Context-Aware Local Information Privacy

Bo Jiang, *Student Member, IEEE*, Mohamed Seif, *Student Member, IEEE*, Ravi Tandon, *Senior Member, IEEE*, and Ming Li, *Senior Member, IEEE*

***Abstract*— In this paper, we study Local Information Privacy (LIP). As a context-aware privacy notion, LIP relaxes the *de facto* standard privacy notion of local differential privacy (LDP) by incorporating prior knowledge and therefore achieving better utility. We study the relationships between LIP and some of the representative privacy notions including LDP, mutual information and maximal leakage. We show that LIP provides strong instance-wise privacy protection compared to other context-aware privacy notions. Moreover, we present some useful properties of LIP, including post-processing, linkage, composability, transferability and robustness to imperfect prior knowledge. Then we study a general utility-privacy tradeoff framework, under which we derive LIP based privacy-preserving mechanisms for both discrete and continuous-valued data. Three types of perturbation mechanisms are studied in this paper: 1) randomized response (RR), 2) random sampling (RS) and 3) additive noise (AN) (e.g., Gaussian mechanism). Our privacy mechanisms incorporate the prior knowledge into the perturbation parameters so as to enhance utility. Finally, we present a comprehensive set of experiments on real datasets to illustrate the advantage of context-awareness and compare the utility-privacy tradeoffs provided by different mechanisms.**

***Index Terms*— Privacy-preserving data aggregation, local information privacy, information-theoretic privacy.**

I. Introduction

PERSONALIZED data collection is becoming pervasive, and data is the key enabler that drives applications spanning all sectors of our society, including e-commerce, social networking, and healthcare. On one hand, collecting data at a fine granularity can provide higher utility (such as in recommendation systems, location-based services and precision medicine). On the other hand, without rigorous privacy-preserving mechanisms, there is a risk of potential privacy breaches which often come with a psychological and socio-economic impact. Such privacy breaches are becoming increasingly commonplace, and could be intentional or unintentional. Massive data breaches (2013 at Yahoo [1], 2015 at Equifax [2], 2018 at Facebook [3], 2019 at Capital One [4]) are just a few examples in the former category. Besides malicious activities, seemingly benign data collected directly from individuals also causes potential privacy leakage by inference [5], [6]. For example, daily activity information collected from wearable health trackers could be used to infer sensitive disease information [7], [8]; mobile location traces collected by large companies can leak people's social relationships [9], [10].

Within the data privacy community, Differential Privacy (DP) [11]–[13] is the *de facto* standard notion for providing rigorous privacy guarantees. DP guarantees that each user's presence in the dataset has minimal statistical influence (measured by the privacy budget ϵ) on the output of queries. While DP has been applied in several applications such as surveying demographics and commuting patterns [14], and the 2020 U.S. Census [15], it assumes a centralized trusted server to collect data and answer queries. In contrast, local differential privacy (LDP), which is a local variant of DP, has gained significant recent attention. In the local setting the server/aggregator who collects data is considered as untrusted by the users, who perturb their own data before sending them to the server. LDP based mechanisms have been successfully adopted by Google's RAPPOR [16] for collecting web browsing behavior, and Apple's MacOS to identify popular emojis and media preferences in Safari [17], [18]. On one hand, DP/LDP based notions provide strong privacy guarantees against adversaries that may possess arbitrary side information. On the other hand, this leads to significant utility degradation, which is more pronounced in the local setting [19], [20], where more noise is needed to achieve the same level of privacy as in a centralized setting. That is, for a summation/count query, with additive noise privacy-preserving mechanism, a lower bound of noise magnitude of $\Omega(\sqrt{N})$ is required for LDP in order to defend against potential coalitions of compromised users, where N is the number of users. In contrast, only $O(1)$ is required for central DP [21]. One of the reasons is that DP/LDP notions do not take the particular contextual knowledge of the data into account.

Contextual information often exists in various applications, which may include prior knowledge about the data distribution, correlations within the data, user privacy expectations/different input sensitivity levels, etc. For instance, in location-based services, people are more likely to be present at some locations than others (e.g., in Paris, people are more likely closer to Eiffel tower than a coffee shop nearby [22]); in mobile-health data, background knowledge such as likelihood of a certain disease are available through previously published

Manuscript received December 4, 2020; revised March 7, 2021; accepted May 10, 2021. Date of publication June 7, 2021; date of current version July 28, 2021. This work was supported by NSF under Grant CNS-1715947, Grant CAREER-1651492, Grant CCF-2100013, and Grant CNS-1731164. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anna Squicciarini. *(Bo Jiang and Mohamed Seif are co-first authors.) (Corresponding author: Bo Jiang.)*

The authors are with the Department of Electrical and Computer Engineering, The University of Arizona, AZ 85718 USA (e-mail: bjiang@email.arizona.edu; mseif@email.arizona.edu; tandonr@email.arizona.edu; lim@email.arizona.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TIFS.2021.3087350, provided by the authors.

Digital Object Identifier 10.1109/TIFS.2021.3087350

medical studies [23]; privacy-sensitive attributes (e.g., sexual orientation, religious or political affiliations) may be different than but correlated with the raw data (such as the web browsing history of a user). While the privacy community has been trying to avoid modeling the background knowledge explicitly, there is a recent trend among the privacy research community to leverage such contextual knowledge or partial contextual knowledge and incorporate them in privacy notions to ultimately provide better utility-privacy tradeoffs. Intuitively, the main advantage is the amount of noise/perturbation can be adjusted depending on the different input instances/prior distributions.

Existing *Context-aware* privacy notions can be broadly classified into two categories: 1) those providing average case guarantees (e.g., mutual information [24]–[26], maximal leakage [27]), or 2) those providing worst-case (or per-instance) guarantees (e.g., information privacy [28], Pufferfish [29], Bayesian DP [30], membership privacy [31]). Notions providing average case guarantees are weaker than the latter since they cannot bound the leakage for all the input and output pairs, which may not be easily adopted by the users who are privacy-sensitive. On the other hand, some of the worst-case context-aware privacy notions (such as Pufferfish and Bayesian DP) still follow the same structure of DP/LDP – the maximum ratio between two likelihoods of a certain output given different input data or secrets.

In this paper, we study local information privacy (LIP), which guarantees that the ratio of the posterior and prior of the input data are bounded. The IP (original cerntralized version of posterior/prior ratio) was proposed in [28], then the local variant called LIP was presented in our previous work in [32], where preliminary results on binary randomized response mechanisms are presented. Then, in our recent work [33], relationships between LIP and some other privacy notions such as maximal leakage and Pufferfish were studied, and several RR based mechanisms for LIP were proposed for discrete-valued data aggregation applications, including weighted sum and frequency estimation. In [34], we present preliminary context-aware additive noise (Laplacian) mechanism for discrete valued data. This notion not only acts as a natural bridge between average and worst-case privacy notions (since mutual information is the expected $log(\cdot)$ of the ratio of posterior to prior), but it can also be related to DP/LDP. The main advantages of LIP are that, 1) it can be readily adapted to account for a variety of scenarios with different contextual/prior knowledge and for any given prior, the leakage of the data in measured by LIP increases linearly with the number of output (linear composability). 2) it leads to simple modular mechanism designs with low complexity: To design a mechanism satisfying LIP for a prior distribution P is sufficient to design K sub-mechanisms, each satisfying LIP for a prior P', where P is a convex combination of P's. On the other hand, for any given prior distribution, LIP needs only $O(n^2)$ constraints for all input-output combinations in contrast to $O(n^3)$ for LDP-like notions. Building on our previous works, the key goal of this paper is to lay down the theoretical foundations of LIP and designing practical context-aware privacy mechanisms in the local setting.

A. Contributions

The contributions of this paper are summarized as follows:

(1) We present and categorize some of the most representative localized privacy notions, including local differential privacy, mutual information privacy, maximal information leakage, differential identifiability, and study the relationship between them and LIP. In this work we generalize ϵ-LIP into (ϵ, δ)-Local Information Privacy, and we prove tighter bounds on the relation between LIP and those definitions than the ones in our previous works [33]. We show that LIP provides a strong instance-wise privacy guarantee compared with other privacy notions, and the structure in the definition of LIP is amenable to the efficient design of privacy-preserving mechanisms.

(2) We present an in-depth study of useful properties of LIP, including post-processing, composability, transferability of mechanisms from one prior to another, and modularity of mechanism design, which is studied for the first time in the context-aware setting. We use these properties to understand how prior knowledge affects information leakage in different scenarios, such as multiple data releases and when the knowledge of the prior may be imperfect.

(3) We present a utility-privacy tradeoff framework that focuses on maximizing a general class of utility functions subject to LIP privacy constraints. We also present *prior-aware* perturbation mechanisms for both discrete and continuous-valued data that satisfy LIP, including randomized response (RR) mechanism, random sampling (RS) mechanism, and additive noise (AN) mechanisms (Gaussian noise and Laplacian noise). Also, we extend our previous mechanism in [32] into continuous data, and input instance-dependent noise distributions.

(4) We present a comprehensive set of experiments conducted on real-world datasets to validate our analysis and compare the utility-privacy tradeoff of our proposed LIP based mechanisms to those based on LDP for both discrete and continuous-valued data. We showed that under each scenario, LIP-based mechanisms provide enhanced utility than those based on context-free LDP.

B. Paper Organization

The remainder of the paper is organized as follows: In Section II, we investigate the definition of LIP along with some other privacy definitions and then derive the relationships among them. In Section III, we study LIP related properties and the impact of context-awareness on privacy leakage. In Section IV, we present different LIP based mechanisms for both discrete and continuous-valued data, including randomized response, random sampling, as well as additive noise mechanisms. Finally, in Section V, we conduct numerical simulations on real data to show the utility-privacy tradeoffs provided by proposed mechanisms with comparisons to LDP based mechanisms. The notations used throughout this paper are listed in Table I.

II. Local Information Privacy and Relationships With Existing Privacy Notions

The system model of the privacy-preserving data release problem shown in Fig. 1 can be summarized as follows:

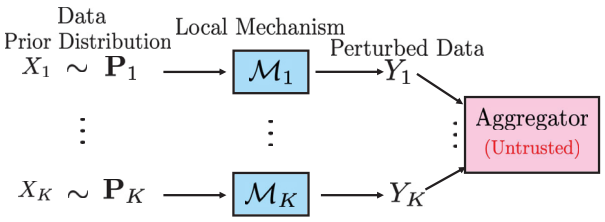


Fig. 1. System model for context-aware local privacy-preserving data release, users privatize data independently according to his/her prior distribution.

TABLE I
LIST OF SYMBOLS

X	Raw data (Input random variable)	$\mathcal{X}$	Input alphabets for X
$\mathbf{P}$	The prior distribution of X	S_x	A subset of $\mathcal{X}$
$\mathscr{P}_{\mathcal{X}}$	Family of prior on X	P_X	PMF of X
$\mathscr{P}_{\mathcal{X}}^{bp}$	A bounded subset of $\mathscr{P}_{\mathcal{X}}$	f_X	PDF of X
Y	Released data (Output random variable)	$\mathcal{Y}$	Output alphabets for Y
$\mathcal{M}$	Privacy preserving mechanism	ϵ	Privacy budget
δ	Failure probability of a mechanism	$\mathcal{L}$	Leakage of a mechanism

each individual possesses data or answer to some query that needs to be published or submitted to an untrusted curator. It is assumed that different users privatize data locally and independently, and we take an arbitrary user as an example to formulate the local problem. Denote $X \in \mathcal{X}$ as the data or query answer, where $\mathcal{X}$ is the support of X. In the context-aware setting, X is regarded as a random variable with a prior distribution of $\mathbf{P} \in \mathscr{P}_{\mathcal{X}}$, where $\mathscr{P}_{\mathcal{X}}$ denotes the set of all possible prior distributions on $\mathcal{X}$. When X is discrete, $\mathbf{P}$ denotes the probability mass function (PMF), and when X is continuous, $\mathbf{P}$ denotes the probability density function (PDF). To maintain the privacy of X, before releasing it, a privacy-preserving mechanism $\mathcal{M}$ perturbs the input X to satisfy a certain privacy guarantee. Denote $Y \in \mathcal{Y}$ as the output of $\mathcal{M}$, or the perturbed version of X. $\mathcal{Y}$ is the support of Y, typically, $\mathcal{Y} = \text{Range}(\mathcal{M})$.

A. Background on Privacy Definitions

We now list some of the related privacy metrics.

