Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

A survey on differential privacy methods for big data privacy protection

Chengliang Liu , Miaomiao Yu & Yong Zhou

To cite this article: Chengliang Liu , Miaomiao Yu & Yong Zhou (27 May 2026): A survey on differential privacy methods for big data privacy protection, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2026.2679084

To link to this article: https://doi.org/10.1080/24754269.2026.2679084

Published online: 27 May 2026.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2026.2679084

A survey on differential privacy methods for big data privacy protection

Chengliang Liu, Miaomiao Yu and Yong Zhou

Key Laboratory of Advanced Theory and Application in Statistics and Data Science (MOE), School of Statistics and Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai, People's Republic of China

ABSTRACT

In the era of big data, ensuring data privacy has emerged as a significant challenge in large-scale data applications. Currently, differential privacy is one of the most promising privacy preserving algorithms, as it provides an explicit measure of the degree of privacy protection. Although the development of differential privacy is still in its early stages within the field of statistics, it is expected to play an integral role in future research. Motivated by this, this paper first provides a review of the development of privacy models, including the detailed introduction and interpretation of the differential privacy framework. In addition, we present the applications of several commonly used noise mechanisms and elaborate on the parallel and sequential composition theorems in differential privacy. Finally, this paper also discusses potential future research on differential privacy for online data analysis and statistical inference.

ARTICLE HISTORY
Received 26 August 2025
Revised 1 April 2026
Accepted 20 May 2026

KEYWORDS
dData security; differential privacy; privacy budget; federated learning; linking attack

1. Introduction

With advances in data collection technology and computational power, it has become increasingly feasible to rapidly collect, store, and analyze large scale datasets, thereby generating complex big data ecosystems. Big data has emerged as a fundamental strategic resource and a core production factor comparable in importance to material assets and human capital. With the rapid development of big data applications, society expects to use big data to stimulate economic growth, strengthen social governance, and improve government services and regulatory capabilities. This has led to an increasing degree of data centralization and an increasing amount of data shared and cross-utilized.

While big data brings tremendous value, it also poses significant challenges. In particular, the processes of data collection, storage, and utilization are subject to strict policy and regulatory constraints, which raise complex management issues. More importantly, as data become increasingly centralized and widely shared across domains, the risks of privacy leakage and misuse grow sharply. These challenges highlight the urgent need for effective privacy preserving methods. Therefore, data security and privacy protection, as critical components

CONTACT Miaomiao Yu mmyu@fem.ecnu.edu.cn Key Laboratory of Advanced Theory and Application in Statistics and Data Science, MOE, and Academy of Statistics and Interdisciplinary Sciences, East China Normal University, 3663 North Zhongshan Road, Putuo District, Shanghai 200062, People's Republic of China

of internet governance, have become key bottlenecks constraining the advancement of big data applications.

To address these pressing concerns, governments worldwide have begun to strengthen data security and privacy protections. Many countries are strengthening data security and privacy protections. The European Union's General Data Protection Regulation (GDPR) implemented in 2018, demonstrates that stricter governance of user data privacy and security is becoming a global trend. This trend poses unprecedented challenges to big data applications and privacy protection efforts. Similarly, China's Cybersecurity Law and the General Rules of the Civil Code stipulate that network operators do not disclose, tamper with, or destroy collected personal information. Additionally, when conducting data transactions with third parties, they ensure that contracts explicitly define the scope of data sharing and the corresponding data protection obligations. The enactment of these regulations has introduced varying degrees of challenges to traditional data processing methods. For instance, the GDPR's 'right to be forgotten' complicates the retention of long-term datasets essential for model training, while China's data localization rules increase operational costs and hinder cross-border data analysis.

A common approach to protecting personal data privacy in data analysis is data anonymization, which refers to the process of removing or transforming personally identifiable information so that individual data subjects cannot be directly identified. However, anonymized data can still be re-identified through linking attack to external datasets or by exploiting distinctive patterns in the released data. A classic example is the Massachusetts voter privacy case (Sweeney, 2002). Figure 1 illustrates this process: the left panel displays a medical dataset containing records of 135,000 Massachusetts state employees and their families, available from the Group Insurance Commission for 20 dollars, while the right panel shows the state's voter registration records. The medical dataset removed direct identifiers through anonymization, retaining only three quasi-identifiers: gender, date of birth, and zip code. By cross-referencing these variables across the two datasets, the attacker successfully identified the state governor's medical records among millions of entries. This type of privacy breach is known as a linking attack. A linking attack refers to the process of combining anonymized datasets with external sources to re-identify individuals based on shared attributes. This demonstrates that data anonymization alone is insufficient to effectively prevent privacy breaches caused by such attacks.

To address the impact of linking attacks on data security, Samarati (2002) and Sweeney (2002) proposed the k-anonymity model as a metric to quantify the level of privacy protection. A dataset satisfies k-anonymity if each record is indistinguishable from at least $k-1$ other records for each combination of identifiers, ensuring that individuals cannot be uniquely identified through record linking. As shown in Table 1, in a medical data study, the full dataset on the left was desensitized to protect patient privacy by suppressing the three sensitive attributes of zip code, age, and nationality, while retaining only the key variables that represent the disease of interest in the study. When the three desensitized variables are used to identify individual records, at least four records share identical values across all three variables, indicating that the mechanism satisfies the 4-anonymity model. Due to its conceptual simplicity, the k-anonymity model has been widely recognized as a practical and accepted definition and measure of privacy in data dissemination (Aggarwal et al., 2004; Bayardo & Agrawal, 2005; LeFevre et al., 2005).

However, when the homogeneity of the data is high, i.e., records contain highly similar information, an attacker may be able to infer attribute information about an individual. In

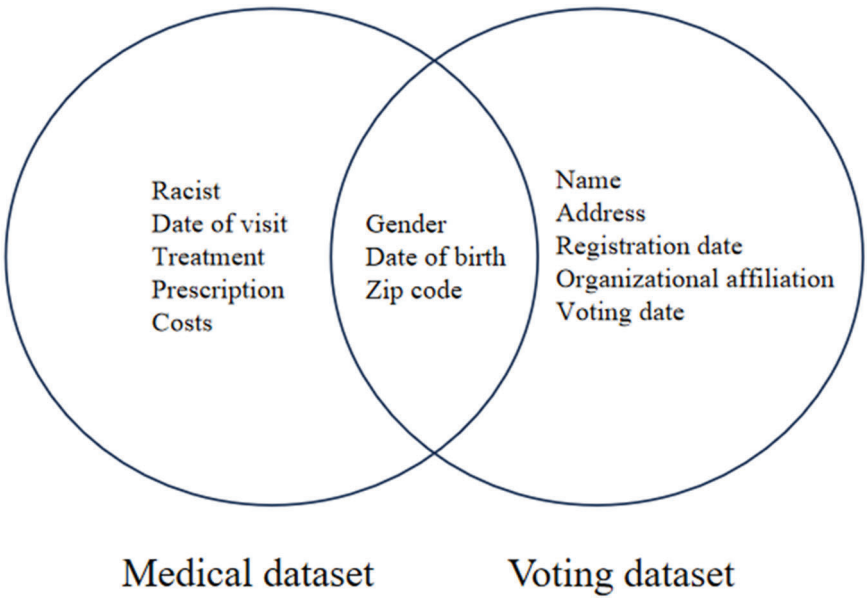

Figure 1. The example of a linking attack.

Table 1. 4-anonymity model in medical data.

	Original data				Anonymized data			
	Non-Sensitive				Non-Sensitive			
No.	Zip Code	Age	Nationality	Sensitive Condition	Zip Code	Age	Nationality	Sensitive Condition
1	13053	28	Russian	Heart Disease	130**	< 30	*	Heart Disease
2	13068	29	American	Heart Disease	130**	< 30	*	Heart Disease
3	13068	21	Japanese	Viral Infection	130**	< 30	*	Viral Infection
4	13053	23	American	Viral Infection	130**	< 30	*	Viral Infection
5	14853	55	Russian	Heart Disease	1485*	≥ 40	*	Heart Disease
6	14850	47	American	Viral Infection	1485*	≥ 40	*	Viral Infection
7	14850	49	American	Viral Infection	1485*	≥ 40	*	Viral Infection
8	14853	50	Indian	Cancer	1485*	≥ 40	*	Cancer
9	13053	31	American	Cancer	130**	3*	*	Cancer
10	13053	37	Indian	Cancer	130**	3*	*	Cancer
11	13068	36	Japanese	Cancer	130**	3*	*	Cancer
12	13068	35	American	Cancer	130**	3*	*	Cancer

Table 1, although it is impossible to accurately identify a patient record using zip code, age and nationality, it is possible to deduce that a patient has cancer for the 8th to 12th records, where the patient's records are determined by guessing the patient's age. Therefore, the k-anonymity model can prevent identity disclosure but does not effectively prevent attribute disclosure. To address this limitation, Machanavajjhala, Kifer, Gehrke et al. (2007) proposed the l-diversity model, which requires that each critical attribute contains at least l distinct values for each sensitive attribute. As shown in Table 2, it conforms to the 3-diversity model if privacy attributes such as age and zip code are desensitized so that the values of diseases and income exhibit at least three distinct values for each combination of these attributes. Although the l-diversity model effectively mitigates information leakage caused by a single value, it may lead to overprotection or underprotection of privacy due to semantic correlations among attribute values and varying sensitivity levels of different attributes. For example,

Table 2. The example of 3-anonymity model.

Original data					Anonymized data			
No.	Zip code	Age	Salary	Disease	ZIP Code	Age	Salary	Disease
1	47677	29	3K	gastric ulcer	476**	2*	3K	gastric ulcer
2	47602	22	4K	gastritis	476**	2*	4K	gastritis
3	47678	27	5K	stomach cancer	476**	2*	5K	stomach cancer
4	47905	43	6K	gastritis	4790*	$\geq$ 40	6K	gastritis
5	47909	52	11K	flu	4790*	$\geq$ 40	11K	flu
6	47906	47	8K	bronchitis	4790*	$\geq$ 40	8K	bronchitis
7	47605	30	7K	bronchitis	476**	3*	7K	bronchitis
8	47673	36	9K	pneumonia	476**	3*	9K	pneumonia
9	47607	32	10K	stomach cancer	476**	3*	10K	stomach cancer

in Table 2, although the values of diseases differ in the first to third records, all these diseases are related to the stomach allowing the inference that the individuals associated with these records have a stomach condition. To further enhance privacy protection, N. Li et al. (2006) introduced the *t*-closeness model, which is a novel concept for preserving privacy. This model limits the amount of specific individual information that an observer can infer by requiring that the distance between the distribution of sensitive attributes in any equivalence class and the distribution in the overall table measured by using Earth Mover's Distance (Rubner et al., 2000), does not exceed a threshold *t*. In addition, several anonymity models have been proposed for securing medical data (Avancha et al., 2012; Francis et al., 2015; Sadki & El Bakkali, 2014).

Federated learning techniques have recently been developed to enhance data privacy protection. Federated learning is a distributed machine learning framework in which multiple machines collaboratively train a shared model while keeping their raw data local (McMahan et al., 2017). This approach enables knowledge sharing without direct data exchange, thereby reducing privacy risks. Since its introduction by Google in 2016, Federated learning has attracted significant research interest and has evolved rapidly. Early studies focussed on communication-efficient algorithms (Konečnỳ et al., 2016), while subsequent works extended federated learning to heterogeneous data (Gu & Chen, 2024; T. Li et al., 2020), robust aggregation against adversarial attacks (Bhagoji et al., 2019), and privacy preserving enhancements through differential privacy and secure multi-party computation (Bonawitz et al., 2017; Geyer et al., 2017). However, similar to many machine learning approaches including neural networks, federated learning can algorithmically preserve data privacy, but its interpretability is limited due to the black box nature of the learning process. In many cases, privacy preserving mechanisms limit access to specific model information, which complicates effective interpretation. Moreover, Nasr et al. (2018) demonstrated that the model training phase remains vulnerable to privacy leakage in federated learning. For example, when training a DenseNet model using a federated learning algorithm on the CIFAR100 dataset, the central server which receives and updates parameters from other servers can achieve a membership inference accuracy of 79.2%, while local participants can achieve 72.2% accuracy by observing aggregated parameter updates from the central server (Tan & Zhang, 2020).

