On the Relation Between Identifiability, Differential Privacy, and Mutual-Information Privacy

Weina Wang, Lei Ying, and Junshan Zhang
School of Electrical, Computer and Energy Engineering
Arizona State University, Tempe, AZ 85287
{weina.wang, lei.ying.2, junshan.zhang}@asu.edu

***Abstract*—This paper investigates the relation between three different notions of privacy: identifiability, differential privacy and mutual-information privacy. Under a privacy–distortion framework, where the distortion is defined to be the expected Hamming distance between the input and output databases, we establish some fundamental connections between these three privacy notions. Given a maximum distortion D, let $\epsilon_{\rm i}^*(D)$ denote the smallest (best) identifiability level, and $\epsilon_{\rm d}^*(D)$ the smallest differential privacy level. Then we characterize $\epsilon_{\rm i}^*(D)$ and $\epsilon_{\rm d}^*(D)$, and prove that $\epsilon_{\rm i}^*(D) - \epsilon_X \le \epsilon_{\rm d}^*(D) \le \epsilon_{\rm i}^*(D)$ for D in some range, where ϵ_X is a constant depending on the distribution of the original database X, and diminishes to zero when the distribution of X is uniform. Furthermore, we show that identifiability and mutual-information privacy are consistent in the sense that given a maximum distortion D in some range, there is a mechanism that optimizes the identifiability level and also achieves the best mutual-information privacy.**

I. Introduction

Privacy has been an increasing concern in the emerging big data era, particularly with the growing use of personal data such as medical records or online activities for big data analysis. Analyzing these data results in new discoveries in science and engineering, but on the flip side of the coin it may put individual's privacy at potential risks. Therefore, privacy-preserving data analysis, where the goal is to preserve the accuracy of data analysis while maintaining individual's privacy, has emerged as one of the main challenges of this big data era. The basic idea of privacy-preserving data analysis is to add randomness in the released information to guarantee that an individual's information cannot be inferred. Intuitively, the higher the randomness is, the better privacy protection individual users get, but the less accurate (useful) the output statistical information is. While randomization seems to be inevitable, for the privacy-preserving data analysis it is of great interest to quantitatively define the notion of privacy. Specifically, we need to understand the amount of randomness needed to protect privacy while preserving the usefulness of the data. To this end, we consider three different notions: identifiability, differential privacy and mutual-information privacy. In particular, identifiability is concerned with the posteriors of recovering the original data from the released data; differential privacy is concerned with the additional information leakage of an individual due to the release of the data; mutual information measures the amount of information about the original database contained in the released data.

It turns out while these three different privacy notions are defined from different perspectives, they are fundamentally related. The focus of this paper is devoted to exploring the fundamental connections between these three different privacy notions in the following setting:

- We consider the non-interactive database releasing approach for privacy-preserving data analysis, where a synthetic database is released to the public. The synthetic database is a sanitized version of the original database, on which queries and operations can be carried out as if it was the original database. It is then natural to assume that the synthetic database and the original database are in the same universe so the entries have the same interpretation. Therefore we focus on mechanisms that map an input database to an output synthetic database in the same universe. Specifically, we consider a database consisting of n rows, each of which takes values from a finite domain $\mathcal{D}$ of size m. In this paper, the database is modeled as a discrete random variable X drawn from $\mathcal{D}^n$ with distribution p_X. A mechanism $\mathcal{M}$ is a mapping from an input database X to an output database Y, which is also a random variable with alphabet $\mathcal{D}^n$.
- We define the *distortion* between the output database and the input database to be the expected Hamming distance. [1] When the input and the output are in the same universe, the Hamming distance measures the number of rows two databases differ on, which directly points to the number of rows that need to be modified in order to guarantee a given privacy level.

In this paper, we use a unified *privacy–distortion* framework to understand the relation between the three privacy notions. Define the privacy–distortion function to be the best privacy level given a distortion constraint. Then we have the following main results, which are also summarized in Fig. 1.

(i) We derive the exact form of the privacy–distortion function $\epsilon_{\rm i}^*(D)$ under the notion of identifiability, for some range of distortion values, by showing that $\epsilon_{\rm i}^*(D) =$

[1] Our study of more general distortion measures is underway.

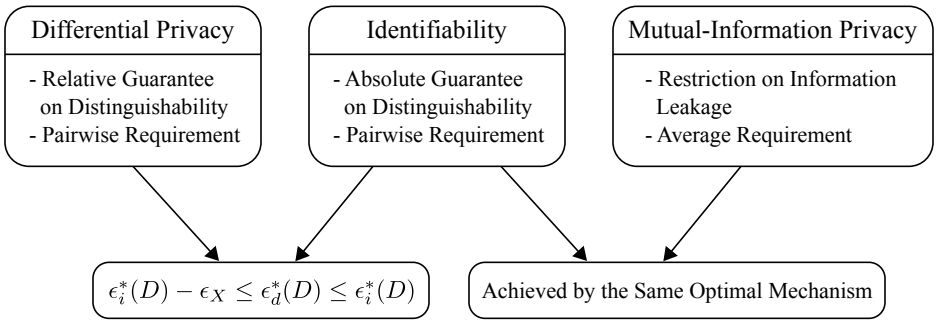

Fig. 1: Relation between identifiability, differential privacy and mutual-information privacy.

