applied sciences

MDPI

Article

Privacy-Preserving Monotonicity of Differential Privacy Mechanisms

Hai Liu [1], Zhenqiang Wu [1,*], Yihui Zhou [1,*], Changgen Peng [2], Feng Tian [1] and Laifeng Lu [3]

[1] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China; liuhai@snnu.edu.cn (H.L.); tianfeng@snnu.edu.cn (F.T.)
[2] Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China; sci.cgpeng@gzu.edu.cn
[3] School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710119, China; lulaifeng@snnu.edu.cn
* Correspondence: zqiangwu@snnu.edu.cn (Z.W.); zhouyihui@snnu.edu.cn (Y.Z.); Tel.: +86-187-9268-1949 (Z.W.); +86-136-3021-8729 (Y.Z.)

Received: 3 September 2018; Accepted: 25 October 2018; Published: 28 October 2018

Abstract: Differential privacy mechanisms can offer a trade-off between privacy and utility by using privacy metrics and utility metrics. The trade-off of differential privacy shows that one thing increases and another decreases in terms of privacy metrics and utility metrics. However, there is no unified trade-off measurement of differential privacy mechanisms. To this end, we proposed the definition of privacy-preserving monotonicity of differential privacy, which measured the trade-off between privacy and utility. First, to formulate the trade-off, we presented the definition of privacy-preserving monotonicity based on computational indistinguishability. Second, building on privacy metrics of the expected estimation error and entropy, we theoretically and numerically showed privacy-preserving monotonicity of Laplace mechanism, Gaussian mechanism, exponential mechanism, and randomized response mechanism. In addition, we also theoretically and numerically analyzed the utility monotonicity of these several differential privacy mechanisms based on utility metrics of modulus of characteristic function and variant of normalized entropy. Third, according to the privacy-preserving monotonicity of differential privacy, we presented a method to seek trade-off under a semi-honest model and analyzed a unilateral trade-off under a rational model. Therefore, privacy-preserving monotonicity can be used as a criterion to evaluate the trade-off between privacy and utility in differential privacy mechanisms under the semi-honest model. However, privacy-preserving monotonicity results in a unilateral trade-off of the rational model, which can lead to severe consequences.

Keywords: differential privacy; trade-off; privacy-preserving monotonicity; semi-honest model; rational model

1. Introduction

In the non-interactive model or interactive model of computation to a database [1], a data curator's sensitive data needs to be protected, and a data analyst can get the availability of data for statistical analysis. For data curators, it is easy to protect the privacy of sensitive data by adding random noise to the query results. However, the availability of data cannot be guaranteed for data analysts. For the research of privacy preservation in statistical queries, a crucial question is the trade-off between privacy and utility [2]. Thus, the key challenge is to provide the trade-off between data privacy and data utility of the computation to a database [3,4], which is achieved by using differential privacy. To demonstrate the trade-off between privacy and utility, privacy metrics and utility metrics are used to denote data privacy and data utility, respectively. In this paper, expected estimation error [5] and

entropy are used as privacy metrics. The modulus of characteristic function and the variant $1 - NE$ of normalized entropy (NE) are used as utility metrics. According to the privacy metrics and utility metrics of this paper, the trade-off means that one thing increases and another thing decreases in terms of privacy and utility in differential privacy.

Differential privacy [6] has emerged as a powerful tool to provide the trade-off between privacy and utility. Differential privacy provides a mathematical foundation for the trade-off between privacy and utility and offers provable privacy preservation in a precise sense [7]. Laplace mechanism and Gaussian mechanism are used to achieve the trade-off for numeric data [1]. Exponential mechanism [8] and randomized response [9] also meet the trade-off for categorical data. There are various methods which can get an optimal trade-off between privacy and utility in differential privacy. Hardt and Talwar [10] used methods of convex geometry to determine a nearly optimal trade-off between privacy and error. Fouad et al. [11] proposed a scalable algorithm that met differential privacy when applying a specific random sampling and resolved the trade-off between data utility and data privacy. Generalized differential privacy definitions can achieve the trade-off. Personalized differential privacy was proposed [12], in which users specified a personal privacy requirement for their data. Soria-Comas et al. [13] proposed individual differential privacy that offered the same privacy guarantees as standard differential privacy to individuals and resulted in less distortion and more analytical accuracy. The definition of differential privacy based on information theory can also achieve the trade-off. Mutual information differential privacy ensured the trade-off between privacy and informational utility to be gained from the output of any query mechanism [14]. The unified privacy definition can also achieve the trade-off. He et al. [15] proposed Blowfish privacy with the goal of seeking a better trade-off between privacy and utility. Multiparty differential privacy protocol can reach the trade-off. Goyal et al. [16] obtained general upper bounds on accuracy of two parties' differential privacy protocol of computing any Boolean function. Kasiviswanathan et al. [17] proposed local differential privacy that could also reach the trade-off by using random response mechanism. Machine learning with differential privacy has obtained the trade-off. Since Boosting used a general method to improve the accuracy of learning algorithms, Dwork et al. [18] used boosting to construct improved privacy-preserving synopses of an input database. Zhang and Zhu [19] proposed the dual variable perturbation and primal variable perturbation to provide differential privacy for the distributed learning algorithms over a network. The former slightly outperformed the latter for balancing the trade-off between privacy and utility. Differential privacy ensures the trade-off in data analysis. Zeng et al. [20] proposed a differential privacy frequent itemsets mining algorithm and had precisely quantified the trade-off between privacy and utility in frequent itemsets mining. Search log using differential privacy can get the trade-off. Hong et al. [21] proposed a sanitization framework to perfectly tackle the trade-off in the context of search logs in a distributed manner. Differential privacy can guarantee the trade-off in statistics analysis. Within Bayesian inference based on differential privacy under certain conditions on the prior, the sampling from the posterior distribution could lead to a desired level of privacy and utility [22]. In data aggregation, keeping trade-off using differential privacy is very important. Sei and Ohsuga [23] proposed two methods: single to randomized multiple dummies, and single to randomized multiple dummies with Bayes, both of which generated a set of disguised values of one sensed value of anonymization and ensured a better trade-off between privacy and utility. Differential privacy can maintain the trade-off in data query. Nikolov et al. [24] studied the trade-off in the context of counting queries and more general linear queries. Data releasing using differential privacy has achieved the trade-off. Xu et al. [25] proposed two novel mechanisms achieving the trade-off between the utility of the query results and the degree of privacy preservation, namely NoiseFirst determining the histogram structure after the noise injection and StructureFirst of deriving the histogram structure before the addition of the noise. Differential privacy can achieve the trade-off in location-based service and recommendation service. Local differential perturbation mechanisms could provide a well-balanced trade-off between location privacy and service accuracy [26]. Guo et al. [27] proposed a mechanism to achieve differential

privacy graph-link analysis-based social recommendation. Differential privacy can also guarantee the trade-off in other applications fields. Fan et al. [28] adopted differential privacy aggregates of web browsing activities that could be realized in real time while preserving the utility of shared data. He et al. [29] proposed an efficient distortion-based privacy-preserving metering scheme that protected an individual customer's privacy and provided the complete power consumption distribution curve of a multitude of customers without privacy invasion. Jin et al. [30] proposed differential privacy crowdsourced spectrum sensing that allowed the spectrum-sensing provider to select mobile users for executing spatiotemporal spectrum-sensing tasks without violating the location privacy of mobile users. Huang and Kannan [31] showed that the exponential mechanism could be implemented as a truthful mechanism while still preserving differential privacy for any mechanism design problem with the objective of maximizing social welfare. However, there is no unified trade-off measurement of differential privacy mechanisms and their applications. This problem is covered in the definition of privacy-preserving monotonicity.

According to the privacy metrics and utility metrics of this paper, the trade-off of differential privacy means that privacy (utility) increases and utility (privacy) decreases. Thus, the increasing of privacy-preserving level leads to the decreasing of data utility as privacy budget decreases in differential privacy mechanisms. Conversely, the decreasing of privacy-preserving level leads to the increasing of data utility as privacy budget increases. Current works have no unified metrics of measuring the trade-off of differential privacy. It is imperative to seek a unified measurement as an evaluation criterion of the trade-off between privacy and utility in differential privacy. Thus, we proposed the definition of privacy-preserving monotonicity of differential privacy to formulate the trade-off between privacy and utility based on privacy metrics and utility metrics of this paper. Thus, privacy-preserving monotonicity shows that the increasing of privacy-preserving level leads to the decreasing of data utility as privacy budget decreases and the decreasing of privacy-preserving level leads to the increasing of data utility as privacy budget increases in differential privacy. Privacy-preserving monotonicity plays an important role in the research and applications of differential privacy.

Since computational indistinguishability is the central theory of cryptography, two probability distributions are computationally indistinguishable if no efficient algorithm can distinguish them [32]. Therefore, the closer two probability distributions are, the better indistinguishability is. Conversely, the less close two probability distributions are, the weaker indistinguishability is. In other words, the stronger indistinguishability is, the bigger distinguishing error or uncertainty of a polynomial time algorithm is. Conversely, the weaker indistinguishability is, the smaller distinguishing error or entropy is. Thus, computational indistinguishability can become the foundation of differential privacy. In Figure 1, we presented the research framework of this paper. First, we proposed the definition of privacy-preserving monotonicity of differential privacy based on computational indistinguishability to measure trade-off under a non-interactive model or interactive model. Second, our theoretical and experimental results showed the privacy-preserving monotonicity of several differential privacy mechanisms based on the privacy metrics of the expected estimation error and entropy, such as Laplace mechanism, Gaussian mechanism, exponential mechanism, and randomized response mechanism. In addition, we theoretically and experimentally analyze that the utility monotonicity of these differential privacy mechanisms based on utility metrics of modulus of characteristic function and variant $1 - NE$ of normalized entropy. Finally, based on privacy-preserving monotonicity of differential privacy, we discussed the method of seeking trade-off under a semi-honest model and analyzed unilateral trade-off under rational model. Our main contributions are as follows.

- We proposed the definition of privacy-preserving monotonicity of differential privacy based on computational indistinguishability to measure the trade-off between privacy and utility.
- We theoretically and experimentally analyzed privacy-preserving monotonicity of existing differential privacy mechanisms, including the Laplace mechanism, Gaussian mechanism,

exponential mechanism, and randomized response mechanism. In addition, our theoretical and experimental results showed the utility monotonicity of these differential privacy mechanisms.

- We presented the method of seeking trade-off under a semi-honest model and analyzed the unilateral trade-off under rational model for differential privacy.

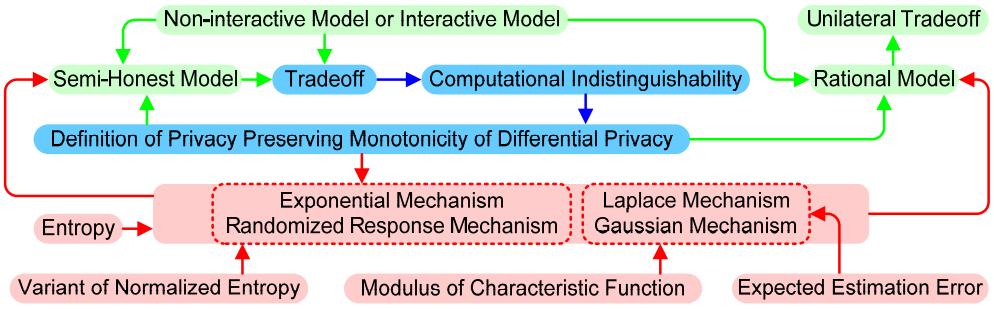

Figure 1. Research framework of this paper.

This paper is organized as follows. Section 2 introduces the related works. Section 3 describes the problem statement. Section 4 introduces the preliminary to differential privacy. Section 5 presents the definition of privacy-preserving monotonicity of differential privacy. In Section 6, we theoretically analyze privacy-preserving monotonicity and utility monotonicity of differential privacy mechanisms. Section 7 covers numerical results of privacy-preserving monotonicity and utility monotonicity of differential privacy mechanisms. Section 8 presents the procedure of seeking trade-off under semi-honest model and analyzes unilateral trade-off under rational model. Section 9 concludes this paper.

2. Related Works

In this section, we summarize and review the existing works to achieve the trade-off in differential privacy. However, these works have not claimed that how to measure the trade-off between privacy and utility. Thus, we proposed the definition of privacy-preserving monotonicity which can measure the trade-off of differential privacy in this paper.

There are some methods that can get the optimal trade-off between privacy and utility in differential privacy. With the independence of user's side information and preferences, Ghosh et al. [33] proposed geometric mechanism which was subject to differential privacy. This mechanism ensured the nearly optimal utility of every potential user; however, not all differential privacy mechanisms and all types of queries had analogous results. Hardt and Talwar [10] used methods of convex geometry to determine a nearly optimal trade-off between privacy and error. However, since this method needed to repeat the privacy and optimality analysis in the presence of approximation errors, this arose some complications. Xiao et al. [34] introduced a differential privacy algorithm for computing answers to reduce relative errors. This algorithm was effective, but it did not indicate the implications of correlations between the answers. Fouad et al. [11] proposed a scalable algorithm that met differential privacy when applying a specific random sampling. This algorithm resolved the trade-off between data utility and data privacy; however, it was difficult that the algorithm was extended to the cases when the risk threshold was small. Tossou and Dimitrakakis [35] optimized trade-off between privacy and utility by using a novel interval-based mechanism and a continual release mechanism. Although these mechanisms significantly improved the state of the art, they might be not suitable for a large database in practice. Furthermore, these methods do not indicate how to measure the trade-off of differential privacy, which is covered in the definition of privacy-preserving monotonicity.

Several novel differential privacy definitions can achieve the trade-off. Dodis et al. [36] successfully achieved differential privacy with perfect randomness, but it required the consistent

sampling that was strictly stronger than differential privacy and could not satisfy by any additive-noise mechanism. Tramèr et al. [37] investigated a relaxation of differential privacy by considering more reasonable amounts of background knowledge. They showed that it could be used to achieve membership privacy for various adversarial settings and had shown that they could achieve a significantly higher utility when protecting against bounded adversaries, but the drawback of membership privacy was unsuitable for other queries to potentially improve the trade-off between privacy and utility of differential privacy. Soria-Comas et al. [13] proposed individual differential privacy that offered the same privacy guarantees as standard differential privacy to individuals and allowed the data controller to adjust the distortion to the actual data set. Although this definition resulted in less distortion and more analytical accuracy, the performance of individual differential privacy for other common queries needed to be further studied in the interactive scenario. More importantly, there is no unified way to measure the trade-off in novel differential privacy definitions, which is addressed by our definition of privacy-preserving monotonicity.

