q(\mathcal{D}) \sim \Pr(\mathcal{D}). \tag{4}$$

This query captures the statistical characteristics of the dataset by directly retrieving an instance from it.

Remark 1 (Extensibility of Random Sampling Query): The random sampling query is a fundamental query that encompasses all possible statistical queries. This is because the random sampling query can capture the statistical distribution of a dataset. Thus, from the post-processing theorem, achieving pDP/DP for random sampling queries can guarantee pDP/DP for all statistical queries, even those utilized in machine learning applications.

C. NASH EQUILIBRIUM

The Nash equilibrium, a foundational concept in a game theory introduced by [31], denotes the ideal state of a game where every player makes their optimal decision based on the choices of their counterparts as below.

Definition 5 (Nash Equilibrium): The Nash equilibrium is a profile of strategies (s_i^*, s_{-i}^*), such that each player's strategy is an optimal response to the other players' strategies: $\Pi_i(s_i^*, s_{-i}^*) \geq \Pi_i(s_i, s_{-i}^*),\ \forall i$ where s_{-i} is the strategy profile of all players except for player i and $\Pi_i(s)$ is a payoff function.

To minimize complexity and structure the problem in a well-known format, our approach frames the challenge as a game to ensure pDP within a dataset and maximize the preservation of statistical features, ultimately reaching an NE point.

IV. NOISE VARIANCE OPTIMIZATION GAME

In this section, our focus is to design a sequential/cooperative game that applies per-instance Laplace noises to the target dataset, ensuring ϵ-pDP. We denote the target dataset by $\mathcal{D}$, and its data instances are represented by $d \in \mathbb{R}$. We assume that the target dataset consists of real-valued data instances, *e.g.*, a regression dataset. The problem we aim to solve using game theory can be defined as a constrained optimization problem:

$$\begin{aligned} &\min_{\mathcal{M}} \ U(\mathcal{D}, \mathcal{M}(\mathcal{D})) \\ &\text{s.t.} \ \left| \ln \frac{\Pr[\mathcal{M}(\mathcal{D}) \in S]}{\Pr[\mathcal{M}(\mathcal{D} \setminus \{d\}) \in S]} \right| \leq \epsilon, \forall d \in \mathcal{D}, \forall S \in \mathcal{R}(\mathcal{M}), \end{aligned} \tag{5}$$

where the function U represents an arbitrary utility function, quantifying the statistical difference between the original dataset $\mathcal{D}$ and the randomized dataset $\mathcal{M}(\mathcal{D})$. In our work, *KL divergence* is used as our utility function. Solving the problem as a conventional optimization problem becomes challenging due to the non-differentiable nature of the constraint function, making it difficult to apply gradient-based methods, and the presence of intricate interdependencies. To address this, we simplify the problem as selecting a variance value among possible options and elucidate it using familiar game theory principles.

An illustration of the proposed game design, including the preprocessing step, is depicted in Fig. 3. Detailed explanations corresponding to this figure will be covered in the subsequent portions of this section.

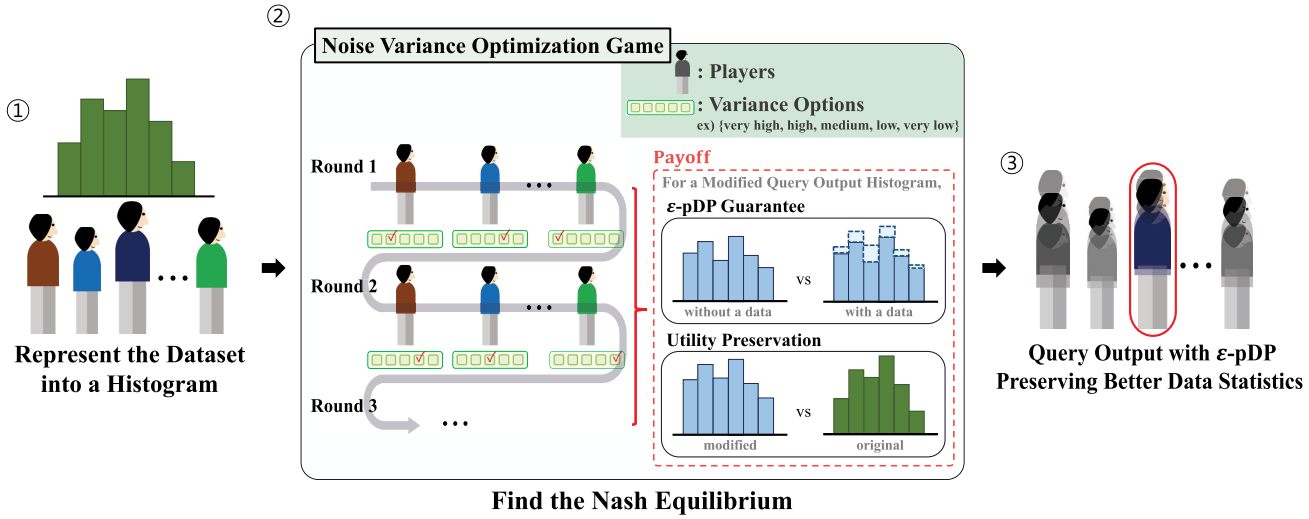

FIGURE 3. The process involves determining the optimal combination of noise variances for the query outputs of individual data instances. In Step 1, we depict the dataset as a histogram through normalization and categorization. In Step 2, we aim to find an NE point for the NVO game, where multiple players iteratively update their variance parameters to ensure ϵ-pDP while preserving statistical utility. Once the prime set of noise variances is determined, Step 3 allows us to formulate queries that assure ϵ-pDP by executing a random sampling query.

A. PREPROCESSING: A HISTOGRAM REPRESENTATION OF THE DATASET

In the context of the NVO game, assessing the probability density function of the mechanism's output for every single point across all data instances is computationally challenging. Additionally, as highlighted by [24], the most precise method for assessing privacy loss involves manual integration within specified intervals, rather than relying solely on theoretical boundaries.

1) NORMALIZATION

For the manual integration, we normalize the dataset into the [0, 1] range by min-max normalization. We conservatively opt to set our integration's target range to encompass p-percentile of the Laplace mechanism with the target ϵ, which is defined by $(-(\Delta q/\epsilon)\ln(2-2p), (\Delta q/\epsilon)\ln(2-2p))$ if $p > 0.5$,[1] where Δq denotes the sensitivity of random sampling query q. For brevity, we denote $(\Delta q/\epsilon)\ln(2-2p)$ as $\Delta_{(\epsilon,p)}$. Then, the min-max linear normalization is executed in the interval $[d_{\min} - \Delta_{(\epsilon,p)}, d_{\max} + \Delta_{(\epsilon,p)}]$, where $d_{\min} = \min_i d_i$ and $d_{\max} = \max_i d_i$.

2) CATEGORIZATION

After normalization, the continuous nature of the domain S poses a unique challenge, differing from that in ϵ-DP. It is necessary to verify the ϵ-pDP condition for each instance and each point within set S, a task impractical to achieve in polynomial time. To address this issue, we make K non-overlapping intervals into the range of the dataset and allocate each data instance $d \in \mathcal{D}$ into categories based on their corresponding intervals. We set the K uniformly divided non-overlapping intervals as follows: the i-th interval is $(\frac{i}{K}, \frac{i+1}{K})$, where the representative value of each interval bin is the midpoint of the interval. The set of the representative values is denoted as $\mathcal{K}$. Through data categorization, the dataset can be represented in the form of a histogram. Segmenting data instances into distinct bins offers an advantage, as it allows certain data instances to inherently ensure a non-infinite ϵ, while the original continuous dataset cannot.[2]

B. DEFINITIONS OF PLAYERS, STRATEGIES, AND PAYOFFS

Here, we introduce the NVO game following data preprocessing. In this game, every data instance possesses data values within the range [0, 1]. The proposed NVO game is classified as: i) *sequential game*, ii) *fully-cooperative game*, and iii) *potential game*.

1) PLAYERS

In this study, each data instance, acting as a player, collaboratively and sequentially participates in the NVO game. The goal of them is to establish ϵ-pDP (their respective payoffs) by designing their strategies (variance optimization), simultaneously maximizing the preservation of original statistical features. We note that the player of the NVO game is represented by $I = \{1, 2, \ldots, |\mathcal{D}|\}$.