The above privacy preserving methods face the inherent trade-off that they cannot simultaneously provide enough privacy protection and maintain model interpretability. Moreover, a common limitation of such methods is the lack of a quantifiable measure to assess the

actual degree of privacy protection they offer. Consequently, it remains a significant challenge to develop statistical methods which integrate privacy protection with interpretable modelling and simultaneously ensure strong privacy guarantees. With increasing data complexity, privacy preserving techniques based on the aforementioned attribute are increasingly inadequate to meet the demands of modern data privacy requirements. Differential privacy frameworks effectively prevent privacy breaches arising from prior knowledge and provide a formal mechanism to quantify the privacy guarantees of an algorithm. In this paper, we provide a detailed introduction to differential privacy and briefly review its applications in data publishing and machine learning. As a landmark real-world application in official statistics, the 2020 U.S. Census demonstrates how differential privacy can be deployed in practice for large-scale statistical data release. Prior work has shown that the Census Bureau adopted differential privacy in response to the limitations of traditional disclosure avoidance methods and the growing risk of reconstruction attacks (Abowd, 2018; M. Li et al., 2023; Su et al., 2025; Zhang et al., 2017). Subsequent studies further examined the Top-Down algorithm and its practical implications, including its effects on redistricting and its accuracy across small-population geographies and demographic groups (Canonne et al., 2020; Cohen et al., 2022; Kenny, Kuriwaki et al., 2021; Kenny, McCartan et al., 2024). This example highlights that applying DP in official statistics requires balancing privacy, statistical accuracy, consistency, and fairness. Finally, we discuss current research and future opportunities for differential privacy in data analysis.

2. Related works

Dwork et al. (2006) first proposed the definition of differential privacy based on the probability framework. This method introduces controlled data perturbation by adding random noise, ensuring that even if an adversary knows all contextual information except for a single record with the maximal scenario for a privacy attack, they cannot determine with high probability whether that record is included in the accessed dataset. The proposed framework renders data security protection quantifiable and amenable to rigorous theoretical verification. Dwork et al. (2006) theoretically demonstrated that adding random noise following specific distributions such as Laplace, Gaussian and exponential, satisfies the requirements of differential privacy. These distributions are characterized by probability density functions with an exponential form, which aligns well with the formal definition of the model. However, the paper primarily focuses on the selection of unit random noise and does not extensively address the multivariate case. Dwork and Roth (2014) provided a more comprehensive discussion of differential privacy methodologies and their construction. The monograph not only elaborates on the theoretical foundations and algorithmic techniques of differential privacy, but also explores their applications in areas such as query release, mechanism design, and machine learning. Nevertheless, the inherent tension between privacy as a property right and the requirements of big data analytics creates a fundamental trade-off that it is challenging to preserve privacy while simultaneously producing highly accurate query results with minimal noise (Dinur & Nissim, 2003; Dwork et al., 2007; Dwork & Yekhanin, 2008).

Building upon the above theoretical foundation, differential privacy has been widely applied in real-world scenarios. Data publication emerges as one of the most critical and extensively studied application domains, where data publishers provide data accessors with perturbed versions of the original data. Yang et al. (2020) proposed a balanced differential privacy framework for data publication, coupled with a multi-indicator decision approach

to assess potential privacy leakage. Wu et al. (2017) utilized the advantages of matrix representations for handling associative queries and proposed a differential privacy protection scheme for streaming data publication based on an exponential decay model. For additional differential privacy models in data publication, Xiong et al. (2014) provided a comprehensive review, categorizing methods into interactive and non-interactive privacy preserving publication approaches.

While data publication has been a primary application domain, the scope of differential privacy has expanded significantly, with machine learning emerging as another critical area where privacy preserving techniques are increasingly required. In the support vector machine method, the original function is projected into a finite dimensional space, which represents the classification function as a finite dimensional vector, and standard techniques are then applied to ensure privacy protection (Chaudhuri et al., 2011; Rubinstein et al., 2009). Hall et al. (2013) proposed a weak-type differential model to protect the classification function. This approach entirely avoids approximations in support vector machines, and the private function does not require representation as any finite dimensional vector. Wang and Zhang (2020) introduced a scheme to achieve differential privacy by adding noise to the outputs of distributed algorithms under a logistic regression framework. To address the limitation that privacy preserving algorithms can become ineffective due to excessive global sensitivity in time series models, Peng et al. (2020) described a privacy preserving model based on sequence lattices. The primary challenge stems from the inherent conflict between the instability of machine learning algorithms and the requirements of differential privacy.

While machine learning illustrates one prominent application of differential privacy, another critical domain involves statistical testing and estimation, where careful design is required to mitigate the impact of added noise on the precision and convergence of estimates. Barak et al. (2007) and Hardt et al. (2012) proposed a scheme that generates a contingency table with added random noise, enabling statistical tests to be conducted on the table while preserving data privacy. Histograms constitute a primary approach for estimating density functions. Y. Xiao et al. (2012) introduced a tree-based partitioning strategy in which noise is added to each region prior to merging with subsequent noise reduction after merging. For additional discussions on histogram estimation under differential privacy, see Chawla et al. (2012) and Wasserman and Zhou (2010). For density estimation methods, the histogram is not optimal due to their inherent lack of smoothness. In contrast, kernel density estimation typically exhibits superior convergence rate. Hall et al. (2013) described incorporating random noise following an exponential distribution into kernel density estimation and demonstrated that this approach is applicable to arbitrary dimensions under a Gaussian kernel. In general, the addition of random noise may degrade the convergence rate of statistical estimates. For instance, histogram estimation fails to attain the optimal convergence rate of $n^{-\frac{2}{2+p}}$, while kernel density estimation methods cannot reach the optimal convergence rate of $n^{-\frac{4}{4+p}}$, where n and p denote the sample size and data dimension (Wasserman, 2006). For more comprehensive treatments of privacy preserving statistical estimation, they are referred to Machanavajjhala, Kifer, Abowd et al. (2008), Smith (2011), and Avella-Medina (2021).

Importantly, the applications discussed above illustrate methods and theoretical developments, but differential privacy is also applicable to a wide range of real-world sensitive datasets. These include coordinate-based data from smart devices and geolocation tags on social media (Errounda & Liu, 2019), personal health records (Dankar & El Emam, 2012, 2013), and data generated by smart home and wireless sensor networks (J.

Wang et al., 2018). In this sense, the frameworks, algorithms, and estimation techniques developed for data publication, machine learning, and statistical estimation provide the foundation for protecting privacy across diverse practical scenarios. For comprehensive reviews of differential privacy, see Ji et al. (2014), L. Xu et al. (2014), Shaikh and Patil (2018), and Jin et al. (2019).

3. Differential privacy methods

Differential privacy methods are one of the latest and most promising privacy protection concepts. This section starts with the definition of the model to demonstrate its specific connotation. Differential privacy methods ensure that when a record in a dataset is changed including deletion, replacement or addition, the analysis based on the data is not affected. Thus, the opponent cannot know whether the record is included in the dataset. The specific definitions of differential privacy are as follows.

Definition 3.1: Two datasets X, X' are called neighboring datasets if and only if they have at most one record in common.

Definition 3.2 ((ϵ, δ)-differential privacy, (ϵ, δ)-DP): A randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ for any neighboring datasets $X, X' \in D^n$ provides (ϵ, δ)-differential privacy if for all output spaces $S \subseteq \mathbb{R}^p$, the following inequality holds:

$$\Pr\left[\mathcal{M}\left(X\right) \in S\right] \leqslant \exp\left(\epsilon\right) \Pr\left[\mathcal{M}\left(X'\right) \in S\right] + \delta. \tag{1}$$

The privacy guarantee in differential privacy is characterized by a pair of parameters (ϵ, δ), which jointly determine the strength of privacy protection. Here, ϵ controls the magnitude of the privacy loss, with smaller values providing stronger privacy guarantees. In the extreme case $\epsilon = 0$, the output distributions of the algorithm $\mathcal{M}$ on any pair of neighboring datasets are identical, making them completely indistinguishable. The parameter δ represents the probability that the privacy guarantee may fail. The lower values of δ correspond to a lower probability of such failure events. When $\delta = 0$, the model reduces to pure ϵ-differential privacy. To ensure meaningful privacy protection, the choice of δ should depend on the dataset size n. δ is typically required to be negligible in n, for example $\delta \leqslant n^{-(1+\gamma)}$ for some constant $\gamma > 0$, and at least $\delta \ll 1/n$ to avoid pathological cases where sensitive data may be disclosed with non-negligible probability, where $\delta \ll 1/n$ indicates that δ is negligibly small compared to $1/n$.

Although (ϵ, δ)-differential privacy is the standard framework for quantifying privacy guarantees, it is not always the most convenient form for privacy accounting, especially under repeated composition. To obtain tighter and more tractable composition bounds, several refined variants of differential privacy have been introduced. Among them, Rényi differential privacy measures privacy loss through the Rényi divergence between the output distributions $\mathcal{M}(X)$ and $\mathcal{M}(X')$ on neighboring datasets. This formulation often leads to a cleaner and tighter composition analysis than working directly with (ϵ, δ)-differential privacy.

Definition 3.3 (Rényi differential privacy, RDP): A randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ for any neighboring datasets $X, X' \in D^n$ provides (α, ϵ)-Rényi differential privacy if for all

output spaces $S \subseteq \mathbb{R}^p$,

$$R_\alpha(\mathcal{M}(X)\|\mathcal{M}(X')) \leqslant \epsilon \tag{2}$$

holds for all $\alpha > 1$, where $R_\alpha(P\|Q) := \frac{1}{\alpha-1}\log\int(\frac{p(x)}{q(x)})^\alpha q(x)\,\mathrm{d}x$ is the Rényi divergence between distributions P and Q.

An important property of Rényi differential privacy is its connection to the standard (ϵ, δ)-differential privacy framework, as stated in the following lemma.

Lemma 3.1: *If $\mathcal{M}$ is an (α, ϵ)-Rényi differential privacy algorithm, it also satisfies $(\epsilon + \frac{\log 1/\delta}{\alpha-1}, \delta)$-differential privacy for any $0 < \delta < 1$.*

Closely related to Rényi differential privacy, zero-concentrated differential privacy (zCDP) is another relaxation of differential privacy based on Rényi divergence. Rather than specifying a privacy guarantee at a fixed order α, zCDP controls the entire family of Rényi divergences through a simple linear upper bound in α.

Definition 3.4 (Zero-concentrated differential privacy, zCDP): A randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ for any neighboring datasets $X, X' \in D^n$ provides (ξ, ρ)-zero-concentrated differential privacy if for all output spaces $S \subseteq \mathbb{R}^p$,

$$R_\alpha(\mathcal{M}(X)\|\mathcal{M}(X')) \leqslant \xi + \rho\alpha, \tag{3}$$

holds for all $\alpha > 1$.

The following lemma shows that zCDP also implies a corresponding (ϵ, δ)-differential privacy.

Lemma 3.2: *If a randomized algorithm $\mathcal{M}$ satisfies $(0, \rho)$-zero-concentrated differential privacy, then $\mathcal{M}$ satisfies $(\rho + 2\sqrt{\rho\log(1/\delta)}, \delta)$-differential privacy for any $\delta > 0$.*

Besides Rényi divergence, the distinguishability between $\mathcal{M}(X)$ and $\mathcal{M}(X')$ is measured using the hypothesis testing formulation and is systematically studied as f-differential privacy. More specifically, consider a hypothesis testing problem

$$H_0 : \text{data} \sim P \quad \text{versus} \quad H_1 : \text{data} \sim Q$$

and a rejection rule $\phi \in [0, 1]$. We define the type I error as $\alpha_\phi = \mathbb{E}_P[\phi]$, which represents the probability of mistakenly rejecting the null hypothesis H_0. The type II error $\beta_\phi := 1 - \mathbb{E}_Q[\phi]$ is the probability of incorrectly accepting the alternative hypothesis H_1. The trade-off function $T(P, Q)$ denotes the minimal type II error at level α of type I error where $T(P, Q)(\alpha) = \inf_\phi\{\beta_\phi : \alpha_\phi \leqslant \alpha\}$.

Definition 3.5 (f-differential privacy, f-DP): A randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ for any neighboring datasets $X, X' \in D^n$ provides (ϵ, ρ)-zero-concentrated differential privacy if for all output spaces $S \subseteq \mathbb{R}^p$,

$$T(f(X), f(X')) \geqslant f \tag{4}$$

holds.

A mechanism satisfying f-differential privacy is considered more private when its associated trade-off function f is larger. In the extreme case where $\mathcal{M}(X)$ and $\mathcal{M}(X')$ are indistinguishable, the trade-off function reaches its maximum: the identity function $\mathrm{Id}(x) := 1 - x$. Therefore, any valid trade-off function must satisfy $f \leqslant \mathrm{Id}$ pointwise. A trade-off function $f = T(P, Q)$ is said to be symmetric if $T(P, Q) = T(Q, P)$. Furthermore, the connection between f-differential privacy and the standard (ϵ, δ)-differential privacy framework is made explicit by the following lemma.