Generalized definitions of differential privacy based on information theory can also achieve the trade-off. Sanker et al. [38] applied information-theoretic tools to provide a rigorous information-theoretic treatment of the trade-off between privacy and utility. This method guaranteeing tight bounds of how much utility was possible for a given level of privacy and vice versa. Cuff and Yu [14] gave an equivalent definition of privacy using mutual information that made plain some of the subtleties of differential privacy. Mutual information differential privacy ensured the trade-off between privacy and informational utility to be gained from the output of any query mechanism. Rényi differential privacy could also reach the trade-off between privacy and utility [39]. For specific situations, this method outperformed existing methods, but itself as a relaxation of differential privacy needed to further investigation. Although these generalized definitions of differential privacy based on information theory can achieve the trade-off, these works cannot provide the measurement method of the trade-off, which is solved by using the definition of privacy-preserving monotonicity.

The unified privacy definition provides a way of achieving the trade-off between privacy and utility in differential privacy. He et al. [15] proposed Blowfish privacy with the goal of seeking a better trade-off between privacy and utility. The key feature of Blowfish was a policy, where users could specify sensitive information that needed to be protected and knowledge about their databases which had been released to potential adversaries. Song et al. [40] proposed Pufferfish which could be used to address privacy of correlated data. Although Pufferfish offered a good solution to privacy of correlated data, modeling privacy of users connected to social networks and privacy of spatiotemporal information gathered from sensors into rigorous privacy framework of Pufferfish and designing novel mechanism for these models were still a key challenge. On the whole, unified privacy definition does not show how to measure the trade-off of differential privacy. Our definition of privacy-preserving monotonicity can address the problem.

Considering multiplayer circumstance, multiparty differential privacy protocol can get the trade-off. Goyal et al. [16] obtained general upper bounds on accuracy of two parties differential privacy protocol of computing any Boolean function. The major result was a new general geometric technique for obtaining non-trivial accuracy bounds for any Boolean functionality. Kairouz et al. [41] proposed multiparty XOR computation framework of addressing the trade-off between privacy and accuracy. This optimality result of the protocol was very general, which it held for all types of functions, heterogeneous privacy conditions on the parties, all types of cost metrics, and both average and worst-case measures of accuracy, but it was no longer true to correlated data and (ε, δ)-differential privacy in multiparty context. Moreover, none of these multiparty differential privacy protocols has offered the way to measure the trade-off. We account this problem with our definition of privacy-preserving monotonicity.

By using different methods, local differential privacy can obtain the trade-off. Kasiviswanathan et al. [17] proposed local differential privacy that provided strong confidentiality guarantees in the contexts where aggregate information was released about a database containing

sensitive information about individuals. This notion could also reach the trade-off between privacy and utility. Kairouz et al. [42] studied the fundamental trade-off between local differential privacy and information-theoretic utility functions. They introduced a combinatorial family of extremal privatization mechanisms called staircase mechanisms; however, solving this linear program could be computationally expensive for multivariate random response. Kairouz et al. [43] investigated the constrained minimization problem and showed that randomized response [9] achieved the optimal trade-off between privacy and utility under minimax distribution estimation. Although they presented the hash k-ary randomized response that empirically met or exceeded the utility of existing mechanisms at all privacy levels, these mechanisms did not consider the correlation between data. Also, there is no way to measure the trade-off of local differential privacy, which is addressed using the definition of privacy-preserving monotonicity.

Machine learning with differential privacy can obtain the trade-off. Dwork et al. [18] used boosting to construct improved privacy-preserving synopses of an input database. Boosting was a general method to improve the accuracy of learning algorithms. Chaudhuri et al. [44] provided general techniques to produce privacy-preserving approximations of classifiers learned via empirical risk minimization. They proposed a new method that objective perturbation was superior to output perturbation in managing the inherent trade-off between privacy and accuracy for privacy-preserving approximations of classifiers learning via empirical risk minimization; however, objective perturbation method did not address the question of finding privacy solutions to more general convex optimization problems. Iyer et al. [45] identified the skew within class ratios of the sets across training data, which controlled the trade-off between learning and privacy preservation. Although the proposed model that could get more accurate, the model was unsuitable for the case of continuous values. Zhang and Zhu [19] proposed the dual variable perturbation and primal variable perturbation to provide dynamic differential privacy to the distributed learning algorithms over a network. The former slight outperformed the latter in balancing the trade-off between privacy and utility. Shokri and Shmatikov [46] presented a practical system that enabled multiparty to jointly learn an accurate neural network model for a given objective without sharing their input databases. This method offered an attractive point in the trade-off space between privacy and utility; however, designing a completely distributed implementation of the parameter storage system was a challenge to the independence of parameters from each other in distributed selective stochastic gradient descent. Bernstein et al. [47] investigated undirected graphical models using collective graphical models in a differential privacy way. They showed that the approach to release noisy sufficient statistics using the Laplace mechanism achieved a good trade-off between privacy, utility, and practicality. Although machine learning with differential privacy achieves the trade-off, there is no advisable method to measure the trade-off. The definition of privacy-preserving monotonicity can address this problem.

Keeping the trade-off is the most important thing to data mining with differential privacy. Friedman and Schuster [48] concluded that the introduction of formal privacy guarantees into a system required the data miner to take a different approach to data mining algorithms. Their goal was to investigate algorithms that could extend the Pareto frontier allowing for better privacy and accuracy trade-off; however, the large variance in the method was clearly a problem, and more stable results were desirable even if they came at a cost. Zeng et al. [20] proposed a differential privacy frequent itemsets mining algorithm. They had precisely quantified the trade-off between privacy and utility in frequent itemsets mining; however, the success of their algorithm relied on the assumption that short transactions dominated the benchmark datasets. Xu et al. [49] designed a differential privacy frequent sequence mining algorithm by leveraging a sampling candidate pruning technique. This algorithm was ε-differential privacy and could privately find frequent sequences of high accuracy. Although data mining with differential privacy can obtain the trade-off, these works did not explain how to measure the trade-off. This problem is covered in the definition of privacy-preserving monotonicity.

In query logs, differential privacy can guarantee the trade-off. Feild et al. [50] described an approach to distributed search logs collection, storage, and mining, which the dual goals preserved

privacy and made the mined information broadly available. This method had the trade-off between the strength of the guarantees of privacy loss and the utility of the released data, but this method did not generalize to the feasibility of realistic settings. Hong et al. [21] proposed a sanitization framework to perfectly tackle the trade-off in the context of search logs in a distributed manner. This framework significantly preserved the utility of the sanitized search logs, but this framework was not an effective approach to publish search logs with the correlation between users' query-url pairs. Korolova et al. [51] precisely characterized the trade-off between privacy and threshold used for publishing a query. They proposed a solution to release the search logs by an algorithm for releasing queries and clicks on a manner that guaranteed user privacy according to a rigorous privacy definition, but it could not best generate a search log from perturbed data. Zhang et al. [52] introduced a framework of anonymous query logs by differential privacy. This framework was able to achieve a good balance between retrieval utility and privacy; however, the framework was only empirically evaluated against multiple search algorithms on their retrieval utility. Sánchez et al. [53] proposed a privacy-preserving method of query logs that joined the flexibility and convenience of privacy-preserving data releasing with the strong privacy guarantees of differential privacy. Although this method produced query logs with differential privacy that was useful for data analysis, the major challenge was that the microaggregation algorithm did adapt to the lack of structure of query logs. In addition, query logs with differential privacy do not provide a way to measure the trade-off. This problem has been addressed based on the definition of privacy-preserving monotonicity.

Differential privacy can ensure the trade-off of statistics analysis. Chaudhuri et al. [54] studied the theoretical and empirical performance of differential privacy approximations to principal components analysis. They proposed a new method which explicitly optimized the utility of the output; however, developing a framework of analyzing general approximations to principal components analysis might be a challenge to machine learning. Balle et al. [55] gave the details of two differential privacy policy evaluation algorithms based on first visit Monte Carlo estimation. This algorithm came with utility guarantees showing that the cost of privacy diminished as training batches get larger but extending the algorithm to other applications was a challenge. Zhang et al. [56] studied a suite of mechanisms how to communicate findings of Bayesian inference to third parties of the strong guarantee of differential privacy. These mechanisms showed the trade-off between privacy and utility. Within Bayesian inference based on differential privacy under certain conditions on the prior, Dimitrakakis et al. [22] presented that the sampling from the posterior distribution could lead to a desired level of privacy and utility. Although they had shown how Bayesian inference could already be differential privacy by appropriately setting the prior, they had not examined how this affects learning. Furthermore, the measurement method of the trade-off was not provided for statistics analysis with differential privacy. Thus, we proposed the definition of privacy-preserving monotonicity to measure the trade-off.

Data aggregation using differential privacy can get the trade-off. Pathak et al. [57] proposed a method of composing an aggregate classifier satisfying differential privacy from classifiers locally trained by multiple untrusted parties. Although the method allowed these parties to interact with an untrusted curator to construct additive shares of a perturbed aggregate classifier, generalizing the theoretical analysis of the perturbed aggregate classifier to the setting in which each party had data generated from a different distribution was also crucial challenge. Sei and Ohsuga [23] proposed two methods: single to randomized multiple dummies and single to randomized multiple dummies with Bayes, both of which generated a set of disguised values of one sensed value of anonymization. These two methods ensure a better trade-off between privacy and utility, but the measurement way of trade-off of data aggregation with differential privacy is not shown in these works. The definition of privacy-preserving monotonicity can measure the trade-off.

Differential privacy maintains the trade-off of data query. Geng and Viswanath [58] characterized the fundamental trade-off between privacy and utility in differential privacy and derived the staircase mechanism for a single real-valued query function under a very general utility maximization framework. Within the classes of mechanisms oblivious of the database and the queries exceeding the

global sensitivity, they showed that adding query-output independent noise with staircase distribution was optimal among all randomized mechanisms that preserved differential privacy, and this mechanism could extend to the multidimensional setting and had been studied for (ε, δ)-differential privacy under a similar optimization framework. Nikolov et al. [24] studied the trade-off in the context of counting queries and more general linear queries. They showed that there was an inherent trade-off between privacy and accuracy when answering a large number of queries. Wang et al. [59] formulated the correlated query results as the non-zero prior knowledge and proposed a novel differential privacy approach to enhance privacy of social network data being inferred. The proposed approach was superior to the state-of-the-art privacy-preserving approaches with respect to data privacy and data utility; however, this approach might be not extended to other types of social graph queries. In addition, data query using differential privacy does not show the way to measure the trade-off, which is accounted by using the definition of privacy-preserving monotonicity.

For data releasing with differential privacy, the trade-off is also very important. Xu et al. [25] proposed two novel mechanisms achieving the trade-off between the utility of the query results and the degree of privacy preservation, namely NoiseFirst determining the histogram structure after the noise injection and StructureFirst of deriving the histogram structure before the addition of the noise. These two proposals generated highly accurate query answers and consistently outperformed existing competitors, but it was a challenge that the proposals were extended to publish multi-dimension histograms under differential privacy. Xu et al. [60] proposed a differential privacy algorithm for high-dimensional data releasing via random projection to maximize utility while guaranteeing privacy. This algorithm substantially outperformed several state-of-the-art solutions in terms of perturbation error and privacy budget on high-dimensional data sets, but the algorithm was not high computation efficiency to data publication using differential privacy of large domain. Furthermore, data releasing with differential privacy does not show that the measurement method is used to measure the trade-off. Thus, we solve this problem using the definition of privacy-preserving monotonicity.

Location privacy-preserving based on differential privacy also needs to keep the trade-off. Obfuscated locations could provide the means to access a location-based service without privacy leakage [26]. With additional knowledge such as the variation in the resemblance due to different perturbed locations, a user could be in better control of balancing location privacy and service accuracy. In response to users' natural personalized privacy requirements for privacy spatial data aggregation problem of the local setting, a novel personalized local differential privacy model provided a better trade-off between location privacy and utility [61]. The model could achieve desirable utility and provide rigorous privacy-preserving. However, location privacy-preserving based on differential privacy does not provide the measurement method to measure the trade-off. Thus, the definition of privacy-preserving monotonicity is imperative to measure the trade-off.

Differential privacy achieves the trade-off in recommendation service. McSherry and Mironov [62] found that several leading approaches in the Netflix Prize competition could be adapted to provide differential privacy without significantly degrading their accuracy. Although the loss in accuracy decreased as more data became available for a fixed value of the privacy parameter, these approaches did not immediately admit applying to privacy-preserving computations of latent factors and incorporation in differential privacy framework of advanced methods for collaborative filtering. Guo et al. [27] proposed a mechanism to achieve differential privacy graph-link analysis-based social recommendation. The proposed mechanism achieved a better trade-off between privacy and accuracy in comparison with existing work. However, recommendation service with differential privacy does not provide an approach to measure the trade-off. Thus, this problem is covered the proposed definition of privacy-preserving monotonicity.

In further application fields, differential privacy can also ensure the trade-off. Fan et al. [28] adopted differential privacy aggregates of web browsing behavior that could be released in real time while preserving the utility of shared data. Although the quality of the privacy of releasing data by this solution closely resembled that of the unperturbed aggregates, the scalability of multivariate

method of this solution became an immediate challenge when the domain of web pages was large. He et al. [29] proposed an efficient distortion-based privacy-preserving metering scheme that protected an individual customer's privacy and provided the complete power consumption distribution curve of a multitude of customers without privacy invasion. This scheme provided a proper trade-off between privacy-preserving and accuracy of power consumption distribution reconstruction in smart grids. Jin et al. [30] proposed a novel framework that allowed the spectrum-sensing provider to select mobile users for executing spatiotemporal spectrum-sensing tasks without violating the location privacy of mobile users. This framework could provide differential location privacy to mobile users while ensuring that the spectrum-sensing provider could accomplish spectrum-sensing tasks with overwhelming probability and the minimum cost. Huang and Kannan [31] showed that the exponential mechanism could be implemented as a truthful mechanism while still preserving differential privacy for any mechanism design problem with the objective of maximizing social welfare. The trade-off between privacy and social welfare was asymptotically optimal in combinatorial public project problem and multiple items auction. However, these application fields with differential privacy do not show any method to measure the trade-off. Thus, the purpose of the definition of privacy-preserving monotonicity measures the trade-off.

To sum up, currently works only achieve the trade-off between privacy and utility in differential privacy and its applications. However, there is no unified measurement of the trade-off between privacy and utility in differential privacy as we know. According to the privacy metrics and utility metrics of this paper, we proposed the definition of privacy-preserving monotonicity based on computational indistinguishability to address this problem of differential privacy. Furthermore, our theoretical and experimental analysis showed that the existing differential privacy mechanisms satisfy the privacy-preserving monotonicity. In addition, we theoretically and experimentally analyzed the utility monotonicity of differential privacy mechanisms. Finally, we discussed the method of seeking trade-off under semi-honest model and analyzed unilateral trade-off under rational model.

3. Notation and Problem Statement

In this section, we present the model of trade-off, privacy metrics, utility metrics and identify the problem. In Table 1, we give the explanations of some notations of this paper.

Table 1. Definition of notations.