2) STRATEGIES

With data instances cast as players, the strategy for player i is defined by the additive noise applied to data d_i. Denoting the variance of additive noises to the i-th data instance d_i as b_i, the action of the player i is written by b_i. That is, from the random sampling query in Definition 4, the per-instance

[1] In our experiments, the value of p is set to be 0.9.

[2] Once categorized, if a category contains three data points, the inherent ϵ-pDP assurance for these instances is given by $\log 3/2 \approx 0.4$.

Laplace mechanism is expressed as

$$\mathcal{M}(d_i) = d_i + y_i, \tag{6}$$

where y_i is a random variable drawn form Lap(b_i). Typically, games with strategy sets of uncountable infinity might not always have an NE solution. As a result, we confine these strategy sets to a discrete domain. In other words, the added variance is chosen from a discrete set $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$.

3) PAYOFFS

In the context of the NVO game, the payoff should induce the player to primarily uphold ϵ-pDP and secondarily preserve the dataset's statistical features. In this domain, there is a trade-off between data statistics and privacy. Emphasizing robust privacy can reduce query output quality, while maximizing the preservation of statistical features might compromise privacy. The goal is to optimize the preservation of statistical features without violating the ϵ-pDP constraint, achievable by carefully adjusting noise variance on query results. With this in mind, we express the overall payoff $P(\mathcal{M}, \mathcal{D})$ as a composite of two objectives: privacy assurance $P_{\mathrm{E}}(\mathcal{M}, \mathcal{D})$, and minimizing utility that measures the statistical difference between the original dataset and the randomized dataset $P_{\mathrm{U}}(\mathcal{M}, \mathcal{D})$. We note that the players in the proposed NVO game cooperate to benefit from the shared payoffs. Simply, the NVO game is characterized as a kind of common interest game.

4) PRIVACY ASSURANCE PAYOFF

The payoff related to privacy assurance, denoted as P_{E}, functions as an indicator of how effectively ϵ-pDP is met for a dataset, assessed from the perspective of pDP. Let us define $p_{\epsilon,i}$ as an indicator for representing whether the i-th data instance's pDP is satisfied or not, *i.e.,*

$$p_{\epsilon,i}(\mathcal{M}, \mathcal{D}) = \begin{cases} 1, & \text{if } d \in \mathcal{D} \text{ satisfies the } \epsilon\text{-pDP condition.} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Then, the privacy assurance payoff P_{E} is defined as the number of data instances satisfying the ϵ-pDP condition as

$$P_{\mathrm{E}}(\mathcal{M}, \mathcal{D}) = \sum_{i=1}^{|\mathcal{D}|} p_{\epsilon,i}(\mathcal{M}, \mathcal{D}). \tag{8}$$

After establishing the privacy assurance payoff, two subsequent questions arise: one regarding the preservation of statistical features, and the other addressing the assurance of privacy at the NE point.

- ***Q1:*** *How can we define the payoff related to the preservation of statistical features?*
- ***Q2:*** *Does the NVO game truly ensure ϵ-pDP for all data instances using the privacy assurance payoff of (8)?*

C. UTILITY PRESERVATION PAYOFF (ANSWER TO Q1)

In response to *Q1*, we formulate the utility preservation payoff $P_{\mathrm{U}}(\mathcal{M}, \mathcal{D})$ to measure the statistical difference between the original dataset $\mathcal{D}$ and the randomized dataset $\mathcal{M}(\mathcal{D})$. Here, a higher value indicates a smaller difference. Furthermore, we ensure the utility preservation payoff does not compromise the assurance of ϵ-pDP by scaling the targeted utility function U into the range [0, 1]. It is important to note that various statistical features can be incorporated into the utility function, encompassing metrics like the difference of n-th order momentum, Kullback-Leibler (KL) divergence, and Jensen-Shannon (JS) divergence, among others.

In this paper, for example, $U(q(\mathcal{D})||q(\mathcal{M}(\mathcal{D}))) = D_{\mathrm{KL}}(q(\mathcal{D})||q(\mathcal{M}(\mathcal{D})))$, i.e., we set the utility function by using KL divergence as in the following remark.

Remark 2 (Examples of the Utility Function): For a dataset $\mathcal{D}$, the output probability distribution $q(\mathcal{D})$ of a query q, and a randomized function $\mathcal{M}$, the utility preservation payoff is defined as

$$P_{\mathrm{U}}(\mathcal{M}, \mathcal{D}) = 1 - \frac{D_{\mathrm{KL}}(q(\mathcal{D})||q(\mathcal{M}(\mathcal{D})))}{\log(K)} \in [0, 1], \tag{9}$$

where the utility function $D_{\mathrm{KL}}(q(\mathcal{D})||q(\mathcal{M}(\mathcal{D})))$ is bounded in $[0, \log(K)]$. The minus sign is used since the KL-divergence is a measure of information-theoretic distance between two probability distributions, where a smaller value indicates greater similarity between the distributions. The bound can be obtained by the fact that $\log(K) \geq \log \frac{q(\mathcal{D})}{q(\mathcal{M}(\mathcal{D}))}$.

D. GUARANTEE OF THE ϵ-PDP (ANSWER TO Q2)

Regarding the earlier question, we present proof illustrating that the NE strategy in the proposed NVO game consistently guarantees ϵ-pDP for a dataset.

Theorem 1: Let us define the minimum variance in the set of possible action $\mathcal{V}$ as $b_{\min} \neq 0$. Then, ϵ-pDP for all data instances upholds if the following condition is satisfied:

$$b_{\min} \geq \frac{1}{\log\left(1 + (|\mathcal{D}| - 1)(\exp(\epsilon) - 1)\right)}. \tag{10}$$

In Theorem 1, we show that an NE point for the NVO game guarantees the ϵ-pDP for all $d \in \mathcal{D}$ if the condition in (10) holds. In the theorem, there always exists a value $b_{\min}$ that makes the NE point of the NVO game ensure the ϵ-pDP for all $\epsilon \geq 0$.

Remark 3 (Intuition of Theorem 1): In Equation 10, if there are more data instances in the dataset, the influence of the individual data point diminishes, thereby allowing us to guarantee ϵ-pDP easily. That is, if the value of $|\mathcal{D}|$ increases, we can guarantee ϵ-pDP with smaller variance noise. On the other hand, if ϵ decreases to zero, query output with and without a data point should be statistically the same. Thus, the variance of the added noise becomes infinite, resembling a uniform query output distribution.

V. ALGORITHM FINDING THE NASH EQUILIBRIUM OF THE NVO GAME

In this section, we delve into an algorithm designed to secure an NE strategy within the framework of the NVO game. We begin by showcasing the BRD algorithm, adapted specifically for this game.

Algorithm 1 BRD for NVO game

Input dataset $\mathcal{D} = \{d_i | i = 1, \ldots, m\}$, variance space $\mathcal{V} = \{v_i | i = 1, \ldots, n\}$, target epsilon value ϵ

$i \leftarrow 1$ and $p^* \leftarrow 0$
▷ Initialize data index and the best payoff
V$[l] \leftarrow$ randomly initiates $v \in \mathcal{V}$, for $l = 1, \ldots, |\mathcal{D}|$
▷ Initialize the best variance set
while p^* converges over the dataset **do**
  PAYOFF$[l] \leftarrow 0$, for $l = 1, \ldots, |\mathcal{V}|$
  for $j \leftarrow 1$ to $|\mathcal{V}|$ **do**
    V_temp $\leftarrow$ V
    V_temp$[i] \leftarrow v_j$
    PAYOFF$[j] \leftarrow$ `GET_PAYOFF`$(\mathcal{D}$, V_temp, $\epsilon)$
    ▷ Explore and store payoffs for all variance options
  end for
end while
$p^* \leftarrow$ max PAYOFF and $j^* \leftarrow$ argmax PAYOFF
V$[i] \leftarrow v_{j^*}$
▷ Get the best variance for a current instance
$i \leftarrow (i + 1)$ mod $|\mathcal{D}|$
return V ▷ Nash equilibrium

*`GET_PAYOFF()` is a function of the proposed payoff by summing up Equations (8) and (9).

A. BRD ALGORITHM

The BRD algorithm is a concept in game theory where players, taking into account the current strategies of their opponents, opt for their most favorable response. During this iterative process, players sequentially decide on their best actions, which is presented in Alg. 1. The choice of values within the variance space can be tailored to encompass all the possible noise variance values. As the cardinality of the variance space $\mathcal{V}$ expands, the algorithm's performance improves, but there is a significant increase in computational complexity. Therefore, it is crucial to define the variance set $\mathcal{V}$ considering the trade-off between computational complexity and utility.