$h^{-1}(D)$ regardless of the prior distribution, where

$$h^{-1}(D) = \ln\Big(\frac{n}{D} - 1\Big) + \ln(m-1).$$

We further show that for the privacy–distortion function $\epsilon_{\rm d}^*(D)$ under the notion of differential privacy,

$$\epsilon_{\rm i}^*(D) - \epsilon_X \le \epsilon_{\rm d}^*(D) \le \epsilon_{\rm i}^*(D),$$

where ϵ_X is a constant depending on the distribution of X only, given by

$$\epsilon_X = \max_{x,x' \in \mathcal{D}^n : x \sim x'} \ln \frac{p_X(x)}{p_X(x')}.$$

It can be seen that when the input database has a uniform distribution, $\epsilon_{\rm i}^* = \epsilon_{\rm d}^*$, i.e., differential privacy is equivalent to identifiability. Note that for ϵ_X to be finite, the distribution p_X needs to have full support on $\mathcal{D}^n$, i.e., $p_X(x) > 0$ for any $x \in \mathcal{D}^n$. When ϵ_X is large, differential privacy provides only weak guarantee on identifiability. In other words, it is possible to identify some entries of the database with non-trivial accuracy even if the differential privacy is guaranteed when ϵ_X is large. This is because differential privacy provides a *relative* privacy guarantee, which ensures that limited *additional* information of an individual is leaked in the released data in addition to the knowledge that an adversary has known. Identifiability, on the other hand, guarantees an *absolute* level of indistinguishability of neighboring databases when being inferred from the output database assuming the prior distribution of p_X and the mechanism are both known to the adversary.

(ii) Given a maximum distortion D in some range, there is a mechanism that minimizes the identifiability level and also minimizes the mutual information between X and Y. In other words, identifiability and mutual-information privacy are consistent under the setting studied in this paper. This is somewhat surprising since identifiability imposes constraints on the distributions of neighboring input databases, which are "local" requirements; whereas the mutual information quantifies the correlation strength between the input database and the output database, which is a "global" measure. While the two notions are not directly comparable, the fact that they can be optimized simultaneously in the setting studied in this paper reveals the fundamental connection between these two privacy notions.

A. Related Work

Differential privacy, as an analytical foundation for privacy-preserving data analysis, was developed by a line of work (see, e.g., [1]–[3]). Dwork et al. [1] proposed the Laplace mechanism which adds Laplace noise to each query result, with noise amplitude proportional to the global sensitivity of the query function. Nissim et al. [4] later generalize the mechanism using the concept of local sensitivity. The notion of (ϵ, δ)-differential privacy [3] has also been proposed as a relaxation of ϵ-differential privacy.

The existing research of differential privacy can be largely classified into two categories: the interactive model where the randomness is added to the result of a query; and the non-interactive model, where the randomness is added to the database before queried. Under the interactive model, a significant body of work has been devoted to privacy–usefulness tradeoff and differentially private mechanisms with accuracy guarantee on each query result have been developed (see, e.g., [5]–[8]). Since the interactive model allows only a limited number of queries to be answered before the privacy is breached, researchers have also studied the non-interactive model, where synthetic databases or contingency tables with differential privacy guarantees were generated. Mechanisms with distortion guarantee for a set of queries to be answered using the synthetic database have been developed (see, e.g., [9]–[13]).

To be consistent with the definitions of privacy in legal rules, identifiability has also been considered as a notion of privacy. Lee and Clifton [14] and Li et al. [15] proposed differential identifiability and membership privacy, respectively. These privacy notions prevent the presence of certain entity from being identified. In this paper, we consider a form of identifiability that prevents the input database from being identified given the observation of the output database, prior knowledge and the mechanism.

Mutual information, as a measure of privacy leakage, has been widely used in the literature (see, e.g., [16]–[21]), mostly under the context of quantitative information flow and anonymity systems. The connection between differential privacy and information theory has been studied recently, e.g., [22], [23]. Alvim et al. [22] showed that differential privacy implies a bound on the min-entropy leakage. Mir [23] pointed out that the mechanism that achieves the optimal rate–distortion also guarantees a certain level of differential privacy. However, whether this differential privacy level is optimal or how far it is from optimal was not answered in [23]. The fundamental connection between differential privacy, mutual information and distortion is not yet clear. The connection between differential privacy and mutual information has also been studied in the two-party setting [24], where mutual information is used as the information cost for the protocol of communication between the two parities.

II. Model

Consider a database consisting of n rows/entries, each of which corresponds to some sensitive information. For example, each row could be an individual's medical records. The database could also be a graph, where each row indicates the existence of some edge. Suppose that rows take values from a domain $\mathcal{D}$. Then $\mathcal{D}^n$ is the set of all possible values of the database. Two databases, denoted by, $x, x' \in \mathcal{D}^n$, are said to be *neighbors* and denoted as $x \sim x'$ if they differ on exactly row. In this paper, we assume that the domain $\mathcal{D}$ is a finite set and model a database as a discrete random variable X with alphabet $\mathcal{D}^n$ and probability mass function (PMF) p_X. Suppose $|\mathcal{D}| = m$, where m is an integer and $m \geq 2$. A (randomized) mechanism $\mathcal{M}$ takes a database x as the input, and outputs a random variable $\mathcal{M}(x)$.