Lemma 3.3: *A randomized algorithm $\mathcal{M}$ satisfies (ϵ, δ)-differential privacy if and only if $\mathcal{M}$ satisfies $f_{\epsilon,\delta}(\gamma)$-differential privacy where*

$$f_{\varepsilon,\delta}(\gamma) = \max\left\{0, 1 - \delta - \mathrm{e}^{\varepsilon}\gamma, \mathrm{e}^{-\varepsilon}(1 - \delta - \gamma)\right\},$$

for $0 \leqslant \gamma \leqslant 1$.

Among the various instances of f-differential privacy, an especially important one is the Gaussian differential privacy, which is obtained by choosing the trade-off function corresponding to the optimal test between two shifted Gaussian distributions. This specialization is particularly useful because it provides an analytically tractable privacy measure and has close connections to the privacy analysis of Gaussian mechanisms and noisy iterative algorithms. The resulting notion is referred to as μ-Gaussian differential privacy.

Definition 3.6 (μ-Gaussian differential privacy, μ-GDP): A randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ for any neighboring datasets $X, X' \in D^n$ provides μ-Gaussian differential privacy if it is G_μ-differential privacy with

$$G_\mu(x) = \Phi(\Phi^{-1}(1 - x) - \mu),$$

where Φ denotes the cumulative distribution function of the standard normal distribution.

The following lemma makes explicit the connection between μ-Gaussian differential privacy and the standard (ϵ, δ)-differential privacy framework.

Lemma 3.4: *A randomized algorithm $\mathcal{M}$ satisfies μ-Gaussian differential privacy if and only if $\mathcal{M}$ satisfies $\epsilon, \delta(\epsilon)$-differential privacy for all $\epsilon \geqslant 0$ where*

$$\delta(\varepsilon) = \Phi\left(-\frac{\varepsilon}{\mu} + \frac{\mu}{2}\right) - \mathrm{e}^{\varepsilon}\Phi\left(-\frac{\varepsilon}{\mu} - \frac{\mu}{2}\right).$$

While the above notions of differential privacy focus on privacy guarantees for randomized mechanisms applied to centralized datasets, another important framework is local differential privacy. Unlike the central model, local differential privacy requires each data provider to privatize his or her observation before sending it to the data collector, thus removing the need for a trusted curator.

Definition 3.7 (Local differential privacy, LDP): For a given privacy parameter $\epsilon \geqslant 0$, the random variable Z_i is ϵ-local differential private of X_i if for all $z_1, \ldots, z_{i-1}$ and $x, x' \in \mathcal{X}$ we have

$$\sup_{S \in \sigma(\mathcal{Z})} \frac{\mathcal{M}_i(Z_i \in S \mid X_i = x, Z_1 = z_1, \ldots, Z_{i-1} = z_{i-1})}{\mathcal{M}_i(Z_i \in S \mid X_i = x', Z_1 = z_1, \ldots, Z_{i-1} = z_{i-1})} \leqslant \exp(\epsilon)$$

where $\sigma(\mathcal{Z})$ denotes an appropriate σ-field on $\mathcal{Z}$.

We say that the privacy mechanism $\mathcal{M}_i$ is ϵ-LDP if each variable Z_i is ϵ-LDP view. In addition, Duchi et al. (2018) systematically investigated statistical estimation problems under LDP constraints. By establishing privacy-preserving versions of lower bounds such as Le Cam, Fano, and Assouad, it precisely characterized the impact of privacy protection on statistical estimation risks. It also proposed privacy mechanisms and estimation methods that achieve minimal optimal convergence rates, covering multiple classical statistical problems such as mean, median, high-dimensional sparse vectors, generalized linear models, and nonparametric density estimation. More recent research on local differential privacy can be found in the literature (Kim et al., 2021; M. Li et al., 2023; Lin et al., 2022).

The differential privacy methods are affected by sensitivity in addition to the privacy parameters ϵ and δ. For convenience, we define the following notations. For any vector $\boldsymbol{v} = (v_1, \ldots, v_p)^\top$, the ℓ_1-norm and ℓ_2-norm are defined by $\|\boldsymbol{v}\|_1 = \sum_{j=1}^{p} |v_j|$ and $\|\boldsymbol{v}\|_2 = \sqrt{\sum_{j=1}^{p} v_j^2}$, respectively. First, we introduce the global sensitivity (Dwork et al., 2006) as follows.

Definition 3.8 (Global sensitivity): For a randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$, the global sensitivity is

$$\Delta\mathcal{M} = \max_{X,X'} \left\| \mathcal{M}(X) - \mathcal{M}(X') \right\|_1, \tag{5}$$

where $X, X' \in D^n$ are any neighboring datasets and $\|\cdot\|_1$ is L_1 norm.

Next, we will use an example to illustrate that the value of global sensitivity can be very large, even tending towards infinity. Let $\mathcal{M}(X) = \sum_{i=1}^{n} X_i$ denote the sum of all elements in the dataset X. Thus,

- when $D^n = \{0, 1\}^n$, the global sensitivity is 1;
- when $D^n = \{0, 100\}^n$, the global sensitivity is 100;
- when $D^n = \mathbb{R}^n$, the global sensitivity is unbounded, since the difference between neighboring datasets can be arbitrarily large.

However, noise increases with global sensitivity and excessive noise can affect the effectiveness of the dataset. Therefore, global sensitivity is not always a good measure of noise control. Based on this, Nissim et al. (2007) proposed the definition of local sensitivity.

Definition 3.9 (Local sensitivity): For a randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ and a given dataset $X \in D^n$, the local sensitivity is

$$\mathrm{LS}_{\mathcal{M}} = \max_{Y} \|\mathcal{M}(X) - \mathcal{M}(Y)\|_1, \tag{6}$$

where $Y \in D^n$ and X are the neighboring datasets.

When $D^n = \{0, 100\}^n$ and $\mathcal{M}(\cdot)$ represents the median of the dataset, the value of $\mathcal{M}(X)$ is 50 for the given dataset $X = \{0, 100\}$. Correspondingly, Y can be taken in the following two cases:

- when $Y = \{0, 0\}$, $\mathcal{M}(Y) = 0$ and local sensitivity is 50;

- when $Y = \{100, 100\}$, $\mathcal{M}(Y) = 100$ and local sensitivity is 50.

Although local sensitivity can effectively reduce the amplitude of noise, it already contains part of the information in the dataset. Specifically, when we know that the local sensitivity of a dataset is 50 with a sample size 2 and only two possible values including 0 and 100, we can conclude that the dataset is $\{0, 100\}$. Therefore, local sensitivity may cause privacy leaks. Based on this, Nissim et al. (2007) proposed β-smooth sensitivity.

Definition 3.10 (β-smooth sensitivity): For a randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$ and a given dataset $X \in D^n$, the β-smooth sensitivity is

$$S_{\mathcal{M},\beta} = \max_{Y} \left(\|\mathcal{M}(X) - \mathcal{M}(Y)\|_1 \exp\{-\beta d(X, Y)\} \right), \tag{7}$$

where $\beta > 0$ and $d(X, Y)$ represents the number of records that differ between datasets Y and X.

From Definition 3.10, β-smooth sensitivity does not specify the difference information between datasets Y and X. For convenience, it can be transformed as follows: let d represent the number of records that differ between datasets Y and X; then Equation (7) is equivalent to

$$S_{f,\beta} = \max_{d=0,1,\ldots,n} \exp\{-\beta d\} \left(\max_{Y:d(X,Y)=d} \|\mathcal{M}(X) - \mathcal{M}(Y)\|_1 \right). \tag{8}$$

Therefore, β-smooth sensitivity contains almost no information about the dataset and provides better privacy protection compared to local sensitivity.

Similarly, when $D^n = \{0, 100\}^n$ and $\mathcal{M}(\cdot)$ represents the median of the dataset, the β-smooth sensitivity is $50 \exp(-\beta)$ for the given dataset $X = \{0, 100\}$. Moreover, the original dataset cannot be determined when $\beta > \ln 2$ (Nissim et al., 2007). Specifically, the β-smooth sensitivities of datasets $\{0, 0\}^n$, $\{100, 100\}^n$ and $\{0, 100\}^n$ are all $50 \exp(-\beta)$. In addition, it is worth noting that when expressing the median of the dataset, the calculation time for β-smooth sensitivity is $O(n^2)$ while the calculation time for local sensitivity is only $O(n)$.

4. Noise mechanism

In this section, we describe some common noise mechanisms and further introduce two composition theorems of differential privacy. First, Dwork (2006) proved that adding random noise that follows a Laplace distribution to the statistics can make the algorithm output satisfy ϵ-differential privacy.

Proposition 4.1 (Laplace mechanism): *For a randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$, the mechanism $\mathcal{M}$ satisfies ϵ-differential privacy by adding independent random noise from* $\mathrm{Lap}(\frac{\Delta\mathcal{M}}{\epsilon})$ *to the output, where* $\mathrm{Lap}(\frac{\Delta\mathcal{M}}{\epsilon})$ *represents a Laplace distribution with the scale parameter* $\frac{\Delta\mathcal{M}}{\epsilon}$.

Example 4.1: Assume that there is a dataset $X = \{0, 1\}^n$ and an algorithm $\mathcal{M}(X) = \sum_{i=1}^{n} X_i$, where visitors want to know the number of variables with a value of 1 in the dataset.

The algorithm $\mathcal{M}(X)$ is considered to add noise as follows:

$$T(X) = \sum_{i=1}^{n} X_i + \eta, \tag{9}$$

where $\eta \sim \text{Lap}(\frac{1}{\epsilon})$ and η is independent of the sample X. Any neighboring datasets X and X' satisfy $|\mathcal{M}(X) - \mathcal{M}(X')| \leqslant 1$. Thus, for any $t \in R$, the following inequality holds

$$\frac{\Pr(T(X) = t)}{\Pr(T(X^{'}) = t)} = \frac{h(t - \mathcal{M}(X))}{h(t - \mathcal{M}(X^{'}))} \leqslant \exp\left(\epsilon|\mathcal{M}(X) - \mathcal{M}(X^{'})|\right) \leqslant \exp(\epsilon), \tag{10}$$

where $h(\cdot)$ is the cumulative distribution function of the Laplace distribution.

In addition to the Laplace distribution, the normal distribution is one of the most common distributions, which is a favourable candidate for noise mechanisms. Generally speaking, a normal distribution with zero mean is used to design a differential privacy mechanism.

Proposition 4.2 (Gaussian mechanism): *For a randomized algorithm $\mathcal{M} : D^n \to \mathbb{R}^p$, the mechanism $\mathcal{M}$ satisfies (ϵ, δ)-differential privacy by adding independent random noise from $N(0, \sigma^2)$ to the output, where $\sigma \geqslant \frac{\sqrt{2\ln(1.25/\delta)}\Delta\mathcal{M}}{\epsilon}$ for any $\delta \in (0, 1)$.*

Example 4.2: Similar to Example 4.2, the density function of noise η is $\frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{y^2}{2\sigma})$. The ratio of this density function at adjacent points x and $x + 1$ is

$$\frac{\frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma}\right)}{\frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x+1)^2}{2\sigma}\right)} = \exp\left(\frac{x}{\sigma} + \frac{1}{2\sigma}\right). \tag{11}$$

If inequality $\exp(\frac{x}{\sigma} + \frac{1}{2\sigma}) \leqslant \exp(\epsilon)$ holds, x satisfies $x \leqslant \epsilon\sigma - 1/2$. Therefore, when $\sigma > \frac{\sqrt{2\ln(1.25/\delta)}\Delta\mathcal{M}}{\epsilon}$ (Dwork et al., 2006), the probability of exceeding point $\epsilon\sigma - 1/2$ satisfies

$$\Pr\left(x \geqslant \epsilon\sigma - \frac{1}{2}\right) \leqslant \delta. \tag{12}$$

Therefore, the noise mechanism of the normal distribution satisfies (ϵ, δ)-differential privacy where the variance of the normal distribution satisfies $\sigma \geqslant \frac{\sqrt{2\ln(1.25/\delta)}\Delta\mathcal{M}}{\epsilon}$.

Both of the above noise mechanisms simply add noise to the numerical output results to achieve differential privacy. However, for non-numerical data, the exponential mechanism does not determine a specific output, but rather returns a result with a certain probability value to achieve differential privacy. McSherry and Talwar (2007) proved that noise mechanisms obeying an exponential distribution satisfy ϵ-differential privacy.

Proposition 4.3 (Exponential mechanism): *Let $X \in D^n$ be the input dataset and R be the output range. There exists a score function $s : D^n \times R \to R$. The output algorithm f after adding*

the noise mechanism is

$$f(D^n, s) = \left\{ r : \Pr(r \in R) \propto \exp\left(\frac{\epsilon s(X, r)}{2\Delta s}\right)\right\}, \tag{13}$$

where Δs *satisfies*

$$\Delta s = \max_{r \in R} \max_{X, X'} |s(X, r) - s(X', r)|. \tag{14}$$

Then algorithm $\mathcal{M}$ *satisfies* ϵ*-differential privacy.*

In addition to involving different noise mechanisms, differential privacy has three important theorems that are used to construct reasonable noise mechanisms (McSherry, 2009). They are sequential composition theorem and parallel composition theorem.