B. COMMON INTEREST GAME AND POTENTIAL GAME

In a common interest game, participants share a unified payoff. A player's strategy change directly impacts both the potential function and their own payoff, classifying it inherently as a potential game. Essentially, every data point acts as a cooperative player aiming for a joint goal. By crafting a shared payoff to maximize and iteratively selecting the optimal noise for each data instance's output, achieving an NE is feasible.

C. CONVERGENCE OF BRD TOWARD NE

As shown by [32], the BRD algorithm always converges into an NE point, if the target game belongs to one of the following games: potential games, weakly acyclic games, aggregative games, and quasi-acyclic games. As noted above, the NVO game is a potential game; thus, the BRD algorithm can obtain an NE point of the NVO game.

From the proof of Theorem 4.1 in the Appendix A, when every variance option exceeds $b_{\min}$, there always exists a choice that consistently increases one pDP assurance at each round. Hence, guaranteeing ϵ-pDP for all data instances is feasible after $|\mathcal{D}|$ rounds. Intuitively, as players choose their optimal responses, either sequentially or simultaneously, the value of the potential function increases, reaching its maximum at some point. The strategy at this peak is the game's NE.

VI. EXPERIMENTS

In this section, we evaluate the NE strategy of the NVO game. Our primary focus is to observe if the proposed NVO game achieves superior preservation of the dataset's statistical features compared to the conventional Laplace mechanism while maintaining the same level of ϵ-pDP. To assess this, we conduct simulations on three publicly available datasets: 1) NBA player dataset [33], 2) personal income dataset [34], and 3) credit profile dataset [35].

Experimental Detail: In our experiments, we configure the target ϵ values in $\{1, 2, 4, 8\}$. Additionally, we experimented on the NBA player's height dataset with extremely low target epsilon values of $\{0.1, 0.3\}$. After normalization and discretization, the height and weight values belong to 101 categories, *i.e.,* $K = 101$. For the action of the players, variance set $\mathcal{V}$ is defined by $\{3 \times \Delta q/\epsilon, 2 \times \Delta q/\epsilon, \Delta q/\epsilon, 0.33 \times \Delta q/\epsilon, 0.2 \times \Delta q/\epsilon\}$. From Theorem 1, ϵ-pDP for the smallest ϵ is achievable with the configured variance set $\mathcal{V}$, since $b_{\min} \approx 0.129$. That is, the NE points in the proposed NVO game ensure ϵ-pDP. For comparison, we additionally implemented an approximated enumeration (AE) algorithm based on the genetic algorithm, with excessive generations. For more details, please refer to Appendix B.

Metrics of Data Statistics: In the experiments, we use the following metrics related to data statistics:

- **KL divergence**: We measure the KL divergence between the probability distribution of the original dataset and the randomized dataset. The lower KL divergence indicates better preservation of the information of the original dataset.
- **L1 loss of standard deviation (SD)**: This metric measures the ℓ_1 error between the standard deviation of the original dataset and the randomized dataset.
- **Jaccard index**: The Jaccard index is calculated by representing values in a probability distribution exceeding a certain threshold as sets and then computing the intersection over the union (IoU) of the two sets. This measure quantifies the similarity between two probability distributions, where a value closer to 0 indicates

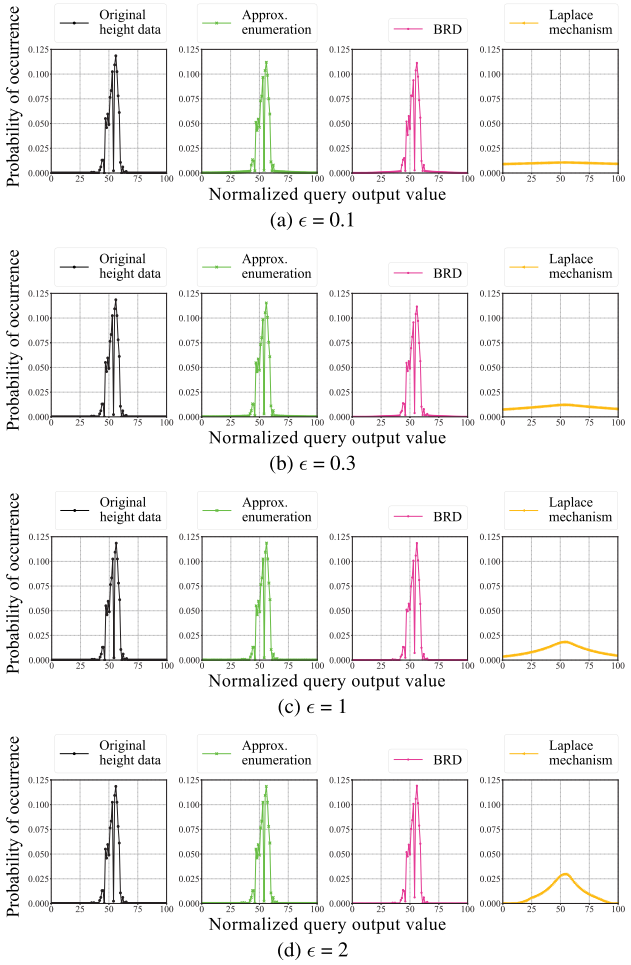

FIGURE 4. Comparison of query output probability distributions for the height data with each algorithm, when $\epsilon = 0.1$, 0.3, 1, and 2. The x-axis represents the index of the categorization bins from 0 to 100 ($K = 101$).

dissimilarity, while a value closer to 1 signifies similarity between the distributions. We set the threshold to 0.001 to examine the probability distribution of query output during experiments and select significant values.

- **Cosine similarity** [36]: The probability mass function can be viewed as a vector with probability values. We leverage the cosine similarity to measure the similarity between two probability distributions represented as vectors.

A. EXPERIMENTAL RESULT 1: HEIGHT DATA

1) DATASET

In the NBA players dataset, we employ 1,307 data instances with the tuple of (height and weight) for players who joined five teams from 1963 to 2021: *Atlanta Hawks*, *Boston Celtics*, *Charlotte Hornets*, *Chicago Bulls*, and *Cleveland Cavaliers*.

2) ANALYSIS 1: PRESERVATION OF STATISTICAL FEATURES

Here, we first analyze the preservation of statistical features after executing randomized mechanisms for the height feature

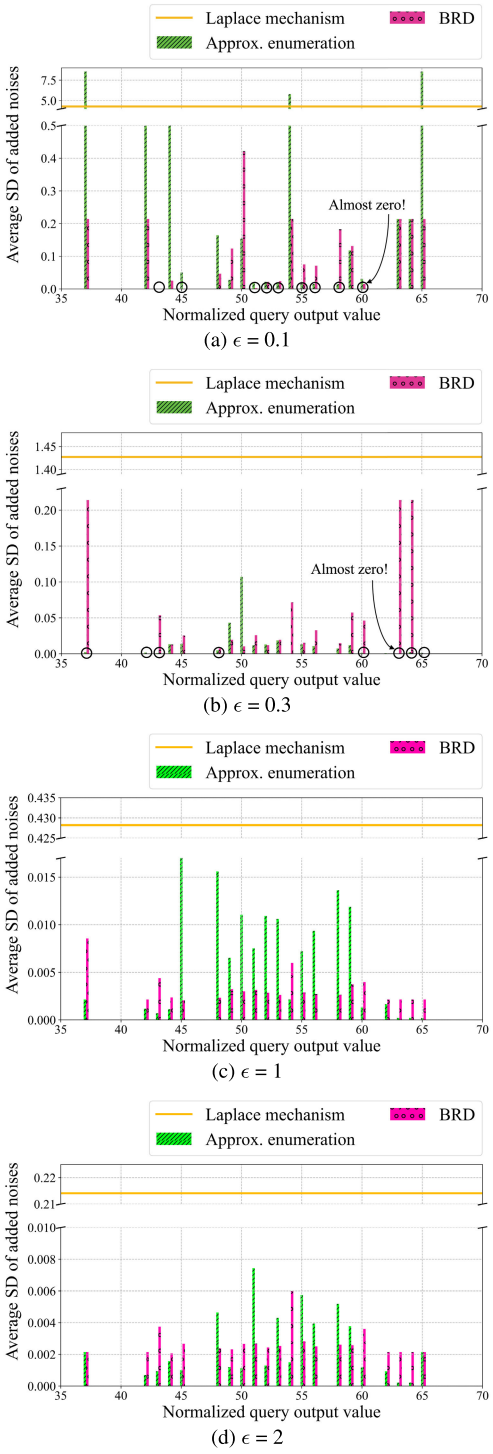

FIGURE 5. Distributions of average noise standard deviation for the height dataset for $\epsilon = 0.1$, 0.3, 1, and 2.

of the NBA player dataset. In Fig. 4, we compare the probability distribution of the randomized mechanisms' output.