Definition 1 (Mechanism)**.** A *mechanism* $\mathcal{M}$ is specified by an *associated mapping* $\phi_{\mathcal{M}} \colon \mathcal{D}^n \to \mathcal{F}$, where $\mathcal{F}$ is the set of multivariate CDF's on some $\mathcal{R} \subseteq \mathbb{R}^r$. Taking database X as the input, the mechanism $\mathcal{M}$ outputs a $\mathcal{R}$-valued random variable Y with $\phi_{\mathcal{M}}(x)$ as the multivariate conditional CDF of Y given $X = x$. □

In this paper, we focus on mechanisms $\mathcal{M}$ of which the range is the same as the alphabet of X, i.e., $\mathcal{R} = \mathcal{D}^n$. Then the output Y is also a discrete random variable with alphabet $\mathcal{D}^n$, and the entries of Y have the same interpretation as the entries of X. Denote the conditional PMF of Y given $X = x$ defined by the CDF $\phi_{\mathcal{M}}(x)$ as $p_{Y|X}(\cdot \mid x)$. Throughout this paper we use the following basic notation. We denote the set of real numbers by $\mathbb{R}$, the set of nonnegative real numbers by $\mathbb{R}^+$, and the set of nonnegative integers by $\mathbb{N}$. Let $\overline{\mathbb{R}}^+ = \mathbb{R}^+ \cup \{+\infty\}$.

A. Different Notions of Privacy

In addition to the output database Y, we assume that the adversary knows the prior distribution $p_X(x)$, which represents the side information the adversary has, and the privacy-preserving mechanism $\mathcal{M}$. The three notions of privacy studied in this paper are defined next.

Definition 2 (Identifiability)**.** A mechanism $\mathcal{M}$ satisfies ϵ-*identifiability* for some $\epsilon \in \overline{\mathbb{R}}^+$ if for any pair of neighboring elements $x, x' \in \mathcal{D}^n$ and any $y \in \mathcal{D}^n$,

$$p_{X|Y}(x \mid y) \leq e^{\epsilon} p_{X|Y}(x' \mid y). \tag{1}$$

□

Different from differential identifiability [14] and membership privacy [15], which are concerned with whether a particular entity occurs in the database, the notion of identifiability considered here provides a posteriori indistinguishability between neighboring x and x', thus preventing the whole database from being identified. Note that this is an *absolute* level of indistinguishability, such that the adversary cannot distinguish x and x' based on the output database y, prior knowledge p_X and the mechanism $\mathcal{M}$.

Definition 3 (Differential Privacy [1], [2])**.** A mechanism $\mathcal{M}$ satisfies ϵ-*differential privacy* for some $\epsilon \in \overline{\mathbb{R}}^+$ if for any pair of neighboring elements $x, x' \in \mathcal{D}^n$ and any $y \in \mathcal{D}^n$,

$$p_{Y|X}(y \mid x) \leq e^{\epsilon} p_{Y|X}(y \mid x'). \tag{2}$$

□

In [1], [2], a mechanism $\mathcal{M}$ satisfies ϵ-*differential privacy* for some $\epsilon \in \overline{\mathbb{R}}^+$ if for any pair of neighboring elements $x, x' \in \mathcal{D}^n$, and any $\mathcal{S} \subseteq \mathcal{R}$,

$$\Pr\{Y \in \mathcal{S} \mid X = x\} \leq e^{\epsilon} \Pr\{Y \in \mathcal{S} \mid X = x'\}, \tag{3}$$

where the conditional probabilities $\Pr\{Y \in \mathcal{S} \mid X = x\}$ and $\Pr\{Y \in \mathcal{S} \mid X = x'\}$ are defined by the multivariate conditional CDF's $\phi_{\mathcal{M}}(x)$ and $\phi_{\mathcal{M}}(x')$, respectively. In the case that the range $\mathcal{R} = \mathcal{D}^n$, which is a discrete set, this is equivalent to the requirement (2). The differential privacy property of a mechanism is fully characterized by the associated mapping. Given any particular database, a mechanism $\mathcal{M}$ provides the same privacy guarantee regardless of the prior p_X, as long as the associated mapping $\phi_{\mathcal{M}}$ has been specified.

Note that the guarantee that the notion of differential privacy provides is a *relative* one, which ensures that limited *additional* information of an individual is leaked due to the presence of this individual in the database in addition to the knowledge that an adversary has known.