Lemma 4.1 (Sequential composition theorem): *Suppose that algorithms* $\mathcal{M}_i : D^n \to \mathbb{R}^p$ *satisfy the differential privacy with privacy budget parameters* (ϵ_i, δ_i) *for* $i = 1, \ldots, m$*. Then, the multivariate mechanism* $\mathcal{M}(\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_m)$ *satisfies* $(\sum_{i=1}^m \epsilon_i, \sum_{i=1}^m \delta_i)$*-differential privacy.*

Example 4.3: There is a looping algorithm $\mathcal{M}$ with m steps, assuming that the privacy budget parameter for each step is (ϵ_i, δ_i) for $i = 1, \ldots, m$. Then algorithm $\mathcal{M}$ satisfies $(\sum_{i=1}^m \epsilon_i, \sum_{i=1}^m \delta_i)$-differential privacy.

Lemma 4.2 (Parallel composition theorem): *Suppose that algorithms* $\mathcal{M}_i : D_i^n \to \mathbb{R}^p$ *satisfy the differential privacy with privacy budget parameters* (ϵ_i, δ_i) *for* $i = 1, \ldots, m$ *where the datasets* $D_1, D_2, \ldots, D_m$ *are disjoint. Then, the multivariate mechanism* $\mathcal{M}(\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n)$ *satisfies* $(\max_i \epsilon_i, \max_i \delta_i)$*-differential privacy.*

Example 4.4: Suppose that there are m machines, and each machine has an algorithm $\mathcal{M}_i : D_i^n \to \mathbb{R}^p$ with the privacy budget parameter (ϵ_i, δ_i) for $i = 1, 2, \ldots, m$ where the datasets $D_1, D_2, \ldots, D_m$ are disjoint. Algorithm $\mathcal{M}$ combines the results from these computers. Then the algorithm $\mathcal{M}$ satisfies $(\max_i \epsilon_i, \max_i \delta_i)$-differential privacy.

Lemma 4.3 (Advanced composition theorem): *Suppose that for each* $i = 1, \ldots, m$*, the algorithm* $\mathcal{M}_i : D^n \to \mathbb{R}^p$ *satisfies* (ϵ, δ)*-differential privacy. Then, for any* $\tilde{\delta} > 0$*, the composed mechanism* $\mathcal{M}(\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_m)$ *satisfies* (ϵ', δ')*-differential privacy under m-fold adaptive composition with*

$$\epsilon' = \sqrt{2m \log(1/\tilde{\delta})} \cdot \epsilon + m\epsilon(e^\epsilon - 1),$$
$$\delta' = m\delta + \tilde{\delta}.$$

The above three composition theorems are useful to ensure that multiple combinations that satisfy differential privacy algorithms still satisfy differential privacy and obtain explicit privacy parameters. In combination algorithms, they reasonably allocate privacy budgets to each algorithm so that the overall algorithm protects privacy.

After characterizing how privacy guarantees behave under composition, we next introduce another basic property of differential privacy, namely the post-processing property.

Lemma 4.4 (Post-processing property): *Let $\mathcal{M}$ be a randomized algorithm that is (ϵ, δ)-differentially private. Let $w : \mathbb{R}^p \to \mathbb{R}^p$ be an arbitrary randomized mapping. Then $w(\mathcal{M})$ is (ϵ, δ)-differentially privacy.*

This lemma formalizes the intuitive fact that privacy cannot be degraded by any data-independent manipulation of the output of a differentially private mechanism.

5. Privacy-utility trade-off

In the differential privacy framework, improving the privacy-utility trade-off requires not only guaranteeing rigorous privacy protection but also minimizing the performance degradation induced by privacy constraints. One important direction is to count the privacy budget as tightly as possible, so that the cumulative privacy loss can be characterized more accurately and excessive noise injection can be avoided. Tighter privacy accounting methods, therefore, help improve the utility of estimation and learning algorithms under the same privacy guaranty. Section 5.1 introduces some privacy accounting methods, especially f-DP-based methods and numerical methods, which can count the privacy budget tightly and thus enhance utility. In Section 5.2, we describe that some data processing procedure may amplify the privacy. Section 5.3 elaborates on some optimal noise-adding mechanisms.

5.1. Count the privacy budget tightly

In practice, (ϵ, δ)-DP composition typically relies on loose upper bounds, leading to a systematic overestimation of privacy loss and consequently unnecessary noise addition. This observation has motivated a line of recent work on tighter privacy accounting, which aims to more accurately characterize cumulative privacy loss and thereby improve utility under the same privacy constraints. Broadly speaking, existing approaches can be divided into two directions: refined analytical frameworks and numerical accounting methods.

On the analytical side, f-differential privacy (f-DP) provides a more expressive framework for tracking privacy loss under composition. Dong et al. (2022) introduced f-DP, which characterizes privacy via the trade-off function and avoids the information loss inherent in (ϵ, δ)-DP. This formulation enables a tight and closed-form composition. In particular, the μ-GDP special case admits an exact composition rule $\mu_{\text{total}} = \sqrt{\mu_1^2 + \cdots + \mu_n^2}$, which prevents overestimation of privacy loss and allows smaller noise injection. Building on this framework, H. Wang et al. (2022) proposed the Edgeworth Accountant, which models the privacy loss random variable and analyzes its sum under composition. By leveraging Edgeworth expansions instead of standard central limit approximations, their method captures higher-order corrections and yields more accurate estimates of (ϵ, δ). Similarly, C. Wang, Su et al. (2023) studied mixture mechanisms within the f-DP framework, directly characterizing privacy at the level of trade-off functions. This avoids the looseness of divergence-based approaches and leads to tighter bounds for complex mechanisms such as sub-sampling and shuffling. Overall, these analytical methods reduce the gap introduced by traditional composition bounds and improve utility by enabling more precise noise calibration.

Complementary to these analytical approaches, numerical methods aim to approximate the exact privacy loss distribution under composition. Koskela et al. (2020) proposed a numerical composition method based on privacy loss random variables, transforming privacy accounting into a convolution problem by exploiting the additivity of privacy loss. Specifically, if Y_i denotes the privacy loss of the i-th mechanism, then the total privacy loss is $Y = \sum_{i=1}^{k} Y_i$, and the privacy curve can be computed from the distribution of Y. Their method approximates this distribution via discretizations and efficiently computes convolutions using the fast Fourier transform (FFT), achieving arbitrarily accurate approximations. Along a similar line, Gopi et al. (2021) introduced the Fourier accountant, which computes (ϵ, δ) guarantees through a numerical evaluation of an integral formula,

$$\delta(\epsilon) = \int_{\epsilon}^{\infty} (1 - \mathrm{e}^{\epsilon - s})(\omega^{*k})(s)\, \mathrm{d}s,$$

where ω is the privacy loss distribution of a single mechanism and ω^{*k} denotes its k-fold convolution. By discretizing the integral and leveraging FFT, this method yields tight privacy bounds with controlled numerical error. In contrast to upper-bound-based techniques such as RDP or the moments accountant, these numerical approaches directly approximate the exact privacy expression, thereby avoiding systematic overestimation.

In summary, both analytical frameworks such as f-DP and numerical accounting methods provide tighter characterizations of cumulative privacy loss. By reducing conservativeness in privacy estimation, these approaches allow for smaller noise injection and consequently lead to improved utility in practical differentially private algorithms.

5.2. Privacy amplification techniques

Another important approach to improving the privacy-utility trade-off is to exploit privacy amplification, where certain data processing procedures inherently strengthen privacy guarantees. The key idea is that by introducing randomness or reducing the influence of individual data points, these procedures effectively decrease the distinguishability between neighboring datasets. As a result, for a fixed privacy requirement, less noise is needed, leading to improved utility. Existing amplification techniques can be broadly categorized into sub-sampling, shuffling, and iterative procedures.

Privacy amplification by sub-sampling captures the effect of applying a differentially private mechanism to a randomly selected subset of the data. Balle et al. (2018) provided a unified analysis of this phenomenon, showing that sub-sampling induces a contraction of the privacy profile and yields strictly improved privacy guarantees, i.e. $\epsilon' < \epsilon$. Intuitively, since each individual is included with a smaller probability, their influence on the output distribution is reduced. This diminished sensitivity allows the same level of privacy to be achieved with less noise, thereby improving statistical accuracy.

Beyond sub-sampling, privacy amplification can also arise from anonymization effects in data aggregation. Feldman et al. (2022) studied privacy amplification by shuffling, where n locally private reports are randomly permuted before aggregation. They show that this procedure yields significantly stronger central privacy guarantees, which substantially improves over the original local privacy parameter. The key mechanism is that shuffling hides each user's contribution among many others, reducing distinguishability and enhancing privacy without increasing noise.

In addition to randomness in data selection or ordering, amplification can also emerge from the dynamics of iterative algorithms. Bok et al. (2024) established privacy amplification by iteration, showing that in convex optimization, the privacy loss of gradient-based methods does not grow linearly with the number of iterations. Instead, due to contractive properties of the optimization dynamics, the privacy loss can decay over time. Their analysis within the f-DP framework provides tight and convergent guarantees for the final iterate, significantly improving upon standard composition-based bounds.

In summary, sub-sampling, shuffling, and iteration represent three complementary mechanisms through which privacy can be amplified. By effectively reducing the contribution of individual data points or leveraging structural properties of algorithms, these techniques decrease the overall privacy loss and enable lower noise injection, thereby leading to improved utility in differentially private systems.

5.3. Optimal noise

In addition to tighter privacy accounting and amplification techniques, another fundamental approach to improving the privacy-utility trade-off is to design optimal noise-adding mechanisms. The key idea is to minimize the distortion introduced by noise while still exactly satisfying the desired privacy constraints. Classical mechanisms such as Laplace and Gaussian are often adopted due to their simplicity, but they are not always optimal under general utility criteria. This motivates the study of noise distributions that are tailored to specific privacy definitions and loss functions.

From an optimization perspective, Geng and Viswanath (2014) studied the problem of designing optimal noise under a general utility (loss) framework. They show that the optimal mechanism is not the standard Laplace distribution, but rather a staircase-shaped distribution, which can be interpreted as a geometric mixture of uniform distributions. This mechanism concentrates the probability mass more efficiently, significantly reducing the magnitude of added noise-especially in the moderate-to-low privacy regime (i.e., large ϵ). As a result, it achieves strictly better accuracy than classical mechanisms under the same privacy constraint.

Furthermore, Awan and Vadhan (2023) provided a more general and unified perspective through the notion of canonical noise distributions within the f-DP framework. Instead of optimizing for a specific loss function, they characterize noise distributions that are intrinsically matched to a given privacy trade-off function f. In particular, they show that for any valid trade-off function, there exists a corresponding canonical noise distribution that achieves the privacy constraint tightly, without introducing unnecessary randomness. This framework unifies several classical mechanisms–including Gaussian, Tulap, and staircase mechanisms by showing that each is optimal under its respective privacy formulation.

In summary, optimal noise design focuses on aligning the noise distribution with the underlying privacy definition and utility objective. By avoiding suboptimal, overly diffuse noise distributions, these methods reduce estimation error while strictly maintaining privacy guarantees, thereby providing a principled way to enhance utility in differentially private algorithms.

6. Application of differential privacy in statistical estimation

In differential privacy framework, it is crucial to add noise encryption to the transmission of relevant data or local results Achieving privacy protection not only requires considering how to make differential encryption methods secure and reliable but ensuring that the encrypted results can still provide accurate parameter estimations. Section 6.1 introduces the impact of differential privacy on the convergence rate of M estimation. In Section 6.2, we describe the impact of differential privacy on the bias of parameter estimation. In addition, Section 6.3 presents the lower bound of accuracy loss in estimation algorithms based on differential privacy protection. Finally, in Section 6.4, we mention some recent research about differential privacy.