Notation	Definition
PM	Privacy metric
PM_e	Expected privacy metric under expected privacy budget ε_e
PM_ε	Privacy metric under privacy budget ε
UM	Utility metric
UM_ε	Utility metric under privacy budget ε
EEE	Expected estimation error
$\phi(t)$	Characteristic function
$\|\phi(t)\|$	Modulus of characteristic function
ENT	Entropy
NE	Normalized entropy
$1-NE$	Utility metric of output of random selection
$\mathbb{N}$	Set of all nonnegative integers
$\mathcal{X}$	Universe set of data types
x, y	Databases represented by histograms, $x, y \in \mathbb{N}^{\|\mathcal{X}\|}$
x_i	The number of elements in the database x of type $i \in \mathcal{X}$
$\mathbb{R}$	Set of real numbers
$\mathcal{R}$	Range of exponential mechanism
r	Exponential mechanism outputting an element $r \in \mathcal{R}$
$\mathcal{A}$	Probabilistic polynomial time algorithm

3.1. Model of Trade-Off

In differential privacy, privacy budget cannot clearly show privacy-preserving level. Thus, privacy metric can show privacy-preserving level, and utility metric shows the availability of data. According to the privacy metrics and utility metrics of this paper, the increasing of privacy metric causes the decreasing of utility metric and the decreasing of privacy metric leads to the increasing of utility metric. Thus, when privacy budget increases, the privacy metric decreases, and the utility metric increases. When privacy budget decreases, the privacy metric increases and the utility metric decreases. This result is consistent with differential privacy using privacy budget controlling the privacy-preserving level of data and the availability of data, when a differential privacy mechanism achieves the trade-off. For example, the model of the trade-off shows that privacy metric decreases and utility metric increases as privacy budget increases in Figure 2. Conversely, the privacy metric increases and utility metric decreases as privacy budget decreases.

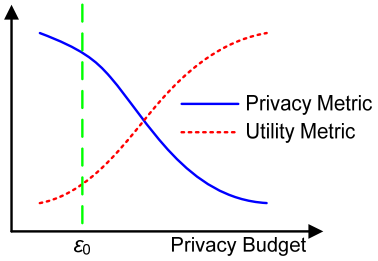

Figure 2. Model of trade-off of differential privacy.

3.2. Threat Model

There are two models of computation, including non-interactive model and interactive model [1]. In these two models, we consider the data curator is fully trusted, while the data analyst is semi-honest. Data analyst is curious about sensitive information on the data analysis. We assume that two entities can mutually authenticate within the application models. Furthermore, we consider both data curator and data analyst to be rational. Specifically, rational data curator does not ensure that an expected privacy preservation cannot suffer any privacy leakage of data releasing, but it does guarantee that any undesirable privacy preservation cannot gain any privacy preservation. Rational data analyst does not ensure that an expected data utility cannot suffer any worthless results from data analysis, but it does guarantee that any undesirable data utility cannot gain any valuable results. In this paper, expected privacy preservation refers to the preference regarding the privacy budget of a differential privacy mechanism. The expected data utility means the preference regarding the data utility of valuable analysis under a differential privacy mechanism. Since data curator and data analyst are rational, both hope to obtain the maximum privacy-preserving level and the best data utility, respectively. However, the protecting strategies of data curator and the analysis strategies of data analyst are mutually restrictive in non-interactive model or interactive model. Therefore, data curator and data analyst hope to get the expected privacy preservation and expected data utility, respectively, which is a better result. Formally, we denote privacy metric and utility metric by using PM and UM, respectively. PM_ε and UM_ε denote privacy metric and utility metric under privacy budget ε, respectively. PM_e and UM_e denote the expected privacy metric and expected utility metric of data curator and data analyst, respectively. PM_e is the expected privacy metric under the expected privacy budget ε_e. Thus, rational data curator hopes to the privacy metric PM_ε satisfying $PM_e \leq PM_\varepsilon$ under privacy budget ε meeting $\varepsilon \leq \varepsilon_e$. When $\varepsilon \leq \varepsilon_e$, data curator may be suffering privacy leakage. However, when $\varepsilon_e \leq \varepsilon$, data curator cannot gain any privacy-preserving. Rational data analyst hopes to the utility metric UM_ε satisfying $UM_e \leq UM_\varepsilon$ under privacy budget ε meeting $\varepsilon_e \leq \varepsilon$. When $UM_e \leq UM_\varepsilon$, data analyst may be suffering worthless analysis result. However, when $UM_\varepsilon \leq UM_e$, data analyst cannot gain any valuable results.

In this paper, considering the above two application models, the specific threat model is as follows. Semi-honest model: data curator is trusted, and data analyst is honest-but-curious. Rational model: data curator and data analyst are rational.

3.3. Privacy and Utility Metrics

To numerical and categorical data, some privacy metrics and utility metrics are necessary for formalizing the model of trade-off. It cannot compare Euclidean distances between categorical data, because the categorical data directly indicate the semantics. Thus, Euclidean distance cannot be used as privacy metric for categorical data. Since differential privacy mechanisms make randomized selection of categorical data, the uncertainty can be used as privacy metric. In this paper, the expected estimation error and entropy are used as privacy metrics. The modulus of characteristic function and the variant $1 - NE$ of normalized entropy are used as utility metrics.

Let $s_i^{'}$ be random perturbation value of the original s_i. The expected estimation error (EEE) of query results is

$$EEE = \sum P(s_i^{'})||s_i^{'} - s_i||_1 \tag{1}$$

Since the characteristic function of any random variable completely defines its probability distribution, we use the modulus of characteristic function as utility metric. When X is a continuous random variable and the probability density function of X is $p(x)$, the characteristic function of X is

$$\phi(t) = \int_{-\infty}^{+\infty} p(x) \exp(itx) dx, -\infty < t < +\infty \tag{2}$$

Thus, the modulus of $\phi(t) = a + ib$ is $|\phi(t)| = \sqrt{a^2 + b^2}$. In this paper, the $|\phi(t)|$ denotes the data utility of a differential privacy mechanism. When use the modulus $|\phi(t)|$ of characteristic function $\phi(t)$ as utility metric, the parameter t is a preset fixed value.

By random selection query results of a database, the entropy (ENT) of output is

$$ENT = -\sum P(s_i) \log_2 P(s_i) \tag{3}$$

We can get the normalized entropy of n outputs is

$$NE = -\frac{\sum P(s_i) \log_2 P(s_i)}{\log_2 n} \tag{4}$$

Since normalized entropy indicates the uncertainty of output of random selection and $0 \leq NE \leq 1$, we use $1 - NE$ denotes the utility metric of output of random selection.

3.4. Problem

Trade-off of differential privacy shows that one increases while the another decreases to privacy metric and utility metric of this paper. The increasing of privacy metric means the decreasing of utility metric. Conversely, the decreasing of privacy metric leads to the increasing of utility metric for differential privacy. In other words, the privacy metric decreases as privacy budget increases, and the utility metric increases as privacy budget increases. Conversely, the privacy metric increases as privacy budget decreases, and the utility metric decreases as privacy budget decreases. In view of the privacy metrics and utility metrics of this paper, we only need a mathematical model of privacy metric changing with privacy budget to formulate the trade-off between privacy-preserving and utility in differential privacy. In this paper, the privacy metric PM denotes the expected estimation error or entropy. PM_ε denotes the privacy metric using differential privacy mechanisms under privacy budget ε. The utility metric is denoted by UM, which is the modulus of characteristic function or the variant $1 - NE$ of normalized entropy. UM_ε denotes the utility metric using differential privacy mechanisms under privacy budget ε. The formalization of this problem is as follows.

In a differential privacy, privacy metric PM decreases and the utility metric UM increases as privacy budget ε increases. Conversely, the privacy metric PM increases and the utility metric UM decreases as privacy budget ε decreases. Thus, the increasing of privacy metric PM leads to the decreasing of utility metric UM as privacy budget ε decreases. Conversely, the decreasing of privacy metric PM leads to the increasing of utility metric UM as privacy budget ε increases.

4. Differential Privacy

In this section, we introduce preliminary to differential privacy [1]. A database x as being collections of records coming from a universe $\mathcal{X}$ of data types. It will be convenient to represent databases by histograms, $x \in \mathbb{N}^{|\mathcal{X}|}$, in which each x_i represents the number of elements in the database x of type $i \in \mathcal{X}$. Symbol $\mathbb{N}$ denotes the set of all nonnegative integers. In this presentation, a natural measurement of distance between two databases x and y will be ℓ_1 norm $||x-y||_1$. If $||x-y||_1 \leq 1$, then two databases x and y is called as adjacent databases.

Definition 1 (Differential Privacy)**.** *A randomized algorithm $\mathcal{M}$ with domain $\mathbb{N}^{|\mathcal{X}|}$ is (ε, δ)-differential privacy if for all $\mathcal{S} \subseteq Range(\mathcal{M})$ and for all $x, y \in \mathbb{N}^{|\mathcal{X}|}$ such that $||x-y||_1 \leq 1$, then*

$$P(\mathcal{M}(x) \in \mathcal{S}) \leq e^{\varepsilon} P(\mathcal{M}(y) \in \mathcal{S}) + \delta \tag{5}$$

where the probability space is over the coin flips of the mechanism $\mathcal{M}$. If $\delta = 0$, $\mathcal{M}$ is ε-differential privacy.

For each record of every individual, the coin flips of the mechanism mean that $\mathcal{M}$ inherently has only two possible and equally likely outcomes. Specifically, differential privacy ensures that any sequence of outputs in response to queries is essentially equally likely to occur, which probability spaces is the coin flips of the mechanism, independent of the presence or absence of any individual. The coin flips of the mechanism have the same mean for the following definition of local differential privacy. According to the definition of (ε, δ)-differential privacy, the mechanism $\mathcal{M}$ is ε-differential privacy with probability at least $1-\delta$ for all adjacent databases x and y. In the follow-up section, symbol $\mathbb{R}$ is the set of real numbers.

Definition 2 (ℓ_1-Sensitivity)**.** *The ℓ_1-sensitivity of a function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$ is*

$$\Delta f = \max_{\substack{x,y \in \mathbb{N}^{|\mathcal{X}|} \\ ||x-y||_1 = 1}} ||f(x) - f(y)||_1 \tag{6}$$

Definition 3 (Laplace Mechanism)**.** *Given any function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$, the Laplace mechanism is defined as*

$$\mathcal{M}_L(x, f(\cdot), \varepsilon) = f(x) + (Y_1, \ldots, Y_k) \tag{7}$$

where Y_i are independent identical distribution random variables drawn from Laplace distribution $Lap(\frac{\Delta f}{\varepsilon})$.

Definition 4 (ℓ_2-Sensitivity)**.** *The ℓ_2-sensitivity of a function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$ is*

$$\Delta_2 f = \max_{\substack{x,y \in \mathbb{N}^{|\mathcal{X}|} \\ ||x-y||_1 = 1}} ||f(x) - f(y)||_2 \tag{8}$$

Definition 5 (Gaussian Mechanism)**.** *Given any function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$, the Gaussian mechanism is defined as*

$$\mathcal{M}_G(x, f(\cdot), \varepsilon) = f(x) + (Y_1, \ldots, Y_k) \tag{9}$$

where Y_i are independent identical distribution random variables drawn from Gaussian distribution $N(0, \sigma^2)$ with parameters $\sigma \geq \frac{\Delta_2 f \sqrt{2\ln(\frac{1.25}{\delta})}}{\varepsilon}$ and $\varepsilon \in (0, 1)$.

While both Laplace mechanism and Gaussian mechanism are suitable for numerical data, exponential mechanism and randomized response mechanism are used for categorical data. Given some arbitrary range $\mathcal{R}$ of exponential mechanism, the exponential mechanism is defined with respect to some utility function $u : \mathbb{N}^{|\mathcal{X}|} \times \mathcal{R} \to \mathbb{R}$, which maps database and output pairs into utility scores. The sensitivity of utility scores u is

$$\Delta u = \max_{r \in \mathcal{R}} \max_{x,y:||x-y||_1 \leq 1} |u(x, r) - u(y, r)| \tag{10}$$

Definition 6 (Exponential Mechanism)**.** *The exponential mechanism $\mathcal{M}_E(x, u, \mathcal{R})$ selects and outputs an element $r \in \mathcal{R}$ with probability proportional to $\exp(\frac{\varepsilon u(x,r)}{2\Delta u})$.*

Definition 7 (Local Differential Privacy)**.** *A randomized algorithm $\mathcal{M}$ is ϵ-local differential privacy if for any possible output $\xi \in Range(\mathcal{M})$ and for any two inputs b_1, b_2, then*

$$P(\mathcal{M}(\xi|b_1)) \leq e^{\varepsilon} P(\mathcal{M}(\xi|b_2)) \tag{11}$$

where the probability space is over the coin flips of the mechanism $\mathcal{M}$.

The randomized response [9] is a major perturbation mechanism of local differential privacy [17]. For example [63], for any possible output $\xi \in \{0, 1\}$ (the top of the matrix of Equation (12)) and any input $b \in \{0, 1\}$(the left side of the matrix of Equation (12)), the conditional probability matrix of the bivariate randomized response mechanism is

$$\begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} \frac{e^{\varepsilon}}{1+e^{\varepsilon}} & \frac{1}{1+e^{\varepsilon}} \\ \frac{1}{1+e^{\varepsilon}} & \frac{e^{\varepsilon}}{1+e^{\varepsilon}} \end{pmatrix} \end{array} \tag{12}$$

For any possible output $\xi \in \{1, 2, \ldots, n\}$ (the top of the matrix of Equation (13)) and any input $b \in \{1, 2, \ldots, n\}$ (the left side of the matrix of Equation (13)), the conditional probability matrix of the multivariate randomized response mechanism is

$$\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & n \end{array} \\ \begin{pmatrix} \frac{e^{\varepsilon}}{n-1+e^{\varepsilon}} & \frac{1}{n-1+e^{\varepsilon}} & \cdots & \frac{1}{n-1+e^{\varepsilon}} \\ \frac{1}{n-1+e^{\varepsilon}} & \frac{e^{\varepsilon}}{n-1+e^{\varepsilon}} & \cdots & \frac{1}{n-1+e^{\varepsilon}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{n-1+e^{\varepsilon}} & \frac{1}{n-1+e^{\varepsilon}} & \cdots & \frac{e^{\varepsilon}}{n-1+e^{\varepsilon}} \end{pmatrix} \end{array} \tag{13}$$

5. Privacy-Preserving Monotonicity

Computational indistinguishability [32] is the foundation of differential privacy. Two probability distributions are computationally indistinguishable if no efficient algorithm can distinguish them. Considering two distributions X and Y over database $x \in \mathbb{N}^{|\mathcal{X}|}$, X and Y each assigns some probabilities to every element in x. X and Y are computationally indistinguishable, if for any probabilistic polynomial time algorithm $\mathcal{A}$ there exists a negligible function $negl(n)$ on security parameter n such that

$$|P(\mathcal{A}(x, X)) - P(\mathcal{A}(x, Y))| \leq negl(n) \tag{14}$$

Equivalently, the variant of computational indistinguishability is

$$1 - negl(n) \leq \frac{P(\mathcal{A}(x, X))}{P(\mathcal{A}(x, Y))} \leq 1 + negl(n) \tag{15}$$

Let $negl(n) = \varepsilon$. Since $e^{\varepsilon} \approx 1 + \varepsilon$ for small ε, we get the equal variant of computational indistinguishability is

$$e^{-\varepsilon} \leq \frac{P(\mathcal{A}(x, X))}{P(\mathcal{A}(x, Y))} \leq e^{\varepsilon} \tag{16}$$

Since ε is small under computational indistinguishability, we say that algorithm $\mathcal{A}$ cannot distinguish these two distributions. This means that $\mathcal{A}$ cannot distinguish whether an element x_i sampled according to distribution X or whether an element x_i sampled according to distribution Y. The smaller ε is, the better indistinguishability is. Thus, the closer to 1 the probability ratio of two distributions X and Y is, the better indistinguishability is. Therefore, the stronger indistinguishability is, the bigger distinguishing error or uncertainty of an algorithm $\mathcal{A}$ is. Conversely, the weaker indistinguishability is, the smaller distinguishing error or uncertainty of an algorithm $\mathcal{A}$ is. The stronger indistinguishability is, the better privacy-preserving level is. The weaker indistinguishability is, the worse privacy-preserving level is.