Definition 1 ((ϵ, δ)-Local Differential Privacy): A mechanism $\mathcal{M}$ satisfies (ϵ, δ)-LDP for some $\epsilon \in \mathbb{R}^+$ and $\delta \in [0, 1]$, if for any two measurable sets $S'_x, S_x \in \mathcal{X}$ and $S_y \in Range(\mathcal{M})$:

$$\Pr(Y \in \mathcal{S}_y | X \in S_x) \leq e^{\epsilon} \Pr(Y \in \mathcal{S}_y | X \in \mathcal{S}'_x) + \delta. \quad (1)$$

When X and Y are both discrete valued, the definition of $(\epsilon, 0)$-LDP (referred to as ϵ-LDP) can be written as:

$$\mathsf{LR}(y, x, x') = \frac{P_{Y|X}(y|x)}{P_{Y|X}(y|x')} \leq e^{\epsilon},$$

where $\mathsf{LR}(y, x, x')$ denotes the ratio of two likelihoods.

Definition 2 (ϵ-Local Mutual Information Privacy (L-MIP) [25]): A mechanism $\mathcal{M}$ satisfies ϵ-L-MIP for some $\epsilon \in \mathbb{R}^+$, if the mutual information between X and Y is bounded by ϵ, i.e., $I(X; Y) \leq \epsilon$.

Definition 3 (ϵ-Differential Identifiability (DI) [35]): A mechanism $\mathcal{M}$ satisfies ϵ-DI for some $\epsilon \in \mathbb{R}^+$, if $\forall S_x, S'_x \in \mathcal{X}$ and $\forall S_y \in Range(\mathcal{M})$ if

$$\frac{\Pr(X \in \mathcal{S}_x | Y \in \mathcal{S}_y)}{\Pr(X \in \mathcal{S}'_x | Y \in \mathcal{S}_y)} \leq e^{\epsilon}.$$

Definition 4 (ϵ-Maximal Information Leakage (MIL) [27]): The maximal information leakage for a mechanism $\mathcal{M}$ is

$$\mathcal{L}_{MIL}(X; Y) = \sup_{U - X - Y - \hat{U}} \ln \frac{\Pr(U = \hat{U})}{\max_{u \in \mathcal{U}} P_U(u)}, \quad (2)$$

where U is a (possibly randomized) function of X, $\hat{U}$ denotes a guess from the adversary. For any joint distribution $P_{X,Y}$ on finite alphabets $\mathcal{X}$ and $\mathcal{Y}$, (2) can be rewritten as

$$\mathcal{L}_{MIL}(X; Y) = \ln \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} P_{Y|X}(y|x)$$

and $\mathcal{M}$ satisfies ϵ-Maximal Information Leakage Privacy if $\mathcal{L}_{MIL}(X; Y) \leq \epsilon$.

B. Local Information Privacy

Next, we introduce the notion of local information privacy [32], which was studied in our prior work. Information Privacy was originally proposed for the central setting in [28], where X denotes the dataset that holds all users' data. The difference between the two settings lies in the meaning of X.

Definition 5 ((ϵ, δ)-Local Information Privacy): A mechanism $\mathcal{M}$ satisfies (ϵ, δ)-LIP for some $\epsilon \in \mathbb{R}^+$ and $\delta \in [0, 1]$, if $\forall S_x \in \mathcal{X}$, $S_y \in Range(\mathcal{M})$:

$$\begin{aligned} \Pr(Y \in \mathcal{S}_y) &\geq e^{-\epsilon} \Pr(Y \in \mathcal{S}_y | X \in \mathcal{S}_x) - \delta, \\ \Pr(Y \in \mathcal{S}_y) &\leq e^{\epsilon} \Pr(Y \in \mathcal{S}_y | X \in \mathcal{S}_x) + \delta. \end{aligned} \quad (3)$$

The operational meaning of LIP is, the output Y provides limited additional information about any possible input X, and the amount of the additional information is measured by the privacy budget ϵ and failure probability δ. Note that, when ϵ is small, the posterior probability of X given Y is close to the prior of X. For discrete valued data, (3) is equivalent to:

$$e^{-\epsilon} P_{Y|X}(y|x) - \delta \leq P_Y(y) \leq e^{\epsilon} P_{Y|X}(y|x) + \delta. \quad (4)$$

Specifically, when $\delta = 0$, pure ϵ-LIP [32] is satisfied if:

$$e^{-\epsilon} \leq \frac{\Pr(X \in \mathcal{S}_x)}{\Pr(X \in \mathcal{S}_x | Y \in \mathcal{S}_y)} \leq e^{\epsilon}. \quad (5)$$

Compared with the pure $(\epsilon, 0)$-LIP, the (ϵ, δ)-LIP allows a certain probability of failure when there exist some subsets in $\mathcal{X}$ and $\mathcal{Y}$ not satisfying the LIP constraints, and δ captures the failure probability.

Corollary 1: The definition of $(\epsilon, 0)$-LIP (referred to as ϵ-LIP) can also be written as the ratio between two distributions:

- *For discrete-valued X, ϵ-LIP is equivalent to: $\forall x \in \mathcal{X}$, $y \in \mathcal{Y}$:*

$$e^{-\epsilon} \leq \frac{P_X(x)}{P_{X|Y}(x|y)} \leq e^{\epsilon}. \quad (6)$$

- *For continuous-valued X, ϵ-LIP is equivalent to: $\forall x \in \mathcal{X}$, $y \in \mathcal{Y}$:*

$$e^{-\epsilon} \leq \frac{f_X(x)}{f_{X|Y}(x|y)} \leq e^{\epsilon}. \quad (7)$$

In the remainder of this paper, we denote $P_X(x)$ as the prior distribution of X for both discrete and continuous-valued data.

Then, we define the leakage of a mechanism $\mathcal{M}$ under LIP when $\delta = 0$ as follows:

$$\mathcal{L}_{\mathsf{LIP}}(X, Y) \triangleq \sup_{x\in\mathcal{X}, y\in\mathcal{Y}} \ln\left(\max\left(\frac{P_{X|Y}(x|y)}{P_X(x)}, \frac{P_X(x)}{P_{X|Y}(x|y)}\right)\right). \tag{8}$$

We next generalize the definition of LIP by considering other scenarios depending on the underlying assumptions on prior knowledge availability.

Definition 6 (Bounded Prior/Worst-Case ϵ-Local Information Privacy (BP/WC-LIP)): Denote $\mathscr{P}_{\mathcal{X}}^{bp}$ as a subset of $\mathscr{P}_{\mathcal{X}}$, then a mechanism $\mathcal{M}$ satisfies ϵ-BP/WC-LIP for some $\epsilon \in \mathbb{R}^+$, if $\forall \boldsymbol{P} \in \mathscr{P}_{\mathcal{X}}^{bp}/\mathscr{P}_{\mathcal{X}}$, $\mathcal{Y} = Range(\mathcal{M})$, (6) is satisfied.

Similar to LDP, WC-LIP also provides context-free privacy protection. On the other hand, BP-LIP provides a connection between context-free and context-aware guarantees by initiating the size of $\mathscr{P}_{\mathcal{X}}^{bp}$. If $\mathscr{P}_{\mathcal{X}}^{bp}$ contains all possible priors, BP-LIP is equivalent to WC-LIP, on the other hand, if $\mathscr{P}_{\mathcal{X}}^{bp}$ contains only one possible prior, BP-LIP is equivalent to fixed prior LIP.

C. Comparison Between LIP and Other Local Privacy Notions

In the following part of this Section, we derive the relationships among different privacy notions and LIP.

1) LIP v.s. LDP for Discrete-Valued Data: We first compare the relationship between LIP and LDP with discrete-valued data for a fixed prior. From our previous work [32], ϵ-LDP implies ϵ-LIP, and ϵ-LIP implies 2ϵ-LDP. Our goal in this subsection is to understand if this relationship can be tightened, and how it depends on the prior of the data. We state our first result in the following Theorem.

Theorem 1 (Relationship Between LIP and LDP): If $\mathcal{M}$ satisfies ϵ-LIP, it satisfies $\tilde{\epsilon}(\mathbf{P})$-LDP, where $\tilde{\epsilon}(\mathbf{P})$ can be calculated by solving the following optimization problem:

$$\tilde{\epsilon}(\mathbf{P}) \triangleq \max_{x,x'\in\mathcal{X}, y\in\mathcal{Y}} \frac{\mathsf{LR}(y, x', 1)}{\mathsf{LR}(y, x, 1)},$$
$$s.t. \begin{cases} \mathsf{LR}(y, x, 1) \geq 0, \quad \forall x \in \mathcal{X}, \ y \in \mathcal{Y}, \\ \sum_{x\in\mathcal{X}} P_X(x)\mathsf{LR}(y, x, 1) \\ \quad \geq e^{-\epsilon} \max_{x\in\mathcal{X}, y\in\mathcal{Y}}\{\mathsf{LR}(y, x, 1)\}, \\ e^{\epsilon} \min_{x\in\mathcal{X}, y\in\mathcal{Y}}\{\mathsf{LR}(y, x, 1)\} \\ \quad \geq \sum_{x\in\mathcal{X}} P_X(x)\mathsf{LR}(y, x, 1). \end{cases} \tag{9}$$

Conversely, if $\mathcal{M}$ satisfies ϵ-LDP, it satisfies $\ln(P_{\min} + e^{\epsilon}(1 - P_{\min}))$-LIP, where $P_{\min} = \min_{x\in\mathcal{X}} P_X(x)$.

In addition, based on the optimization problem in Eq. (9), we derive a looser closed-form bound which corresponds to the following corollary.

Proposition 1: If $\mathcal{M}$ satisfies ϵ-LIP, it satisfies $\min\{2\epsilon, \ln\frac{e^{\epsilon}-1+P_{\min}}{P_{\min}}\}$-LDP, where $P_{\min} = \min_{x\in\mathcal{X}} P_X(x)$.

Detailed proofs of Theorem 1 and Proposition 1 are presented in Appendix A of the supplementary document. We next visualize the leakage of LDP given ϵ-LIP by simulation on synthetic data with $|\mathcal{X}| = 4$ under two possible prior distributions: $\mathbf{P}_1 = [0.01, 0.33, 0.33, 0.33]$, $\mathbf{P}_2 = [0.25, 0.25, 0.25, 0.25]$. We compare the maximum $\mathsf{LR}(y, x', x)$ derived in Theorem 1 with the loose bounds in Proposition 1 under $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively. The results are presented in Fig. 2(a). Observe that, when a mechanism satisfies ϵ-LIP, then the leakage of the mechanism under LDP is always sandwiched between ϵ and 2ϵ under different priors. When the prior is uniformly distributed (i.e., $\mathbf{P}_2$), the curve of the theoretical bound perfectly overlaps with the maximal LR. When the prior is more skewed ($P_{\min}$ is smaller), the leakage of LDP increases as a result. Intuitively, ϵ-LIP guarantees the ratio of prior to posterior to be bounded for a fixed prior, but the leakage of LDP examines the ratio of prior to posterior for any arbitrary prior.

Next, we show the relationship between BP-LIP and LDP in the following Theorem.

Theorem 2 (Relationship Between BP-LIP and LDP): If a mechanism $\mathcal{M}$ satisfies ϵ-LDP, it satisfies

$$\ln\left\{\min P_{\min}^{bp} + e^{\epsilon}(1 - \min P_{\min}^{bp})\right\}\text{-BP-LIP}.$$

Conversely, if a mechanism $\mathcal{M}$ satisfies ϵ-BP-LIP, it satisfies

$$\min\left\{2\epsilon, \ln\frac{e^{\epsilon} - 1 + \max P_{\min}^{bp}}{\max P_{\min}^{bp}}\right\}\text{-LDP},$$

where $\min P_{\min}^{bp} = \min_{x\in\mathcal{X}, \boldsymbol{P}\in\mathscr{P}_{\mathcal{X}}^{bp}} P_X(x)$ *and* $\max P_{\min}^{bp} = \max_{\boldsymbol{P}\in\mathscr{P}_{\mathcal{X}}^{bp}} \min_{x\in\mathcal{X}} P_X(x)$.

The proof is provided in Appendix B of the supplementary document.

Notice that $\min P_{\min}^{bp} \leq P_{\min} \leq \max P_{\min}^{bp}$, for any $\mathbf{P} \in \mathscr{P}_{\mathcal{X}}^{bp}$. Thus, $\min P_{\min}^{bp} + e^{\epsilon}(1 - \min P_{\min}^{bp}) \leq P_{\min} + e^{\epsilon}(1 - P_{\min})$, also, $\min\left\{2\epsilon, \ln\frac{e^{\epsilon}-1+\max P_{\min}^{bp}}{\max P_{\min}^{bp}}\right\} \leq \min\left\{2\epsilon, \ln\frac{e^{\epsilon}-1+P_{\min}}{P_{\min}}\right\}$. This observation implies that the privacy guarantee provided by ϵ-BP-LIP approaches that of ϵ-LDP as the size of the set of $\mathscr{P}_{\mathcal{X}}^{bp}$ increases. As an extreme case, we have the following corollary.