As shown in the figure, the proposed NVO game (BRD and AE) has more similar shapes of distribution to the original one than the conventional Laplace mechanism, by executing

TABLE 1. Each algorithm's average computation time, KL divergence, L1 loss of SD, Jaccard index with a threshold 0.001, and cosine similarity are evaluated for the height data, for $\epsilon = 0.1, 0.3, 1, 2, 4$, and 8. The modified query output distributions for all algorithms satisfy ϵ-pDP.

Algorithm	ϵ	Comp. time (minutes) ↓	KL divergence ↓	L1 SD loss ↓	Jaccard index (threshold=0.001) ↑	Cosine similarity ↑
BRD	0.1	**8**	0.0512	0.0126	0.4878	0.9989
	0.3	**5**	0.0393	0.0122	0.5882	0.9994
	1.0	**4**	**0.0066**	**0.0049**	**0.9523**	0.9992
	2.0	**5**	**0.0045**	0.0084	0.9523	0.9997
	4.0	**5**	**0.0005**	0.0016	**1.0000**	**0.9999**
	8.0	**5**	0.0006	0.0017	**1.0000**	**0.9999**
Approx. enum.	0.1	287	**0.0475**	0.0123	0.5000	0.9995
	0.3	294	**0.0248**	**0.0108**	**0.8261**	**0.9998**
	1.0	392	0.0176	0.0058	**0.9523**	**0.9998**
	2.0	354	0.0047	**0.0080**	**1.0000**	**0.9999**
	4.0	182	0.0006	**0.0015**	**1.0000**	**0.9999**
	8.0	155	**0.0001**	**0.0007**	**1.0000**	**0.9999**
Laplace mechanism (baseline)	0.1	-	0.0475	0.0279	0.1980	0.3732
	0.3		0.0475	0.0279	0.1980	0.3732
	1.0		1.3991	0.0261	0.1980	0.5656
	2.0		0.9480	0.0247	0.2631	0.7040
	4.0		0.5064	0.0216	0.4444	0.8299
	8.0		0.2401	0.0170	0.6451	0.9074

*Best: **bold**, second-best: underline.

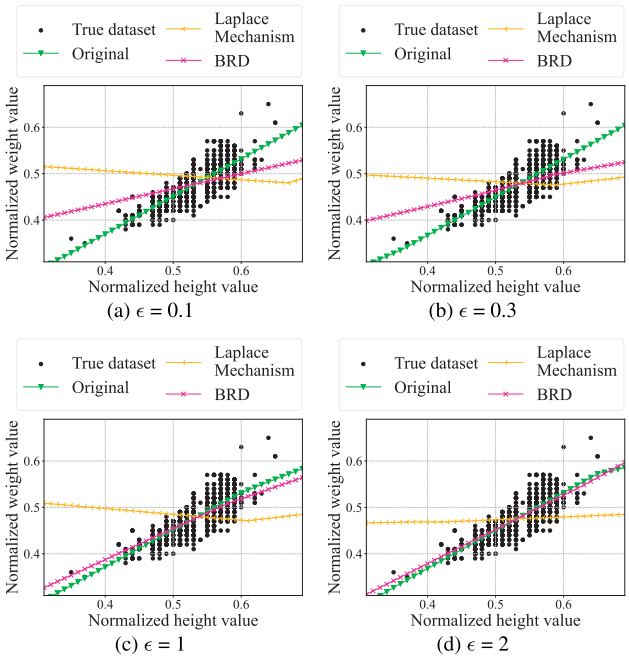

FIGURE 6. Linear regression result of 500 sampled data for each algorithm for $\epsilon = 0.1, 0.3, 1$, and 2.

per-instance noise optimization. We can observe that the probability distribution of the Laplace mechanism is getting similar to the original query output distribution as ϵ increases; however, the proposed NVO game better preserves the shape than with only using $\epsilon = 0.1$ than the Laplace mechanism of $\epsilon = 8$.

In Table. 1, we quantitatively measure various statistical utility functions: 1) KL divergence, 2) L1 loss of standard deviation (SD), 3) Jaccard index, and 4) cosine similarity of the distribution. In the table, the NVO game-related algorithms have superior statistical utility than the Laplace mechanism at 99.53%. Despite its superior performance, the AE algorithm requires a much longer computation time than the BRD algorithm.

In Fig. 5, the SD of the added noise in each categorization bin is depicted. As we can observe, compared to the conventional Laplace mechanism, the BRD and AE algorithms add relatively small variance noises, thereby achieving superior

TABLE 2. The average RMSE loss for regression task for the entire dataset, where the samples were generated using the noise associated with the pDP/DP algorithms.

Algorithm	Average RMSE					
	$\epsilon = 0.1$	$\epsilon = 0.3$	$\epsilon = 1$	$\epsilon = 2$	$\epsilon = 4$	$\epsilon = 8$
Original data (reference)	0.0218					
NVO game (BRD)	**0.0278**	**0.0273**	**0.0227**	**0.0221**	**0.0219**	**0.0218**
Laplace mechanism	0.0516	0.0449	0.0444	0.0380	0.0335	0.0272

*Best: **bold**.

preservation of statistical features. Furthermore, we note that as the guaranteed ϵ value decreases, significant noise is introduced into the query outputs of rare data, specifically statistical outliers.

3) ANALYSIS 2: REGRESSION TASK

In order to evaluate the practical usefulness of the randomized mechanisms, we conduct a simulation of a regression task to estimate the weight feature from the height feature of the NBA player dataset. For the regression task, we configure a multi-layer neural network, which consists of three layers with ten parameters activated by the Rectified Linear Unit (ReLU) function. There are three different neural networks trained with 1) the original dataset, 2) a randomized dataset with the NVO game, and 3) a randomized dataset with the conventional Laplace mechanism. In Fig. 6, we show the scatter diagram of the preprocessed original dataset, and the height-weight regression curve of the datasets. Compared to the conventional Laplace mechanism, the regression curve of the NVO game is more similar to that of the original data.

Quantitatively, as shown in Table 2, even with the case of low epsilon, 1-DP, the average RMSE of the BRD algorithm is apart from only 8.6% from that of the original dataset. For 8-DP, the average RMSE for the prediction of the BRD algorithm and original regression are almost the same.

B. EXPERIMENTAL RESULT 2: INCOME DATA

Dataset: We utilize the test dataset of the personal income dataset, crafted by UC Irvine. The number of data in the test dataset is 899. Similar to the NBA player dataset, the income values belong to one of the 101 categorization bins. We note that the single feature analysis is conducted here because there is no continuous feature in the dataset except income.

The qualitative results can be found in Table 3. The AE algorithm can preserve nearly equivalent data statistics but at a computational cost around 100-150 times higher, involving approximately 270 generations. The BRD algorithm achieves similar performance much more efficiently. The BRD algorithm achieved up to a 99.71 % improvement in KL utility than the Laplace mechanism, while guaranteeing 4-DP for every element, on the income dataset.

For the comparison of the per-instance noise variance, we depict the average standard deviation of the added noise in Fig. 7. As similar to the results in the NBA player dataset, the proposed NVO game allocates different amounts of noise

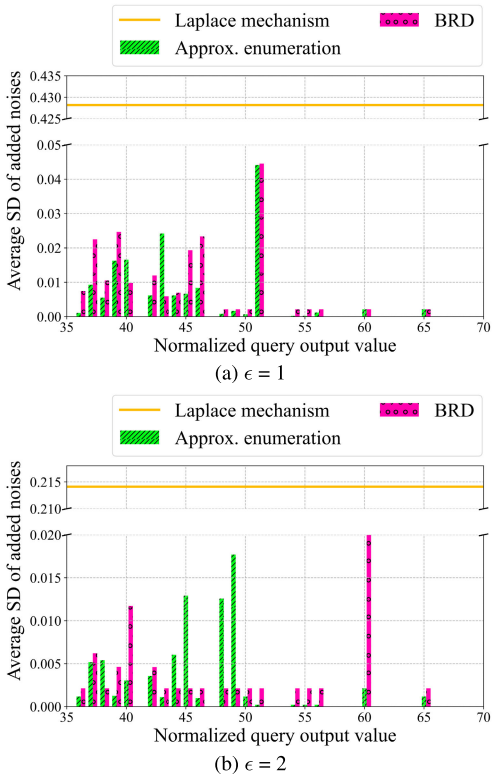

FIGURE 7. Distributions of average noise standard deviation for the income dataset for $\epsilon = 1$ and 2.

by considering its probability mass, thereby having better statistical utility.