Definition 4 (Mutual-Information Privacy)**.** A mechanism $\mathcal{M}$ satisfies ϵ-*mutual-information privacy* for some $\epsilon \in \overline{\mathbb{R}}^+$ if the mutual information between X and Y satisfies $I(X;Y) \leq \epsilon$, where

$$I(X;Y) = \sum_{x,y \in \mathcal{D}^n} p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x) p_Y(y)}. \tag{4}$$

□

Mutual information is widely used to quantify information leakage in the literature, mostly under the context of quantitative information flow and anonymity systems. Under our setting, the information leakage we need to quantify is between the input database X and the output database Y. Note that the notion of mutual information is an information theoretic notion of privacy, which measures the *average* amount of information about X contained in Y. When X and Y are independent, $I(X;Y) = 0$. The mutual information is maximized and equal to $H(X)$ when $Y = X$.

B. Distortion

In this paper, we measure the usefulness of a mechanism by the distortion between the database X and the output Y, where smaller distortion yields greater usefulness. Consider the Hamming distance $d\colon \mathcal{D}^n \times \mathcal{D}^n \to \mathbb{N}$. Viewing elements in $\mathcal{D}^n$ as vectors of n rows, the distance $d(x, x')$ between two elements $x, x' \in \mathcal{D}^n$ is the number of rows they differ on. We define the distortion between X and Y to be the expected Hamming distance

$$\mathbb{E}[d(X,Y)] = \sum_{x\in\mathcal{D}^n}\sum_{y\in\mathcal{D}^n} p_X(x)p_{Y|X}(y \mid x)d(x,y). \quad (5)$$

The Hamming distance also characterizes the neighboring relation on $\mathcal{D}^n$. Two elements $x, x' \in \mathcal{D}^n$ are neighbors if and only if $d(x, x') = 1$.

C. Privacy–Distortion Function

A privacy–distortion pair (ϵ, D) is said to be *achievable* if there exists a mechanism $\mathcal{M}$ with output Y such that $\mathcal{M}$ satisfies ϵ-privacy level and $\mathbb{E}[d(X,Y)] \le D$. The *privacy–distortion function* $\epsilon^*\colon \mathbb{R}^+ \to \overline{\mathbb{R}^+}$ is defined by

$$\epsilon^*(D) = \inf\{\epsilon\colon (\epsilon, D) \text{ is achievable}\}, \quad (6)$$

which is the smallest (best) privacy level given the distortion constraint $\mathbb{E}[d(X,Y)] \le D$. We are only interested in the range $[0, n]$ for D since this is the meaningful range for distortion. The privacy–distortion function depends on the prior p_X, which reflects the impact of the prior on the privacy–distortion tradeoff. To characterize the privacy–distortion function, we also consider the *distortion–privacy function* $D^*\colon \overline{\mathbb{R}^+} \to \mathbb{R}^+$ defined by

$$D^*(\epsilon) = \inf\{D\colon (\epsilon, D) \text{ is achievable}\}, \quad (7)$$

which is the smallest achievable distortion given a privacy level ϵ.

In this paper we consider three different notions of privacy: identifiability, differential privacy and mutual-information privacy, so we denote the privacy–distortion functions under these three notions by ϵ_{i}^*, ϵ_{d}^* and ϵ_{m}^*, respectively.

III. Identifiability vs. Differential Privacy

In this section, we establish a fundamental connection between identifiability and differential privacy. Given privacy level ϵ_{i} and ϵ_{d}, the minimum distortion level is the solution to the following optimization problems.

The Privacy–Distortion Problem under Identifiability (PD-I):

$$\begin{aligned}
\min_{p_{X|Y},\, p_Y} \quad & \sum_{x\in\mathcal{D}^n}\sum_{y\in\mathcal{D}^n} p_Y(y)p_{X|Y}(x \mid y)d(x,y) \\
\text{subject to} \quad & p_{X|Y}(x \mid y) \le e^{\epsilon_{\mathrm{i}}} p_{X|Y}(x' \mid y), \\
& \qquad \forall x, x' \in \mathcal{D}^n\colon x \sim x', y \in \mathcal{D}^n, && (8)\\
& \sum_{x\in\mathcal{D}^n} p_{X|Y}(x \mid y) = 1, \qquad \forall y \in \mathcal{D}^n, && (9)\\
& p_{X|Y}(x \mid y) \ge 0, \qquad \forall x, y \in \mathcal{D}^n, && (10)\\
& \sum_{y\in\mathcal{D}^n} p_{X|Y}(x \mid y)p_Y(y) = p_X(x), \\
& \qquad \forall x \in \mathcal{D}^n, && (11)\\
& p_Y(y) \ge 0, \qquad \forall y \in \mathcal{D}^n. && (12)
\end{aligned}$$

The Privacy–Distortion Problem under Differential Privacy (PD-DP):

$$\begin{aligned}
\min_{p_{Y|X}} \quad & \sum_{x\in\mathcal{D}^n}\sum_{y\in\mathcal{D}^n} p_X(x)p_{Y|X}(y \mid x)d(x,y) \\
\text{subject to} \quad & p_{Y|X}(y \mid x) \le e^{\epsilon_{\mathrm{d}}} p_{Y|X}(y \mid x'), \\
& \qquad \forall x, x' \in \mathcal{D}^n\colon x \sim x', y \in \mathcal{D}^n, && (13)\\
& \sum_{y\in\mathcal{D}^n} p_{Y|X}(y \mid x) = 1, \qquad \forall x \in \mathcal{D}^n, && (14)\\
& p_{Y|X}(y \mid x) \ge 0, \qquad \forall x, y \in \mathcal{D}^n. && (15)
\end{aligned}$$