We can see that differential privacy definition is a generalization of computational indistinguishability, when algorithm $\mathcal{A}$ is replaced with randomized algorithm $\mathcal{M}$ and the ℓ_1 distance is $||x - y||_1 \leq 1$. Therefore, the smaller privacy budget ε is, the stronger indistinguishability is. Conversely, the bigger privacy budget ε is, the weaker indistinguishability is. According to the privacy metrics and utility metrics of this paper, the smaller privacy budget ε is, the bigger the privacy metric is and the smaller the utility metric is. Conversely, the bigger privacy budget ε is, the smaller the privacy metric is and the bigger the utility metric is. Thus, the expected estimation error or entropy of a differential privacy mechanism increases as privacy budget decreases, but the modulus of characteristic function or the variant $1 - NE$ of normalized entropy decreases as privacy budget decreases. Conversely, the expected estimation error or entropy of a differential privacy mechanism decreases as privacy budget increases, but the modulus of characteristic function or the variant $1 - NE$ of normalized entropy increases as privacy budget increases.

Therefore, the expected estimation error or entropy decreases and the modulus of characteristic function or the variant $1 - NE$ of normalized entropy increases as privacy budget increases. Conversely, the expected estimation error or entropy increases and the modulus of characteristic function or the variant $1 - NE$ of normalized entropy decreases as privacy budget decreases. In other words, the smaller privacy budget is, the better privacy-preserving level is and the worse the availability of data is. Conversely, the bigger privacy budget is, the worse privacy-preserving is and the better data utility is. According to the privacy metrics and utility metrics of this paper, the increasing of privacy metric leads to the decreasing of utility metric as privacy budget decreases. Conversely, the decreasing of privacy metric causes the increasing of utility metric as privacy budget increases. Thus, we only give the monotonicity property of privacy metric of differential privacy as privacy budget changing based on privacy metrics and utility metrics of this paper. In view of this, we can know the monotonicity property of utility metric of differential privacy as privacy budget changing. That is to say, the monotonicity of privacy metric can show the trade-off between privacy and utility of differential privacy based on privacy metric and utility metric of this paper. In other words, privacy-preserving monotonicity can show the trade-off between privacy and utility of differential privacy. According to the model of the trade-off of differential privacy, the privacy metrics, the utility metrics, and the identification of the problem, we have the following definition of privacy-preserving monotonicity.

Definition 8 (Privacy-Preserving Monotonicity)**.** *A privacy metric PM of differential privacy, the privacy-preserving monotonicity is defined as* $PM_{\varepsilon_1} \geq PM_{\varepsilon_2}$*, when* $\varepsilon_1 \geq 0$*,* $\varepsilon_2 \geq 0$ *and* $\varepsilon_1 \leq \varepsilon_2$*.*

According to Definition 8 and the privacy and utility metrics of this paper, given $\varepsilon_1 \geq 0$, $\varepsilon_2 \geq 0$ and $\varepsilon_1 \leq \varepsilon_2$, we have $PM_{\varepsilon_1} \geq PM_{\varepsilon_2}$, then $UM_{\varepsilon_1} \leq UM_{\varepsilon_2}$. Conversely, given $\varepsilon_1 \geq 0$, $\varepsilon_2 \geq 0$ and $\varepsilon_1 \geq \varepsilon_2$, we have $PM_{\varepsilon_1} \leq PM_{\varepsilon_2}$, then $UM_{\varepsilon_1} \geq UM_{\varepsilon_2}$. Thus, the increasing of privacy metrics leads to the decreasing of utility metrics as privacy budget decreases. Conversely, the decreasing of privacy metrics leads to the increasing of utility metrics as privacy budget increases. Thus, the definition of privacy-preserving monotonicity can describe the trade-off between privacy and utility, and it also shows that privacy-preserving level decreases as privacy budget increases and that data utility increases as privacy budget increases. Conversely, privacy-preserving monotonicity shows that privacy-preserving level increases as privacy budget decreases and that data utility decreases as privacy budget decreases.

6. Privacy-Preserving Monotonicity of Several Differential Privacy Mechanisms

Since privacy budget cannot directly show the privacy-preserving level of differential privacy mechanisms for sensitive data, the expected estimation error and entropy are used as privacy metrics to show privacy-preserving level. Thus, we analyze the privacy-preserving monotonicity of serval differential privacy mechanisms based on expected estimation error and entropy. Also, we analyze the utility monotonicity of these differential privacy mechanisms based on utility metrics of modulus of characteristic function and variant $1 - NE$ of the normalized entropy.

6.1. Privacy-Preserving Monotonicity Based on Expected Estimation Error

Theorem 1. *The EEE of Laplace mechanism decreases as privacy budget ε increases.*

Proof. The probability density function of Laplace distribution is

$$p(x) = \frac{1}{2b} \exp(-\frac{|x|}{b}) \tag{17}$$

where $b = \frac{\Delta f}{\varepsilon}$.

Then

$$EEE = \int_{-\infty}^{+\infty} \frac{1}{2b} \exp(-\frac{|x|}{b})|x|dx = -\int_{-\infty}^{0} \frac{1}{2b} \exp(\frac{x}{b})xdx + \int_{0}^{+\infty} \frac{1}{2b} \exp(\frac{-x}{b})xdx = b = \frac{\Delta f}{\varepsilon} \tag{18}$$

Therefore, $\frac{dEEE}{d\varepsilon} = -\frac{\Delta f}{\varepsilon^2} < 0$. The EEE of Laplace mechanism decreases as privacy budget ε increases. □

The characteristic functions of Laplace mechanism can refer to literature [64] for analysis of utility metric with privacy budget varying.

Corollary 1. *The modulus $|\phi(t)|$ of characteristic function of Laplace mechanism increases as privacy budget ε increases.*

Proof. Since the characteristic function of Laplace mechanism is

$$\phi(t) = \frac{\varepsilon^2}{\varepsilon^2 + \Delta f^2 t^2} \tag{19}$$

the modulus of characteristic function of Laplace mechanism is

$$|\phi(t)| = \frac{\varepsilon^2}{\varepsilon^2 + \Delta f^2 t^2} \tag{20}$$

Thus, we can get

$$\frac{d(|\phi(t)|)}{d\varepsilon} = \frac{2\varepsilon\Delta f^2 t^2}{(\varepsilon^2 + \Delta f^2 t^2)^2} \geq 0 \tag{21}$$

Therefore, the modulus $|\phi(t)|$ of characteristic function of Laplace mechanism increases as privacy budget ε increases. □

By Theorem 1, the expected estimation error of Laplace mechanism decreases as privacy budget increases. The privacy-preserving level of Laplace mechanism decreases as privacy budget increases. Corollary 1 shows that the modulus of characteristic function of Laplace mechanism increases as privacy budget increases. Thus, the data utility using Laplace mechanism increases as privacy budget increases. According to Definition 8, Theorem 1 shows the privacy-preserving monotonicity of Laplace mechanism. Privacy-preserving monotonicity of Laplace mechanism shows the trade-off between privacy and utility according to the Theorem 1 and Corollary 1.

Theorem 2. *The EEE of Gaussian mechanism decreases as privacy budget ε increases.*

Proof. The probability density function of Gaussian distribution is

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2}) \tag{22}$$

where $\sigma = \frac{\Delta_2 f\sqrt{2\ln(\frac{1.25}{\delta})}}{\varepsilon}$.

Then

$$EEE = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2})|x|dx = \frac{\sqrt{2}\sigma}{\sqrt{\pi}} = \frac{2\Delta_2 f\sqrt{\ln(\frac{1.25}{\delta})}}{\varepsilon\sqrt{\pi}} \tag{23}$$

Therefore, $\frac{dEEE}{d\varepsilon} = -\frac{2\Delta_2 f\sqrt{\ln(\frac{1.25}{\delta})}}{\varepsilon^2\sqrt{\pi}} < 0$. The EEE of Gaussian mechanism decreases as privacy budget ε increases. □

The characteristic functions of Gaussian mechanism can refer to literature [64] for analysis of utility metric with privacy budget varying.

Corollary 2. *The modulus $|\phi(t)|$ of characteristic function of Gaussian mechanism decreases as privacy budget ε increases.*

Proof. Since the characteristic function of Gaussian mechanism is

$$\phi(t) = \exp(-\frac{\Delta_2 f^2 t^2 \ln(\frac{1.25}{\delta})}{\varepsilon^2}) \tag{24}$$

the modulus of characteristic function of Gaussian mechanism is

$$|\phi(t)| = \exp(-\frac{\Delta_2 f^2 t^2 \ln(\frac{1.25}{\delta})}{\varepsilon^2}) \tag{25}$$

Thus, we can get

$$\frac{d(|\phi(t)|)}{d\varepsilon} = \frac{2\Delta_2 f^2 t^2 \ln(\frac{1.25}{\delta})}{\varepsilon^3} \exp(-\frac{\Delta_2 f^2 t^2 \ln(\frac{1.25}{\delta})}{\varepsilon^2}) \geq 0 \tag{26}$$

Therefore, the modulus $|\phi(t)|$ of characteristic function of Gaussian mechanism increases as privacy budget ε increases. □

According to the Theorem 2, the expected estimation error of Gaussian mechanism decreases as privacy budget increases. The privacy-preserving level of Gaussian mechanism decreases as privacy budget increases. By Corollary 2, the modulus of characteristic function of Gaussian mechanism increases as privacy budget increases. Thus, the data utility using Gaussian mechanism increases as privacy budget increases. According to Definition 8, Theorem 2 demonstrates the privacy-preserving monotonicity of Gaussian mechanism. Privacy-preserving monotonicity of Gaussian mechanism demonstrates the trade-off between privacy and utility based on Theorem 2 and Corollary 2.

6.2. Privacy-Preserving Monotonicity Based on Entropy

To prove privacy-preserving monotonicity of exponential mechanism, we introduce the following Cauchy-Schwarz inequality [65].

Lemma 1 (Cauchy-Schwarz Inequality)**.** *Let* $(a_1, a_2, \ldots, a_n)$ *and* $(b_1, b_2, \ldots, b_n)$ *be two sequences of real numbers. Then*

$$(\sum_{i=1}^{n} a_i b_i)^2 \geq \sum_{i=1}^{n} a_i^2 \times \sum_{i=1}^{n} b_i^2 \tag{27}$$

with equality if and only if there is a $k \in \mathbb{R}$ *such that* $a_i = kb_i$ *for each* $i \in \{1, \ldots, n\}$.

Theorem 3. *The ENT of output of exponential mechanism decreases as privacy budget* ε *increases.*

Proof. The exponential mechanism $\mathcal{M}_E(x, u, \mathcal{R})$ selects and outputs an element $r \in \mathcal{R}$ with probability proportional to $exp(\frac{\varepsilon u(x,r)}{2\Delta u})$ to utility function $u : \mathbb{N}^{|\mathcal{X}|} \times \mathcal{R} \to \mathbb{R}$. For convenience, we set $u_i = \frac{u(x,r_i)}{2\Delta u}$.

Let

$$H_i(\varepsilon) = -\frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \log_2 \frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \tag{28}$$

Thus, the entropy of output is

$$H(\varepsilon) = -\sum_{i=1}^{n} H_i(\varepsilon) = -\sum_{i=1}^{n} \frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \log_2 \frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \tag{29}$$

We have

$$\frac{dH_i}{d\varepsilon} = -\frac{1}{\ln 2} \frac{e^{\varepsilon u_i} \sum_{j=1}^{n} (u_i - u_j) e^{\varepsilon u_j}}{(\sum_{j=1}^{n} e^{\varepsilon u_j})^2} (\varepsilon u_i - \ln \sum_{j=1}^{n} e^{\varepsilon u_j} + 1) \tag{30}$$

and

$$\begin{aligned} \frac{dH}{d\varepsilon} &= \sum_{i=1}^{n} \frac{dH_i}{d\varepsilon} \\ &= -\sum_{i=1}^{n} \frac{e^{\varepsilon u_i} \sum_{j=1}^{n} (u_i - u_j) e^{\varepsilon u_j}}{(\sum_{j=1}^{n} e^{\varepsilon u_j})^2 \ln 2} (\varepsilon u_i - \ln \sum_{j=1}^{n} e^{\varepsilon u_j} + 1) \\ &= -\frac{1}{(\sum_{j=1}^{n} e^{\varepsilon u_j})^2 \ln 2} (\sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} \varepsilon u_i (u_i - u_j) e^{\varepsilon u_j} + \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} (u_i - u_j) e^{\varepsilon u_j} (1 - \ln \sum_{j=1}^{n} e^{\varepsilon u_j})) \end{aligned} \tag{31}$$

Since

$$\begin{aligned} \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} (u_i - u_j) e^{\varepsilon u_j} &= \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} (u_i e^{\varepsilon u_j} - u_j e^{\varepsilon u_j}) = \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} u_i e^{\varepsilon u_j} - \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} u_j e^{\varepsilon u_j} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} e^{\varepsilon u_i} u_i e^{\varepsilon u_j} - \sum_{i=1}^{n} \sum_{j=1}^{n} e^{\varepsilon u_i} u_j e^{\varepsilon u_j} = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i e^{\varepsilon u_i} e^{\varepsilon u_j} - \sum_{i=1}^{n} \sum_{j=1}^{n} u_i e^{\varepsilon u_i} e^{\varepsilon u_j} \\ &= 0 \end{aligned} \tag{32}$$

we have

$$\begin{aligned}\frac{dH(\varepsilon)}{d\varepsilon} &= -\frac{1}{(\sum_{j=1}^{n} e^{\varepsilon u_j})^2 \ln 2} \sum_{i=1}^{n} e^{\varepsilon u_i} \sum_{j=1}^{n} \varepsilon u_i (u_i - u_j) e^{\varepsilon u_j} \\ &= -\frac{\varepsilon}{(\sum_{j=1}^{n} e^{\varepsilon u_j})^2 \ln 2} (\sum_{i=1}^{n} u_i^2 e^{\varepsilon u_i} \sum_{i=1}^{n} e^{\varepsilon u_i} - (\sum_{i=1}^{n} u_i e^{\varepsilon u_i})^2)\end{aligned} \tag{33}$$