6.1. The impact of differential privacy on convergence rate

M estimation is one of the most important parameter estimation methods commonly used in statistics and econometrics. It has good statistical properties and a wide range of applications. It is defined as follows:

$$\theta^* = \arg\min_{\theta\in\Theta} \int m(\theta, x)g(x)\,\mathrm{d}x, \tag{15}$$

where $g(x)$ is the density function and Θ is the parameter space. When $g(x)$ is unknown, sample frequency estimation is commonly used to estimate the density function. Specifically, suppose that X is a dataset containing n independent and identically distributed samples, and the observed values are $x_1, x_2, \ldots, x_n$. Without loss of generality, assume that a sample space $\mathcal{X} = [0, 1]^p$ where p is the dimension of the variable. The sample space is uniformly divided into small blocks and each small block is denoted as $B_{\boldsymbol{r}} = \otimes_{j=1}^{p}[(r_j - 1)h_n, r_j h_n)]$, where h_n is the bandwidth and $\boldsymbol{r} = (r_1, r_2, \ldots, r_p) \in (1, 2, \ldots, \frac{1}{h_n})^p$. Thus, the density function is estimated as follows:

$$\hat{f}_{\mathrm{hist}}(x) = h_n^{-p} \sum_{\boldsymbol{r}} \frac{n_{\boldsymbol{r}}}{n} I_{\{x\in B_{\boldsymbol{r}}\}}, \tag{16}$$

where $n_{\boldsymbol{r}} = \sum_{i=1}^{n} I_{\{x_i\in\mathcal{B}_{\boldsymbol{r}}\}}$. In density function estimation that only n depends on sample information, Lei (2011) proposed the following privacy protection algorithm:

$$\widehat{n}_{\boldsymbol{r}} = n_{\boldsymbol{r}} + z_{\boldsymbol{r}}, \tag{17}$$

where $z_{\boldsymbol{r}}$ is independent of $\mathcal{X}$ and the density function is $f(z) = \epsilon \exp(\frac{-\epsilon|z|}{2})/4$. Since $n_{\boldsymbol{r}} \geqslant 0$ and the differential privacy estimator $\widehat{n}_{\boldsymbol{r}}$ satisfies $\widehat{n}_{\boldsymbol{r}} \geqslant 0$, the noise here comes from the transformed Laplace mechanism. The privacy protection algorithm satisfies ϵ-differential privacy. Then, the perturbed histogram is estimated as

$$\hat{f}_{\mathrm{PH}}(x) = h_n^{-p} \sum_{r} \frac{\hat{n}_r}{n} I_{\{x\in\mathcal{B}_r\}}. \tag{18}$$

The M estimator obtained from the perturbation histogram is

$$\hat{\theta}_{\mathrm{PH}} = \arg\min_{\theta\in\Theta} \int m(\theta, x)\hat{f}_{\mathrm{PH}}(x)\,\mathrm{d}x. \tag{19}$$

Lei (2011) proved that when $h_n = (\frac{\sqrt{\ln(n)}}{n})^{\frac{2}{p+2}}$, the convergence rate of the local optimal estimate is

$$\left|\hat{\theta}_{\mathrm{PH}} - \theta^*\right| = O_p\left(n^{-\frac{1}{2}} \vee \left(\frac{\sqrt{\ln(n)}}{n}\right)^{\frac{2}{p+2}}\right). \tag{20}$$

Generally, the convergence speed of M estimation is $O_p(n^{-\frac{1}{2}})$. From Equation (20), although the addition of the differential privacy protection algorithm greatly enhances data security, it affects the convergence rate of the estimation.

6.2. The impact of differential privacy algorithms on bias

In the above subsection, Lei (2011) theoretically explained the impact of differential privacy on the convergence speed of parameter estimation. This section discusses the impact of this method on the bias of parameter estimation. A linear regression model was constructed for house prices in the San Francisco Bay Area from 2003 to 2006, with independent variables including the house area, the year, and the county of residence. $m(\theta, x)$ is selected as the loss function of L_2 norm, which uses the least squares method for parameter estimation. There are a total of 250,070 samples. The indicator to measure the bias loss caused by differential privacy is

$$\alpha = \left|\frac{\hat{\theta}_{\mathrm{PH}}}{\hat{\theta}_{\mathrm{OLS}}} - 1\right|, \tag{21}$$

where $\hat{\theta}_{\mathrm{OLS}}$ is the optimal least square estimator. Note that this metric is not intended to directly measure the statistical bias or variance of the estimator. Instead, it serves as a relative deviation indicator that quantifies the accuracy loss of the privatized estimator compared to the non-private benchmark. It should therefore be interpreted as a measure of relative estimation error rather than a formal statistical risk. To accurately measure the randomness of the added noise, 100 experiments are repeated to obtain the mean square error of the metric $\bar{\alpha} = (\sum_{i=1}^{100} \alpha_i^2/100)^{1/2}$. As can be seen from Table 3, the differential privacy algorithm has a significant impact on the accuracy of parameter estimation. On the one hand, when the privacy budget is small ($\epsilon = 0.1$), the parameter estimation based on differential privacy has a maximum deviation of 29.8% compared to the least squares method. Even when the data security protection level is very weak ($\epsilon = 1$), the maximum deviation between the two methods is still 7.2%, clearly indicating that the differential privacy algorithm has a significant negative impact on model estimation and prediction. On the other hand, the differential privacy algorithm has different effects on different variables, especially when the data security level is high, with the impact ranging from 4.2% to 29.8%, indicating that this method is extremely unstable.

In summary, differential privacy has a significant impact on parameter estimation and statistical inference. How to protect data security while ensuring the effectiveness of traditional statistical methods is something we need to further consider.

6.3. Minimax risk lower bound for differential privacy algorithm

Lei (2011) demonstrated the impact of differential privacy algorithms on the converge nce speed of parameter estimation, while Cai et al. (2021) proved the lower bound of accuracy loss

Table 3. The impact of differential privacy algorithms on parameter estimation (%).

	$\bar{\alpha}$	
Variables	$\epsilon = 0.1$	$\epsilon = 1$
Intercept	10.6	7.2
Area	4.7	3.6
Year	4.6	1.0
County 2	8.0	1.5
County 3	4.2	0.8
County 4	29.8	2.8
County 5	9.8	1.4
County 6	7.1	1.0

in estimation algorithms based on differential privacy protection. Specifically, let $\mathcal{P}$ denote the family of distributions supported on a set $\mathcal{X}$, and let $\theta : \mathcal{P} \to \Theta \subset \mathbb{R}^p$ denote the parameter of interest. We have a sample set $X = (x_1, x_2, \ldots, x_n) \in \mathcal{D}^n$ containing n independent and identically distributed samples from a statistical model $P \in \mathcal{P}$, where the empirical distribution of the sample set is $\widehat{P}_X$. Our goal is to use the estimator $f(X)$ as an estimate of the true value $\theta(P)$. Let $f_{\epsilon,\delta}$ denote any (ϵ, δ)-differential privacy algorithm that satisfies the definition. Let ℓ denote a monotonically increasing loss function, and then the lower bound of the loss for parameter estimation under differential privacy is

$$\inf_{f \in f_{\epsilon,\delta}} \sup_{P \in \mathcal{P}} E[\ell(\|f(X) - \theta(P)\|_2)] \geqslant \ell\left(\frac{\psi(\mathcal{P}, \delta) \cdot \lambda(\mathcal{P}, \delta) \cdot \ln\left(\frac{1}{\delta}\right)}{n\epsilon}\right), \tag{22}$$

where $n \gtrsim \psi(\mathcal{P}, \delta) \cdot \frac{\ln(\frac{1}{\delta})}{\epsilon}$ and $\|f(X) - \theta(\widehat{P}_X)\|_2 \leqslant \lambda(\mathcal{P}, \delta)$. In addition, $a_n \gtrsim b_n$ indicates that there exists a constant C such that $a_n \geqslant Cb_n$ holds for any n.

As concluded by Cai et al. (2021), due to the monotonic increasing nature of the loss function, as the sum of the differential privacy parameters decreases, the deviation between the parameter estimates based on the differential privacy and the true values increases. As the sample size increases, the estimation bias introduced by the differential privacy algorithm is effectively reduced. Equation (22) provides a lower bound on the risk, which suggests that different differential privacy algorithms cause loss at least $\ell(\frac{\psi(\mathcal{P},\delta)\cdot\lambda(\mathcal{P},\delta)\cdot\ln(\frac{1}{\delta})}{n\epsilon})$. Therefore, it is necessary to carefully design the structure of the protection algorithm to minimize interference with parameter estimation while meeting data protection requirements.

6.4. Recent research on differential privacy

In this section, we briefly introduce some recent research on differential privacy. Section 6.4.1 describes a novel differential privacy algorithm that distributes privacy data in the same way as the original data, which mitigates this trade-off between high statistical accuracy and strict differential privacy. In Section 6.4.2, an improved communication-efficient distributed algorithm with differential privacy is presented. Both non-private and private schemes, in which only local estimators are passed from the local machine to the central machine, are more theoretically and practically accurate and efficient than traditional distributed algorithms. Section 6.4.3 establishes an asymptotic inference framework for randomized SGD and DP-SGD, derives the limiting distribution of the averaged private estimator, and develops valid confidence interval construction methods under differential privacy. In Section 6.4.4,

we introduce recent developments on privately fine-tuning foundation models. Section 6.4.5 discusses some recent developments in federated learning, where privacy could be enhanced through decentralization. In Section 6.4.6, we focus on recent developments of privately fine-tuning foundation models.

6.4.1. Distribution-invariant differential privacy

Traditional differential privacy mechanisms inevitably change the sample distribution of the original data when adding noise to ensure privacy. This distribution shift results in downstream statistical analysis of privatized data such as regression, classification and mean estimation, which may differ significantly from the results of the original data. That is, there is a trade-off between privacy protection and statistical accuracy.

To address this limitation, Bi and Shen (2023) proposed the distribution-invariant differential privacy (DIP). The core idea of DIP is to maintain the distribution characteristics of the original data while meeting strict differential privacy constraints. Specifically, suppose $(x_1, \ldots, x_n)$ is a random sample of a given cumulative distribution function F which is continuous. The DIP's privatized sample $(\tilde{x}_1, \ldots, \tilde{x}_n)$ is obtained via a formula:

$$\tilde{x}_i = H\left(F(x_i) + \eta_i\right), \quad H(\cdot) = F^{-1} \circ G(\cdot), \tag{23}$$

where $\circ$ denotes function composition and η_i is randomly sampled from a Laplace distribution $\mathrm{Lap}(0, 1/\epsilon)$. Meanwhile, the explicit expression of $G(\cdot)$ is

$$G(x) = \begin{cases} \dfrac{1}{2\epsilon}\mathrm{e}^{\epsilon x}(1 - \mathrm{e}^{-\epsilon}), & x < 0, \\ x + \dfrac{1}{2\epsilon}\mathrm{e}^{-\epsilon x} - \dfrac{1}{2\epsilon}\mathrm{e}^{\epsilon(x-1)}, & 0 \leqslant x \leqslant 1, \\ 1 - \dfrac{1}{2\epsilon}\mathrm{e}^{-\epsilon(x-1)}(1 - \mathrm{e}^{-\epsilon}), & x > 1. \end{cases} \tag{24}$$

When the sample distribution is known, convert the raw data to a sample on a uniform distribution $[0, 1]$ through its cumulative distribution function F. Then, Laplace noise is added to the evenly distributed sample, ensuring to satisfy ϵ-differential privacy. Design a specific nonlinear transform function $H = F^{-1} \circ G$ to transform the data after adding noise back to a sample space that follows the original distribution of F.

Proposition 6.1: *DIP in Equation* (23) *is not only ϵ-differential privacy, but the privatized sample also follows the original distribution F when F is known.*

When the sample distribution is unknown, DIP randomly splits the original sample into hold-out samples and to-be-privatized samples. The hold-out sample is used to estimate the empirical distribution $\hat{F}$ and is not published. In addition, to-be-privatized samples are added to noise for privacy protection. DIP method is used to reconcile both high statistical accuracy and strict differential privacy.

Proposition 6.2: *DIP in Equation* (23) *is ϵ-differential privacy, and the privatized sample follows the original distribution F asymptotically as $m \to \infty$ when F is unknown and m is the hold-out sample size.*

DIP method reconciles both high statistical accuracy and strict differential privacy. As a result, any downstream statistical or machine learning task yields essentially the same conclusion as when one uses the original data.

6.4.2. Distributed algorithm with differential privacy

In big data analysis, due to the enormous amount of data stored on different machines, commonly used statistical analysis methods cannot be directly applied. Therefore, distributed computing is required. Whether it is divide-and-conquer or multi-center data, intermediate data results need to be transmitted. Data transmission not only requires data privacy protection but also high efficiency. Meanwhile, excessive transmission not only affects computing efficiency but also poses challenges to data privacy protection.