Corollary 2: ϵ-WC-LIP is equivalent to ϵ-LDP.

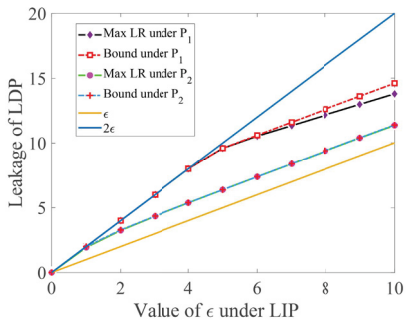
The equivalence of these two definitions means that in the worst-case setting, the LIP privacy notion (if satisfied for all possible priors), is equivalent to LDP. By Corollary 2, BP-LIP connects the setting of context-aware to context-free through the notion of priors, $\mathscr{P}_{\mathcal{X}}^{bp}$.

We next present the relationship between LIP and LDP for the approximate case in the following Corollary.

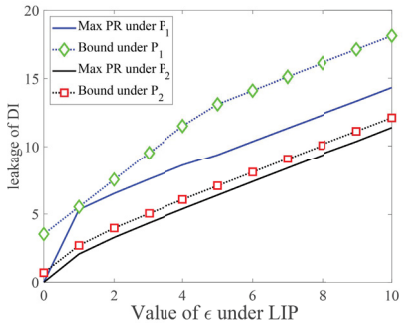
Corollary 3 (Relationship Between Approximate LIP and Approximate LDP): If a mechanism $\mathcal{M}$ satisfies (ϵ, δ)-LDP, then it also satisfies (ϵ, δ)-LIP. Conversely, if $\mathcal{M}$ satisfies (ϵ, δ)-LIP, then it satisfies $(2\epsilon, (e^{\epsilon} + 1)\delta)$-LDP.

The proof steps follow on similar lines from our previous work [34].

2) LIP v.s. Other Privacy Notions: The following Theorem states the relationships between LIP and other context-aware privacy notions.



(a) Bounds and Leakages of LDP under ϵ-LIP with different priors.



(b) Bounds and Leakages of DI under ϵ-LIP with different priors.

Fig. 2. Comparison between theoretical bounds and exact leakage under two different priors: $\mathbf{P}_1 = [0.01, 0.33, 0.33, 0.33]$, $\mathbf{P}_2 = [0.25, 0.25, 0.25, 0.25]$.

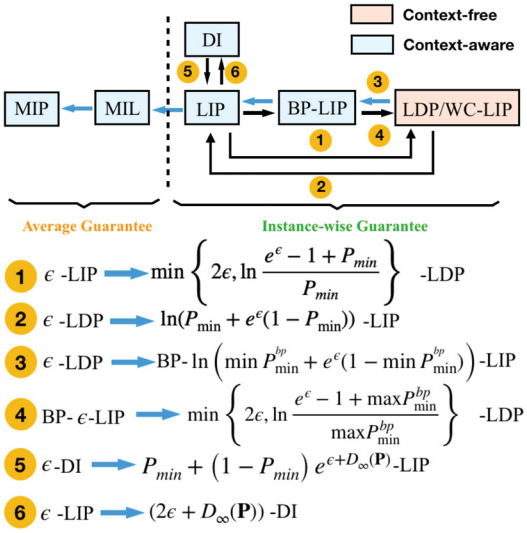


Fig. 3. Relationships between LIP and other privacy notions.

Theorem 3: Relationships between LIP and other context-aware privacy notions.

1) *LIP v.s. DI: ϵ-LIP implies $(2\epsilon + D_\infty(\mathbf{P}))$-DI, conversely, ϵ-DI implies $\ln[P_{\min} + (1 - P_{\min})e^{\epsilon + D_\infty(\mathbf{P})}]$-LIP, where $D_\infty(\mathbf{P}) = \max_{x,x' \in \mathcal{X}} \ln \frac{P_X(x)}{P_X(x')}$.*
2) *LIP v.s MIP: ϵ-LIP implies ϵ-MIP, conversely, ϵ-MIP does not necessarily imply ϵ'-LIP for any $\epsilon' \geq 0$.*
3) *LIP v.s. MIL: ϵ-LIP implies ϵ-MIL, conversely, ϵ-MIL does not necessarily imply ϵ'-LIP for any $\epsilon' \geq 0$.*

The proof is provided in Appendix C of the supplementary document. The relationships between LIP and other related privacy notions are summarized in Fig. 3. LDP provides the strongest context-free instance-wise privacy protection compared with context-aware privacy notions. Among context-aware privacy notions, the relationship of ϵ-MIL implying ϵ-MIP is studied in [27], and they both provide relatively weak average privacy guarantee. For those notions providing instance-wise privacy guarantee, LIP and DI assume the exact prior distribution is available. The exact maximal posterior ratio of $\Pr(X = x|Y = y)/\Pr(X = x'|Y = y)$ when $\mathcal{L}_{\mathsf{LIP}} \leq \epsilon$ can be calculated numerically through the following optimization problem (subject to constraints similar to (9)):

$$\max_{x,x' \in \mathcal{X}, y \in \mathcal{Y}} \frac{\mathsf{LR}(y, x', 1) P_X(x')}{\mathsf{LR}(y, x, 1) P_X(x)}. \tag{10}$$

The comparison between the theoretical bound is described in Theorem 3, and the exact maximal posterior ratio is shown in Fig. 2. Observe that the gap between theoretical bound and the exact leakage of DI under ϵ-LIP is larger than that of the leakage of LDP given ϵ-LIP, because the objective function in (10) contains a prior ratio which is amplified by taking the maximal value in deriving the loose bound presented in Theorem 3.

We next use the following mechanism as an example to demonstrate that MIP and MIL provide weaker privacy guarantees compared to LIP or LDP.

Example 1: Consider releasing a binary valued data $X \in \{0, 1\}$ through a privacy preserving mechanism $\mathcal{M}$. The mechanism $\mathcal{M}$ is described by $P_{Y|X}(0|1) = P_{Y|X}(1|0) = q$, where q denotes the perturbation parameter. We derive the leakages under different privacy notions as functions of the prior $P_X(1) \triangleq P$:

- $\mathcal{L}_{LDP} = \max \ln \left\{ \frac{q}{1-q}, \frac{1-q}{q} \right\}$,
- $\mathcal{L}_{LIP} = \max \left\{ \frac{P_Y(1)}{q}, \frac{P_Y(1)}{1-q}, \frac{P_Y(0)}{q}, \frac{P_Y(0)}{1-q} \right\}$,
- $\mathcal{L}_{MIL} = \ln(2 \max\{1 - q, q\})$,
- $\mathcal{L}_{MIP} = P(1-q) \ln \left(\frac{1-q}{P_Y(1)} \right) + Pq \ln \left(\frac{q}{P_Y(0)} \right) + (1 - P)(1-q) \ln \left(\frac{1-q}{P_Y(0)} \right) + (1-P)q \ln \left(\frac{q}{P_Y(1)} \right)$.

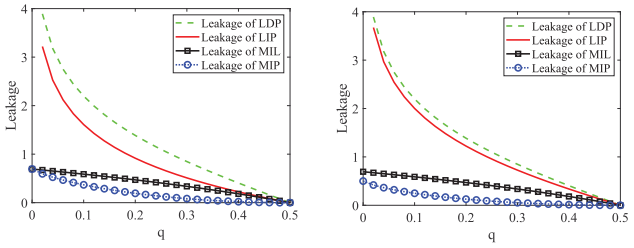
The marginal probability of $P_Y(1)$ and $P_Y(0)$ can be expressed as functions of P and q. The leakages of LDP, LIP, MIP, and MIL as functions of q are plotted in Fig. 4(a) and Fig. 4(b) (for two sets of priors $P = 0.5$ and $P = 0.2$). Observe that, the leakages under LIP and LDP for different mechanisms (i.e., different values of q) are always larger than that of MIP or MIL, as a special case, when $q = 0$, $\mathcal{L}_{LIP} = \infty$. Obviously, there exist some $\epsilon \geq 0$ such that, $\mathcal{M}$ satisfies ϵ-MIP or ϵ-MIL, but does not satisfy ϵ-LIP.

III. Properties of LIP

In this Section, we present and discuss several fundamental properties of LIP.

A. Extreme Values of Leakage, Post-Processing, Linkage Inequality and Modular Property

In this Section, we assume that the output of the mechanism $\mathcal{M}$ is Y, and thus together with the data prior on X, it induces a joint distribution P_{XY} on the r.v.'s (X, Y).



(a) Leakages of LDP, LIP, MIP and MIL of a binary perturbation mechanism when $P_X(1) = 0.5$.

(b) Leakages of LDP, LIP, MIP and MIL of a binary perturbation mechanism when $P_X(1) = 0.2$.

Fig. 4. In (a) and (b): Comparison of the leakages under symmetric perturbation mechanism among different privacy notions with two prior distributions.

Lemma 1: For any joint distribution $\mathbf{P}_{X,Y}$*:*

1) $\mathcal{L}_{LIP}(X;Y)$ *is symmetric, i.e.,* $\mathcal{L}_{LIP}(X;Y) = \mathcal{L}_{LIP}(Y;X)$.
2) $\mathcal{L}_{LIP}(X;Y) = 0$ *if and only if* Y *is independent of* X.
3) *Given the input and output data domain* $\mathcal{X}$ *and* $\mathcal{Y}$*, if there exist* $x \in \mathcal{X}$ *and* $y \in \mathcal{Y}$ *such that* $\Pr(Y = y|X = x) = 0$*, then* $\mathcal{L}_{LIP}(X;Y) = \infty$.
4) *(Post-processing property) The leakage of LIP cannot be increased by any post-processing of the output* Y *i.e.,*

$$\mathcal{L}_{LIP}(X;Z) \leq \mathcal{L}_{LIP}(X;Y),$$

where $X \to Y \to Z$ *forms a Markov chain.*
5) *(Linkage inequality) The local information leakage of input data is larger than the leakage of any correlated data, i.e.,*

$$\mathcal{L}_{LIP}(X;Y) \geq \mathcal{L}_{LIP}(S;Y),$$

where $S \to X \to Y$ *forms a Markov chain.*

The proof of Lemma 1 is presented in Appendix D in the supplementary material. The linkage inequality captures the notion that if there were primary (e.g., measurements) and secondary (e.g., latent variable) sensitive data, X and S, respectively that are correlated and the release was independently generated from only the primary sensitive data X, then the privacy-leakage for the secondary sensitive data S is bounded by the privacy-leakage for the primary sensitive data.

We next present modular property for LIP under mixture distributions in the following Lemma with proof presented in Appendix E of the supplementary document.

Lemma 2 (Modular Property): Given a mechanism $\mathcal{M}$ *satisfying* (ϵ, δ)*-LIP for each prior distribution* P_i*, where* $i = 1, 2, \ldots, K$*. Then* $\mathcal{M}$ *satisfies* (ϵ, δ)*-LIP for the composite prior distribution* $P_X(x) = \sum_{i=1}^{K} \alpha_i P_i(x)$ *where* $\alpha_i \in (0,1)$ *and* $\sum_{i=1}^{K} \alpha_i = 1$.

In other words, if a mechanism $\mathcal{M}$ was designed to satisfy (ϵ, δ)-LIP for each prior $P_i(x), i = 1, \cdots, K$, then the same LIP privacy guarantee is achieved for the mixture distribution $P_X(x) = \sum_{i=1}^{K} \alpha_i P_i(x)$.