For convenience, we first define two constants ϵ_X and $\widetilde{\epsilon}_X$ that only depend on the prior p_X. Let

$$\epsilon_X = \max_{x,x'\in\mathcal{D}^n\colon x\sim x'} \ln\frac{p_X(x)}{p_X(x')}, \quad (16)$$

which is the maximum prior probability difference between two neighboring databases. For ϵ_X to be finite, the distribution p_X needs to have full support on $\mathcal{D}^n$, i.e., $p_X(x) > 0$ for any $x \in \mathcal{D}^n$. To define $\widetilde{\epsilon}_X$, note that the prior p_X puts some constraints on the posterior probabilities. We say $\{p_{X|Y}(x \mid y), x, y \in \mathcal{D}^n\}$ is *feasible* if there exists a PMF p_Y such that it is the marginal PMF of Y. Let $\widetilde{\epsilon}_X$ be the smallest ϵ such that the following posterior probabilities are feasible:

$$p_{X|Y}(x \mid y) = \frac{e^{-\epsilon d(x,y)}}{\big(1 + (m-1)e^{-\epsilon}\big)^n}, \quad x, y \in \mathcal{D}^n.$$

For any p_X, $\widetilde{\epsilon}_X$ is finite since when $\epsilon \to +\infty$, the PMF $p_Y = p_X$ is the marginal PMF of Y. Finally we consider the function

$$h^{-1}(D) = \ln\big(\frac{n}{D} - 1\big) + \ln(m-1). \quad (17)$$

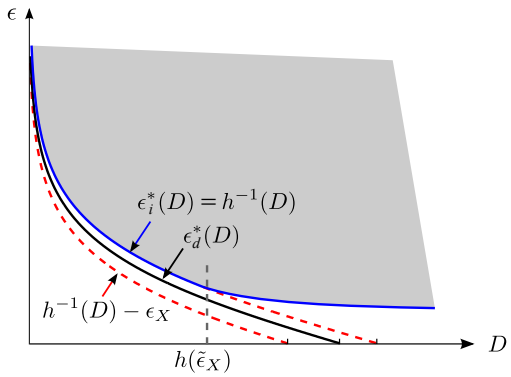

Fig. 2: The privacy–distortion function ϵ_{i}^* under identifiability and ϵ_{d}^* under differential privacy satisfy $\epsilon_{\mathrm{i}}^*(D) - \epsilon_X \le \epsilon_{\mathrm{d}}^*(D) \le \epsilon_{\mathrm{i}}^*(D)$ for D in some range.

Recall that $\epsilon_{\mathrm{i}}^*(D)$ denotes the best identifiability level under a maximum distortion D, and $\epsilon_{\mathrm{d}}^*(D)$ denotes the best differential privacy level under a maximum distortion D. The connection between the privacy–distortion functions ϵ_{i}^* and ϵ_{d}^* is established in the following theorem. See Fig. 2 for an illustration.

Theorem 1. *For identifiability, the privacy–distortion function ϵ_{i}^* of a database X with $\epsilon_X < +\infty$ satisfies*

$$\begin{cases} \epsilon_{\mathrm{i}}^*(D) = h^{-1}(D), & 0 \le D \le h(\widetilde{\epsilon}_X), \\ \epsilon_{\mathrm{i}}^*(D) \ge \max\{h^{-1}(D), \epsilon_X\}, & h(\widetilde{\epsilon}_X) < D \le n. \end{cases} \tag{18}$$

For differential privacy, the privacy–distortion function ϵ_{d}^* of a database X satisfies the following bounds for any D with $0 \le D \le n$:

$$\max\{h^{-1}(D) - \epsilon_X, 0\} \le \epsilon_{\mathrm{d}}^*(D) \le \max\{h^{-1}(D), 0\}. \tag{19}$$

□

From the theorem above, we can see that $0 \le \epsilon_{\mathrm{i}}^*(D) - \epsilon_{\mathrm{d}}^*(D) \le \epsilon_X$ when $0 \le D \le h(\widetilde{\epsilon}_X)$. A detailed proof of this theorem can be found in the complete version of our work [25]. Here we give a sketch of the proof, which consists of the following key steps:

- The first key step is to show that both PD-I and PD-DP, through (respective) relaxations as shown in Fig. 3, boil down to the same optimization problem.
  Relaxed Privacy–Distortion (R-PD):

$$\begin{aligned} \min_{p_{X|Y},\, p_Y} \quad & \sum_{x\in\mathcal{D}^n} \sum_{y\in\mathcal{D}^n} p_Y(y) p_{X|Y}(x \mid y) d(x,y) \\ \text{subject to} \quad & p_{X|Y}(x \mid y) \le e^{\epsilon} p_{X|Y}(x' \mid y), \\ & \qquad \forall x, x' \in \mathcal{D}^n \colon x \sim x', y \in \mathcal{D}^n, && (20) \\ & \sum_{x\in\mathcal{D}^n} p_{X|Y}(x \mid y) = 1, \qquad \forall y \in \mathcal{D}^n, && (21) \\ & p_{X|Y}(x \mid y) \ge 0, \qquad \forall x, y \in \mathcal{D}^n, && (22) \end{aligned}$$