According to Lemma 1, we get

$$\sum_{i=1}^{n} u_i^2 e^{\varepsilon u_i} \sum_{i=1}^{n} e^{\varepsilon u_i} = \sum_{i=1}^{n} (\sqrt{u_i^2 e^{\varepsilon u_i}})^2 \sum_{i=1}^{n} (\sqrt{e^{\varepsilon u_i}})^2 \geq (\sum_{i=1}^{n} (\sqrt{u_i^2 e^{\varepsilon u_i}} \sqrt{e^{\varepsilon u_i}}))^2 = (\sum_{i=1}^{n} u_i e^{\varepsilon u_i})^2 \tag{34}$$

Therefore, $\frac{dH(\varepsilon)}{d\varepsilon} \leq 0$. The ENT of output of exponential mechanism decreases as privacy budget ε increases. □

Corollary 3. *The utility metric* $1 - NE$ *of output of exponential mechanism increases as privacy budget* ε *increases.*

Proof. According to Equation (29), the normalized entropy of output of exponential mechanism is

$$NE = \frac{H(\varepsilon)}{\log_2 n} = -\frac{1}{\log_2 n} \sum_{i=1}^{n} \frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \log_2 \frac{e^{\varepsilon u_i}}{\sum_{j=1}^{n} e^{\varepsilon u_j}} \tag{35}$$

Since $\frac{dH(\varepsilon)}{d\varepsilon} \leq 0$, the normalized entropy NE of output of exponential mechanism decreases as privacy budget ε increases. Therefore, the utility metric $1 - NE$ of output of exponential mechanism increases as privacy budget ε increases. □

By Theorem 3, the entropy of output of exponential mechanism decreases as privacy budget increases. The privacy-preserving level of output of exponential mechanism decreases as privacy budget increases. According to the Corollary 3, the utility metric $1 - NE$ of output of exponential mechanism increases as privacy budget increases. Thus, the data utility using exponential mechanism increases as privacy budget increases. According to Definition 8, Theorem 3 shows the privacy-preserving monotonicity of exponential mechanism. Privacy-preserving monotonicity shows the trade-off between privacy and utility based on Theorem 3 and Corollary 3.

Theorem 4. *The ENT of output of bivariate randomized response mechanism decreases as privacy budget* ε *increases.*

Proof. For any possible output $\xi \in \{0, 1\}$ and any input $b \in \{0, 1\}$, the conditional probability matrix of the bivariate randomized response mechanism is Equation (12). Let π_1 denotes the proportion of value 1 in $\{0, 1\}$. Then, the proportion of value 0 is $\pi_0 = 1 - \pi_1$.

Then, we have

$$P(\xi = 0) = (1 - \pi_1) \frac{e^{\varepsilon}}{1 + e^{\varepsilon}} + \pi_1 \frac{1}{1 + e^{\varepsilon}} = \frac{(1 - \pi_1) e^{\varepsilon} + \pi_1}{1 + e^{\varepsilon}} \tag{36}$$

and

$$P(\xi = 1) = (1 - \pi_1) \frac{1}{1 + e^{\varepsilon}} + \pi_1 \frac{e^{\varepsilon}}{1 + e^{\varepsilon}} = \frac{1 - \pi_1 + \pi_1 e^{\varepsilon}}{1 + e^{\varepsilon}} \tag{37}$$

Thus, we get

$$\begin{aligned} H_{\tilde{\xi}}(\varepsilon) &= -P(\xi=0)\log_2 P(\xi=0) - P(\xi=1)\log_2 P(\xi=1) \\ &= -\frac{(1-\pi_1)e^{\varepsilon}+\pi_1}{1+e^{\varepsilon}}\log_2\frac{(1-\pi_1)e^{\varepsilon}+\pi_1}{1+e^{\varepsilon}} - \frac{(1-\pi_1)+\pi_1 e^{\varepsilon}}{1+e^{\varepsilon}}\log_2\frac{(1-\pi_1)+\pi_1 e^{\varepsilon}}{1+e^{\varepsilon}} \end{aligned} \tag{38}$$

and

$$\frac{dH_{\tilde{\xi}}(\varepsilon)}{d\varepsilon} = \frac{(2\pi_1-1)e^{\varepsilon}}{(1+e^{\varepsilon})^2}\log_2\frac{e^{\varepsilon}(1-\pi_1)+\pi_1}{\pi_1 e^{\varepsilon}+1-\pi_1} \tag{39}$$

Since $\pi_1 e^{\varepsilon}+1-\pi_1 > 0$, we have

$$e^{\varepsilon}(1-\pi_1)+\pi_1 > \pi_1 e^{\varepsilon}+1-\pi_1 \Leftrightarrow 2\pi_1(1-e^{\varepsilon}) > 1-e^{\varepsilon} \tag{40}$$

Obviously, $1-e^{\varepsilon} < 0$. Thus, we have

$$2\pi_1(1-e^{\varepsilon}) > 1-e^{\varepsilon} \Leftrightarrow \pi_1 < \frac{1}{2} \tag{41}$$

Similarly, $0 < \frac{e^{\varepsilon}(1-\pi_1)+\pi_1}{\pi_1 e^{\varepsilon}+1-\pi_1} < 1 \Leftrightarrow \frac{1}{2} \le \pi_1 \le 1$.

Therefore, when $0 \le \pi_1 \le \frac{1}{2}$, we have $2\pi_1 - 1 \le 0$ and $\log_2 \frac{e^{\varepsilon}(1-\pi_1)+\pi_1}{\pi_1 e^{\varepsilon}+1-\pi_1} \ge 0$, and so we can get $\frac{dH_{\tilde{\xi}}(\varepsilon)}{d\varepsilon} \le 0$. When $\frac{1}{2} \le \pi_1 \le 1$, we have $2\pi_1 - 1 \ge 0$ and $\log_2 \frac{e^{\varepsilon}(1-\pi_1)+\pi_1}{\pi_1 e^{\varepsilon}+1-\pi_1} \le 0$, and so we can obtain $\frac{dH_{\tilde{\xi}}(\varepsilon)}{d\varepsilon} \le 0$. The ENT of output of bivariate randomized response mechanism decreases as privacy budget ε increases. □

Corollary 4. *The utility metric $1-NE$ of output of bivariate randomized response mechanism increases as privacy budget ε increases.*

Proof. According to Equation (38), the normalized entropy of output of bivariate randomized response mechanism is

$$NE = H_{\tilde{\xi}}(\varepsilon) = -\frac{(1-\pi_1)e^{\varepsilon}+\pi_1}{1+e^{\varepsilon}}\log_2\frac{(1-\pi_1)e^{\varepsilon}+\pi_1}{1+e^{\varepsilon}} - \frac{(1-\pi_1)+\pi_1 e^{\varepsilon}}{1+e^{\varepsilon}}\log_2\frac{(1-\pi_1)+\pi_1 e^{\varepsilon}}{1+e^{\varepsilon}} \tag{42}$$

Since $\frac{dH_{\tilde{\xi}}(\varepsilon)}{d\varepsilon} \le 0$, the normalized entropy NE of output of bivariate randomized response mechanism decreases as privacy budget ε increases. Therefore, the utility metric $1-NE$ of output of bivariate randomized response mechanism increases as privacy budget ε increases. □

According to the Theorem 4, the entropy of output of bivariate randomized response mechanism decreases as privacy budget increases. The privacy-preserving level of output of bivariate randomized response mechanism decreases as privacy budget increases. By Corollary 4, the utility metric of $1-NE$ of output of bivariate randomized response mechanism increases as privacy budget increases. Thus, the data utility using bivariate randomized response mechanism increases as privacy budget increases. According to Definition 8, Theorem 4 demonstrates the privacy-preserving monotonicity of bivariate randomized response mechanism. Privacy-preserving monotonicity shows the trade-off between privacy and utility based on Theorem 4 and Corollary 4.

We introduce the following Chebyshev inequality [65] to prove the privacy-preserving monotonicity of multivariate randomized response mechanism.

Lemma 2 (Chebyshev Inequality). *Let* $(a_1, a_2, \ldots, a_n)$ *and* $(b_1, b_2, \ldots, b_n)$ *be two sequences of real numbers. If* $a_1 \leq a_2 \leq \ldots \leq a_n$, $b_1 \leq b_2 \leq \ldots \leq b_n$, *then*

$$\sum_{i=1}^{n} a_i b_i \geq \frac{1}{n} \sum_{i=1}^{n} a_i \times \sum_{i=1}^{n} b_i \tag{43}$$

Theorem 5. *The ENT of output of multivariate randomized response mechanism decreases as privacy budget* ε *increases.*

Proof. For any possible output $\xi \in \{1, 2, \ldots, n\}$ and any input $b \in \{1, 2, \ldots, n\}$, the conditional probability matrix of the multivariate randomized response mechanism is Equation (13). Let $\pi_i (i = 1, 2, \ldots, n)$ denote the proportion of value i in $\{1, 2, \ldots, n\}$. Then, $\sum_{i=1}^{n} \pi_i = 1$. Moreover, $\sum_{j=1, j\neq i}^{n} \pi_i = 1 - \pi_i$ for $i \in \{1, 2, \ldots, n\}$.

For any $i \in \{1, 2, \ldots, n\}$, we have

$$P(\xi = i) = \pi_i \frac{e^{\varepsilon}}{n - 1 + e^{\varepsilon}} + \sum_{j=1, j\neq i}^{n} \pi_j \frac{1}{n - 1 + e^{\varepsilon}} = \frac{\pi_i e^{\varepsilon} + 1 - \pi_i}{n - 1 + e^{\varepsilon}} \tag{44}$$

Thus, we have

$$H_{\xi}(\varepsilon) = -\sum_{i=1}^{n} \frac{\pi_i e^{\varepsilon} + 1 - \pi_i}{n - 1 + e^{\varepsilon}} \log_2 \frac{\pi_i e^{\varepsilon} + 1 - \pi_i}{n - 1 + e^{\varepsilon}} \tag{45}$$

Since $\frac{d(\frac{\pi_i e^{\varepsilon} + 1 - \pi_i}{n - 1 + e^{\varepsilon}})}{d\varepsilon} = \frac{e^{\varepsilon}(n\pi_i - 1)}{(n - 1 + e^{\varepsilon})^2}$, then

$$\begin{aligned} \frac{dH_{\xi}(\varepsilon)}{d\varepsilon} = -\frac{e^{\varepsilon}}{(n - 1 + e^{\varepsilon})^2} [&\sum_{i=0}^{n} (n\pi_i - 1) \log_2 (\pi_i e^{\varepsilon} + 1 - \pi_i) \\ &- \sum_{i=1}^{n} (n\pi_i - 1) \log_2 (n - 1 + e^{\varepsilon}) - \frac{1}{\ln 2} \sum_{i=1}^{n} (n\pi_i - 1)] \end{aligned} \tag{46}$$

Here, $\sum_{i=1}^{n} (n\pi_i - 1) = \sum_{i=1}^{n} n\pi_i - \sum_{i=1}^{n} 1 = n \sum_{i=1}^{n} \pi_i - n = n \times 1 - n = 0$. Thus, we have

$$\frac{dH_{\xi}(\varepsilon)}{d\varepsilon} = -\frac{ne^{\varepsilon}}{(n - 1 + e^{\varepsilon})^2} [\sum_{i=0}^{n} \pi_i \log_2 (\pi_i e^{\varepsilon} + 1 - \pi_i) - \frac{1}{n} \sum_{i=1}^{n} \log_2 (\pi_i e^{\varepsilon} + 1 - \pi_i)] \tag{47}$$

Let $a_i = \pi_i$ and $b_i = \log_2 (\pi_i (e^{\varepsilon} - 1) + 1)$. Since the addition has the commutative law, we assume $a_1 \leq a_2 \leq \cdots \leq a_n$. Moreover, the function $\log_2 x$ is monotonically increasing and $e^{\varepsilon} - 1$ is also monotonically increasing, we have $b_1 \leq b_2 \leq \cdots \leq b_n$. According to Lemma 2, we can get

$$\frac{\sum_{i=0}^{n} \pi_i \log_2 (\pi_i e^{\varepsilon} + 1 - \pi_i)}{n} \geq \frac{\sum_{i=1}^{n} \pi_i}{n} \times \frac{\sum_{i=1}^{n} \log_2 (\pi_i (e^{\varepsilon} - 1) + 1)}{n} \tag{48}$$

Thus,

$$\sum_{i=0}^{n} \pi_i \log_2 (\pi_i e^{\varepsilon} + 1 - \pi_i) \geq \frac{\sum_{i=1}^{n} \log_2 (\pi_i (e^{\varepsilon} - 1) + 1)}{n} \tag{49}$$

Therefore, $\frac{dH_{\xi}(\varepsilon)}{d\varepsilon} \leq 0$. The *ENT* of output of multivariate randomized response mechanism decreases as privacy budget ε increases. □

Corollary 5. *The utility metric* $1 - NE$ *of output of multivariate randomized response mechanism increases as privacy budget* ε *increases.*

Proof. According to Equation (45), the normalized entropy of output of multivariate randomized response mechanism is

$$NE = \frac{H_{\tilde{\xi}}(\varepsilon)}{\log_2 n} = -\frac{1}{\log_2 n}\sum_{i=1}^{n}\frac{\pi_i e^{\varepsilon}+1-\pi_i}{n-1+e^{\varepsilon}}\log_2\frac{\pi_i e^{\varepsilon}+1-\pi_i}{n-1+e^{\varepsilon}} \tag{50}$$

Since $\frac{dH_{\tilde{\xi}}(\varepsilon)}{d\varepsilon} \leq 0$, the normalized entropy NE of output of multivariate randomized response mechanism decreases as privacy budget ε increases. Therefore, the utility metric $1-NE$ of output of multivariate randomized response mechanism increases as privacy budget ε increases. □

By Theorem 5, the entropy of output of multivariate randomized response mechanism decreases as privacy budget increases. The privacy-preserving level of output of multivariate randomized response mechanism decreases as privacy budget increases. According to the Corollary 5, the utility metric of output of multivariate randomized response mechanism increases as privacy budget increases. Thus, the data utility using multivariate randomized response mechanism increases as privacy budget increases. According to Definition 8, Theorem 5 shows the privacy-preserving monotonicity of multivariate randomized response mechanism. Privacy-preserving monotonicity of multivariate randomized response mechanism demonstrates the trade-off between privacy and utility based on Theorem 5 and Corollary 5.

Furthermore, Laplace mechanism and Gaussian mechanism are applicable for categorical data based on the entropy of continuous random variables. Since the entropy of output of Laplace mechanism and Gaussian mechanism may be negative, we do not consider the variant $1-NE$ of normalized entropy as the utility metric of Laplace mechanism and Gaussian mechanism. We only use the modulus of characteristic function as the utility metric of Laplace mechanism and Gaussian mechanism. Thus, we can get the following theorems.