Considering the efficiency of data transmission, M. Yu et al. (2026) proposed a communication-efficient distributed algorithm that simultaneously implements differential privacy protection during the transmission of intermediate statistics. Specifically, suppose that there are N data points distributed across K computers. To simplify the model, assume that each computer has the sample size n_k. The observation $\boldsymbol{X}_{ik} = (y_{ik}, x_{ik})$ represents the i-th sample in the k-th machine for $i = 1, 2, \ldots, n_k$ and $k = 1, 2, \ldots, K$. Let $\mathcal{L}(\cdot, \cdot)$ be the loss function, and we define the local loss function and global loss function as follows:

$$\ell_k(\boldsymbol{\theta}) := \frac{1}{n_k} \sum_{i=1}^{n_k} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{X}_{ik}), \quad k = 1, 2, \ldots, K,$$

$$\ell_N(\boldsymbol{\theta}) := \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{X}_{ik}) = \frac{1}{N} \sum_{k=1}^{K} n_k \ell_k(\boldsymbol{\theta}).$$

Motivated by the approximation of $\ell_N(\boldsymbol{\theta})$, the communication-efficient likelihood estimator is

$$\tilde{\boldsymbol{\theta}} := \arg\min_{\boldsymbol{\theta} \in \Theta} \tilde{\ell}_N(\boldsymbol{\theta}), \tag{25}$$

where $\tilde{\ell}_N(\boldsymbol{\theta}) = \frac{1}{2N} \sum_{k=1}^{K} < n_k \nabla^2 \ell_1(\hat{\boldsymbol{\theta}}_k)(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k), \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_k >$ and $\hat{\theta}_k := \arg\min_{\boldsymbol{\theta} \in \Theta} \ell_k(\boldsymbol{\theta})$.

From Equation (25), $K-1$ machines need to transport the local estimators $\hat{\boldsymbol{\theta}}_k$ to the center machine. However, privacy leaks are likely to occur during the transmission of estimates, so privacy protection measures are taken during the transmission process. Therefore, the local estimators $\hat{\boldsymbol{\theta}}_k$ add Laplace noise $\boldsymbol{\eta}_k$ where $\eta_k = (\eta_{k1}, \eta_{k2}, \ldots, \eta_{kp})^\top$ and $\eta_{kj} \overset{\text{i.i.d.}}{\sim} \text{Lap}(0, \frac{2Cp\mu}{n_k \epsilon_k})$ for some constants $C > 0$ and $\mu > 0$, which construct a new loss function that is ϵ-differential privacy:

$$\tilde{\ell}_{N,\text{PV}}(\boldsymbol{\theta}) = \frac{1}{2N} \sum_{k=1}^{K} < n_k \nabla^2 \ell_1(\hat{\boldsymbol{\theta}}_{k,\text{PV}})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{k,\text{PV}}), \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{k,\text{PV}} > \tag{26}$$

$$= \frac{1}{2N} \sum_{k=1}^{K} < n_k \nabla^2 \ell_1(\hat{\boldsymbol{\theta}}_k + \boldsymbol{\eta}_k)[\boldsymbol{\theta} - (\hat{\boldsymbol{\theta}}_k + \boldsymbol{\eta}_k)], \boldsymbol{\theta} - (\hat{\boldsymbol{\theta}}_k + \boldsymbol{\eta}_k) > . \tag{27}$$

Then the corresponding privacy preserving estimator is $\tilde{\boldsymbol{\theta}}_{\text{PV}} := \arg\min_{\boldsymbol{\theta} \in \Theta} \tilde{\ell}_{N,\text{PV}}(\boldsymbol{\theta})$.

The above algorithm guarantees the risk bound and asymptotic normality of the estimator $\tilde{\boldsymbol{\theta}}_{\text{PV}}$ as follows.

Proposition 6.3: *Under some mild assumptions, the risk bound is*

$$\mathbb{E}\left[\|\tilde{\boldsymbol{\theta}}_{\mathrm{PV}}-\boldsymbol{\theta}^*\|_2^2\right] \leqslant \frac{2\mathrm{tr}(\boldsymbol{\Sigma})}{N}+O\left(\frac{K^2}{N^2}\right),$$

where $\boldsymbol{\theta}^*=\arg\min_{\boldsymbol{\theta}\in\Theta}\mathbb{E}_z[\mathcal{L}(\boldsymbol{\theta},X)]$ *and* $\boldsymbol{\Sigma}=[\boldsymbol{I}(\boldsymbol{\theta}^*)]^{-1}E_{\boldsymbol{x}}[\nabla\mathcal{L}(\boldsymbol{\theta}^*,\boldsymbol{X})\nabla\mathcal{L}(\boldsymbol{\theta}^*,\boldsymbol{X})^\top][\boldsymbol{I}(\boldsymbol{\theta}^*)]^{-1}$. *Meanwhile, the privacy preserving estimator satisfies*

$$\sqrt{N}(\tilde{\boldsymbol{\theta}}_{\mathrm{PV}}-\boldsymbol{\theta}^*)\to N(\mathbf{0},\boldsymbol{\Sigma}).$$

From Proposition 6.3, MSE of the privacy preserving estimator is identical to the non-private one; that is, not only the privacy of the data is protected, but also the accuracy of the algorithm remains the same. In addition, under the same assumptions of the non-private algorithm, the privacy preserving algorithm still remains the same rate of convergence without efficacy loss. It illustrates the superiority of the privacy preserving algorithm.

In addition, we consider the extra loss of the privacy-preserving estimator caused by the random noise, especially when the sample size is small. Thus, the traditional plug-in method to estimate $\boldsymbol{\Sigma}$ performs ineffectively. This paper proposed a bootstrap method to estimate $\boldsymbol{\Sigma}$, which reduces the negative influence caused by volatility via averaging. In detail, for the i-th step, K random numbers that are from $N(1,1)$ and independent of raw data are generated, which are denoted as $\xi_{i1},\xi_{i2},\ldots,\xi_{iK}$. Thus, $\frac{N}{G}\sum_{i=1}^{G}\boldsymbol{U}_i\boldsymbol{U}_i^\top$ is an estimator of the middle part of $\boldsymbol{\Sigma}$ for a constant $G>0$, where $\boldsymbol{U}_i=\frac{1}{N}\sum_{k=1}^{K}n_k\xi_{ik}\nabla\hat{\ell}_k(\boldsymbol{\theta}^*)$ and $\nabla\hat{\ell}_k(\boldsymbol{\theta}^*)=\nabla^2\ell_1(\hat{\boldsymbol{\theta}}_{k,\mathrm{PV}})(\tilde{\boldsymbol{\theta}}_{\mathrm{PV}}-\hat{\boldsymbol{\theta}}_{k,\mathrm{PV}})$. Therefore, the global estimator of covariance $\boldsymbol{\Sigma}$ with differential privacy is

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{PV}}=[\hat{I}(\boldsymbol{\theta}^*)]^{-1}\left(\frac{N}{G}\sum_{i=1}^{G}\boldsymbol{U}_i\boldsymbol{U}_i^\top\right)[\hat{I}(\boldsymbol{\theta}^*)]^{-1}, \tag{28}$$

where $\hat{I}(\boldsymbol{\theta}^*)=\nabla^2\ell_1(\tilde{\boldsymbol{\theta}}_{\mathrm{PV}})$.

In summary, the bootstrap scheme to estimate the covariance matrix of the parametric estimators is beneficial to effective inference, which reduces the negative effects of noise in estimating the covariance matrix of parameters. More recent research on distributed algorithms with differential privacy can be found in the literature (Auddy et al., 2024; Avella-Medina et al., 2023; M. Yu et al., 2023).

6.4.3. Statistical inference for differentially private stochastic gradient descent

Important early advances in privacy-preserving deep learning mainly followed two representative directions. First, Shokri and Shmatikov (2015) studied collaborative deep learning without direct data sharing. Their framework enables multiple participants to train neural networks on their local datasets while selectively sharing a small subset of parameters or gradients during stochastic gradient descent, thereby achieving a favourable privacy-utility trade-off. Empirical results on MNIST and SVHN datasets showed that such distributed training can approach the accuracy of centralized learning while reducing direct exposure of training data. Second, Abadi et al. (2016) established differential privacy as a practical tool for deep neural network training by introducing the DP-SGD algorithm. The method clips example gradients and adds Gaussian noise to the aggregated update, thus bounding the influence of any individual training example. They further proposed the moment

accountant to obtain significantly tighter estimates of the cumulative privacy loss over many training steps. This work demonstrated that even deep models with non-convex objectives can be trained under rigorous differential privacy guarantees with manageable utility degradation. Taken together, these studies laid the foundation for subsequent research on privacy-preserving deep learning from both distributed collaborative learning and formal differential privacy perspectives.

Building upon these algorithmic advances, a natural and important question is whether valid statistical inference remains possible in privacy-preserving deep learning frameworks. X. Xia et al. (2025) investigated statistical inference for the output of differentially private stochastic gradient descent (DP-SGD). Unlike most existing studies, which primarily analyze privacy guarantees, optimization error, or excess risk, this paper focuses on whether valid uncertainty quantification remains possible after privacy-preserving noise is injected during iterative optimization. It develops its results in a clear three-step structure, first establishing an asymptotic theory for randomized SGD, then extending this theory to DP-SGD, and finally proposing confidence interval construction methods. For the averaged SGD iterate

$$\bar{\theta}_T = \frac{1}{T}\sum_{t=1}^{T}\theta^{(t)}. \tag{29}$$

Theory results show that under randomized sub-sampling, the estimator remains asymptotically normal. Let

$$A = \nabla^2 L(\theta^*), \quad S = \mathbb{E}\left[\nabla l(Z;\theta^*)\nabla l(Z;\theta^*)^\top\right].$$

When the total number of iterations is $T = kn$ and the mini-batch size is m, the asymptotic variance of randomized SGD is inflated by a factor $1 + 1/(km)$ relative to the classical cyclic-SGD setting, so that

$$\{1 + 1/(km)\}^{-1/2}\sqrt{n}\,(\bar{\theta}_T - \theta^*) \xrightarrow{d} N(0, A^{-1}SA^{-1}).$$

This result provides the inferential foundation for randomized SGD, which is precisely the sampling mechanism required by DP-SGD. Building on this non-private analysis, the authors then study the DP-SGD update

$$\theta^{(t)} = \theta^{(t-1)} - \eta_t\big(g^{(t)} + \xi_t\big), \quad \xi_t \sim N(0, \sigma_1^2 I),$$

where $g^{(t)}$ is the mini-batch gradient and ξ_t is the Gaussian noise added to guarantee privacy. Their main theoretical contribution is to prove the asymptotic normality of the averaged DP-SGD estimator and to show that its asymptotic fluctuation can be decomposed into three distinct components:

$$\sqrt{n}\,(\bar{\theta}_T - \theta^*) \xrightarrow{d} \phi_{\text{stat}} + \phi_{\text{sam}} + \phi_{\text{privacy}}. \tag{30}$$

Here, ϕ_{stat} denotes the ordinary statistical error, ϕ_{sam} captures the extra variation introduced by randomized sub-sampling, and ϕ_{privacy} represents the randomness induced by privacy noise. This three-way decomposition is the key insight of the paper, since it shows that the

uncertainty of DP-SGD is jointly determined by the statistical model, the stochastic sampling mechanism, and privacy perturbation, rather than by privacy noise alone.

Based on the limiting theory, the authors propose two approaches for confidence interval construction. The first is a plug-in method, which privately estimates the Hessian matrix A and the score covariance matrix S, and then forms Wald-type confidence intervals using the estimated asymptotic variance. The second is a random scaling pivotal method, motivated by a functional central limit theorem for the partial-sum process of DP-SGD. Its main advantage is that it avoids explicitly releasing a covariance estimator, thereby reducing additional privacy budget consumption, although the resulting intervals may be wider.

Overall, the contribution of this paper is threefold. First, it establishes the first systematic asymptotic inference theory for randomized SGD, extending the existing literature beyond the cyclic setting. Second, it derives the limiting distribution of averaged DP-SGD and reveals its characteristic decomposition into statistical, sampling, and privacy components. Third, it develops practical confidence interval procedures for DP-SGD, thereby moving differential privacy research beyond optimization and privacy accounting toward formal statistical inference and uncertainty quantification. In this sense, the paper builds an important bridge between modern private machine learning and classical asymptotic statistics. More recent research on DP-SGD can be found in the literature (Altschuler & Talwar, 2022; Bu, Dong et al., 2020; Chourasia et al., 2021).

6.4.4. DP fine-tuning of modern AI models

Recent advances in differential privacy have increasingly shifted from training deep models from scratch to privately fine-tuning modern pretrained or foundation models. This paradigm is motivated by the observation that standard DP-SGD exhibits severe dimension dependence in high-dimensional settings, which often leads to substantial utility degradation. In contrast, representations learned through large-scale public pretraining can significantly alleviate this issue, thereby improving the privacy-utility trade-off during downstream adaptation.

In particular, C. Wang, Zhu et al. (2024) analyzed DP fine-tuning through the lens of Neural Collapse. They show that when the pretrained features at the final layer are sufficiently close to an ideal separable structure, the misclassification error of noisy gradient-based methods can become nearly dimension-independent. This result provides a theoretical explanation for the effectiveness of DP fine-tuning in overparameterized models. However, they also demonstrate that such favourable properties are sensitive to perturbations. To address this issue, they propose practical techniques such as feature normalization and principal component analysis (PCA) to enhance robustness under privacy noise.