From these results, we confirm that the proposed NVO game concisely outperforms the conventional Laplace mechanism.

VII. CONCLUSION

In conclusion, our research optimizes noise on a per-instance basis to achieve ϵ-pDP using a Laplace distribution, enhancing statistical utility over traditional methods. This approach is framed as an NVO game, proving that the Nash equilibrium point assures ϵ-pDP for all instances. Our experiments validate that this method significantly outperforms conventional Laplace mechanisms across various utility metrics. This framework can be universally applied to all statistical queries under differential privacy, as demonstrated by its extension to random sampling queries. However, the method has limitations, such as reliance on the Laplace distribution and computational constraints due to large datasets. The choice of noise variances is confined to discrete options, even though we demonstrate convergence to ϵ-pDP. Future research should focus on optimizing noise variances within discrete spaces, exploring alternative noise distributions, and extending the applicability of NVO games to various domains. Additionally, to apply our proposed mechanism to high-dimensional data used in machine learning, research to reduce computational complexity is essential.

TABLE 3. Each algorithm's average computation time, KL divergence, L1 loss of standard deviation, Jaccard index with a threshold of 0.015, and cosine similarity are evaluated for the income data. The modified query output distributions for all algorithms satisfy ϵ-pDP.

Algorithm	ϵ	Comp. time (minutes) ↓	KL divergence ↓	L1 SD loss ↓	Jaccard index (threshold=0.001) ↑	Cosine similarity ↑
BRD	1.0	**1**	0.0223	0.0068	**0.9565**	0.9997
	2.0	**2**	0.0087	**0.0003**	0.8695	**0.9999**
	4.0	**2**	0.0014	**0.0003**	**1.0000**	**0.9999**
	8.0	**1**	0.0013	0.0007	**1.0000**	0.9999
Approx. enum.	1.0	148	**0.0202**	**0.0061**	**0.9565**	**0.9998**
	2.0	204	**0.0061**	0.0026	**1.0000**	**0.9999**
	4.0	178	**0.0010**	0.0017	**1.0000**	**0.9999**
	8.0	164	**0.0000**	**0.0000**	**1.0000**	**1.0000**
Laplace mechanism (baseline)	1.0	-	1.3952	0.0274	0.2277	0.5529
	2.0		0.9373	0.0261	0.3066	0.6959
	4.0		0.4758	0.0230	0.4782	0.8369
	8.0		0.1736	0.0178	0.6363	0.9378

*Best: **bold**, second-best: underline.

APPENDIX A PROOF OF THEOREM 1

Theorem 1. *Let us define the minimum variance in the set of possible action $\mathcal{B}$ as $b_{\min} \neq 0$. Then, ϵ-pDP upholds if the following condition is satisfied:*

$$b_{\min} \geq \frac{1}{\log\left(1 + (|\mathcal{D}| - 1)(\exp(\epsilon) - 1)/K\right)}. \tag{11}$$

For simplicity of the proof, we tackle the situation for the scalar dataset case of the NVO game, *i.e.,* $d = 1$. We note that the proof can be extended to the vector version by considering each element separately.

A. NOTATIONS

In this proof, we use the following notations:

- b_i: Action of the player i, the variance of noise added to data d_i.
- (b_i, b_{-i}): Set of each players' strategy.
- $b^* = (b_i^*, b_{-i}^*)$: An NE point of the NVO game.
- $b_{\min} = \arg\min_{b \in \mathcal{B}} b$.
- $\mathcal{K}$: The set of possible query output values, s.t. $\mathcal{K} \subset [0, 1]$.
- $m_{i,x}$: The overall probability mass added to $x \in \mathcal{K}$ by the noise assigned to the i-th data instance.
- $M_{-i,x} = \sum_{j \in [|\mathcal{D}|] \setminus \{i\}} m_{j,x}$
- $v_{\min}, v_{\max}$: The minimum and maximum values of the probability mass $m_{i,x}$ added to $x \in \mathcal{K}$ by additive noise to the i-th data instance.
- $\Pi_i(b_i, b_{-i})$: The i-th player's payoff for the strategy (b_i, b_{-i}).
- $P_{i,\mathrm{E}}, P_{i,\mathrm{U}}$: The i-th player's P_{E} and P_{U} for the strategy.
- $\Delta P_{i,\mathrm{E}}, \Delta P_{i,\mathrm{U}}$: The change of P_{E} and P_{U} values for the i-th player, s.t. $\Delta P_{i,\mathrm{E}} = P_{i,\mathrm{E}} - P_{i-1,\mathrm{E}}$ and $\Delta P_{i,\mathrm{U}} = P_{i,\mathrm{U}} - P_{i-1,\mathrm{U}}$.

B. ASSUMPTION

Let us assume that we implement the NVO games with a continuous variance space $\mathcal{B} = [b_{\min}, \infty)$ for $b_{\min} \neq 0$ and the set of possible query output values $\mathcal{X} = [0, 1]$. We do not add noise with a probability of occurring outside the target range of the integration (*i.e.,* $[0, 1]$); thus, the probability density function of the Laplace noise is normalized as in (16).

C. PROOF

1) THE WORST CASE TO ENSURE ϵ-PDP FOR AN DATA INSTANCE

For the proof of the theorem, we start with the worst case of the ϵ-pDP of an arbitrary data instance. In order to satisfy ϵ-pDP for an element d_i, the following condition should be satisfied:

$$\max_x \frac{m_{i,x} + M_{-i,x}}{M_{-i,x} \cdot \frac{|\mathcal{D}|}{|\mathcal{D}-1|}} < \max_x \frac{m_{i,x} + M_{-i,x}}{M_{-i,x}} \leq \exp(\epsilon) \tag{12}$$

$$\Rightarrow \max_x \frac{m_{i,x} + M_{-i,x}}{M_{-i,x}} \leq \max_x \frac{m_{i,x} + \min M_{-i,x}}{\min M_{-i,x}} \tag{13}$$

$$= \max_x \frac{m_{i,x} + (|\mathcal{D}| - 1)v_{\min}}{(|\mathcal{D}| - 1)v_{\min}} \tag{14}$$

$$= \frac{m_{i,q(d_i)} + (|\mathcal{D}| - 1)v_{\min}}{(|\mathcal{D}| - 1)v_{\min}} \leq \exp(\epsilon), \tag{15}$$

where (12) is initialized from the definition of ϵ-pDP.

2) FIND THE V_{MIN}

The minimum value of the $m_{i,x}$, represented by $v_{\min}$ is obtained by

$$v_{\min} = \min_{i,x} m_{i,x} = \min_{\substack{\mu,x \in [0,1] \\ b \geq b_{\min}}} \frac{\frac{1}{2b}\exp(-\frac{|x-\mu|}{b})}{\int_0^1 \frac{1}{2b}\exp(-\frac{|t-\mu|}{b})dt} \tag{16}$$

$$= \min_{\substack{\mu,x \in [0,1] \\ b \geq b_{\min}}} \frac{\frac{1}{2b}\exp(-\frac{|x-\mu|}{b})}{1 - \frac{1}{2}\exp(\frac{\mu-1}{b}) - \frac{1}{2}\exp(\frac{-\mu}{b})} \tag{17}$$

$$= \min_{\substack{\mu,x \in [0,1] \\ b \geq b_{\min}}} V(\mu, b, x). \tag{18}$$

In (16), the definition of $v_{\min}$ is rewritten by the Laplace distribution $f(x|\mu, b) = \frac{1}{2b}\exp(-\frac{|x-\mu|}{b})$. For brevity, in (18), we newly define a function $V(\mu, b, x)$.