B. Composability

Next, we discuss the composability of LIP. In particular, consider multiple (say n) queries over the same data, and we use the mechanism independently over time. Recall that for LDP, basic composition results say that applying a mechanism independently n times has leakage no more than $n\epsilon$ [11]. In contrast to LDP, where the leakage over multiple queries is additive, the composition of LIP is more nuanced and depends on the underlying prior. Given the input data X and an output sequence $\mathbf{Y}_1^n = \{Y_1, Y_2, \ldots, Y_n\}$ which are generated by n independent mechanisms, where each mechanism $\mathcal{M}_k$ satisfies ϵ_k-LIP. Thus, given the raw data, the output Y at each time are independent of each other: $Y_i \perp\!\!\!\perp Y_j | X, \forall j \neq i$. Then the leakage about X given $\mathbf{Y}_1^n$, $\mathcal{L}_{\mathsf{LIP}}(X; \mathbf{Y}_1^n)$ becomes:

$$\sup_{x \in \mathcal{X}, \mathbf{y}_1^n \in \mathcal{Y}^n} \ln\left(\max\left(\frac{P_X(x)}{P_{X|\mathbf{Y}_1^n}(x|\mathbf{y}_1^n)}, \frac{P_{X|\mathbf{Y}_1^n}(x|\mathbf{y}_1^n)}{P_X(x)}\right)\right).$$

For the maximum leakage of the mechanism after n independent outputs, we have the following Theorem with detailed proof provided in Appendix F of the supplementary document.

Theorem 4: $\mathcal{L}_{LIP}(X; \mathbf{Y}_1^n)$ *is upper bounded by*

$$\ln\left(P_{\min} + \exp\left(\sum_{k=1}^{n} \min\left\{2\epsilon_k, \ln\frac{e^{\epsilon_k} - 1 + P_{\min}}{P_{\min}}\right\}\right)(1 - P_{\min})\right),$$

where $P_{\min} = \min_{x \in \mathcal{X}} P_X(x)$.

Corollary 4: In the above setting, if each mechanism $\mathcal{M}_k$ *satisfies* ϵ_k*-BP-LIP for a bounded prior set of* $\mathscr{P}_{\mathcal{X}}^{bp}$*, then* $\mathcal{L}_{BP\text{-}LIP}(X; \mathbf{Y}_1^n)$ *is upper bounded by*

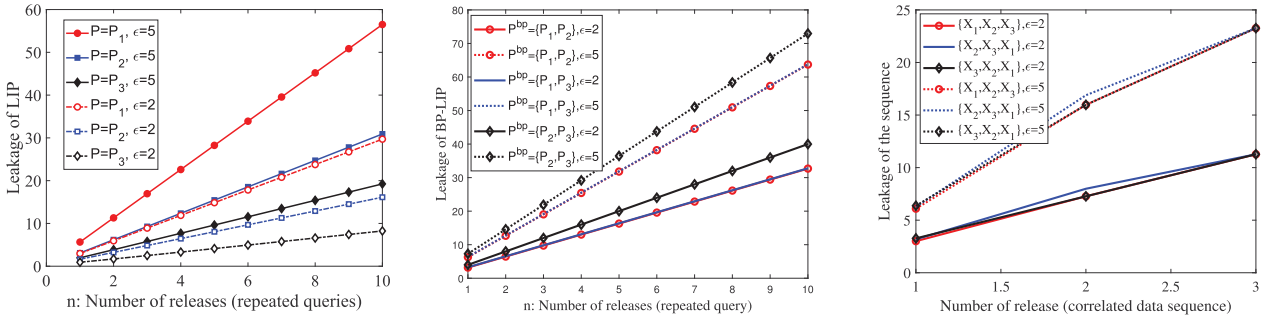
$$\ln\left(\min P_{\min}^{bp} + \exp\left(\sum_{k=1}^{n} \omega_k\right)(1 - \min P_{\min}^{bp})\right),$$

where $\omega_k = \min\left\{2\epsilon_k, \ln\frac{e^{\epsilon_k} - 1 + \max P_{\min}^{bp}}{\max P_{\min}^{bp}}\right\}$, $\min P_{\min}^{bp} = \min_{x \in \mathcal{X}, \boldsymbol{P} \in \mathscr{P}_{\mathcal{X}}^{bp}} P_X(x)$, *and* $\max P_{\min}^{bp} = \max_{\boldsymbol{P} \in \mathscr{P}_{\mathcal{X}}^{bp}} \min_{x \in \mathcal{X}} P_X(x)$.

The proof for Corollary 4 follows the same steps with the proof for Theorem 4.

Fig. 5(a) and 5(b) numerically show the leakage of LIP (BP-LIP) as a function of the number of repeated queries under three sets of prior distributions: $\mathbf{P}_1 = [0.25, 0.25, 0.25, 0.25]$, $\mathbf{P}_2 = [0.1, 0.1, 0.1, 0.7]$ and $\mathbf{P}_3 = [0.01, 0.33, 0.33, 0.33]$ (for BP-LIP, we let $\mathbf{P}_1^{bp} = \{\mathbf{P}_1, \mathbf{P}_2\}$, $\mathbf{P}_2^{bp} = \{\mathbf{P}_1, \mathbf{P}_2\}$ and $\mathbf{P}_3^{bp} = \{\mathbf{P}_2, \mathbf{P}_3\}$) for $\epsilon = 2$ and $\epsilon = 5$. We can observe that, under each possible prior (set) and each value of ϵ, the LIP (BP-LIP) leakage increases linearly with the number of releases. Other than the number of releases, a larger ϵ allows larger information leakage. Intuitively, the increment of leakage as a function n is related to the underlying prior. For LIP, when $P_{\min}$ is small, the prior knowledge is more skewed; in this case, we observe a decrease in terms of the maximum leakage. Intuitively, when $P_{\min}$ is small, the adversary already possesses significant knowledge on the data, after multiple outputs, the additional information leaked about X is relatively small. On the contrary, when $P_{\min}$ is large, the prior is close to a uniform distribution, we observe an increase in the maximum leakage of X.

Sequential Composability: Next, we consider releasing a data vector in a sequential manner: denote



(a) Illustration of the composability of LIP. Leakage increases linearly with number of releases under any prior distribution.

(b) Illustration of the composability of BP-LIP. Leakage increases linearly with number of releases under any bounded prior set.

(c) Illustration of the sequential composability of LIP.

Fig. 5. In (a) Composition for LIP for different priors, (b) BP-LIP for different priors sets, (c) sequential composition for LIP for correlated data.

$\mathbf{X}_1^n = \{X_1, X_2, \ldots, X_n\}$ as the data sequence, where the index denotes the timestamp. Let $P_{\mathbf{X}_1^n}(\mathbf{x}_1^n)$ be the joint distribution of the sequence. Suppose at each time k, Y_k is released independently through mechanism $\mathcal{M}_k$, i.e., $P_{Y_k|X_k,X_i}(y|x,x') = P_{Y_k|X_k}(y|x)$ and $\mathcal{L}_{LIP}(X_k, Y_k) \leq \epsilon_k$, $\forall k = \{1, 2, ..n\}$. We next derive an upper bound of the leakage of the whole data sequence after time n, defined as:

$$\mathcal{L}_{LIP}(\mathbf{X}_1^n; \mathbf{Y}_1^n) \triangleq \sup_{\mathbf{x}_1^n \in \mathcal{X}^n, \mathbf{y}_1^n \in \mathcal{Y}^n} \ln\left(\max\left(\frac{P_{\mathbf{X}_1^n}(\mathbf{x}_1^n)}{P_{\mathbf{X}_1^n|\mathbf{Y}_1^n}(\mathbf{x}_1^n|\mathbf{y}_1^n)}, \frac{P_{\mathbf{X}_1^n|\mathbf{Y}_1^n}(\mathbf{x}_1^n|\mathbf{y}_1^n)}{P_{\mathbf{X}_1^n}(\mathbf{x}_1^n)}\right)\right).$$

Corollary 5: If each mechanism $\mathcal{M}_k$ satisfies ϵ_k-LIP, then the total leakage of the sequence is upper bounded by

$$\ln\left(P^{\min}_{\mathbf{X}_1^n} + \exp\left(\sum_{k=1}^{n} \min\left\{2\epsilon_k, \ln\frac{e^{\epsilon_k} - 1 + P^k_{\min}}{P^k_{\min}}\right\}\right)(1 - P^{\min}_{\mathbf{X}_1^n})\right), \tag{11}$$

where $P^{\min}_{\mathbf{X}_1^n} = \min_{\mathbf{x}_1^n \in \mathcal{X}^n} P_{\mathbf{X}_1^n}(\mathbf{x}_1^n)$, $P^k_{\min} = \min_{x \in \mathcal{X}} P_{X_k}(x)$.

The proof of Corollary 5 is provided in Appendix M of the supplementary material.

To numerically demonstrate the leakage of releasing a correlated data sequence in terms of LIP, we consider releasing three data X_1, X_2 and X_3 in a sequential manner, suppose the prior of X_1 is $\mathbf{P}_1 = [0.25, 0.25, 0.25, 0.25]$, X_2 is $\mathbf{P}_2 = [0.1, 0.1, 0.1, 0.7]$ and $\mathbf{P}_3 = [0.01, 0.33, 0.33, 0.33]$ for X_3. Now we consider releasing the data in different orders: $\{X_1, X_2, X_3\}$ or $\{X_2, X_3, X_1\}$ or $\{X_3, X_2, X_1\}$ with different $\epsilon = 2$ or $\epsilon = 5$. Fig. 5(c) numerically shows the leakage of the whole sequence for different orders and ϵs. Observe that the leakage of the data sequence does not increase linearly with the number of releases (individual leakage increment can be different from data to data), however, for a given ϵ, the total leakage after 3 independent releases with different order are identical to each other.

C. Transferability

The basic question of transferability of a context-aware mechanism is the following: suppose we design a mechanism $\mathcal{M}(\mathbf{P}_1)$ assuming a prior $\mathbf{P}_1$, and achieve a leakage $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_1)$. How much does this mechanism $\mathcal{M}(\mathbf{P}_1)$ leak if we use it to release data with a different prior $\mathbf{P}_2$, i.e., how can we relate $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_1)$ to $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_2)$? It is intuitive to expect that if the distributions $\mathbf{P}_1$ and $\mathbf{P}_2$ are statistically close, then we should expect similar leakage. We next define metrics, to capture the statistical distance between distributions, and then present our result on transferability of leakage under LIP.

Definition 7: Given two pmfs $\mathbf{P}_1$ and $\mathbf{P}_2$, the total variation distance between $\mathbf{P}_1$ and $\mathbf{P}_2$ is defined as

$$D_{TV}(\mathbf{P}_1, \mathbf{P}_2) = \sup_{\mathcal{S} \subseteq \mathcal{X}} |\mathbf{P}_1(S) - \mathbf{P}_2(S)|.$$

Next, we present our main result on the transferability.

Lemma 3: For any distributions $\mathbf{P}_1$ and $\mathbf{P}_2$ and a mechanism $\mathcal{M}$, we have

$$|\mathcal{L}_{LIP}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_1) - \mathcal{L}_{LIP}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_2)| \leq \eta,$$

where $\eta = \ln\left(1 + \frac{\Delta}{c}\right)$, $\Delta = D_{TV}(\mathbf{P}_1, \mathbf{P}_2)$ *and* $c = \min\left\{P^1_{\min}, P^2_{\min}\right\}$.

The proof is presented in Appendix G of the supplementary document. The main idea behind the above result is to first relate the ratio $P_1(x)/P_2(x)$ to the TV distance $D_{\mathsf{TV}}(\mathbf{P}_1, \mathbf{P}_2)$ between distributions. Subsequently, we use this relationship to bound the leakage of the mechanism $\mathcal{M}(\mathbf{P}_1)$ under $\mathbf{P}_1$ to that of the leakage of the same mechanism $\mathcal{M}(\mathbf{P}_1)$ under $\mathbf{P}_2$. The key feature of the above result is that the transferability gap between the leakage measure is a monotonic function of the TV distance, i.e., a smaller TV distance implies smaller difference in leakage and vice-versa.

Remark 1: Note that, the leakage of $\mathcal{L}_{LIP}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_2)$ is larger than $\mathcal{L}_{LIP}(\mathcal{M}(\mathbf{P}_1), \mathbf{P}_1)$ by at most η, which depends on the TV distance between $\mathbf{P}_1$ and $\mathbf{P}_2$. Also note that, the value of $c = \min\left[P^1_{\min}, P^2_{\min}\right]$ could be small, which makes the increased amount very large. To this end, we can tighten the upper bound by considering the relationship between LIP and LDP. The Leakage defined in Eq. 8 is upper bounded by:

$$\min\left[\eta + \epsilon, 2\epsilon, \ln\frac{e^{\epsilon} - 1 + P^1_{\min}}{P^1_{\min}}\right]. \tag{12}$$

The bound follows the relationship between LIP and LDP; the optimal mechanism with $\mathbf{P}_1$ implies at

least $\min\left[2\epsilon, \ln\frac{e^{\epsilon}-1+P^1_{\min}}{P^1_{\min}}\right]$*-LDP, which further implies* $\min\left[2\epsilon, \ln\frac{e^{\epsilon}-1+P^1_{\min}}{P^1_{\min}}\right]$*-LIP for any prior* $\mathbf{P}_2$.