Complementing this theoretical perspective, Zhao et al. (2025) studied private fine-tuning of Vision Transformers from a representation learning viewpoint. Rather than focussing solely on the final linear head, they emphasize the role of intermediate representations during fine-tuning. Their empirical results indicate that an improperly chosen hyperparameter can cause DP noise to severely degrade learned representations, particularly in high-privacy regimes. In contrast, with careful tuning, the impact of privacy noise on representation quality can be substantially mitigated.

Taken together, these works highlight that the effectiveness of DP fine-tuning for modern AI models critically depends on the structure and robustness of pretrained representations,

as well as on the design of the fine-tuning procedure. More recent developments along this direction can be found in Bu, Wang et al. (2022), De et al. (2022), and D. Yu et al. (2022).

6.4.5. Enhanced privacy in federated learning

Recent work has shown that, beyond communication efficiency, decentralization itself can strengthen privacy guarantees in federated learning. In particular, Cyffers and Bellet (2022) showed for the first time that formal privacy amplification can arise from full decentralization, because in a peer-to-peer network each participant only observes a local subset of communications rather than all exchanged messages. To capture this partial observability, they introduced network differential privacy, and proved that decentralized protocols based on token passing or random walks can achieve significantly better privacy-utility trade-offs than standard local differential privacy, in some cases nearly matching the trusted-curator model.

Building on this idea, Cyffers et al. (2024) studied differentially private decentralized learning with random walks on arbitrary graphs. Using pairwise network differential privacy (PNDP), they derived closed-form expressions for the privacy loss between pairs of nodes and showed that the communication topology plays a fundamental role in determining privacy leakage. Their results further indicate that random-walk-based methods can provide stronger privacy guarantees than gossip-based methods for nearby nodes, highlighting that privacy in decentralized learning depends not only on the amount of injected noise but also on the graph structure itself.

More recently, X. Li et al. (2025) further advanced this line of work by developing an f-DP framework for decentralized federated learning. Their analysis introduced pairwise network f-DP and secret-based f-local DP, which enable tighter privacy accounting for decentralized communication and correlated noise injection. Compared with previous RDP analyzes, their framework yields tighter (ϵ, δ)-DP guarantees and improves the privacy-utility trade-off. Taken together, these works suggest that decentralization in federated learning should be viewed not only as a system design choice for scalability and robustness, but also as a mechanism that can intrinsically enhance privacy.

6.4.6. Enhanced utility with relaxation of DP

Some relaxations of standard differential privacy can substantially improve utility. A representative example is label differential privacy (label-DP), which assumes that the feature vector is public while only the label is sensitive. This relaxation is particularly suitable for applications such as online advertising and recommendation systems, where the input context is observable but the user outcome or preference should remain private. Under this framework, Ghazi et al. (2021) proposed a multi-stage deep learning method based on randomized response, showing that incorporating prior information across stages can significantly improve predictive performance under the same privacy guarantee. S. Xu et al. (2023) further studied binary classification under local label-DP from a theoretical perspective, proving that the randomized response preserves the Bayes classifier but slows the convergence of excess risk through an ϵ-dependent factor, thereby clarifying the privacy-utility trade-off. In addition, Esmaeili et al. (2021) proposed two representative alternatives, namely the private aggregation of teacher ensembles (PATE) framework and the additive Laplace noise coupled with Bayesian inference (ALIBI), where the latter combines additive Laplace noise on one-hot labels with iterative Bayesian inference for denoising. Their results show

that label-DP can achieve a substantially better privacy-utility trade-off than standard DP in learning tasks where only labels need protection.

7. Further research

As the privacy protection algorithm with the most potential for development, differential privacy models have been widely applied in various fields. However, there are still many gaps in research on the impact of this algorithm on statistical inference and the construction of effective estimation algorithms under this algorithm. We discuss the following three aspects in another article.

First, the research on the effective estimation properties of data transmission based on privacy protection is considered. On the one hand, the addition of random noise provides an effective guarantee for data security; on the other hand, noise introduces estimation bias into data analysis, undermining the validity of the estimation. Moreover, most existing differential privacy algorithms are primarily designed for vector-valued data. When extending these methods to matrix-valued data, a common approach is to flatten the matrix into a vector and then apply vector-based differential privacy mechanisms. However, if the algorithm involves matrix inversion operations (e.g., Newton-type iterative methods), an additional challenge arises that the inversion of a matrix perturbed by noise may amplify the injected noise, thereby further exacerbating the estimation bias. Therefore, it is necessary to develop reliable measurement techniques to accurately determine the impact of noise, prove the asymptotic normality and convergence speed of estimates after noise is added, and effectively estimate the impact of noise on the variance of estimates and eliminate it in the final estimation problem.

Second, we can investigate online data issues under the lens of data privacy. Online data has become increasingly common in applications such as real-time recommendation systems, financial markets, sensor networks, and healthcare monitoring. Compared with traditional offline settings where differential privacy mechanisms are designed based on the entire dataset, online data arrive sequentially. Typically, historical data are compressed into statistical summaries while the original records are discarded. This streaming nature brings new challenges for differential privacy: how to allocate and update privacy parameters in a dynamic environment, and how to design mechanisms that achieve rigorous privacy guarantees when only aggregated statistics and new incoming data are available. Moreover, the sequential arrival of data raises fundamental difficulties, such as the cumulative effect of noise addition. Thus, developing differential privacy mechanisms tailored for online data remains a significance problem.

Third, the research on the impact of random noise on test power is worth to study. Since the test statistic contains a large amount of effective information from the data, it is highly susceptible to information leakage, so it needs to be protected by adding random noise. In addition to white noise, noise following exponential distributions represented by the Laplace distribution has also attracted extensive research. These distributions, due to their simpler form, demonstrate unparalleled computational advantages. However, their heavy-tailed characteristics have a significant impact on the probability of Type I errors and test power. Based on this, we can analyze the extent to which different noise types and privacy protection parameters affect test power, providing theoretical guidance for the reasonable design of privacy protection schemes.

Author contributions

CRediT: **Chengliang Liu:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing; **Miaomiao Yu:** Conceptualization, Methodology, Supervision, Writing – original draft; **Yong Zhou:** Supervision

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported in part by funds from the Program of National Natural Science Foundation of China [Grant no. 72301108], National Key R&D Program of China [Grant nos. 2021YFA1000100, 2021YFA1000101], Shanghai Pujiang Program [Grant no. 23PJC040], State Key Program of National Natural Science Foundation of China [Grant nos. 72331005, 92046005], Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China [Grant no. JYB2025XDXM904], Shanghai Municipal Education Commission [Grant no. 2024AI01002], and Shanghai Pilot Program for Basic Research [Grant no. TQ20240201].

ORCID


Miaomiao Yu http://orcid.org/0000-0002-4536-2078

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 308–318). Association for Computing Machinery (ACM).

Abowd, J. M. (2018). The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2867–2867). Association for Computing Machinery (ACM).

Aggarwal, G., Feder, T., Kenthapadi, K., Motwani, R., Panigrahy, R., Thomas, D., & Zhu, A. (2004). *k-Anonymity: Algorithms and Hardness*. Stanford University.

Altschuler, J., & Talwar, K. (2022). Privacy of noisy stochastic gradient descent: More iterations without more privacy loss. *Advances in Neural Information Processing Systems*, *35*, 3788–3800. https://doi.org/10.52202/068431

Auddy, A., Cai, T. T., & Chakraborty, A. (2024). Minimax and adaptive transfer learning for nonparametric classification under distributed differential privacy constraints. arXiv:2406.20088

Avancha, S., Baxi, A., & Kotz, D. (2012). Privacy in mobile technology for personal healthcare. *ACM Computing Surveys (CSUR)*, *45*(1), 1–54. https://doi.org/10.1145/2379776.2379779

Avella-Medina, M. (2021). Privacy-preserving parametric inference: A case for robust statistics. *Journal of the American Statistical Association*, *116*(534), 969–983. https://doi.org/10.1080/01621459.2019.1700130

Avella-Medina, M., Bradshaw, C., & Loh, P. L. (2023). Differentially private inference via noisy optimization. *The Annals of Statistics*, *51*(5), 2067–2092. https://doi.org/10.1214/23-AOS2321

Awan, J., & Vadhan, S. (2023). Canonical noise distributions and private hypothesis tests. *The Annals of Statistics*, *51*(2), 547–572. https://doi.org/10.1214/23-AOS2259

Balle, B., Barthe, G., & Gaboardi, M. (2018). Privacy amplification by subsampling: Tight analyses via couplings and divergences. In *Advances in Neural Information Processing Systems* (Vol. 31, pp. 6277–6287). Curran Associates, Inc.

Barak, B., Chaudhuri, K., Dwork, C., Kale, S., McSherry, F., & Talwar, K. (2007). Privacy, accuracy, and consistency too: A holistic solution to contingency table release. In *Proceedings of the Twenty-Sixth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (pp. 273–282). Association for Computing Machinery (ACM).

Bayardo, R. J., & Agrawal, R. (2005). Data privacy through optimal k-anonymization. In *21st International Conference on Data Engineering (ICDE'05)* (pp. 217–228). IEEE Computer Society.

Bhagoji, A. N., Chakraborty, S., Mittal, P., & Calo, S. (2019). Analyzing federated learning through an adversarial lens. In *International Conference on Machine Learning* (pp. 634–643). Proceedings of Machine Learning Research (PMLR).

Bi, X., & Shen, X. (2023). Distribution-invariant differential privacy. *Journal of Econometrics*, *235*(2), 444–453. https://doi.org/10.1016/j.jeconom.2022.05.004

Bok, J., Su, W. J., & Altschuler, J. M. (2024). Shifted interpolation for differential privacy. In *Proceedings of the 41st International Conference on Machine Learning*. Proceedings of Machine Learning Research (PMLR).

Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175–1191). Association for Computing Machinery (ACM).

Bu, Z., Dong, J., Long, Q., & Su, W. J. (2020). Deep learning with Gaussian differential privacy. *Harvard Data Science Review*, *2020*(23), 10–1162.

Bu, Z., Wang, Y. X., Zha, S., & Karypis, G. (2022). Differentially private bias-term only fine-tuning of foundation models. arXiv:2210.00036

Cai, T. T., Wang, Y., & Zhang, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, *49*(5), 2825–2850. https://doi.org/10.1214/21-AOS2058

Canonne, C. L., Kamath, G., & Steinke, T. (2020). The discrete Gaussian for differential privacy. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 15676–15688). Curran Associates, Inc.

Chaudhuri, K., Monteleoni, C., & Sarwate, A. D. (2011). Differentially private empirical risk minimization. *Journal of Machine Learning Research*, *12*(3), 1069–1109.

Chawla, S., Dwork, C., McSherry, F., & Talwar, K. (2012). On privacy-preserving histograms. arXiv:1207.1371

Chourasia, R., Ye, J., & Shokri, R. (2021). Differential privacy dynamics of Langevin diffusion and noisy gradient descent. *Advances in Neural Information Processing Systems*, *34*, 14771–14781.

Cohen, A., Duchin, M., Matthews, J., & Suwal, B. (2022). Private numbers in public policy: Census, differential privacy, and redistricting. *Harvard Data Science Review*, (Special Issue 2), 1–43.

Cyffers, E., & Bellet, A. (2022). Privacy amplification by decentralization. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics* (pp. 5334–5353). Proceedings of Machine Learning Research (PMLR).

Cyffers, E., Bellet, A., & Upadhyay, J. (2024). Differentially private decentralized learning with random walks. In *Proceedings of the 41st International Conference on Machine Learning*. Proceedings of Machine Learning Research (PMLR).

Dankar, F. K., & El Emam, K. (2012). The application of differential privacy to health data. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops* (pp. 158–166). Association for Computing Machinery (ACM).

Dankar, F. K., & El Emam, K. (2013). Practicing differential privacy in health care: A review. *Transactions on Data Privacy*, *6*(1), 35–67.

De, S., Berrada, L., Hayes, J., Smith, S. L., & Balle, B. (2022). Unlocking high-accuracy differently private image classification through scale. arXiv:2204.13650

Dinur, I., & Nissim, K. (2003). Revealing information while preserving privacy. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (pp. 202–210). Association for Computing Machinery (ACM).

Dong, J., Roth, A., & Su, W. J. (2022). Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *84*(1), 3–37. https://doi.org/10.1111/rssb.12454

Duchi, J. C., Jordan, M. I., & Wainwright, M. J. (2018). Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, *113*(521), 182–201. https://doi.org/10.1080/01621459.2017.1389735

Dwork, C. (2006). Differential privacy. In *International Colloquium on Automata, Languages, and Programming* (pp. 1–12). Springer.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., & Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (pp. 486–503). Springer.

Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference* (pp. 265–284). Springer.