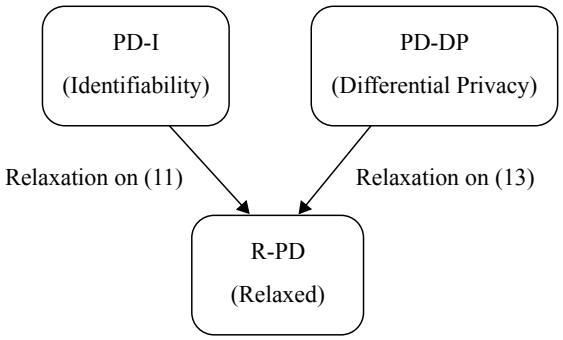

Fig. 3: Both PD-I and PD-DP boil down to R-PD through different relaxations.

$$\sum_{y\in\mathcal{D}^n} p_Y(y) = 1, \tag{23}$$

$$p_Y(y) \ge 0, \qquad \forall y \in \mathcal{D}^n. \tag{24}$$

Relaxing the constraint (11) in PD-I to the constraint (23) gives R-PD. Now consider PD-DP. For any neighboring $x, x' \in \mathcal{D}^n$, $p_X(x) \le e^{\epsilon_X} p_X(x')$ according to the definition of ϵ_X, and a necessary condition for the constraint (13) to be satisfied is

$$p_X(x) p_{Y|X}(y \mid x) \le e^{\epsilon_{\mathrm{d}} + \epsilon_X} p_X(x') p_{Y|X}(y \mid x'). \tag{25}$$

Therefore, replacing constraint (13) with (25) and letting $\epsilon = \epsilon_{\mathrm{d}} + \epsilon_X$, we obtain R-PD. So R-PD can be regarded as a relaxation of both PD-I and PD-DP.

- The minimum distortion in R-PD is showed to be $D_{\mathrm{relaxed}}^*(\epsilon) = h(\epsilon)$, which gives lower bounds on the distortion–privacy functions under identifiability and under differential privacy. By the connection between distortion–privacy function and privacy–distortion function, this minimum distortion implies that $\epsilon_{\mathrm{i}}^*(D) \ge h^{-1}(D)$ and $\epsilon_{\mathrm{d}}^*(D) \ge h^{-1}(D) - \epsilon_X$ for any D with $0 \le D \le n$. For ϵ_{i}^*, we find another lower bound: $\epsilon_{\mathrm{i}}^*(D) \ge \epsilon_X$ for any D with $0 \le D \le n$. These together give the lower bounds in Theorem 1.
- The upper bound on ϵ_{i}^* can be obtained by the following mechanism. Consider the mechanism $\mathcal{E}_i$ specified by

$$p_{Y|X}(y \mid x) = \frac{p_Y(y) e^{-\epsilon d(x,y)}}{p_X(x)\big(1 + (m-1)e^{-\epsilon}\big)^n}, \quad x, y \in \mathcal{D}^n, \tag{26}$$

  where $\epsilon \ge \widetilde{\epsilon}_X$ and p_Y is the corresponding PMF of Y. It can be shown that the mechanism $\mathcal{E}_i$ guarantees an identifiability level of ϵ with distortion $h(\epsilon)$ when $\epsilon \ge \widetilde{\epsilon}_X$. Therefore $\epsilon_{\mathrm{i}}^*(D) \le h^{-1}(D)$ for any D with $0 \le D \le h(\widetilde{\epsilon}_X)$. This yields (18) when combining with the lower bounds above.
- The upper bound on ϵ_{d}^* can be obtained by the following mechanism. Consider the mechanism $\mathcal{E}_d$ specified by the

conditional probabilities

$$p_{Y|X}(y \mid x) = \frac{e^{-\epsilon d(x,y)}}{\left(1+(m-1)e^{-\epsilon}\right)^n}, \quad x, y \in \mathcal{D}^n, \quad (27)$$

where $\epsilon \geq 0$. This is the exponential mechanism with score function $q = -d$ [26]. It can be shown that the mechanism $\mathcal{E}$ satisfies ϵ-differential privacy with distortion $h(\epsilon)$, which provides the upper bound in (19).

IV. Identifiability vs. Mutual-Information Privacy

In this section, we discuss the connection between identifiability and mutual-information privacy. Intuitively, mutual information can be used to quantify the information about X by observing a correlated random variable Y. Recall that the privacy–distortion function under mutual-information privacy denotes the smallest achievable mutual information without exceeding a maximum distortion. Note that this formulation has the same form as the formulation in the rate–distortion theory [27], and thus the privacy–distortion function under mutual-information privacy is identical to the rate–distortion function in this setting. We will show that identifiability and mutual-information privacy are consistent under the privacy–distortion framework in the sense that given a maximum distortion D, there is a mechanism that minimizes the identifiability level $\epsilon_{\mathrm{i}}^*(D)$ and also achieves the minimum mutual information, for some range of D.