Corollary 6: By comparing the bounds in (12) *together, it can be readily shown that the bound* $\eta+\epsilon$ *is the smallest when*

$$D_{TV}(\mathbf{P}_1,\mathbf{P}_2) \leq \min(P^1_{\min}, P^2_{\min})(e^{\min(\epsilon,\Phi)}-1),$$

where $\Phi \triangleq \frac{e^{\epsilon}-1+P^1_{\min}}{P^1_{\min}} - \epsilon$.

The proof is presented in Appendix N. In practice, obtaining an accurate prior distribution is hard. Instead, the data owner uses samples coming from the true distribution. The goal is to design a privacy-preserving mechanism as a function of the empirical distribution estimated from these samples. However, the privacy guarantees for the empirical distribution, and the true distribution might be different. We next quantify the discrepancy between the privacy guarantees for the empirical distribution, used to design privacy mechanism, and for the true distribution experienced by the privacy mechanism used in practice.

Leakage of LIP using an estimated prior from samples: We invoke the above result in Lemma 3 to quantify the discrepancy of the privacy guarantees as follows: Consider a data owner using an empirical distribution $\mathbf{P}_1 = \hat{P}_X$, estimated from n i.i.d. samples drawn from a distribution $\mathbf{P}_2 = P_X$. Given n i.i.d. samples $x_1, x_2, \cdots, x_n$, the empirical probability distribution defined on $\mathcal{X}$ is defined as

$$\hat{P}_X(x) \triangleq \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{x_i=x\}}. \tag{13}$$

The leakage of $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\hat{P}_X), \hat{P}_X)$ in this case is

$$\sup_{x\in\mathcal{X}, y\in\mathcal{Y}} \ln\left(\max\left(\frac{P^{\hat{\mathcal{M}}}_{X|Y}(x|y)}{\hat{P}_X(x)}, \frac{\hat{P}_X(x)}{P^{\hat{\mathcal{M}}}_{X|Y}(x|y)}\right)\right),$$

where $P^{\hat{\mathcal{M}}}_{X|Y}$ is the posterior distribution when using the mechanism $\mathcal{M}(\hat{P}_X)$ and $\hat{P}_X$. We next relate the leakage $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\hat{P}_X), \hat{P}_X)$ with $\mathcal{L}_{\mathsf{LIP}}(\mathcal{M}(\hat{P}_X), P_X)$, i.e., the leakage under the true prior distribution P_X for a given mechanism, i.e., $\mathcal{M}(\hat{P}_X)$. We show a probabilistic bound on the difference between the two leakages.

Corollary 7: Let $\hat{P}_X$ *be the empirical distribution obtained from n i.i.d. samples drawn from* P_X. *Then, with probability* $1-\beta$, *we have*

$$\Delta_n \triangleq |\mathcal{L}_{LIP}(\mathcal{M}(\hat{P}_X), \hat{P}_X) - \mathcal{L}_{LIP}(\mathcal{M}(\hat{P}_X), P_X)| \leq \eta, \tag{14}$$

where $\eta = \ln\left(1+\frac{\bar{\Delta}}{2c}\right)$, $c = \min\left[\hat{P}_{\min}, P_{\min}\right]$ *and* $\bar{\Delta} = \sqrt{\frac{2}{n}(|\mathcal{X}| - \ln\beta)}$.

The proof is presented in Appendix H of the supplementary document. Note that the upper bound η in (14) behaves like $\mathcal{O}\left(\ln\left(1+\sqrt{1/n}\right)\right)$. Therefore, as the number of samples n goes to infinity, the privacy guarantee provided by the mechanism using the estimate $\hat{P}_X$ converges to the ideal case if we know the true prior P_X.

Numerical Evaluations: We give a numerical example for the transferability property for a perturbation mechanism. Specifically, in Fig.(s) 6(a) and 6(b), we compare between the bounds obtained in (12). In Fig. 6(c), we depict the leakage discrepancy for a context-aware binary randomized response perturbation mechanism $\mathcal{M}(\hat{P}_X)$ designed as a function of the empirical prior distribution (the mechanism is provided in Eqn. (19)). As shown in the figure, when the number of samples increases, the leakage discrepancy Δ_n decreases with n, and converges to zero which is consistent with the result in Corollary 7.

Corollary 8 (Generalization to Family of Priors): The Transeferability result can be invoked in a setting, when we have a known family of priors $\{\mathbf{P}_i\}_{i=1}^n$ *and a true prior* $\mathbf{P}_{True}$. *In this case, it is straightforward to show that the bound on the leakage discrepancy is obtained as*

$$\eta = \log\left[1+\frac{\max_i D_{TV}(\mathbf{P}_{True}, \mathbf{P}_i)}{\min_i \min(P^{True}_{\min}, P^i_{\min})}\right]. \tag{15}$$

Proof Sketch: The Transferability result for this case follows on similar lines as before: we bound the worst case among the family of priors (including the true prior or not) and the true prior, i.e.,

$$\begin{aligned}
\frac{P_{\text{True}}(x)}{P_i(x)} &\leq \max_x 1 + \frac{\|P_{\text{True}}(x) - P_i(x)\|_1}{2P_i(x)} \triangleq \delta^i_1,\\
\frac{P_i(x)}{P_{\text{True}}(x)} &\leq \max_x 1 + \frac{\|P_{\text{True}}(x) - P_i(x)\|_1}{2P_{\text{True}}(x)} \triangleq \delta^i_2,\\
&\Rightarrow \max_i \max\left[\frac{P_{\text{True}}(x)}{P_i(x)}, \frac{P_i(x)}{P_{\text{True}}(x)}\right]\\
&\leq \max_i \max(\delta^i_1, \delta^i_2).
\end{aligned} \tag{16}$$

Following the same set of steps, it is straightforward to show that leakage discrepancy is obtained as

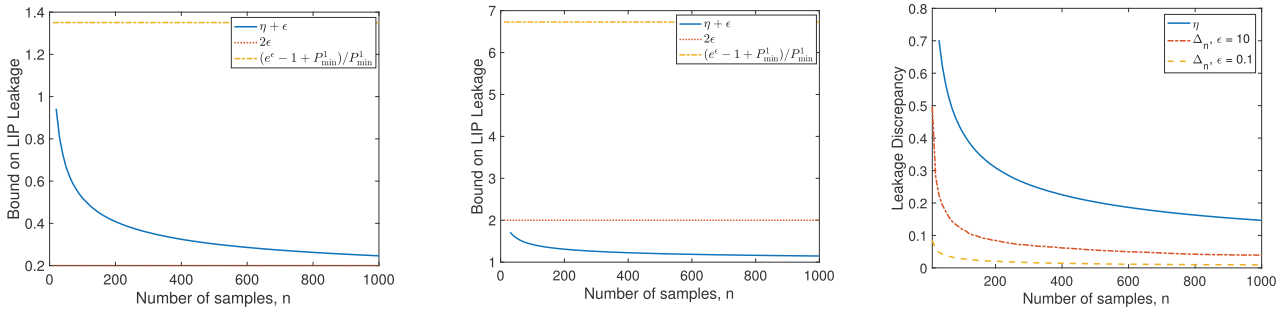
$$\eta = \log\left[1+\frac{\max_i D_{\text{TV}}(\mathbf{P}_{\text{True}}, \mathbf{P}_i)}{\min_i \min(P^{\text{True}}_{\min}, P^i_{\min})}\right]. \tag{17}$$

IV. LIP Mechanisms Under General Utility-Privacy Framework

In this Section, we present LIP based mechanisms for different data modalities. We start with the general utility-privacy framework. Then, under the general framework, we study the case where input data is *discrete*, where we present general randomized response mechanisms. Then, we devise LIP mechanisms where the input data is *continuous* valued. We first study random sampling mechanism that satisfies (ϵ,δ)-LIP. We also devise context-aware Gaussian mechanism and context-aware Laplacian mechanism that satisfy (ϵ,δ)-LIP.

A. A General Utility-Privacy framework

In this Section, we focus on characterizing tradeoffs between utility and context-aware privacy. To this end, we present a general framework for designing context-aware



(a) $\epsilon = 0.1$. (b) $\epsilon = 1$. (c)

Fig. 6. In (a) and (b): Bound on LIP Leakage vs. number of samples n. In (c): Discrepancy between privacy guarantees vs. number of samples n, for an RR mechanism where $\beta = 0.01$, $|\mathcal{X}| = 2$ and $P_X(0) = 0.3$.

privacy-preserving mechanisms that satisfy LIP while maximizing data utility.

The expected utility of a mechanism that releases Y as the output when the raw data is X can be described as

$$\text{Utility} = -E[D(Q(X), Q(Y))], \tag{18}$$

where Q is query function of X and Y that depends on the particular application. $D(a, b) : (\mathbb{R}, \mathbb{R}) \to \mathbb{R}^+$ denotes a distortion/distance measure between a, b, and the expectation $E[\cdot]$ is taken over both the underlying distribution of the data $P_X(x)$, as well as over the randomness of the mechanism, i.e., $P_{Y|X}(y|x)$. Note that the expected distance between $Q(X)$ and $Q(Y)$ stands for a general type of utility measurement. For example, in Location-Based Service (LBS): $Q(X) = X$, and typically Euclidean distance between X and Y is deployed to measure the performance: $\text{Utility} = -E[(X - Y)^2]$. Another example is histogram estimation: To estimate how many people possess each of the data categories, or classification according to users' data value, then $Q(X)$ is an indicator function, with the absolute distance the utility function can be written as $\text{Utility} = -\sum_{i=1}^{K} E[|\mathbb{1}_{\{X\in S_i\}} - \mathbb{1}_{\{Y\in S_i\}}|]$. The above examples illustrate that for a variety of applications, the utility function defined in (18) can be adapted by modifying the Q function and the distortion function $D(\cdot, \cdot)$. In general, mechanism design can be formulated as an optimization problem, maximizing the application/problem specific utility functions given local information privacy constraints:

$$\min E[D(Q(X), Q(Y))], \quad \text{s.t. } \mathcal{L}_{\text{LIP}}(X, Y) \le \epsilon.$$

B. LIP Mechanisms for Discrete-Valued Data

1) RR Mechanism for a Fixed Prior: In [33], our previous work has studied the optimal RR mechanism satisfying LIP. We include this result here for completeness. Denote $q_{xy} = \Pr(Y = y|X = x)\ \forall x \in \mathcal{X}, y \in \mathcal{Y}$ as the perturbation parameters in the RR mechanism, and the optimal solutions satisfying LIP constraints are described as follows:

Optimal RR Mechanism of LIP for a Fixed Prior:

$$q^*_{xy} = \begin{cases} \dfrac{P_X(y)}{e^\epsilon}, & x \ne y,\ x, y \in \mathcal{X}, \\ 1 - \dfrac{(1 - P_X(x))}{e^\epsilon}, & x = y,\ x, y \in \mathcal{X}. \end{cases} \tag{19}$$

Intuitively, the optimal solutions which maximize the utility are achieved when each $P_{Y|X}(y|x)$ $(y \ne x)$ is minimized, and the above results are derived by making the privacy constraints satisfy at the boundary of the convex polytope. An insight of the optimal solution for the RR-LIP mechanism is: less noise is required for more certain data.

2) RR Mechanism for Uncertain Prior: In [36], for the uncertain prior model, we have derived the optimal parameters under the RR mechanism for a binary model where the input and output take value from $\{0, 1\}$. Firstly, specify $\mathscr{P}_{\mathcal{X}}$ as $P_X(1) = \Pr(X = 1) \in [a, b]$, where $0 \le a \le b \le 1$. The optimal solutions are described as follows:

Optimal RR Mechanism of LIP for an Uncertain Prior:

$$q^*_{01} = \frac{b}{b - a + e^\epsilon} \quad \text{and} \quad q^*_{10} = \frac{1 - a}{b - a + e^\epsilon}.$$

Observe that when $a = b = P_X(1)$, the prior knowledge is certain and fixed. Then $q^*_{01} = \frac{P_X(1)}{e^\epsilon}$ and $q^*_{10} = \frac{1-P_X(1)}{e^\epsilon}$ which are identical to the optimal solutions of proposition IV-B1; When $a = 0$, $b = 1$, we have the optimal solutions for the WC-LIP (LDP): $q^*_{01} = q^*_{10} = \frac{1}{1+e^\epsilon}$, which is independent of prior. In Section V-A, we simulate with real data (data with discrete value that from a M-ary domain), and use build-in optimization solver to numerically obtain the optimal perturbation parameters. Then compare among the performance of BP-LIP, LIP and LDP.