Dwork, C., McSherry, F., & Talwar, K. (2007). The price of privacy and the limits of LP decoding. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (pp. 85–94). Association for Computing Machinery (ACM).

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, *9*(3-4), 211–407. https://doi.org/10.1561/TCS

Dwork, C., & Yekhanin, S. (2008). New efficient attacks on statistical disclosure control mechanisms. In *Annual International Cryptology Conference* (pp. 469–480). Springer.

Errounda, F. Z., & Liu, Y. (2019). An analysis of differential privacy research in location data. In *2019 IEEE 5th INTL Conference on Big Data Security on Cloud (BigDataSecurity), IEEE INTL Conference on High Performance and Smart Computing,(HPSC) and IEEE INTL Conference on Intelligent Data and Security (IDS)* (pp. 53–60). Institute of Electrical and Electronics Engineers (IEEE).

Esmaeili, M. M., Mironov, I., Prasad, K., Shilov, I., & Tramèr, F. (2021). Antipodes of label differential privacy: PATE and ALIBI. In *Advances in Neural Information Processing Systems* (Vol. 34, pp. 6934–6945). Curran Associates, Inc.

Feldman, V., McMillan, A., & Talwar, K. (2022). Hiding among the clones: A simple and nearly optimal analysis of privacy amplification by shuffling. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (pp. 954–964). Institute of Electrical and Electronics Engineers (IEEE).

Francis, T., Madiajagan, M., & Kumar, V. (2015). Privacy issues and techniques in E-health systems. In *Proceedings of the 2015 ACM SIGMIS Conference on Computers and People Research* (pp. 113–115). Association for Computing Machinery (ACM).

Geng, Q., & Viswanath, P. (2014). The optimal mechanism in differential privacy. In *2014 IEEE International Symposium on Information Theory* (pp. 2371–2375). Institute of Electrical and Electronics Engineers (IEEE).

Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially private federated learning: A client level perspective. arXiv:1712.07557

Ghazi, B., Golowich, N., Kumar, R., Manurangsi, P., & Zhang, C. (2021). Deep Learning with label differential privacy. In *Advances in Neural Information Processing Systems* (Vol. 34, pp. 27131–27145). Curran Associates, Inc.

Gopi, S., Lee, Y. T., & Wutschitz, L. (2021). Numerical composition of differential privacy. In *Advances in Neural Information Processing Systems* (Vol. 34, pp. 11631–11642). Curran Associates, Inc.

Gu, J., & Chen, S. X. (2024). Statistical inference for decentralized federated learning. *The Annals of Statistics*, *52*(6), 2931–2955. https://doi.org/10.1214/24-AOS2452

Hall, R., Rinaldo, A., & Wasserman, L. (2013). Differential privacy for functions and functional data. *The Journal of Machine Learning Research*, *14*(1), 703–727.

Hardt, M., Ligett, K., & McSherry, F. (2012). A simple and practical algorithm for differentially private data release. In *Advances in Neural Information Processing Systems* (Vol. 25). Curran Associates, Inc.

Ji, Z., Lipton, Z. C., & Elkan, C. (2014). Differential privacy and machine learning: A survey and review. arXiv:1412.7584

Jin, H., Luo, Y., Li, P., & Mathew, J. (2019). A review of secure and privacy-preserving medical data sharing. *IEEE Access*, *7*, 61656–61669. https://doi.org/10.1109/Access.6287639

Kenny, C. T., Kuriwaki, S., McCartan, C., Rosenman, E. T. R., Simko, T., & Imai, K. (2021). The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census. *Science Advances*, *7*(41), eabk3283. https://doi.org/10.1126/sciadv.abk3283

Kenny, C. T., McCartan, C., Kuriwaki, S., Simko, T., & Imai, K. (2024). Evaluating bias and noise induced by the U.S. Census Bureau's privacy protection methods. *Science Advances*, *10*(18), eadl2524. https://doi.org/10.1126/sciadv.adl2524

Kim, M., Günlü, O., & Schaefer, R. F. (2021). Federated learning with local differential privacy: Trade-offs between privacy, utility, and communication. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2650–2654). Institute of Electrical and Electronics Engineers (IEEE).

Konečnỳ, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. arXiv:1610.02527

Koskela, A., Jälkö, J., & Honkela, A. (2020). Computing tight differential privacy guarantees using FFT. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics* (pp. 2560–2569). Proceedings of Machine Learning Research (PMLR).

LeFevre, K., DeWitt, D. J., & Ramakrishnan, R. (2005). Incognito: Efficient full-domain *k*-anonymity. In *Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data* (pp. 49–60). Association for Computing Machinery (ACM).

Lei, J. (2011). Differentially private *m*-estimators. In *Advances in Neural Information Processing Systems* (Vol. 24). Curran Associates, Inc.

Li, M., Berrett, T. B., & Yu, Y. (2023). On robustness and local differential privacy. *The Annals of Statistics*, *51*(2), 717–737. https://doi.org/10.1214/23-AOS2267

Li, N., Li, T., & Venkatasubramanian, S. (2006). *t*-closeness: Privacy beyond *k*-anonymity and *l*-diversity. In *2007 IEEE 23rd International Conference on Data Engineering* (pp. 106–115). Institute of Electrical and Electronics Engineers (IEEE).

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50–60. https://doi.org/10.1109/MSP.79

Li, X., Su, B., Wang, C., Long, Q., & Su, W. J. (2025). Mitigating privacy-utility trade-off in decentralized federated learning via *f*-differential privacy. arXiv:2510.19934

Lin, W., Li, B., & Wang, C. (2022). Towards private learning on decentralized graphs with local differential privacy. *IEEE Transactions on Information Forensics and Security*, *17*, 2936–2946. https://doi.org/10.1109/TIFS.2022.3198283

Machanavajjhala, A., Kifer, D., Abowd, J., Gehrke, J., & Vilhuber, L. (2008). Privacy: Theory meets practice on the map. In *2008 IEEE 24th International Conference on Data Engineering* (pp. 277–286). Institute of Electrical and Electronics Engineers (IEEE).

Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. (2007). *L*-diversity: Privacy beyond *k*-anonymity. *ACM Transactions on Knowledge Discovery from Data*, *1*(1), 3–es. https://doi.org/10.1145/1217299.1217302

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics* (pp. 1273–1282). Proceedings of Machine Learning Research (PMLR).

McSherry, F., & Talwar, K. (2007). Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)* (pp. 94–103). Institute of Electrical and Electronics Engineers (IEEE).

McSherry, F. D. (2009). Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data* (pp. 19–30). Association for Computing Machinery (ACM).

Nasr, M., Shokri, R., & Houmansadr, A. (2018). Comprehensive privacy analysis of deep learning. In *Proceedings of the 2019 IEEE Symposium on Security and Privacy (SP)* (Vol. 2018, pp. 1–15). Institute of Electrical and Electronics Engineers (IEEE).

Nissim, K., Raskhodnikova, S., & Smith, A. (2007). Smooth sensitivity and sampling in private data analysis. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (pp. 75–84). Association for Computing Machinery (ACM).

Peng, H., Jin, K., Fu, C., Fu, N., & Zhang, X. (2020). Private time series pattern mining with sequential lattice. *Acta Electronica Sinica*, *48*(1), 153–163.

Rubinstein, B. I., Bartlett, P. L., Huang, L., & Taft, N. (2009). Learning in a large function space: Privacy-preserving mechanisms for SVM learning. arXiv:0911.5708

Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, *40*(2), 99–121. https://doi.org/10.1023/A:1026543900054

Sadki, S., & El Bakkali, H. (2014). Enhancing privacy on mobile health: An integrated privacy module. In *2014 International Conference on Next Generation Networks and Services (NGNS)* (pp. 245–250). Institute of Electrical and Electronics Engineers (IEEE).

Samarati, P. (2002). Protecting respondents identities in microdata release. *IEEE Transactions on Knowledge and Data Engineering*, *13*(6), 1010–1027. https://doi.org/10.1109/69.971193

Shaikh, A., & Patil, S. (2018). Role of differential privacy in a new age data privacy environment. *International Journal of Pure and Applied Mathematics*, *118*, 24.

Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (pp. 1310–1321). Association for Computing Machinery (ACM).

Smith, A. (2011). Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (pp. 813–822). Association for Computing Machinery (ACM).

Su, B., Su, W. J., & Wang, C. (2025). The 2020 US Decennial Census is more private than you (might) think. *Proceedings of the National Academy of Sciences*, *122*(45), e2500337122. https://doi.org/10.1073/pnas.2500337122

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(5), 557–570. https://doi.org/10.1142/S0218488502001648

Tan, Z., & Zhang, L. (2020). Survey on privacy preserving techniques for machine learning. *Journal of Software*, *31*(7), 2127–2156.

Wang, C., Su, B., Ye, J., Shokri, R., & Su, W. J. (2023). Unified enhancement of privacy bounds for mixture mechanisms via f-differential privacy. In *Advances in Neural Information Processing Systems* (Vol. 36, pp. 55051–55063). Curran Associates, Inc.

Wang, C., Zhu, Y., Su, W. J., & Wang, Y. X. (2024). Neural collapse meets differential privacy: Curious behaviors of NoisyGD with near-perfect representation learning. In *Proceedings of the 41st International Conference on Machine Learning*. Proceedings of Machine Learning Research (PMLR).

Wang, H., Gao, S., Zhang, H., Shen, M., & Su, W. J. (2022). Analytical composition of differential privacy via the edgeworth accountant. arXiv:2206.04236

Wang, J., Zhu, R., Liu, S., & Cai, Z. (2018). Node location privacy protection based on differentially private grids in industrial wireless sensor networks. *Sensors*, *18*(2), 410. https://doi.org/10.3390/s18020410

Wang, P., & Zhang, H. (2020). Distributed privacy-preserving logistic regression. *Scientia Sinica Informationis*, *50*(10), 18.

Wasserman, L (2006). *All of Nonparametric Statistics*. Springer.

Wasserman, L., & Zhou, S. (2010). A statistical framework for differential privacy. *Journal of the American Statistical Association*, *105*(489), 375–389. https://doi.org/10.1198/jasa.2009.tm08651

Wu, Y., Ge, C., Zhang, L., & Sun, L. (2017). Differential privacy stream data publishing algorithm based on matrix mechanism under exponential decay model. *Scientia Sinica Informationis*, *47*(11), 17.

Xia, X., Zhang, L., & Cai, Z. (2025). Statistical inference for differentially private stochastic gradient descent. arXiv:2507.20560

Xiao, Y., Gardner, J., & Xiong, L. (2012). Dpcube: Releasing differentially private data cubes for health information. In *2012 IEEE 28th International Conference on Data Engineering* (pp. 1305–1308). Institute of Electrical and Electronics Engineers (IEEE).

Xiong, P., Zhu, T., & Wang, X. (2014). A survey on differential privacy and applications. *Chinese Journal of Computers*, *37*(1), 22.

Xu, L., Jiang, C., Wang, J., Yuan, J., & Ren, Y. (2014). Information security in big data: Privacy and data mining. *IEEE Access*, *2*, 1149–1176. https://doi.org/10.1109/ACCESS.2014.2362522

Xu, S., Wang, C., Sun, W. W. Y., & Cheng, G. (2023). Binary classification under local label differential privacy using randomized response mechanisms. *Transactions on Machine Learning Research*, 1–26.

Yang, X., Gao, L., Wang, H., Guo, H., & Zheng, J. (2020). Balanced correlation differential privacy protection method for histogram publishing. *Chinese Journal of Computers*, *43*(8), 19.

Yu, D., Naik, S., Backurs, A., Gopi, S., Inan, H. A., Kamath, G., Kulkarni, J., Lee, Y. T., Manoel, A., Wutschitz, L., Yekhanin, S., & Zhang, H. (2022). Differentially private fine-tuning of language models. In *International Conference on Learning Representations*. OpenReview.

Yu, M., Li, J., & Zhou, Y. (2026). Enhancements of communication-efficient distributed statistical inference and its privacy preservation. *Journal of Econometrics*, *253*,106125. https://doi.org/10.1016/j.jeconom.2025.106125

Yu, M., Li, Z., & Zhou, Y. (2023). Privacy-preserving parameter estimation in distributed cases. *Acta Mathematicae Applicatae Sinica*, *46*(2), 145–165.

Zhang, J., Cormode, G., Procopiuc, C. M., Srivastava, D., & Xiao, X. (2017). PrivBayes: Private data release via Bayesian networks. *ACM Transactions on Database Systems*, *42*(4), 25–41. https://doi.org/10.1145/3134428

Zhao, Y., Xia, Y., & Wang, C. (2025). Why does private fine-tuning resist differential privacy noise? A representation learning perspective. In *ICLR 2025 Workshop on Navigating and Addressing Data Problems for Foundation Models*. OpenReview.