It has been pointed out in [23] that the mechanism that achieves the optimal rate–distortion also guarantees a certain level of differential privacy. However, whether this differential privacy level is optimal or how far it is from optimal was not answered. Our result on the connection between identifiability and mutual-information privacy indicates that given a maximum distortion D, there is a mechanism that achieves the optimal rate–distortion and guarantees a differential privacy level ϵ such that $\epsilon_{\mathrm{D}}^*(D) \leq \epsilon \leq \epsilon_{\mathrm{D}}^*(D) + \epsilon_X$.

Given a maximum distortion D, the privacy–distortion function $\epsilon_{\mathrm{m}}^*(D)$ for input X with PMF $p_X(\cdot)$ is given by the optimal value of the following convex optimization problem.

The Privacy and Distortion Problem under Mutual-Information Privacy (PD-MIP):

$$\begin{aligned}
\min_{p_{Y|X}} \quad & I(X;Y) \\
\text{subject to} \quad & \sum_{x\in\mathcal{D}^n}\sum_{y\in\mathcal{D}^n} p_X(x)p_{Y|X}(y \mid x)d(x,y) \leq D, && (28)\\
& \sum_{y\in\mathcal{D}^n} p_{Y|X}(y \mid x) = 1, \qquad \forall x \in \mathcal{D}^n, && (29)\\
& p_{Y|X}(y \mid x) \geq 0, \qquad \forall x, y \in \mathcal{D}^n. && (30)
\end{aligned}$$

Theorem 2. *For any D with $0 \leq D \leq h(\widetilde{\epsilon}_X)$, the identifiability optimal mechanism $\mathcal{E}_i$ defined in* (26) *is also mutual-information optimal.* □

Proof. Consider the Lagrangian and the KKT conditions of the optimization problem PD-MIP. For any $0 \leq D \leq h(\widetilde{\epsilon}_X)$, consider the conditional probabilities $\{p_{Y|X}(y \mid x), x, y \in \mathcal{D}^n\}$ in (26) under $\mathcal{E}_i$ with $\epsilon = h^{-1}(D)$. Then it is easy to verify that $\{p_{Y|X}(y \mid x), x, y \in \mathcal{D}^n\}$ is primal feasible. Let $\lambda = h^{-1}(D)$, $\mu(x) = p_X(x)\ln\left[p_X(x)\left(1+(m-1)e^{-\epsilon}\right)^n\right]$ with $x \in \mathcal{D}^n$ and $\eta(x,y) = 0$ with $x, y \in \mathcal{D}^n$ be the Lagrange multipliers for (28), (29) and (30), respectively. Then these multipliers are dual feasible. The stationarity condition $p_X(x)\ln p_{Y|X}(y \mid x) - p_X(x)\ln p_Y(y) + \lambda p_X(x)d(x,y) + \mu(x) - \eta(x,y) = 0$ (derived in the complete version of our work [25]) and the complementary slackness condition are also satisfied. Therefore, the above $(p_{Y|X}, \lambda, \mu, \eta)$ satisfies the KKT conditions of PD-MIP.

Slater's condition holds for the problem PD-MIP since all the inequality constraints are affine [28]. By convexity, the KKT conditions are sufficient for optimality. Therefore the mechanism $\mathcal{E}_i$ with $\epsilon = h^{-1}(D)$ also gives the smallest mutual information. □

V. Conclusions

In this paper, we investigated the relation between three different notions of privacy: identifiability, differential privacy and mutual-information privacy, where identifiability provides absolute indistinguishability, differential privacy guarantees limited additional information leakage, and mutual information is an information theoretic notion of privacy. Under a unified privacy–distortion framework, where the distortion is defined to be the Hamming distance between the input and output databases, we established some fundamental connections between these three privacy notions. Given a maximum distortion D within some range, the smallest identifiability level $\epsilon_{\mathrm{i}}^*(D)$ and the smallest differential privacy level $\epsilon_{\mathrm{d}}^*(D)$ are proved to satisfy $\epsilon_{\mathrm{i}}^*(D) - \epsilon_X \leq \epsilon_{\mathrm{d}}^*(D) \leq \epsilon_{\mathrm{i}}^*(D)$, where ϵ_X is a constant depending on the distribution of the original database, and is equal to zero when the distribution is uniform. Furthermore, we showed that identifiability and mutual-information privacy are consistent in the sense that given the same maximum distortion D within some range, there is a mechanism that simultaneously optimizes the identifiability level and the mutual-information privacy.

Our findings in this study reveal some fundamental connections between the three notions of privacy. With these three notions of privacy being defined, many interesting issues deserve further attention. For example, in some cases, the prior p_X is imperfect, and it is natural to ask how we can protect privacy with robustness over the prior distribution. Some other interesting directions include the generalization from "pairwise" privacy to "group" privacy, which arises from the pairwise requirements that both identifiability and differential privacy impose on neighboring databases. The connections between membership privacy [15] and these three notions also need to be explored, since membership

privacy has been proposed as a unifying framework for privacy definitions.