Theorem 6. *The ENT of output of Laplace mechanism decreases as privacy budget ε increases.*

Proof. The entropy of output of Laplace mechanism is

$$ENT = -\int_{-\infty}^{+\infty} p(x)\log_2 p(x)dx = \frac{1}{\ln 2}(\ln(2b)+1) = \frac{1}{\ln 2}(\ln(\frac{2\Delta_1 f}{\varepsilon})+1) \tag{51}$$

Therefore, $\frac{dENT}{d\varepsilon} = -\frac{1}{\varepsilon \ln 2} < 0$. The ENT of output of Laplace mechanism decreases as privacy budget ε increases. □

According to the Theorem 6, the entropy of output of Laplace mechanism decreases as privacy budget increases. The privacy-preserving level of output of Laplace mechanism decreases as privacy budget increases. According to Definition 8, Theorem 6 demonstrates the privacy-preserving monotonicity of Laplace mechanism. Similarly, privacy-preserving monotonicity of Laplace mechanism demonstrates the trade-off between privacy and utility based on Theorem 6 and Corollary 1.

Theorem 7. *The ENT of output of Gaussian mechanism decreases as privacy budget ε increases.*

Proof. The entropy of output of Gaussian mechanism is

$$ENT = -\int_{-\infty}^{+\infty} p(x)\log_2 p(x)dx = \frac{1}{\ln 2}(\ln(\sqrt{2\pi}\sigma)+\frac{1}{2}) = \frac{1}{\ln 2}(\ln(\frac{2\Delta_2 f\sqrt{\pi\ln(\frac{1.25}{\delta})}}{\varepsilon})+\frac{1}{2}) \tag{52}$$

Therefore, $\frac{dENT}{d\varepsilon} = -\frac{1}{\varepsilon \ln 2} < 0$. The ENT of output of Gaussian mechanism decreases as privacy budget ε increases. □

By Theorem 7, the entropy of output of Gaussian mechanism decreases as privacy budget increases. The privacy-preserving level of output of Gaussian mechanism decreases as privacy budget increases. According to Definition 8, Theorem 7 demonstrates the privacy-preserving monotonicity of Gaussian mechanism. Similarly, privacy-preserving monotonicity of Gaussian mechanism demonstrates the trade-off between privacy and utility based on Theorem 7 and Corollary 2.

7. Numerical Results

In this section, we conduct numerical experiments for these conclusions of Section 6. We conduct these numerical experiments by implementing them with MATLAB (R2013b) and run our experiments on a desktop computer with Intel i5-2400 3.10 GHz processor, 4GB RAM, and Windows 7 platform. The detailed parameter settings are shown in Table 2.

Table 2. Parameter setting.

Mechanism	Parameter	Value
Laplace mechanism	Δf	$\Delta f = 1, \Delta f = 2, \Delta f = 3$
	t	$t = 1$
Gaussian mechanism	δ	$\delta = 0.1$
	$\Delta_2 f$	$\Delta_2 f = 1, \Delta_2 f = 2, \Delta_2 f = 3$
	t	$t = 1$
Exponential mechanism	u_1, u_2, u_3, u_4, u_5	$u_1 = 2, u_2 = 2, u_3 = 2, u_4 = 2, u_5 = 2$
		$u_1 = 1, u_2 = 3, u_3 = 2, u_4 = 3, u_5 = 1$
		$u_1 = 1, u_2 = 4, u_3 = 3, u_4 = 1, u_5 = 1$
Bivariate randomized response	π_1	$\pi_1 = 0.5, \pi_1 = 0.7, \pi_1 = 0.9$
	ε	$\varepsilon = 0.1, \varepsilon = 0.2, \varepsilon = 0.3$
Multivariate randomized response	$\pi_1, \pi_2, \pi_3, \pi_4$	$\pi_1 = 0.25, \pi_2 = 0.25, \pi_3 = 0.25, \pi_4 = 0.25$
		$\pi_1 = 0.1, \pi_2 = 0.2, \pi_3 = 0.3, \pi_4 = 0.4$
		$\pi_1 = 0.1, \pi_2 = 0.4, \pi_3 = 0.4, \pi_4 = 0.1$

Δf: ℓ_1-sensitivity of Laplace mechanism. t: Parameter of characteristic function of Laplace mechanism and Gaussian mechanism. δ: Gaussian mechanism is ε-differential privacy with probability at least $1 - \delta$ for all adjacent databases x and y. $\Delta_2 f$: ℓ_2-sensitivity of Gaussian mechanism. u_1, u_2, u_3, u_4, u_5: Utility scores of four arbitrary outputs of exponential mechanism. π_1: Proportion of input 1 of bivariate randomized response mechanism. ε: Entropy of output of bivariate randomized response mechanism for different π_1 under a fixed privacy budget ε. $\pi_1, \pi_2, \pi_3, \pi_4$: Probabilities of input 1, 2, 3, 4 of multivariate randomized response mechanism.

In Figure 3, we can observe that the expected estimation error decreases as privacy budget ε increases when using Laplace mechanism for numeric data. This is consistent with Theorem 1. Figure 4 shows that the modulus of characteristic function of query results of Laplace mechanism increases as privacy budget increases. This verifies Corollary 1. In Figure 5, we know that the expected estimation error of query results using Gaussian mechanism decreases as privacy budget ε increases, which is consistent with Theorem 2. Conversely, we observe that the modulus of characteristic function of query results of Gaussian mechanism increases as privacy budget increases from Figure 6. This further demonstrates Corollary 2. We see that the expected estimation error increases as sensitivity increases to Laplace mechanism and Gaussian mechanism from Figures 3 and 5. We can also conclude that expected estimation error increases as sensitivity increases to Laplace mechanism and Gaussian mechanism from Equations (18) and (23). Conversely, we observe that the modulus of characteristic function of query results using Laplace mechanism and Gaussian mechanism decreases as sensitivity increases from Figures 4 and 6. We can also conclude that the modulus of characteristic function of query results using Laplace mechanism and Gaussian mechanism decreases as sensitivity increases from Equations (20) and (25).

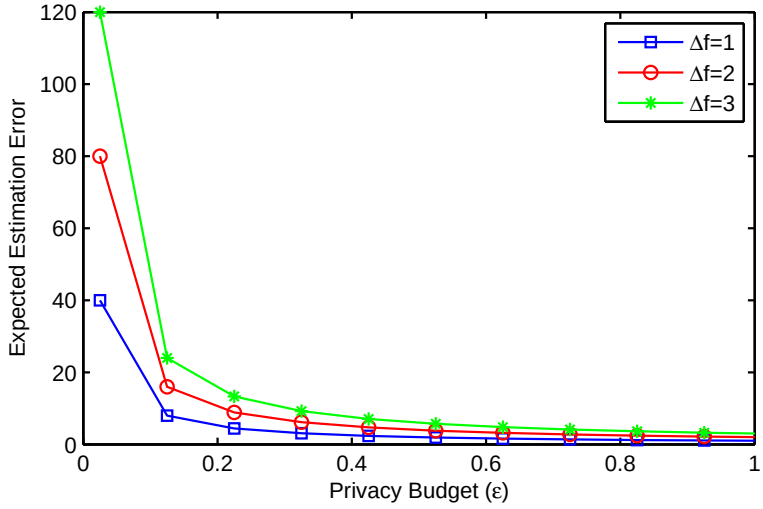

Figure 3. Expected estimation error of query results of Laplace mechanism.

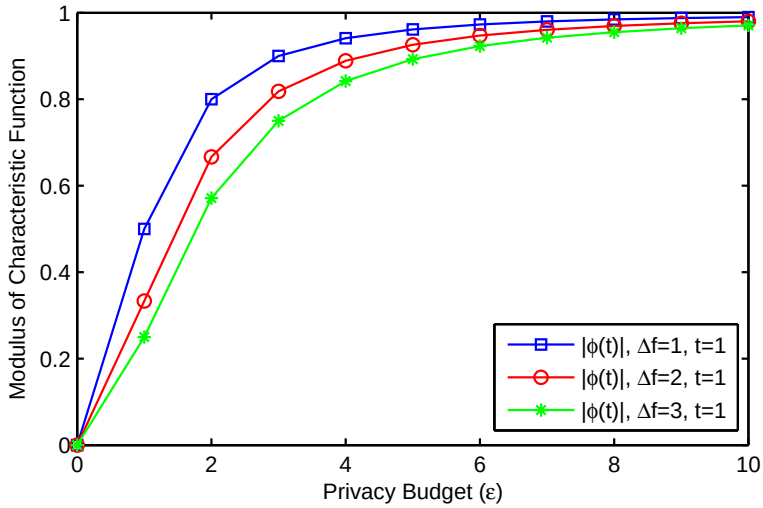

Figure 4. Modulus of characteristic function of query results of Laplace mechanism.

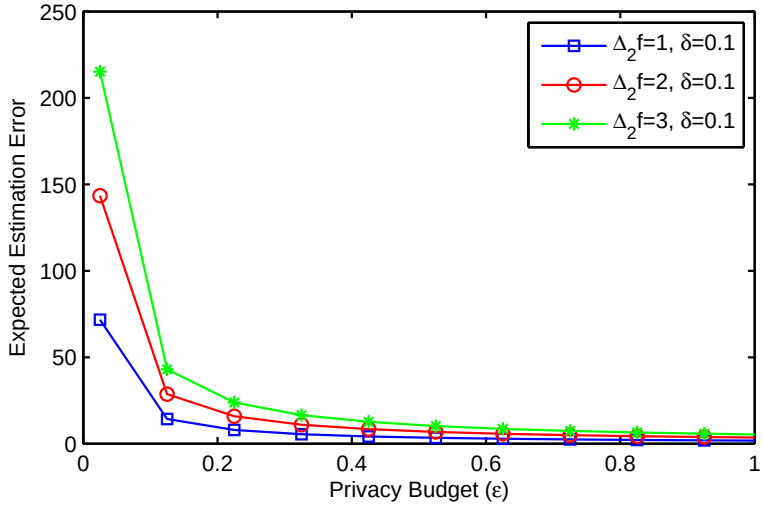

Figure 5. Expected estimation error of query results of Gaussian mechanism.

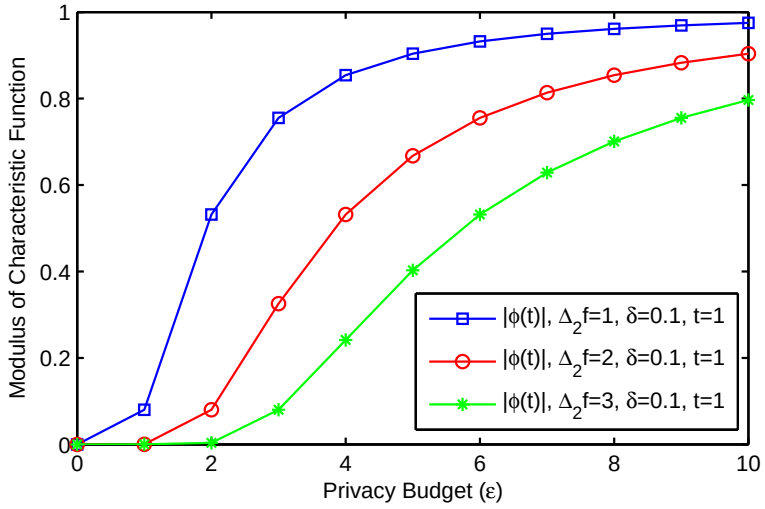

Figure 6. Modulus of characteristic function of query results of Gaussian mechanism.

In Figure 7, we can get the monotonically decreasing curve of entropy of output using exponential mechanism. The result verifies Theorem 3. Figure 8 shows that the utility metric $1 - NE$ of data using exponent mechanism increases as privacy budget increases. This result is covered in Corollary 3. By Theorem 3, when $\varepsilon = 0$ or $u_1(x, r_1) = u_2(x, r_2) = \ldots = u_n(x, r_n) = 0$, $H(\varepsilon)$ has the maximal value $\log_2 n$. At the same time, all random variables have the maximum uncertainty. Thus, this leads to the maximum entropy and gets the best privacy preservation. At the same time, the utility metric $1 - NE$ is equal to 0 in Figure 8, which results in the worst availability of data.

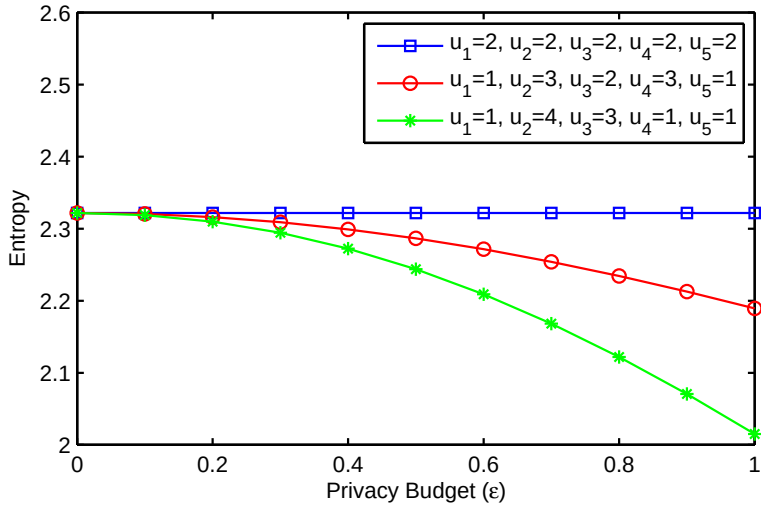

Figure 7. Entropy of output of exponential mechanism.

Entropy of output using bivariate randomized response mechanism decreases as privacy budget ε increases from Figure 9. This illustrates Theorem 4. In Figure 10, the utility metric $1 - NE$ of data using bivariate randomized response mechanism increases as privacy budget increases. This verifies Corollary 4. In Figure 11, we observe that $H_{\tilde{\zeta}}(\pi_1)$ (Equation (38)) monotonically increases as π_1 increases to bivariate randomized response mechanism, when $0 \leq \pi_1 \leq \frac{1}{2}$. When $\frac{1}{2} \leq \pi_1 \leq 1$, $H_{\tilde{\zeta}}(\pi_1)$ is monotonically decreasing as π_1 increases. Thus, $H_{\tilde{\zeta}}(\pi_1)$ has the maximal value 1 when $\pi_1 = \frac{1}{2}$. $H_{\tilde{\zeta}}(\pi_1)$ is symmetry on $\pi_1 = \frac{1}{2}$. That is to say, $H_{\tilde{\zeta}}(\pi_1) = H_{\tilde{\zeta}}(1 - \pi_1)$ for all $\pi_1 \in [0, 1]$. This reason is $P_0(\pi_1) = P_1(1 - \pi_0)$ and $P_1(\pi_1) = P_0(1 - \pi_0)$. In Figure 10, the utility metric $1 - NE$ of data using bivariate randomized response mechanism is equal to 0 when $\pi_1 = \frac{1}{2}$. Thus, this results in the worst availability of data.