C. Mechanism Design for Continuous-Valued Data

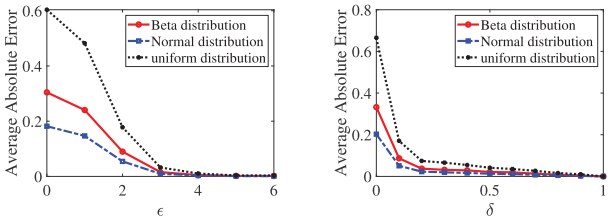
1) Sampling Mechanism: In this subsection we consider the case where the data X is drawn from a continuous distribution, i.e., $X \sim f_X(x)$. The sampling mechanism is defined as follows:

$$Y = \begin{cases} X, & \text{w.p. } \lambda, \\ \tilde{X}, & \text{w.p. } 1 - \lambda, \end{cases} \tag{20}$$

where $\tilde{X}$ is drawn independently from f_X. For a sampling mechanism, the utility function defined in (18) can be expressed as:

$$\begin{aligned} &E[D(Q(X), Q(Y))] \\ &= \int_x \int_y D[Q(x), Q(y)] f_X(x) f_{Y|X}(y|x) dx dy \\ &= (1 - \lambda) \int_x \int_{x'} D[Q(x), Q(x')] f_X(x) f_X(x') dx dx'. \end{aligned} \tag{21}$$

The last step is due to the fact that when $Y = X$, $D[Q(X), Q(X)]$ is zero, and when $Y \ne X$, the distribution



(a) $\delta = 0.01$ (b) $\epsilon = 2$

Fig. 7. Utility-Privacy tradeoff of the sampling mechanism under (ϵ, δ)-LIP for three different distributions.

of Y is identical to that of X. Observe that, the term of the integral is non-negative. The above argument implies that maximizing utility is equivalent to maximizing λ. Note that, for continuous-valued data, we are able to find a mechanism that satisfies a relaxed version of LIP. We next show a feasible choice of λ which satisfies (ϵ, δ)-LIP in the following proposition.

Proposition 2: A feasible λ satisfying (ϵ, δ)-LIP is $\lambda = \min\{\delta e^{\epsilon}, 1 - e^{-\epsilon} + \delta e^{-\epsilon}\}$.

The proof is provided in Appendix I of the supplementary document. Note that, when $\delta \geq \frac{1-e^{-\epsilon}}{e^{\epsilon}-e^{-\epsilon}}$, $\lambda = 1 - e^{-\epsilon} + e^{-\epsilon}\delta$, when $\delta < \frac{1-e^{-\epsilon}}{e^{\epsilon}-e^{-\epsilon}}$, $\lambda = e^{\epsilon}\delta$. Another observation is that there does not exist a sampling mechanism satisfying pure ϵ-LIP. As when $\delta = 0$, $\lambda = \min\{0, 1 - e^{-\epsilon}\} = 0$, which means the mechanism would always sample Y from the distribution of X.

Numerical Evaluations: We consider $N = 50000$ users in the system. Each user locally generates private data X_i (i denotes the user's index) drawn from a fixed distribution, and he/she releases a perturbed version Y_i through a sampling mechanism satisfying (ϵ, δ)-LIP. The utility of the system is measured by the averaged absolute error between X_i and Y_i over N users:

$$\text{Utility} = \frac{1}{N}\sum_{i=1}^{N} |X_i - Y_i|.$$

We consider three different distributions of X_i, including Beta distribution $B(\alpha, \beta)$, Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ and uniform distribution $\text{Unif}(a, b)$ respectively. We first fix the value of δ to be 0.01, and vary ϵ to calculate the utility. The result is depicted in Fig. 7(a). Then we fix the value of ϵ to be 2, and compare the utility of different distributions as functions of δ. The result is shown in Fig. 7(b). From the results, we have the following observations: (1) Different distributions of X_i do not affect each curve's trend in the plot and the utility increases with the value of ϵ, and δ; (2) The utility increases faster with δ when the value of δ is small, and increases slowly when δ approaches 1.

2) Context-Aware Additive Noise Mechanisms: We now consider the design of additive noise mechanism, where a user perturbs the continuous-valued data X by adding a random noise N drawn from a certain distribution. The noise-adding mechanism will output

$$Y = X + N(X),$$

where the distribution of the noise $N(X)$ can also be a function of the raw data X.

a) Context-aware Gaussian mechanism:: Next, we design context-aware Gaussian mechanism whose variance is calibrated directly using the Gaussian cumulative density function (alternatively, Q-function). We would like to highlight that bounding the LIP ratio using tail bound approximation is known to be challenging specially for continuous-valued data [37]. Our goal is to design Gaussian mechanism that satisfies (ϵ, δ)-LIP. We assume a Gaussian prior on X with mean μ_X and variance σ_X^2. Therefore, we have $f_{Y|X} = \mathcal{N}(x, \sigma_N^2)$ and $f_Y = \mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2)$.

For data utility, it is readily seen that under Gaussian mechanism, a small σ_N^2 implies a high utility under zero mean ($\mu_N = 0$). In order to analyze the mechanism, we use a particular divergence metric between to distributions (i.e., f_Y and $f_{Y|X}$) called E-divergence [38], which captures the outage events of the mechanism where we cannot guarantee pure ϵ-LIP. It is worth noting that E-divergence, defined next, is considered as a generalization of the total variation distance.

Definition 8 (E-Divergence [38]): Given two probability distributions f and g defined over the same support set $\mathcal{Y}$ and $\gamma \geq 0$, the E-divergence is defined as follows:

$$E_{\gamma}(f||g) \triangleq \sup_{\mathcal{S}\subseteq\mathcal{Y}} f(\mathcal{S}) - \gamma\, g(\mathcal{S}) = \int_{\{y: f(y)\geq \gamma\, g(y)\}} (f(y) - \gamma\, g(y))\, dy.$$

Notice that for any γ, $E_{\gamma}(P||Q) \leq 1$ for any two distributions P and Q. Next, we show the equivalence between the aforementioned definition and (ϵ, δ)-LIP when $\gamma = e^{\epsilon}$.

Corollary 9: A mechanism satisfies (ϵ, δ)-LIP if

$$E_{e^{\epsilon}}(f_{Y|X}||f_Y) \leq \delta \quad \& \quad E_{e^{\epsilon}}(f_Y||f_{Y|X}) \leq \delta. \tag{22}$$

We next show the relationship between $E_{e^{-\epsilon}}(f_{Y|X}||f_Y)$ and $E_{e^{\epsilon}}(f_Y||f_{Y|X})$ which will be useful later for the mechanism design in the following Proposition with proof provided in Appendix J of the supplementary document.

Proposition 3: For two distributions $f_{Y|X}$ and f_Y defined over the same support $\mathcal{Y}$, we have

$$E_{e^{\epsilon}}(f_{Y|X}||f_Y) = e^{\epsilon} E_{e^{-\epsilon}}(f_Y||f_{Y|X}) - e^{\epsilon} + 1.$$

We next compute the E-divergence for two Gaussian distributions in the following Lemma.

Lemma 4: For two Gaussian distributions, $f \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $g \sim \mathcal{N}(\mu_2, \sigma_2^2)$, $E_{\gamma}(f||g)$ is given as

$$E_{\gamma}(f||g) = 1 + Q\left(\frac{y_u - \mu_1}{\sigma_1}\right) - Q\left(\frac{y_l - \mu_1}{\sigma_1}\right) - \gamma\left[1 + Q\left(\frac{y_u - \mu_2}{\sigma_2}\right) - Q\left(\frac{y_l - \mu_2}{\sigma_2}\right)\right], \tag{23}$$

where $\sigma_1 \geq \sigma_2$, y_l and y_u are the points where $f = \gamma\, g$ and $y_l \leq y_u$ (please refer to Appendix K in the supplementary document for more details).

Having computed $E_{e^{-\epsilon}}(f_Y||f_{Y|X})$ using Lemma 4 by setting $f = f_Y$, $g = f_{Y|X}$ and $\gamma = e^{-\epsilon}$, it is straightforward to compute $E_{e^{\epsilon}}(f_{Y|X}||f_Y)$ using Proposition 3.

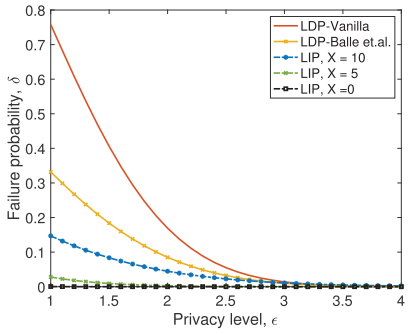


Fig. 8. Numerical comparison for (ϵ, δ) between LDP and LIP: $\mu_X = 0$, $\sigma_X = 5$, $\Delta X = 20$ and $\sigma_N = 10$. For fair comparisons, we compare with 2ϵ-LDP.

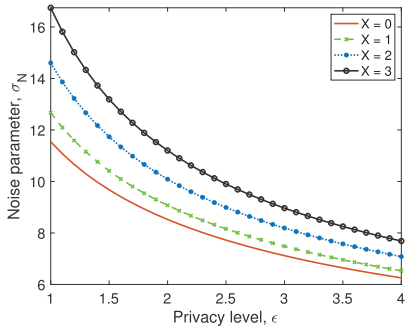


Fig. 9. Numerical comparisons for context-aware Gaussian mechanism under different values of X for $\mu_X = 0$, $\sigma_X = 5$, $\delta = 10^{-4}$. High density instances require less noise and vice versa.

We next summarize the results on the Gaussian mechanism in the following Theorem.

Theorem 5: Suppose $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, then the Gaussian mechanism with noise variance σ_N^2 satisfies (ϵ, δ)-LIP where

$$\delta = \sup_x \max\{E_{e^\epsilon}\big(\mathcal{N}(x, \sigma_N^2)||\mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2)\big), E_{e^\epsilon}\big(\mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2)||\mathcal{N}(x, \sigma_N^2)\big)\}.$$

In Fig. 9, we show the impact of the value of the data X on the choice of σ_N. As we see from the figure, the noise parameter σ_N is adjusted to the priors, i.e., less density (higher value of X) requires more noise and vice versa.

Alternatively, for a given x, we can find the smallest amount of noise as follows.

Corollary 10: For a given privacy level (ϵ, δ)-LIP and a data point x, the optimum noise parameter of the Gaussian mechanism can be numerically obtained from the following optimization problem:

$$\sigma_{LIP}^2(x) = \min_{\sigma_N} \sigma_N^2(x),$$
$$s.t.\ E_{e^\epsilon}\big(\mathcal{N}(x, \sigma_N^2)||\mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2)\big) \le \delta,$$
$$E_{e^\epsilon}\big(\mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2)||\mathcal{N}(x, \sigma_N^2)\big) \le \delta.$$

Numerical Evaluations: We next compare with some results from the literature that design Gaussian mechanism under LDP notion. First, the vanilla Gaussian Mechanism [39]. For a given ϵ, σ_N and ΔX we have the following upper bound on δ:

$$\delta \le 1.25 \times \exp\left[-\frac{\epsilon^2 \sigma_N^2}{2(\Delta X)^2}\right].$$

In [40], the authors improved the Vanilla Gaussian mechanism by using the E-divergence metric, they showed that

$$E_{e^\epsilon}(\mathcal{N}(x, \sigma_N^2)||\mathcal{N}(x', \sigma_N^2)) = \Phi\left(\frac{\Delta X}{2\sigma_N} - \frac{\epsilon\sigma_N}{\Delta X}\right) - e^\epsilon \Phi\left(-\frac{\Delta X}{2\sigma_N} - \frac{\epsilon\sigma_N}{\Delta X}\right), \quad (24)$$

where $\Phi(\cdot)$ is the normal CDF and $\Delta X = \max_{x,x'} |x - x'|$, i.e., the local sensitivity. Note that we can write the above equation in terms of the Q-function, where $Q(x) = 1 - Q(-x) = 1 - \Phi(x)$.

We compare our context-aware mechanism with the vanilla Gaussian mechanism [39] and its improved version in [40]. Specifically, we plot the outage probability δ vs the privacy level ϵ for different data realizations $X = x$, and noise variance σ_N^2 in Fig. 9. We notice that under the same noise variance σ_N^2, we outperform the two mechanisms designed for differential privacy in terms of the outage probability δ.