Then, our focus is to find a value of μ for the $v_{\min}$, and check the critical points with following conditions:

$$\frac{\partial V}{\partial \mu} = 0 \tag{19}$$

$$\Rightarrow \frac{1}{2b^2}\exp(\frac{\mu - x}{b})\left[1 - \frac{1}{2}\exp(\frac{\mu - 1}{b}) + \frac{1}{2}\exp(\frac{-\mu}{b})\right] \tag{20}$$

$$- \frac{1}{2b}\exp(\frac{\mu - x}{b})\left[-\frac{1}{2b}\exp(\frac{\mu - 1}{b}) + \frac{1}{2b}\exp(\frac{-\mu}{b})\right] \tag{21}$$

$$= \frac{1}{2b^2} = 0, \tag{22}$$

where (21) holds because of the Laplace distribution's symmetry, thereby making us consider $x \geq \mu$. Then, from the result of (22), we confirm that there is no critical point that makes $\partial V/\partial \mu = 0$. Also, when $x \geq \mu$, we confirm that the sign of $\partial V/\partial \mu$ is always positive; thus, the minimizer μ of the function $V(\mu, b, x)$ is zero as follows:

$$\text{sign}\left(\frac{\partial V}{\partial \mu}\right) = \text{sign}\left(\frac{1}{2b}\right) = \frac{1}{2b^2} \geq \frac{1}{2b_{\min}^2} > 0 \tag{23}$$

$$\Rightarrow \arg\min_{\mu} V(\mu, b, x) = 0. \tag{24}$$

Then, by substituting $\mu = 0$ into $V(\mu, b, x)$, we confirm that the minimizer x of the function is one as follows:

$$\frac{\partial V}{\partial x} = \frac{-\frac{1}{b^2}\exp(-\frac{x}{b})}{1 - \exp(-\frac{1}{b})} < 0 \Rightarrow \arg\min_{x\in[0,1]} V(0, b, x) = 1. \tag{25}$$

Up to here, we obtained the minimizers $\mu = 0$ and $x = 1$. By substituting the minimizers, we can obtain the minimizer b as

$$\arg\min_{b\geq b_{\min}} V(0, b, 1) = \arg\min_{b\geq b_{\min}} \frac{1}{\exp(\frac{1}{b}) - 1} \tag{26}$$

$$= \arg\max_{b\geq b_{\min}} \exp(\frac{1}{b}) = b_{\min} \tag{27}$$

$$\therefore v_{\min} = \frac{1}{\exp(\frac{1}{b_{\min}}) - 1}. \tag{28}$$

3) SUBSTITUTE THE OBTAINED V_{MIN} FOR GETTING THE WORST CASE

From the result of (28) and (15), we have the bound of ϵ, which always guarantee ϵ-pDP. Here, we assume the case the i-th player does his best to guarantee ϵ-pDP and choose $b_i = \infty$. Then, we have

$$\frac{\min\left(m_{i,q(d_i)}\right) + \frac{|\mathcal{D}|-1}{\exp(1/b_{\min})-1}}{\frac{|\mathcal{D}|-1}{\exp(\frac{1}{b_{\min}})-1}} = \frac{1 + \frac{|\mathcal{D}|-1}{\exp(1/b_{\min})-1}}{\frac{|\mathcal{D}|-1}{\exp(1/b_{\min})-1}} \tag{29}$$

$$\leq \exp(\epsilon) \tag{30}$$

$$\therefore \epsilon \geq \ln\left(\frac{1 + \frac{|\mathcal{D}|-1}{\exp(1/b_{\min})-1}}{\frac{|\mathcal{D}|-1}{\exp(1/b_{\min})-1}}\right), \tag{31}$$

which can be equivalently written by

$$b_{\min} \geq \frac{1}{\log\left(1 + (|\mathcal{D}| - 1)(\exp(\epsilon) - 1)\right)}. \tag{32}$$

In (31), we have $m_{i,q(d_i)} \geq 1$, where the equality holds when $b_i = \infty$ and the PDF is uniform. Finally, there always exists at least one choice to improve the DP guarantee payoff for all elements.

4) THE STRATEGY IS IMPROVED TO FINALLY GUARANTEE ϵ-PDP FOR ALL ELEMENTS

Before the strategy set satisfies the ϵ-pDP for all elements, we have

$$\min_{b_i} \Delta P_{i,\text{E}} \geq 1 > \max_{b_i} \Delta P_{i,\text{U}}. \tag{33}$$

(33) proves that there exists at least a choice to improve the ϵ-pDP guarantee for an element, when the ϵ is bounded like (31), and by the definition of P_{U}. Therefore, players choose a strategy to improve ϵ-pDP until guaranteeing for all elements.

5) THE NASH EQUILIBRIUM ENSURES ϵ-PDP FOR ALL ELEMENTS

Assume that the Nash equilibrium point $\left(b_i^*, b_{-i}^*\right)$ does not satisfy the ϵ-pDP for all elements,

$$\Pi_i(b_i^*, b_{-i}^*) \geq \Pi_i(b_i, b_{-i}^*) \tag{34}$$

$$\Rightarrow |\mathcal{D}| > \Pi_i(b_i^*, b_{-i}^*) \geq \max \Pi_i(b_i, b_{-i}^*) = \max_{b_i}\left(P_{i,\text{E}} + P_{i,\text{U}}\right) \tag{35}$$

$$= \max_{b_i}\left(P_{i-1,\text{E}} + P_{i-1,\text{U}} + \Delta P_{i,\text{E}} + \Delta P_{i,\text{U}}\right) \tag{36}$$

$$= P_{i-1,\text{E}} + P_{i-1,\text{U}} + \max_{b_i}\left(\Delta P_{i,\text{E}} + \Delta P_{i,\text{U}}\right) \tag{37}$$

$$\geq P_{i-1,\text{E}} + P_{i-1,\text{U}} + 1 \geq P_{i-2,\text{E}} + P_{i-2,\text{U}} + 2 \geq \ldots$$

$$\geq \min_{i,(b_i,b_{-i})}\left(P_{i,\text{E}} + P_{i,\text{U}}\right) + |\mathcal{D}| = |\mathcal{D}|, \tag{38}$$

where (35) follows the definition of Nash equilibrium and the definition of NVO game's payoff. Because the result in Equations 33 to 38 contradicts ($|\mathcal{D}| > |\mathcal{D}|$), we show that the assumption in this paragraph is false. That is, the Nash equilibrium point $\left(b_i^*, b_{-i}^*\right)$ must satisfy the ϵ-pDP for all elements.

APPENDIX B APPROXIMATED ENUMERATION FOR NVO GAME VIA GENETIC ALGORITHM

Enumerating the proposed game precisely proved to be computationally challenging. Instead of that, we adopted an approach grounded in evolutionary game theory. We conducted an approximated enumeration (AE) algorithm by running simulations across numerous generations. This technique monitors strategy evolution over time, revealing promising approaches without the need for an exhaustive exploration of every possible option.

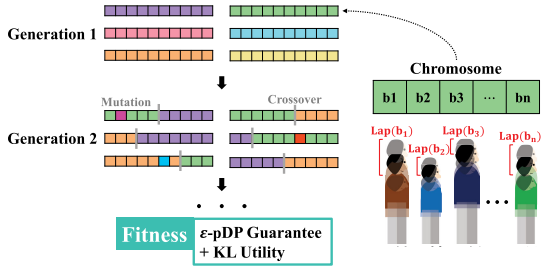

FIGURE 8. AE for the NVO game via genetic algorithm to find an NE point.

A. CHROMOSOME

Chromosomes typically symbolize solutions to the specific optimization challenge being addressed. In the framework of the NVO games, each gene is representative of the variance variable $\mathbf{b}_i$ tied to the noise introduced to the query output for every sequential element.

B. FITNESS FUNCTION

Fitness function serves as the criterion for selecting the most suitable chromosomes that fulfill the specified criteria and can pass down their traits to offspring. Hence, we adopt the payoff $P(\mathcal{M}_{\mathrm{I}}, \mathcal{D})$ as our fitness function.

The initial generation's chromosomes are created by randomly selecting values within the variance space $\mathcal{V}$ for each gene. A larger population broadens the solution search space, minimizing the risk of local optima. Some high-fitness parents are retained in the offspring generation to avoid local optima. During offspring generation, random crossover points are used, and their optimal number can be determined via simulation. Mutation probability is set to balance between avoiding local optima and ensuring trait transfer. If the best fitness value remains constant across generations, it suggests an NE approximation. The current chromosome may be optimal, but due to randomness, other solutions might emerge.