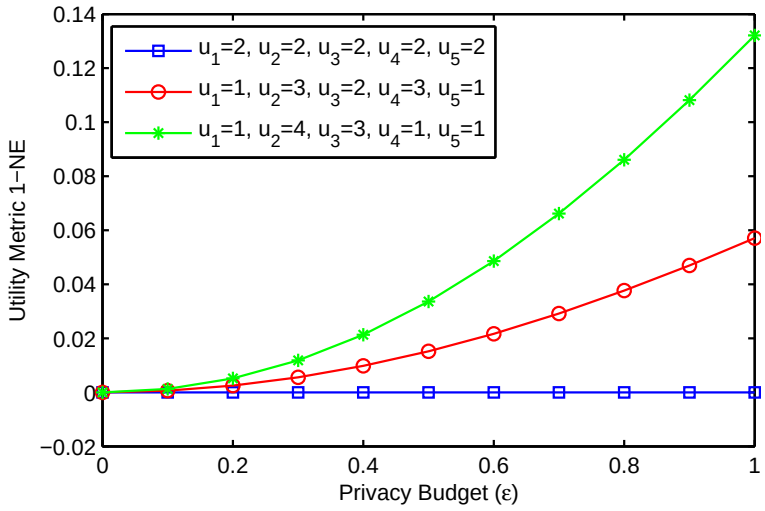

Figure 8. Utility metric $1 - NE$ of output of exponential mechanism.

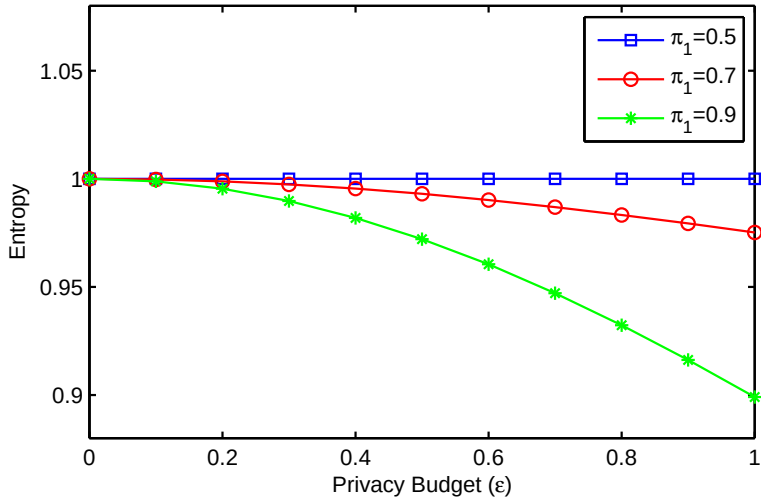

Figure 9. Entropy of output of bivariate randomized response mechanism.

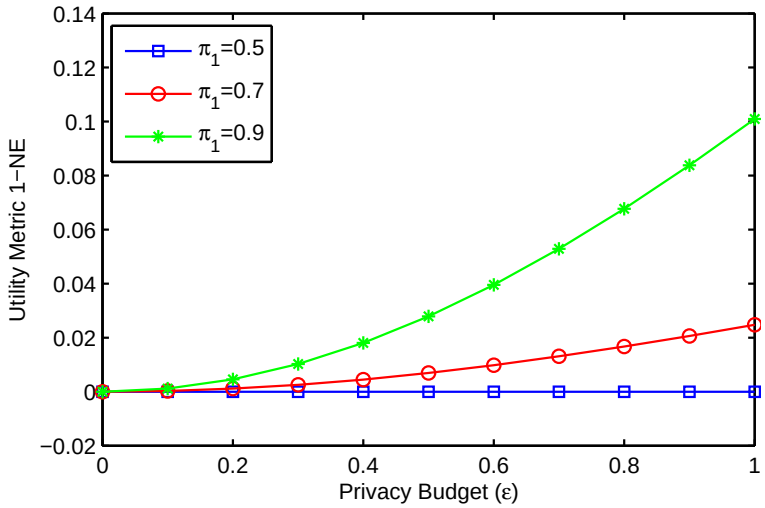

Figure 10. Utility metric $1 - NE$ of output of bivariate randomized response mechanism.

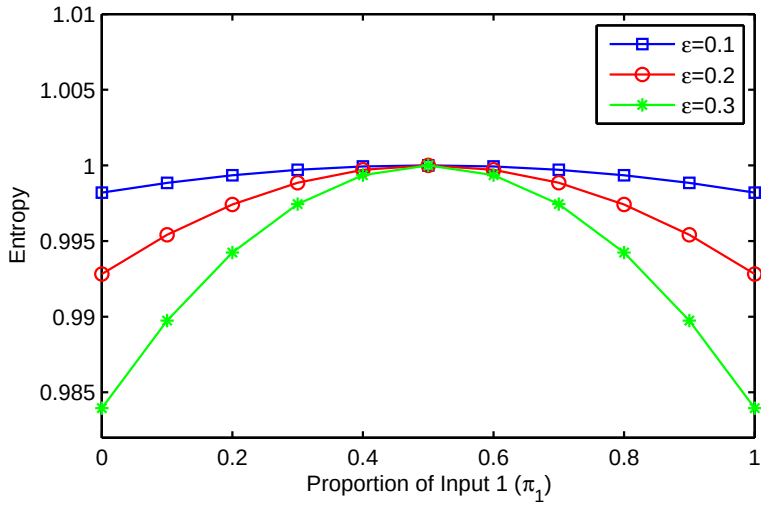

Figure 11. Entropy of output of bivariate randomized response mechanism under different π_1.

For privacy-preserving monotonicity of multivariate randomized response mechanism, we can also see that entropy of output decreases as privacy budget ε increases from Figure 12. This is consistent with Theorem 5. We can obtain that the utility metric $1 - NE$ of data using multivariate randomized response mechanism increases as privacy budget increases from Figure 13. This demonstrates Corollary 5. We can conclude that entropy of multivariate randomized response mechanism is maximal value $\log_2 n$ when $\pi_1 = \pi_2 = \ldots = \pi_n = \frac{1}{n}$ from Figure 12. At the same time, the utility metric $1 - NE$ of data using multivariate randomized response mechanism is equal to 0 in Figure 13, which results in the worst availability of data.

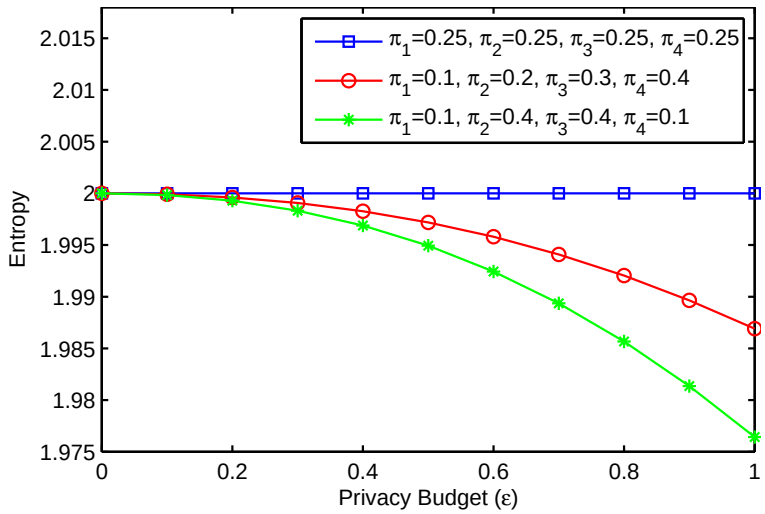

Figure 12. Entropy of output of multivariate randomized response mechanism.

We observe that the entropy of output using Laplace mechanism deceases as privacy budget ε increases from Figure 14, which conforms to Theorem 6. In Figure 15, the entropy of output using Gaussian mechanism decreases as privacy budget ε increases. The result further demonstrates Theorem 7. For Laplace mechanism and Gaussian mechanism, we can see that the entropy increases as sensitivity increases from Figures 14 and 15. Moreover, we also get that the entropy increases as sensitivity increases to Laplace mechanism and Gaussian mechanism from Equations (51) and (52). Conversely, the availability of data using Laplace mechanism and Gaussian mechanism decreases as sensitivity increases based on Equations (20) and (25).

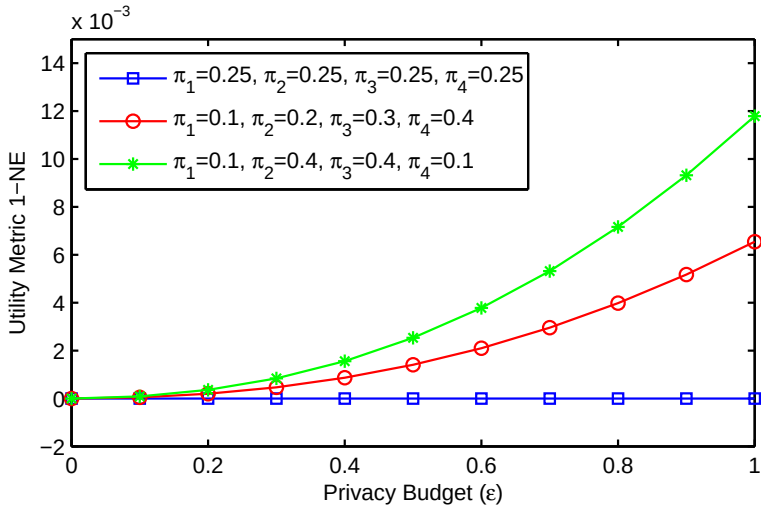

Figure 13. Utility metric $1 - NE$ of output of multivariate randomized response mechanism.

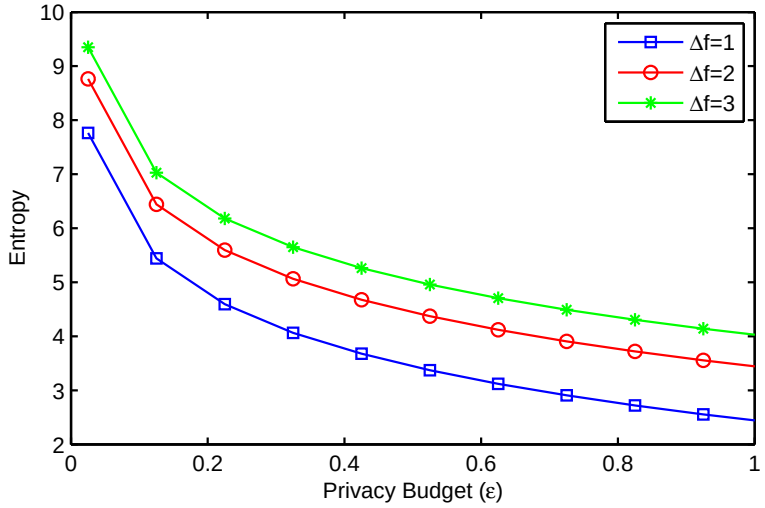

Figure 14. Entropy of output of Laplace mechanism.

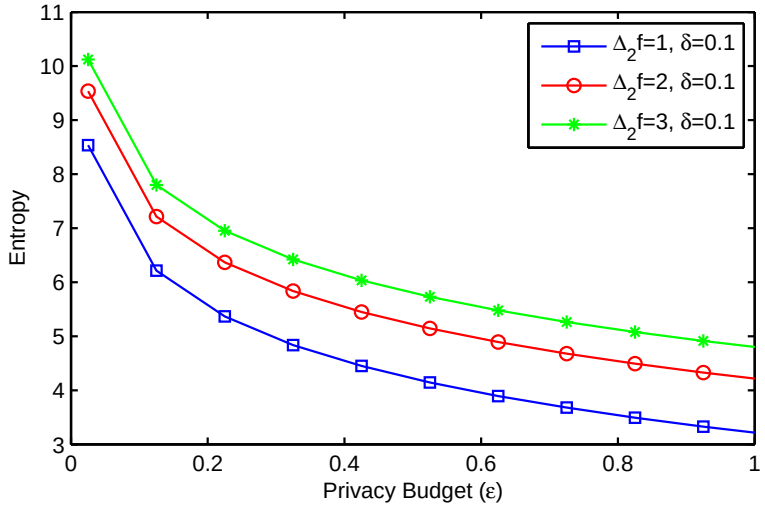

Figure 15. Entropy of output of Gaussian mechanism.

For the research and applications of differential privacy, we can theoretically and numerically show the privacy-preserving monotonicity of privacy metrics of the differential privacy mechanisms to measure the trade-off between privacy and utility. According to the theoretical and numerical

results, privacy-preserving monotonicity shows a double mean, then (i) the privacy metric decreases as privacy budget increases, and the utility metric increases as privacy budget increases, and (ii) the privacy metric increases as privacy budget decreases, and the utility metric decreases as privacy budget decreases. Thus, we can conclude that a differential privacy mechanism can reach a trade-off between privacy and utility by using privacy-preserving monotonicity.

8. Discussion

We have presented the definition of privacy-preserving monotonicity of differential privacy to formulate the trade-off between privacy and utility. Our theoretical and experimental results show that several existing differential privacy mechanisms can ensure privacy-preserving monotonicity. In the following discussion, we discuss the method of seeking trade-off under semi-honest model and analyze unilateral trade-off under rational model.

We have verified privacy-preserving monotonicity of differential privacy mechanisms. To provide a guideline on solving trade-off, we discuss the method of differential privacy under the semi-honest model. We denote the utility metric by using UM_ε under privacy budget ε. In Figure 2, since differential privacy mechanisms ensure privacy-preserving monotonicity, the specific privacy metric PM_{ε_0} and utility metric UM_{ε_0} have one-to-one relationship to privacy budget ε_0. Thus, data curator can ensure the trade-off between data privacy and data utility under the semi-honest model. The procedure is as follows.

- Data curator evaluates privacy metrics PM_{ε_0} under different privacy budget ε_0 for a differential privacy mechanism.
- Since data analyst is semi-honest, data curator chooses desirable privacy budget ε_0 according to the corresponding privacy metric PM_{ε_0}.
- Data curator adding noise to the queries results using differential privacy mechanism under desirable privacy budget ε_0.

From the above steps, a differential privacy mechanism can achieve the trade-off between privacy and utility under semi-honest model. Therefore, privacy-preserving monotonicity of formulating trade-off is beneficial to data curator under semi-honest model. However, we see the data are completely not available under a tiny privacy budget for data analyst. Data curator benefits to protect data privacy, but data analyst has no choice on the data utility. Thus, trade-off of a differential privacy mechanism may lead to utility disaster of engineering implementation under semi-honest model.

Thus, we need to consider the rational model. Next, we discuss whether the trade-off between privacy and utility exits under rational model. Rational data curator and rational data analyst hope to the maximum privacy-preserving level and the maximum data utility, respectively. However, the protection strategies of data curator and the analysis strategies of data analyst are mutually restrictive in non-interactive model or interactive model. Considering rational data curator and rational data analyst, both of them hope to get expected privacy-preserving and expected data utility. Thus, based on the privacy-preserving monotonicity of differential privacy mechanism, data curator can achieve expected privacy preservation, but this may lead to utility disaster. Similarly, data analyst can achieve expected data utility by negotiation with data curator, but this may lead to privacy leakage of data curator. Data curator and data analyst can only achieve unilateral trade-off under rational model. Therefore, the trade-off between privacy of data curator and utility of data analyst cannot be achieved under the rational model.