Remark 2: It is worth mentioning that the proposed Gaussian mechanism is not limited to Gaussian priors only. In fact, any arbitrary distribution f_X can be fitted by a Gaussian mixture model (GMM), i.e., $f_X = \sum_{i=1}^K \alpha_i f_{X_i}$, where $f_{X_i} = \mathcal{N}(\mu_i, \sigma_i^2)$, α_i is the proportional weight and $\sum_{i=1}^K \alpha_i = 1$. The optimal parameters $\{(\mu_i, \sigma_i^2)\}_{i=1}^K$ can be obtained by using the expectation-maximization algorithm [41]. Also, the Gaussian mechanism will guarantee (ϵ, δ)-LIP under GMM using the modularity property of LIP (see Lemma 2).

Next, we present our results on the context-aware Laplacian mechanism.

b) Context-aware Laplacian mechanism: In this subsection, we present our main result on the context-aware Laplacian mechanism which satisfies ϵ-LIP. We perturb the data X by adding a random noise $N(X)$ drawn from Laplacian distribution, where $N(X) \sim \text{Lap}(0, b(X))$ is a random variable with probability density function and $b(X)$ is the noise parameter that describes the Laplacian distribution. Note that $b(X)$ controls the width of the distribution, and the variance is $2\ b(X)^2$. In particular, we show how to design the noise parameter b as a function of the prior knowledge about the user's input data X. We notice that bounding the density ratio $\frac{f_Y}{f_{Y|X}}$ under Laplacian mechanism is challenging. Instead, we can employ a data pre-processing (i.e., discretization) on the continuous-valued data [42], [43]. It has been shown in [13] that if we pick the noise parameter b as $b_{\text{LDP}} = \frac{\Delta X}{\epsilon}$ satsifies ϵ-LDP, where ΔX is the local sensitivity, i.e., $\Delta X \triangleq \max_{x_i, x_j} |X_i - X_j| = x_{\max} - x_{\min}$.

Theorem 6: Let $\mathcal{M}$ be the Laplacian mechanism and let $\epsilon \ge 0$. Then, $\mathcal{M} : X \to Y, Y \in [l, u]$ satisfies ϵ-LIP if

$$b_{LIP}^{indep.} = \begin{cases} \dfrac{\Delta X}{\ln\left(\frac{e^\epsilon - P_{\min}}{1 - P_{\min}}\right)}, & \epsilon < \ln(\frac{1}{P_{\min}}) \\ \dfrac{\Delta X}{\epsilon}, & otherwise, \end{cases} \quad (25)$$

$$b_{LIP}^{dep.}(x) = \frac{\Delta X}{\alpha_\epsilon P_X(x) + \epsilon}, \quad \forall x \in \mathcal{X}. \quad (26)$$

where $P_{\min}$ is the minimum probability value of the distribution $P_X(x)$.

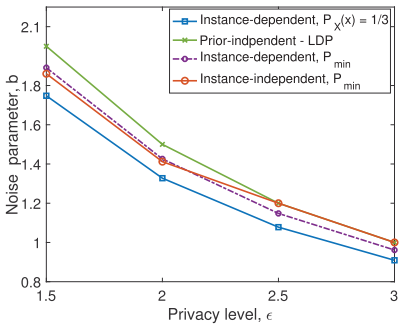


Fig. 10. Comparison between context-free and context-aware Laplacian mechanisms. The prior distribution is $\mathscr{P}_X = \{\frac{1}{3} + \frac{s}{2}, \frac{1}{3}, \frac{1}{3} - \frac{s}{2}\}$ and $s = 0.2$.

The first part of (25) is presented in the Appendix of [34]. The proof of this Theorem is in Appendix L of the supplementary document.

Numerical Evaluations: Fig. 10 shows a comparison between the three mechanisms, i.e., 1) independent mechanism, b_{LDP}, 2) instance independent mechanism, b_{LIP} and 3) instance dependent mechanism, $b_{\text{LIP}}(X)$. We can see that our mechanism outperforms the conventional Laplacian mechanism, this is due to the relaxed definition of LIP where the noise parameter b can be a function of the data prior.

V. Evaluation With Real Data

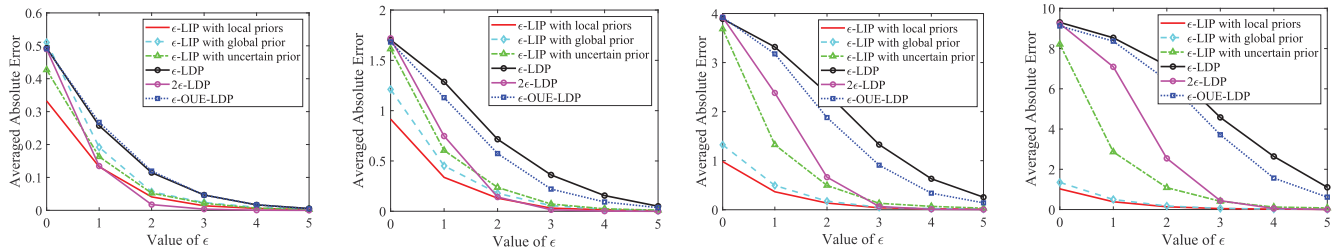
In this Section, we provide numerical results on real datasets to demonstrate the utility-privacy tradeoffs provided by different mechanisms. The first dataset is *Students Performance Assessment* dataset [44], which reports student achievement in secondary education of two Portuguese schools (395 students in total). The data attributes include student grades, demographic, social, and school-related features, which were collected by using school reports and questionnaires. Two datasets are provided regarding the performance in two distinct subjects: Mathematics (mat) and Portuguese language (por), the content of the *students performance dataset*. We use this dataset to test the performance of different mechanisms for discrete data release leveraging different scenarios of the context-awareness. The second dataset is *Baltimore city employee's salary for year 2019* [45], which captures the gross salary of 13812 employees from July 1, 2018, through June 30, 2019. This dataset is used to test mechanisms for continuous-valued data release. Note that in the following experiments, it is assumed that each user locally possesses one row of the dataset.

A. Comparison Among Different Mechanisms for Discrete-Valued Data Release

Observe from the *student performance* dataset, each row stands for a student's entry, and for each student, 30 different personal behaviors, the grades for two periods and the final grade are listed. In this experiment, it is assumed that a subset of the students' full records and other students' records except for the last column (final grade) have been collected. The other students are submitting their final grade (denoted as X) to the collector independently in a privacy-preserving way. Denote set $\mathcal{A}$ as the set of students whose full records have been published and $\mathcal{B}$ as the set of students who are releasing their final grade. We are comparing among the following mechanisms: (a) Each student in $\mathcal{B}$ uses RR-LIP mechanism to perturb their final grade with a local prior. Each local prior is calculated as follows: the data in $\mathcal{A}$ are treated as training data used to learn the correlation between students' personal information (including two-period grade) and the final grade. Then $\mathcal{B}$ is treated as testing data, and the distribution of their final grades are calculated by predicting using their personal information and two-period grades. (b) Each student in $\mathcal{B}$ uses RR-LIP mechanism to perturb with the same global prior (the same mechanism for each student). The global prior is estimated by the grade collected from $\mathcal{A}$. (c) Each student in $\mathcal{B}$ uses RR mechanism satisfying BP-LIP constraints with perturbation parameters numerically solved by build-in optimizer: notice that some personal information is very relevant to the final grade, such as "study time", "go out", "absence" and first and second-period grades, but some are not, such as "gender", "address", "nursery", "parents job", etc. It is assumed that the relationships between different combinations of information and the final grades are learned from $\mathcal{A}$, and each student in $\mathcal{B}$ uses these relationships to predict the distribution of their final grade (assign each value in the support with a probability), the set of different predictions becomes the uncertainty: $\mathscr{P}^{bp}$. (d) Each student in $\mathcal{B}$ uses Optimal LDP mechanism to perturb data, the mechanism is described as follows: $q_{xx} = \frac{e^\epsilon}{e^\epsilon + |\mathcal{X}| - 1}$, $q_{xy} = \frac{1}{e^\epsilon + |\mathcal{X}| - 1}$, where $x \neq y$ [46], we consider both ϵ-LDP and 2ϵ-LDP. (e) Each student in $\mathcal{B}$ uses Optimal Unary Encoding LDP mechanism [46] to perturb data, his/her final grade is firstly mapped into a vector, wherein the value of the i-th bit is 1 if his/her final grade is i, and 0 if not. Then, each bit is independently perturbed under binary LDP mechanism: $q_{01} = \frac{1}{e^\epsilon + 1}$, $q_{10} = \frac{1}{2}$.

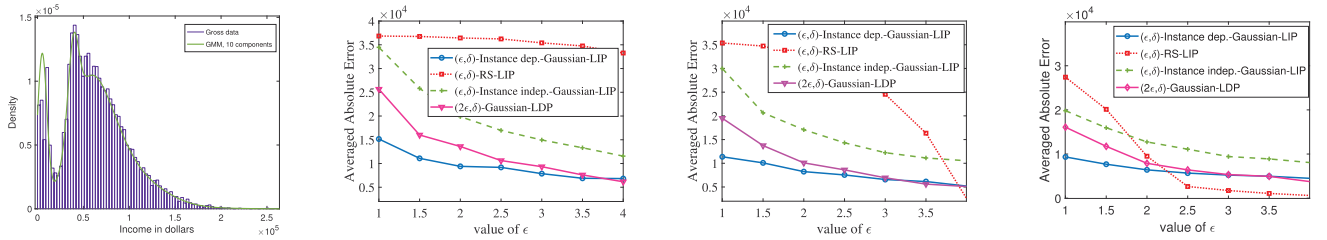
The *Utility* of different mechanisms described above is measured by the averaged absolute distance between the real final grade and the perturbed final grade of each student in $\mathcal{B}$. The relationships between personal information and final grade in training dataset $\mathcal{A}$ are estimated through a two-layer neural network. Note that the domain of X is $\mathcal{X} = \{0, 1, \ldots, 20\}$. To illustrate the impact on the utility-privacy tradeoff from the data domain provided by different mechanisms. We consider the final grade in four cases with the cardinality of 2, 5, 10, and 21, respectively. For example, when the cardinality is 2, the final grade of each student is either above 10 or below 10; when the cardinality is 21, each student's grade ranges from 0 to 20. The privacy of different mechanisms is measured by the privacy budget of ϵ, which takes value from 0 to 5. The curves of different mechanisms with different cardinalities are plotted in Fig. 11.

Observe that LIP with local priors provides the most enhanced utility under each ϵ compared with other cases. As the training results with all features offer the highest accuracy and the estimated prior is close to the real one. When $|\mathcal{X}| = 2$, BP-LIP provides better utility than LIP with global priors. But as $|\mathcal{X}|$ increases, LIP with global prior outperforms BP-LIP. This is because BP-LIP has more constraints than LIP with a fixed prior, and when the cardinality increases,



(a) $|\mathcal{X}| = 2$. (b) $|\mathcal{X}| = 5$. (c) $|\mathcal{X}| = 10$. (d) $|\mathcal{X}| = 21$.

Fig. 11. Utility-privacy tradeoff comparison for discrete-valued data release with *student performance* dataset.



(a) Data distribution and different fitted density distribution. (b) Comparison when $\delta = 10^{-4}$. (c) Comparison when $\delta = 10^{-3}$. (d) Comparison when $\delta = 10^{-2}$.

Fig. 12. Utility-privacy tradeoff comparison for continuous-valued data release with *Baltimore city employee salaries* dataset.

the feasible region of BP-LIP decreases as a result. Hence, the utility decreases. At last, LDP provides the worst utility since no prior knowledge is utilized during perturbation, OUE-LDP performs better than LDP when $|\mathcal{X}| > 2$, as the utility provided by LDP is largely influenced by the cardinality of input data, thus transferring a large domain into a binary vector can significantly mitigate it.

B. Utility-Privacy Tradeoff Comparisons Among Mechanisms for Continuous-Valued Data

Next, we compare different Gaussian mechanisms with the RR mechanism when releasing continuous-valued data. The dataset we use in this part is "Baltimore city employee salaries," where 13812 employee's personal information is listed, including Name, Job tile, Department ID, Job description, Hire date, net/gross income. In this experiment, we assume each user releases his/her gross income, which is from a continuous-valued data domain. We first estimate the distribution of all users' income by fitting the income density with a probability distribution. Observe that there are two peaks in the density of income. To fit the underlying prior distribution, we use Gaussian mixture model (GMM) fitting technique (see Fig. 12(a)). In the experiment, it is assumed that each individual uses Gaussian mechanism with GMM fitted for the prior distribution and sampling mechanism with two fitted models for the prior distribution to release data satisfying (ϵ, δ)-LIP. More specifically, for the Gaussian mechanism, we consider two cases: 1) *instance-independent*, where we design σ_{LIP} as a function of the prior distribution by making the variance proportional to $\min_x f_X(x)$. Note that each instance is perturbed by the same amount of noise. 2) *instance-dependent*, where we design $\sigma_{\text{LIP}}(x)$ as a function of every input instance x such that we add less noise for the high density instances, and vice versa. We also consider Gaussian mechanism under LDP, i.e., σ_{LDP} [40]. We fix δ to be 10^{-4}, 10^{-3} and 10^{-2} respectively, and range the value of ϵ from 1 to 4. The utility of different mechanisms is measured by the averaged absolute error, and the comparison results are shown in Fig. 12.