With ample time, the AE algorithm has the potential to match the performance of the exact enumeration algorithm and attain an NE point. Our experiments continued for an extended period to ensure convergence. Nevertheless, there is no theoretical guarantee that an NE point is achievable within polynomial time.

C. HYPERPARAMETERS

Our proposed BRD algorithm does not require specific hyperparameter settings. In the AE, we initially set the number of chromosomes in the population to 500, and for each generation, we involve 10 chromosomes in the mating process. We randomly designate 2 crossover points, and we introduce a 5 % probability for each gene to undergo mutation. We employ a steady-state selection approach, retaining the top 5 parents with the highest fitness values for the next generation. We utilized PyGad [37] library for the implementation.

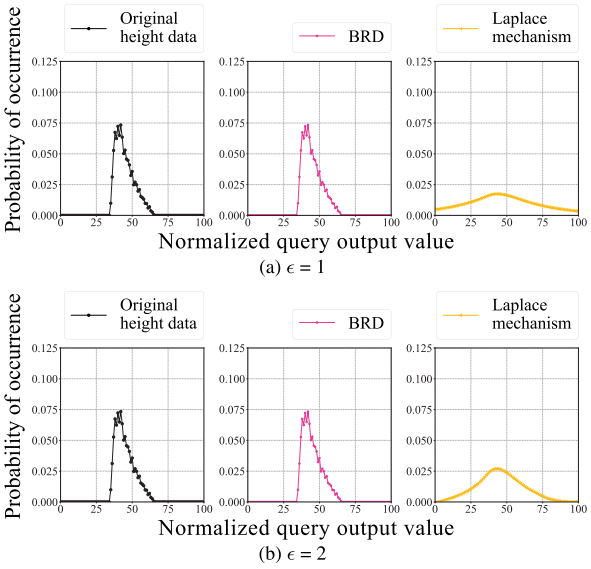

FIGURE 9. Comparison of query output probability distributions for the large income data with each algorithm, when $\epsilon = 1$ and 2.

D. HARDWARE ENVIRONMENT

We conduct experiments using an AMD Ryzen Threadripper 1920X 12-Core Processor and 32 GB of RAM. Since there is no need for extensive parallel computations, GPU utilization is not required. To conserve computing resources and facilitate a fair comparison in terms of execution time on the same evaluation criteria, we exclusively relied on CPU computations.

APPENDIX C ADDITIONAL EXPERIMENTAL RESULTS

A. EXPERIMENTAL RESULT 3: LARGE INCOME DATA

Essentially, individual privacy is harder to guarantee with smaller datasets due to the increased significance of a data point. Hence, opting for smaller datasets makes privacy assurance more challenging. As datasets grow larger, ensuring privacy becomes comparatively easier. For these reasons, we conduct experiments on a dataset of approximately 1000 in size, but to demonstrate scalability, we also perform experiments on a dataset ten times larger, comprising 10,000 instances.

1) DATASET

We utilize the test dataset of the credit profile dataset. We conduct our experiments by randomly sampling a cohort of ten thousand individuals. We note the correlation between age and income, and we exploit those two features in our experiments. Similar to the NBA player dataset, the income and age values belong to one of the 101 categorization bins.

2) ANALYSIS 1: PRESERVATION OF STATISTICAL FEATURES

In Fig. 9, the random sampling query output of the original data, conventional Laplace mechanism, and NVO game (BRD) is depicted. As similar to the result of the main manuscript, the NVO game better preserves the probabil-

TABLE 4. Each algorithm's average computation time, KL divergence, L1 loss of standard deviation, Jaccard index with a threshold 0.001, and cosine similarity are evaluated for the large income data, for $\epsilon = 1$ and 2. The modified query output distributions for all algorithms satisfy ϵ-pDP.

Algorithm	Target ϵ	Comp. time ↓ (minutes)	KL divergence ↓	L1 SD loss ↓	Jaccard index ↑ (threshold=0.001)	Cos. similarity ↑
BRD	ϵ=1	**331**	**0.0000**	**0.0000**	**1.0000**	**1.0000**
	ϵ=2	**633**	**0.0001**	**0.0004**	**1.0000**	**1.0000**
Laplace mechanism (baseline)	ϵ=1	-	0.9490	0.0190	0.3069	0.6876
	ϵ=2		0.5668	0.0173	0.3333	0.8224

*Best: **bold**.

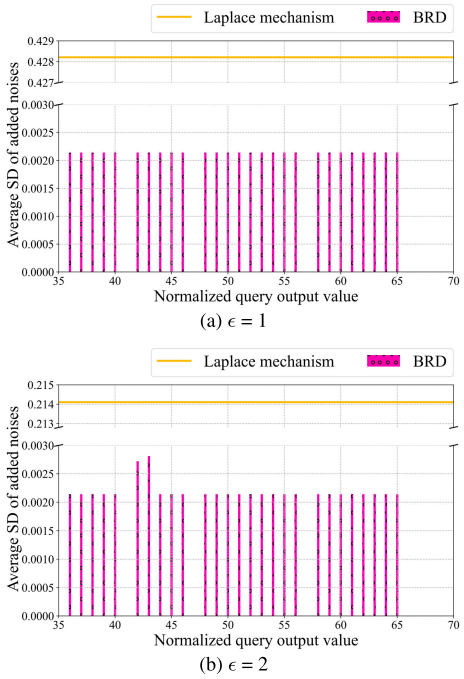

FIGURE 10. Distributions of average noise standard deviation for the large income dataset for $\epsilon = 1$ and 2.

ity distribution than the conventional Laplace mechanism, by executing per-instance noise optimization. As the dataset size increased, the AE took an excessively long time to converge, preventing us from confirming its convergence within a reasonable timeframe. Consequently, we omitted its results from our findings.

In Table 4, the proposed NVO game and the Laplace mechanism are quantitatively evaluated in various statistical metrics. Similar to the results in the main manuscript, the proposed NVO game-based algorithm (BRD) outperforms the Laplace mechanism. Furthermore, for larger datasets, due to the lower individual instance contribution, privacy is better preserved, allowing us to ensure a more robust statistical utility while maintaining the same ϵ-pDP guarantee.

For extremely large datasets, our proposed method incurs high-order computational complexity for the ϵ-pDP guarantee, scaling as $O(|\mathcal{D}|^2)$. To mitigate this, one approach could be to group data points with identical query outputs, allowing for computational reduction through the addition of uniform noise.

In Fig. 10, we compare the average standard deviation of the added noise to each categorization bin of the conventional

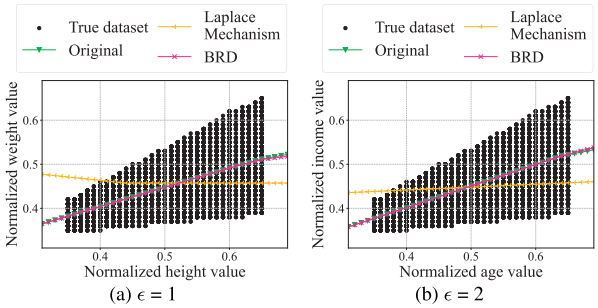

FIGURE 11. Linear regression result of 500 sampled data for each algorithm for $\epsilon = 1$ and 2.

Laplace mechanism and the NVO-game-based algorithm (BRD). As depicted in the figure, the NVO game adds lower variance at all bins, thereby having better data statistical utility.

3) ANALYSIS 2: REGRESSION TASK

To evaluate the regression task of the proposed NVO game, we run the regression network with three layers. Similarly to the main manuscript, the network consists of ten parameters and ReLU activation functions. Figure 11 depicts the scatter diagram of the original dataset and the trained regression line, where the neural network input is age and the output is income. For the value of $\epsilon = 1$ and 2, the proposed NVO game closely preserves the regression line after applying the randomized algorithm (BRD). In regression tasks as well, we observe improved data characteristics for the same ϵ-pDP when dealing with larger dataset sizes.

ACKNOWLEDGMENT

(Sehyun Ryu and Jonggyu Jang contributed equally to this work.)

REFERENCES

[1] N. O. Attoh-Okine, "Differential privacy," in *Automata, Languages and Programming*. Berlin, Germany: Springer, 2017, pp. 241–247.

[2] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *Proc. IEEE Symp. Secur. Privacy (SP)*. Washington, DC, USA: IEEE Computer Society, May 2017, pp. 3–18.

[3] A. Narayanan and V. Shmatikov, "How to break anonymity of the Netflix prize dataset," 2006, *arXiv:cs/0610105*.