Privacy-preserving monotonicity shows the trade-off between privacy and utility under the semi-honest model. However, unilateral trade-off would create a unilateral benefit of the data curator or data analyst under rational model. As a result, unilateral trade-off boosts (or degrades) privacy-preserving of data curator and reduces (or improves) data utility of data analyst. Therefore, it is needed to modify differential privacy with multilateral negotiation, which can achieve expected privacy-preserving and expected data utility of differential privacy under rational model. This resolves

the restriction of unilateral trade-off and ensures the expected privacy-preserving of data curator and expected data utility of data analyst.

9. Conclusions

We proposed the definition of privacy-preserving monotonicity of differential privacy based on privacy metrics and utility metrics of this paper, which is a measurement of the trade-off between privacy and utility. According to the privacy metrics of expected estimation error and entropy, our theoretical and numerical results showed that several differential privacy mechanisms, such as Laplace mechanism, Gaussian mechanism, exponential mechanism, bivariate randomized response mechanism, and multivariate randomized response mechanism, ensure privacy-preserving monotonicity. In addition, we theoretically and numerically analyzed the utility monotonicity of these differential privacy mechanisms based on utility metrics of modulus of characteristic function and variant of normalized entropy. Finally, based on privacy-preserving monotonicity of differential privacy, we presented the method of seeking trade-off under a semi-honest model and discussed unilateral trade-off of differential privacy under rational model. Privacy-preserving monotonicity is helpful to explore privacy-preserving mechanisms requiring the trade-off between privacy and utility under the semi-honest model. However, privacy-preserving monotonicity resulting in the unilateral trade-off can lead to bad results of differential privacy under a rational model. In future work, we will resolve the restriction of unilateral trade-off by modifying differential privacy definitions using multilateral negotiation.

Author Contributions: H.L. wrote the main concepts of the manuscript and implemented numerical experiment, H.L. and Y.Z. contributed the formal analysis, H.L., Z.W., Y.Z., C.P., F.T. and L.L. discussed collaboratively the results, H.L., Z.W., Y.Z., C.P., F.T. and L.L. checked collaboratively the English writing and organization of the manuscript.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (No. 61602290, 61662009), the Natural Science Basic Research Program of Shaanxi Province (No. 2017JQ6038), the Fundamental Research Founds for the Central Universties (No. GK201704016, 2016CBY004, GK201603093, GK201501008, GK201402004), the Program of Key Science and Technology Innovation Team in Shaanxi Province (No. 2014KTC-18) and the Open Project Fund of Guizhou Provincial Key Laboratory of Public Big Data (No. 2017BDKFJJ026).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dwork, C.; Roth, A. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* **2014**, *9*, 211–407.
2. Hardt, M.; Rothblum, G.N. A multiplicative weights mechanism for privacy-preserving data analysis. In Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science, Las Vegas, NV, USA, 23–26 October 2010; pp. 61–70.
3. Chaudhuri, K.; Hsu, D. Convergence rates for differentially private statistical estimation. In Proceedings of the 29th International Conference on Machine Learning, Edinburgh, UK, 26 June–1 July 2012; pp. 1327–1334.
4. Blocki, J.; Blum, A.; Datta, A.; Sheffet, O. The johnson-lindenstrauss transform itself preserves differential privacy. In Proceedings of the 53th Annual IEEE Symposium on Foundations of Computer Science, New Brunswick, NJ, USA, 20–23 October 2012; pp. 410–419.
5. Liu, H.; Wu, Z.; Peng, C.; Tian, F.; Lu, L. Adaptive gaussian mechanism based on expected data utility under conditional filtering noise. *KSII Trans. Internet Inf. Syst.* **2018**, *12*, 3497–3515.
6. Dwork, C.; McSherry, F.; Nissim, K.; Smith, A. Calibrating noise to sensitivity in private data analysis. In Proceedings of the 3rd Theory of Cryptography Conference, New York, NY, USA, 4–7 March 2006; pp. 265–284.
7. Roth, A.; Roughgarden, T. Interactive privacy via the median mechanism. In Proceedings of the 42nd ACM Symposium on Theory of Computing, Cambridge, MA, USA, 5–8 June 2010; pp. 765–774.
8. McSherry, F.; Talwar, K. Mechanism design via differential privacy. In Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, Providence, RI, USA, 20–23 October 2007; pp. 94–103.

9. Warner, S.L. Randomized response: A survey technique for eliminating evasive answer bias. *J. Am. Stat. Assoc.* **1965**, *60*, 63–69.
10. Hardt, M.; Talwar, K. On the geometry of differential privacy. In Proceedings of the 42nd ACM Symposium on Theory of Computing, Cambridge, MA, USA, 5–8 June 2010; pp. 705–714.
11. Fouad, M.R.; Elbassioni, K.; Bertino, E. A supermodularity-based differential privacy preserving algorithm for data anonymization. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1591–1601.
12. Jorgensen, Z.; Yu, T.; Cormode, G. Conservative or liberal? Personalized differential privacy. In Proceedings of the 31th IEEE International Conference on Data Engineering, Seoul, Korea, 13–17 April 2015; pp. 1023–1034.
13. Soria-Comas, J.; Domingo-Ferrer, J.; Sánchez, D.; Megías, D. Individual differential privacy: A utility-preserving formulation of differential privacy guarantees. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 1418–1429.
14. Cuff, P.; Yu, L. Differential privacy as a mutual information constraint. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 43–54.
15. He, X.; Machanavajjhala, A.; Ding, B. Blowfish privacy: Tuning privacy-utility trade-offs using policies. In Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, Snowbird, UT, USA, 22–27 June 2014; pp. 1447–1458.
16. Goyal, V.; Mironov, I.; Pandey, O.; Sahai, A. Accuracy-privacy tradeoffs for two-party differentially private protocols. In Proceedings of the 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, 18–22 August 2013; pp. 298–315.
17. Kasiviswanathan, S.P.; Lee, H.K.; Nissim, K.; Raskhodnikova, S.; Smith, A. What can we learn privately? In Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, Philadelphia, PA, USA, 25–28 October 2008; pp. 793–826.
18. Dwork, C.; Rothblum, G.N.; Vadhan, S. Boosting and differential privacy. In Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science, Las Vegas, NV, USA, 23–26 October 2010; pp. 51–60.
19. Zhang, T.; Zhu, Q. Dynamic differential privacy for ADMM-based distributed classification learning. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 172–187.
20. Zeng, C.; Naughton, J.F.; Cai, J.Y. On differentially private frequent itemset mining. *Proc. VLDB Endow.* **2012**, *6*, 25–36.
21. Hong, Y.; Vaidya, J.; Lu, H.; Karras, P.; Goel, S. Collaborative search log sanitization: Toward differential privacy and boosted utility. *IEEE Trans. Dependable Secur. Comput.* **2015**, *12*, 504–518.
22. Dimitrakakis, C.; Nelson, B.; Zhang, Z.; Mitrokotsa, A.; Rubinstein, B.I.P. Differential privacy for Bayesian inference through posterior sampling. *J. Mach. Learn. Res.* **2017**, *18*, 343–381.
23. Sei, Y.; Ohsuga, A. Differential private data collection and analysis based on randomized multiple dummies for untrusted mobile crowdsensing. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 926–939.
24. Nikolov, A.; Talwar, K.; Zhang, L. The geometry of differential privacy: The small database and approximate cases. *SIAM J. Comput.* **2016**, *45*, 575–616.
25. Xu, J.; Zhang, Z.; Xiao, X.; Yang, Y.; Yu, G.; Winslett, M. Differentially private histogram publication. *VLDB J.* **2013**, *22*, 797–822.
26. Dewri, R. Local differential perturbations: Location privacy under approximate knowledge attackers. *IEEE Trans. Mob. Comput.* **2013**, *12*, 2360–2372.
27. Guo, T.; Luo, J.; Dong K.; Yang, M. Differentially private graph-link analysis based social recommendation. *Inf. Sci.* **2018**, *463–464*, 214–226.
28. Fan, L.; Bonomi, L.; Xiong, L.; Sunderam, V. Monitoring web browsing behavior with differential privacy. In Proceedings of the 23th International World Wide Web Conference, Seoul, Korea, 7–11 April 2014; pp. 177–188.
29. He, X.; Zhang, X.; Kuo, C.C.J. A distortion-based approach to privacy-preserving metering in smart grids. *IEEE Access* **2013**, *1*, 67–78.
30. Jin, X.; Zhang, R.; Chen, Y.; Zhang, Y. DPSense: Differentially private crowdsourced spectrum sensing. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016; pp. 296–307.

31. Huang, Z.; Kannan, S. The exponential mechanism for social welfare: Private, truthful, and nearly optimal. In Proceedings of the 53th Annual IEEE Symposium on Foundations of Computer Science, New Brunswick, NJ, USA, 20–23 October 2012; pp. 140–149.
32. Katz, J.; Lindell, Y. *Introduction to Modern Cryptography*, 2nd ed.; CRC Press: Boca Raon, FL, USA, 2014.
33. Ghosh, A.; Roughgarden, T.; Sundararajan, M. Universally utility-maximizing privacy mechanisms. *SIAM J. Comput.* **2012**, *41*, 1673–1693.
34. Xiao, X.; Bender, G.; Hay, M.; Gehrke, J. iReduct: Differential privacy with reduced relative errors. In Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, Athens, Greece, 12–16 June 2011; pp. 229–240.
35. Tossou, A.C.Y.; Dimitrakakis, C. Algorithms for differentially private multi-armed bandits. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 2087–2093.
36. Dodis, Y.; López-Alt, A.; Mironov, I.; Vadhan, S. Differential privacy with imperfect randomness. In Proceedings of the 32th Annual Cryptology Conference, Santa Barbara, CA, USA,19–23 August 2012; pp. 497–516.
37. Tramèr, F.; Huang, Z.; Hubaux, J.P.; Ayday, E. Differential privacy with bounded priors: Reconciling utility and privacy in genome-wide association studies. In Proceedings of the 22th ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015; pp. 1286–1297.
38. Sankar, L.; Rajagopalan, S.R.; Poor, H.V. Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 838–852.
39. Geumlek, J.; Song, S.; Chaudhuri, K. Renyi differential privacy mechanisms for posterior sampling. In Proceedings of the 31th Annual Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5289–5298.
40. Song, S.; Wang, Y.; Chaudhuri, K. Pufferfish privacy mechanisms for correlated data. In Proceedings of the 2017 ACM SIGMOD International Conference on Management of Data, Chicago, IL, USA, 14–19 May 2017; pp. 1291–1306.
41. Kairouz, P.; Oh, S.; Viswanath, P. Secure multi-party differential privacy. In Proceedings of the 29th Annual Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 2008–2016.
42. Kairouz, P.; Oh, S.; Viswanath, P. Extremal mechanisms for local differential privacy. *J. Mach. Learn. Res.* **2016**, *17*, 1–51.
43. Kairouz, P.; Bonawitz, K.; Ramage, D. Discrete distribution estimation under local privacy. In Proceedings of the 33th International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016.
44. Chaudhuri, K.; Monteleoni, C.; Sarwate, A.D. Differentially private empirical risk minimization. *J. Mach. Learn. Res.* **2011**, *12*, 1069–1109.
45. Iyer, A.S.; Nath, J.S.; Sarawagi, S. Privacy-preserving class ratio estimation. In Proceedings of the 22th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 925–934.
46. Shokri, R.; Shmatikov, V. Privacy-preserving deep learning. In Proceedings of the 22th ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015; pp. 1310–1321.
47. Bernstein, G.; McKenna, R.; Sun, T.; Sheldon, D.; Hay, M.; Miklau, G. Differentially private learning of undirected graphical models using collective graphical models. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017.
48. Friedman, A.; Schuster, A. Data mining with differential privacy. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 25–28 July 2010; pp. 493–502.
49. Xu, S.; Cheng, X.; Su, S.; Xiao, K.; Xiong, L. Differentially private frequent sequence mining. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 2910–2926.
50. Feild, H.A.; Allan, J.; Glatt, J. CrowdLogging: Distributed, private, and anonymous search logging. In Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, Beijing, China, 25–29 July 2011; pp. 375–384.

51. Korolova, A.; Kenthapadi, K.; Mishra, N.; Ntoulas, A. Releasing search queries and clicks privately. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, 20–24 April 2009; pp. 171–180.
52. Zhang, S.; Yang, H.; Singh, L. Anonymizing query logs by differential privacy. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, Pisa, Italy, 17–21 July 2016; pp. 753–756.
53. Sánchez, D.; Batet, M.; Viejo, A.; Rodríguez-Garcíac, M.; Castellà-Roca, J. A semantic-preserving differentially private method for releasing query logs. *Inf. Sci.* **2018**, *460–461*, 223–237.
54. Chaudhuri, K.; Sarwate, A.D.; Sinha, K. A near-optimal algorithm for differentially-private principal components. *J. Mach. Learn. Res.* **2013**, *14*, 2905–2943.
55. Balle, B.; Gomrokchi, M.; Precup, D. Differentially private policy evaluation. In Proceedings of the 33th International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2130–2138.
56. Zhang, Z.; Rubinstein, B.I.P.; Dimitrakakis, C. On the differential privacy of Bayesian inference. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 2365–2371.
57. Pathak, M.; Rane, S.; Raj, B. Multiparty differential privacy via aggregation of locally trained classifiers. In Proceedings of the 24th Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–9 December 2010; pp. 1876–1884.
58. Geng, Q.; Viswanath, P. The optimal noise-adding mechanism in differential privacy. *IEEE Trans. Inf. Theory* **2016**, *62*, 925–951.
59. Wang, Y.; Yang, L.; Chen, X.; Zhang, X.; He, Z. Enhancing social network privacy with accumulated non-zero prior knowledge. *Inf. Sci.* **2018**, *445*, 6–21.
60. Xu, C.; Ren, J.; Zhang, Y.; Qin, Z.; Ren, K. Dppro: Differentially private high-dimensional data release via random projection. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 3081–3093.
61. Chen, R.; Li, H.; Qin, A.K.; Kasiviswanathan, S.P.; Jin, H. Private spatial data aggregation in the local setting. In Proceedings of the 32th IEEE International Conference on Data Engineering, Helsinki, Finland, 16–20 May 2016; pp. 289–300.
62. McSherry, F.; Mironov, I. Differentially private recommender systems: Building privacy into the netflix prize contenders. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 627–636.
63. Holohan, N.; Leith, D.J.; Mason, O. Optimal differentially private mechanisms for randomised response. *IEEE Trans. Inf. Forensic Secur.* **2017**, *12*, 2726–2735.
64. Ushakov, N.G. *Selected Topics in Characteristic Functions*; VSP: Utrecht, The Netherlands, 1999.
65. Hung, P.K. *Secrets in Inequalities*; GIL Publishing House: Zalău, Romania, 2008; Volume 1.