We observe from Fig. 12, when δ is small, Gaussian mechanism provides improved utility than the sampling mechanism. Another observation is that the sampling with different density fits has a very slight influence on the performance of the mechanism, as the density distribution does not affect the value of λ, but only determines how the output is sampled. But on the other hand, releasing data with an inaccurate density could result in additional privacy leakage, as we studied in Section III-C. Finally, we observe that LDP-based Gaussian mechanism performs worse than the *instance-dependent* LIP-based Gaussian mechanism. However, LDP-based mechanism outperforms the *instance-independent* LIP-based mechanism (i.e., sampling mechanism and Gaussian mechanism) in terms of utility. In turn, this observation shows the benefits of designing instance dependent perturbation mechanism to achieve higher utility.

VI. Conclusion

In this paper, we study the local information privacy (LIP) as a relaxation of the *de facto* standard privacy notion, i.e., local differential privacy. For the LIP notion, we proved that the metric satisfies desirable properties such as post-processing, modularity, composability and transferability. We then compare the relationship of LIP with some of the most representative existing privacy notions and show that LIP provides a context-aware instance-wise privacy guarantee. We proposed private mechanisms that satisfy ϵ-LIP or (ϵ, δ)-LIP for both

discrete and continuous-valued data, including randomized response mechanism, random sampling mechanism, additive noise mechanism. We have conducted numerical simulations and evaluations with real data that shows utility improvements over the LDP based mechanisms.

REFERENCES

[1] A. Fitzpatrick. (2016). *What to do After the Massive Yahoo Hack*. [Online]. Available: http://time.com/4504901/yahoo-hack-passwords-logins/

[2] W. Primoff and S. Kess, "The equifax data breach: What CPAS and firms need to know now," *CPA J.*, vol. 87, no. 12, pp. 14–17, 2017.

[3] C. Cadwalladr and E. Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge analytica in major data breach," *Guardian*, vol. 17, p. 22, Mar. 2018.

[4] J. Lu, "Assessing the cost, legal fallout of Capital One data breach," Law360 Expert Anal., Aug. 2019. [Online]. Available: https://ssrn.com/abstract=3438816

[5] J. Krumm, "Inference attacks on location tracks," in *Proc. 5th Int. Conf. Pervas. Comput.* New York, NY, USA: Springer-Verlag, 2007, pp. 127–143.

[6] T. Dalenius, "Towards a methodology for statistical disclosure control," *Statistik Tidskrift*, vol. 15, nos. 429–444, pp. 1–2, 1977.

[7] O. Arias, J. Wurm, K. Hoang, and Y. Jin, "Privacy and security in Internet of Things and wearable devices," *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 1, no. 2, pp. 99–109, Apr. 2015.

[8] S. Wang, L. Huang, Y. Nie, P. Wang, H. Xu, and W. Yang, "PrivSet: Set-valued data analyses with locale differential privacy," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2018, pp. 1088–1096.

[9] B. Aronov *et al.*, "Are friends of my friends too social?: Limitations of location privacy in a socially-connected world," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jun. 2018, pp. 280–289.

[10] S. Boukoros, M. Humbert, S. Katzenbeisser, and C. Troncoso, "On (the lack of) location privacy in crowdsourcing applications," in *Proc. 28th USENIX Secur. Symp.*, 2019, pp. 1859–1876.

[11] C. Dwork, F. McSherry, and K. Nissim, "Calibrating noise to sensitivity in private data analysis," in *Proc. 3rd Theory Cryptogr. Conf.*, 2006, pp. 265–284.

[12] C. Dwork, "Differential privacy," in *Proc. 33rd Int. Colloq. Automata, Lang. Program. (ICALP)*, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, Eds., 2006, pp. 1–12.

[13] C. Dwork, "Differential privacy: A survey of results," in *Proc. 5th Int. Conf. Theory Appl. Models Comput. (TAMC)*, M. Agrawal, D. Du, and Z. Duan, Eds., 2008, pp. 1–19.

[14] A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber, "Privacy: Theory meets practice on the map," in *Proc. IEEE 24th Int. Conf. Data Eng.* Washington, DC, USA: IEEE Computer Society, Apr. 2008, pp. 277–286.

[15] J. M. Abowd, "The U.S. census bureau adopts differential privacy," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, London, U.K., Jul. 2018, p. 2867.

[16] Ú. Erlingsson, V. Pihur, and A. Korolova, "RAPPOR: Randomized aggregatable privacy-preserving ordinal response," in *Proc. 21st ACM CCS*, Nov. 2014, pp. 1054–1067.

[17] G. Cormode, S. Jha, T. Kulkarni, N. Li, D. Srivastava, and T. Wang, "Privacy at scale: Local differential privacy in practice," in *Proc. Int. Conf. Manage. Data*, New York, NY, USA, May 2018, pp. 1655–1658.

[18] Differential Privacy Team. *Learning With Privacy at Scale*. Accessed: Dec. 2017. [Online]. Available: https://machinelearning.apple.com/2017/12/06/learning-with-privacy-at-scale.html

[19] J. Tang, A. Korolova, X. Bai, X. Wang, and X. Wang, "Privacy loss in Apple's implementation of differential privacy on MacOS 10.12," 2017, *arXiv:1709.02753*. [Online]. Available: http://arxiv.org/abs/1709.02753

[20] R. Bassily, K. Nissim, U. Stemmer, and A. G. Thakurta, "Practical locally private heavy hitters," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 2288–2296.

[21] T.-H. H. Chan, E. Shi, and D. Song, "Optimal lower bound for differentially private multi-party aggregation," in *Proc. 20th Annu. ECA*, 2012, pp. 277–288.

[22] M. E. Andrés, N. E. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, "Geo-indistinguishability: Differential privacy for location-based systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*, Nov. 2013, pp. 901–914.

[23] F. Tramèr, Z. Huang, J.-P. Hubaux, and E. Ayday, "Differential privacy with bounded priors: Reconciling utility and privacy in genome-wide association studies," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2015, pp. 1286–1297.

[24] S. Asoodeh, F. Alajaji, and T. Linder, "Notes on information-theoretic privacy," in *Proc. 52nd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2014, pp. 1272–1278.

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

[26] S. Asoodeh, F. Alajaji, and T. Linder, "On maximal correlation, mutual information and data privacy," in *Proc. IEEE 14th Can. Workshop Inf. Theory (CWIT)*, Jul. 2015, pp. 27–31.

[27] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *Proc. Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2016, pp. 234–239.

[28] F. du Pin Calmon and N. Fawaz, "Privacy against statistical inference," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2012, pp. 1401–1408.

[29] D. Kifer and A. Machanavajjhala, "A rigorous and customizable framework for privacy," in *Proc. 31st Symp. Princ. Database Syst. (PODS)*, May 2012, pp. 77–88.

[30] B. Yang, I. Sato, and H. Nakagawa, "Bayesian differential privacy on correlated data," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, May 2015, pp. 747–762.

[31] N. Li, W. Qardaji, D. Su, Y. Wu, and W. Yang, "Membership privacy: A unifying framework for privacy definitions," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*, Nov. 2013, pp. 889–900.

[32] B. Jiang, M. Li, and R. Tandon, "Context-aware data aggregation with localized information privacy," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, May 2018, pp. 1–9.

[33] B. Jiang, M. Li, and R. Tandon, "Local information privacy and its application to privacy-preserving data aggregation," *IEEE Trans. Depend. Sec. Comput.*, early access, Dec. 1, 2020, doi: 10.1109/TDSC.2020.3041733.

[34] M. Seif, R. Tandon, and M. Li, "Context aware Laplacian mechanism for local information privacy," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Aug. 2019, pp. 1–5.

[35] J. Lee and C. Clifton, "Differential identifiability," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2012, pp. 1041–1049.

[36] B. Jiang, M. Li, and R. Tandon, "Local information privacy with bounded prior," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.

[37] A. D. Sarwate and K. Chaudhuri, "Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 86–94, Sep. 2013.

[38] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[39] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014.

[40] B. Balle and Y.-X. Wang, "Improving the Gaussian mechanism for differential privacy: Analytical calibration and optimal denoising," 2018, *arXiv:1805.06530*. [Online]. Available: http://arxiv.org/abs/1805.06530

[41] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[42] P. Perner and S. Trautzsch, "Multi-interval discretization methods for decision tree learning," in *Proc. Joint IAPR Int. Workshops Stat. Techn. Pattern Recognit. (SPR), Struct. Syntactic Pattern Recognit.* Cham, Switzerland: Springer, 1998, pp. 475–482.

[43] M. X. Ribeiro, M. R. P. Ferreira, C. Traina, and A. J. M. Traina, "Data pre-processing: A new algorithm for feature selection and data discretization," in *Proc. 5th Int. Conf. Soft Comput. Transdisciplinary Sci. Technol. (CSTST)*, 2008, pp. 252–257.

[44] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in *Proc. 5th Future Bus. Technol. Conf. (FUBUTEC)*, A. Brito and J. Teixeira, Eds. Porto, Portugal: EUROSIS, Apr. 2008, pp. 5–12.

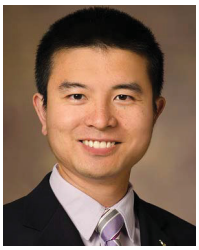
[45] *Baltimore City Employee Salaries*. Accessed: Oct. 3, 2019. [Online]. Available: https://data.baltimorecity.gov/City-Government/Baltimore-City-Employee-Salaries-FY2019/6xv6-e66h

[46] T. Wang, J. Blocki, N. Li, and S. Jha, "Locally differentially private protocols for frequency estimation," in *Proc. 26th USENIX Secur.*, USENIX Association, 2017, pp. 729–745.

Bo Jiang (Student Member, IEEE) received the bachelor's degree and first master's degree from the Harbin Institute of Technology, China, in 2013 and 2015, respectively, and the second master's degree from Worcester Polytechnic Institute, MA, USA, in 2017. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, The University of Arizona. He is currently a Research Assistant with the Department of Electrical and Computer Engineering, The University of Arizona. His current research interests include security and privacy enhancing technologies, machine learning, radar image processing, and signal processing.

Mohamed Seif (Student Member, IEEE) received the B.Sc. degree in electrical engineering from Alexandria University, Alexandria, Egypt, in 2014, and the M.Sc. degree in wireless technologies from Nile University, Giza, Egypt, in 2016. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, The University of Arizona. He joined as a Graduate Research Assistant with The University of Arizona in 2017. His research interests include information theory, machine learning, and wireless communications.

Ravi Tandon (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur (IIT Kanpur) in 2004 and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park (UMCP), in 2010. From 2010 to 2012, he was a Post-Doctoral Research Associate with Princeton University. He is currently an Associate Professor with the Department of ECE, The University of Arizona. Prior to joining The University of Arizona in Fall 2015, he was a Research Assistant Professor with Virginia Tech, with various positions in the Bradley Department of ECE, Hume Center for National Security and Technology, and the Department of Computer Science, Discovery Analytics Center. His current research interests include information theory and its applications to wireless networks, signal processing, communications, security and privacy, machine learning, and data mining. He was a recipient of the 2018 Keysight Early Career Professor Award, the NSF CAREER Award in 2017, and the Best Paper Award at IEEE GLOBECOM 2011. He is currently serves as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON COMMUNICATIONS.



Ming Li (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering (ECE) from Worcester Polytechnic Institute, MA, USA, in 2011. From 2011 to 2015, he was an Assistant Professor with the Computer Science Department, Utah State University. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, The University of Arizona, and also affiliated with the Computer Science Department. His main research interests are wireless and cyber security, with current emphases on cross-layer optimization and machine learning in wireless networks, wireless physical layer security, privacy enhancing technologies, and cyber-physical system security. He received the NSF Early Faculty Development (CAREER) Award in 2014 and the ONR Young Investigator Program (YIP) Award in 2016. He is a member of ACM.