[4] J.-W. Lee, H. Kang, Y. Lee, W. Choi, J. Eom, M. Deryabin, E. Lee, J. Lee, D. Yoo, Y.-S. Kim, and J.-S. No, "Privacy-preserving machine learning with fully homomorphic encryption for deep neural network," *IEEE Access*, vol. 10, pp. 30039–30054, 2022.

[5] X. Wu, F. Huang, Z. Hu, and H. Huang, "Faster adaptive federated learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, Jun. 2023, pp. 10379–10387.
[6] J. Tang, A. Korolova, X. Bai, X. Wang, and X. Wang, "Privacy loss in Apple's implementation of differential privacy on MacOS 10.12," 2017, *arXiv:1709.02753*.
[7] Ú. Erlingsson, V. Pihur, and A. Korolova, "RAPPOR: Randomized aggregatable privacy-preserving ordinal response," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.* New York, NY, USA: Association for Computing Machinery, Nov. 2014, pp. 1054–1067.
[8] A. Blanco-Justicia, D. Sánchez, J. Domingo-Ferrer, and K. Muralidhar, "A critical review on the use (and misuse) of differential privacy in machine learning," *ACM Comput. Surv.*, vol. 55, no. 8, pp. 1–16, Dec. 2022.
[9] M. Min, H. Zhu, J. Ding, S. Li, L. Xiao, M. Pan, and Z. Han, "Personalized 3D location privacy protection with differential and distortion geo-perturbation," *IEEE Trans. Depend. Sec. Comput.*, vol. 21, no. 4, pp. 3629–3643, Jul. 2024.
[10] Y. Wei, J. Jia, Y. Wu, C. Hu, C. Dong, Z. Liu, X. Chen, Y. Peng, and S. Wang, "Distributed differential privacy via shuffling versus aggregation: A curious study," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2501–2516, 2024.
[11] L. Huang, J. Wu, D. Shi, S. Dey, and L. Shi, "Differential privacy in distributed optimization with gradient tracking," *IEEE Trans. Autom. Control*, early access, Jan. 10, 2024, doi: 10.1109/TAC.2024.3352328.
[12] B. Chen and M. Hale, "The bounded Gaussian mechanism for differential privacy," *J. Privacy Confidentiality*, vol. 14, no. 1, Feb. 2024.
[13] Y.-X. Wang, "Per-instance differential privacy," *J. Privacy Confidentiality*, vol. 9, no. 1, Mar. 2019.
[14] Q. Geng and P. Viswanath, "The optimal mechanism in differential privacy," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 2371–2375.
[15] Q. Geng, P. Kairouz, S. Oh, and P. Viswanath, "The staircase mechanism in differential privacy," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1176–1184, Oct. 2015.
[16] Q. Geng, W. Ding, R. Guo, and S. Kumar, "Optimal noise-adding mechanism in additive differential privacy," in *Proc. 22nd Int. Conf. Artif. Intell. Statist.*, vol. 89, Apr. 2019, pp. 11–20.
[17] N. Holohan, S. Antonatos, S. Braghin, and P. M. Aonghusa, "The bounded Laplace mechanism in differential privacy," *J. Privacy Confidentiality*, vol. 10, no. 1, Dec. 2019.
[18] I. Mironov, "Rényi differential privacy," in *Proc. IEEE 30th Comput. Secur. Found. Symp. (CSF)*, Aug. 2017, pp. 263–275.
[19] J. Geumlek, S. Song, and K. Chaudhuri, "Renyi differential privacy mechanisms for posterior sampling," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30. Red Hook, NY, USA: Curran Associates, 2017, pp. 5295–5304.
[20] A. M. Girgis, D. Data, S. Diggavi, A. T. Suresh, and P. Kairouz, "On the Renyi differential privacy of the shuffle model," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 2321–2341.
[21] J. T. Wang, S. Mahloujifar, S. Wang, R. Jia, and P. Mittal, "Renyi differential privacy of propose-test-release and applications to private and robust machine learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35. Red Hook, NY, USA: Curran Associates, 2022, pp. 38719–38732.
[22] Y. Zhu and Y.-X. Wang, "Improving sparse vector technique with Renyi differential privacy," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33. Red Hook, NY, USA: Curran Associates, 2020, pp. 20249–20258.
[23] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014.
[24] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep learning with differential privacy," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*. New York, NY, USA: Association for Computing Machinery, 2016, pp. 308–318.
[25] J. Ding, G. Liang, J. Bi, and M. Pan, "Differentially private and communication efficient collaborative learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, May 2021, pp. 7219–7227.
[26] M. Moreau and T. Benkhelif, "DPSGD strategies for cross-silo federated learning," in *Proc. Int. Conf. Commun., Comput., Cybersec., Informat. (CCCI)*, Oct. 2021, pp. 1–5.
[27] S. Truex, L. Liu, K.-H. Chow, M. E. Gursoy, and W. Wei, "LDP-fed: Federated learning with local differential privacy," in *Proc. 3rd ACM Int. Workshop Edge Syst., Anal. Netw.* New York, NY, USA: Association for Computing Machinery, Apr. 2020, pp. 61–66.
[28] J. V. Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 1944.
[29] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Math. Biosci.*, vol. 40, nos. 1–2, pp. 145–156, Jul. 1978.
[30] F. Zaman, S. M. Elsayed, T. Ray, and R. A. Sarkerr, "Evolutionary algorithms for finding Nash equilibria in electricity markets," *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 536–549, Aug. 2018.
[31] J. F. Nash, "Non-cooperative games," *Ann. Math.*, vol. 54, no. 2, pp. 286–295, 1951.
[32] V. Boucher, "Selecting equilibria using best-response dynamics," *Econ. Bull.*, vol. 37, no. 4, pp. 2728–2734, 2017. [Online]. Available: https://ssrn.com/abstract=3175335
[33] J. Cirtautas. (2023). *NBA Players*. Accessed: Apr. 20, 2024. [Online]. Available: https://www.kaggle.com/datasets/justinas/nba-players-data
[34] M. Fatakdawala. (2019). *Income Dataset*. Accessed: Apr. 20, 2024. [Online]. Available: https://www.kaggle.com/datasets/mastmustu/income
[35] Yashkmd. (2023). *Credit Profile (two-wheeler Loan) Dataset*. Accessed: Apr. 20, 2024. [Online]. Available: https://www.kaggle.com/datasets/yashkmd/credit-profile-two-wheeler-loan-dataset/
[36] G. W. Furnas, S. Deerwester, S. T. Dumais, T. K. Landauer, R. A. Harshman, L. A. Streeter, and K. E. Lochbaum, "Information retrieval using a singular value decomposition model of latent semantic structure," in *Proc. 11th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.* New York, NY, USA: Association for Computing Machinery, 1988, pp. 465–480.
[37] A. F. Gad, "PyGAD: An intuitive genetic algorithm Python library," 2021, *arXiv:2106.06158*.

SEHYUN RYU received the B.S. degree in electrical engineering from Pohang University of Science and Technology (POSTECH), Pohang, Republic of Korea, in 2023, where he is currently pursuing the integrated M.S. and Ph.D. degree with the Department of Electrical Engineering. His research interests include the theory and development of wireless communications, privacy-preserving deep learning, and wireless networking with deep neural networks.

JONGGYU JANG (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea, in 2017 and 2021, respectively. Since March 2021, he has been a Postdoctoral Researcher with the Future IT Innovation Laboratory, POSTECH. His research interests include theory and applications of explainable machine learning and wireless communications.

HYUN JONG YANG (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2004, 2006, and 2010, respectively. From August 2010 to August 2011, he was a Research Fellow with Korea Institute of Ocean Science and Technology (KIOST), Daejeon. From October 2011 to October 2012, he was a Postdoctoral Researcher with the Electrical Engineering Department, Stanford University, Stanford, CA, USA. From October 2012 to August 2013, he was a Staff II Systems Design Engineer with Broadcom Corporation, Sunnyvale, CA, USA, where he developed physical-layer algorithms for LTE-A MIMO receivers. In addition, he was a delegate of Broadcom in 3GPP standard meetings for RAN1 Rel-12 technologies. From September 2013 to July 2020, he was an Assistant/Associate Professor with the School of Electrical and Computer Engineering, UNIST, Ulsan, South Korea. Since July 2020, he has been an Associate Professor with the Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, South Korea. His research interests include theory and development of wireless communications, radar signal processing, and deep learning-inspired information systems.