VI. Acknowledgement

This work was supported in part by NSF Grant ECCS-1255425.

References

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Conf. Theory of Cryptography (TCC)*, New York, NY, 2006, pp. 265–284.

[2] C. Dwork, "Differential privacy," in *Proc. Int. Conf. Automata, Languages and Programming (ICALP)*, Venice, Italy, 2006, pp. 1–12.

[3] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: privacy via distributed noise generation," in *Proc. Annu. Int. Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT)*, St. Petersburg, Russia, 2006, pp. 486–503.

[4] K. Nissim, S. Raskhodnikova, and A. Smith, "Smooth sensitivity and sampling in private data analysis," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, San Diego, CA, 2007, pp. 75–84.

[5] A. Ghosh, T. Roughgarden, and M. Sundararajan, "Universally utility-maximizing privacy mechanisms," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, Bethesda, MD, 2009, pp. 351–360.

[6] A. Roth and T. Roughgarden, "Interactive privacy via the median mechanism," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, Cambridge, MA, 2010, pp. 765–774.

[7] M. Hardt and G. N. Rothblum, "A multiplicative weights mechanism for privacy-preserving data analysis," in *Proc. Ann. IEEE Symp. Found. Comput. Sci. (FOCS)*, Las Vegas, NV, 2010, pp. 61–70.

[8] S. Muthukrishnan and A. Nikolov, "Optimal private halfspace counting via discrepancy," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, New York, NY, 2012, pp. 1285–1292.

[9] A. Blum, K. Ligett, and A. Roth, "A learning theory approach to non-interactive database privacy," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, Victoria, Canada, 2008, pp. 609–618.

[10] C. Dwork, M. Naor, O. Reingold, G. N. Rothblum, and S. Vadhan, "On the complexity of differentially private data release: efficient algorithms and hardness results," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, Bethesda, MD, 2009, pp. 381–390.

[11] S. Kasiviswanathan, M. Rudelson, A. Smith, and J. Ullman, "The price of privately releasing contingency tables and the spectra of random matrices with correlated rows," in *Proc. Ann. ACM Symp. Theory of Computing (STOC)*, Cambridge, MA, 2010, pp. 775–784.

[12] J. Ullman and S. Vadhan, "PCPs and the hardness of generating private synthetic data," in *Proc. Conf. Theory of Cryptography (TCC)*, Providence, RI, 2011, pp. 400–416.

[13] M. Hardt, K. Ligett, and F. McSherry, "A simple and practical algorithm for differentially private data release," in *Advances Neural Information Processing Systems (NIPS)*, Lake Tahoe, NV, 2012, pp. 2348–2356.

[14] J. Lee and C. Clifton, "Differential identifiability," in *Proc. Ann. ACM SIGKDD Conf. Knowledge Discovery and Data Mining (KDD)*, Beijing, China, 2012, pp. 1041–1049.

[15] N. Li, W. Qardaji, D. Su, Y. Wu, and W. Yang, "Membership privacy: A unifying framework for privacy definitions," in *Proc. ACM Conf. Computer Communications Security (CCS)*, Berlin, Germany, 2013, pp. 889–900.

[16] D. Clark, S. Hunt, and P. Malacaria, "Quantitative information flow, relations and polymorphic types," *J. Log. and Comput.*, vol. 15, no. 2, pp. 181–199, Apr. 2005.

[17] G. Smith, "On the foundations of quantitative information flow," in *Proc. Int. Conf. Foundations of Software Science and Computational Structures (FSSACS)*, York, UK, 2009, pp. 288–302.

[18] K. Chatzikokolakis, T. Chothia, and A. Guha, "Statistical measurement of information leakage," in *Tools and Algorithms for the Constr. and Anal. of Syst.*, 2010, vol. 6015, pp. 390–404.

[19] Y. Zhu and R. Bettati, "Anonymity vs. information leakage in anonymity systems," in *Proc. IEEE Int. Conf. Distributed Computing Systems (ICDCS)*, Columbus, OH, 2005, pp. 514–524.

[20] K. Chatzikokolakis, C. Palamidessi, and P. Panangaden, "Anonymity protocols as noisy channels," in *Proc. Int. Conf. Trustworthy Global Computing (TGC)*, Lucca, Italy, 2007, pp. 281–300.

[21] L. Sankar, S. Rajagopalan, and H. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Trans. Inf. Forens. Security*, vol. 8, no. 6, pp. 838–852, Jun. 2013.

[22] M. S. Alvim, M. E. Andrés, K. Chatzikokolakis, P. Degano, and C. Palamidessi, "Differential privacy: On the trade-off between utility and information leakage," in *Formal Aspects of Security and Trust*, ser. Lecture Notes in Comput. Sci., 2012, vol. 7140, pp. 39–54.

[23] D. J. Mir, "Information-theoretic foundations of differential privacy," in *Found. and Practice of Security*, ser. Lecture Notes in Comput. Sci., 2013, vol. 7743, pp. 374–381.

[24] A. McGregor, I. Mironov, T. Pitassi, O. Reingold, K. Talwar, and S. Vadhan, "The limits of two-party differential privacy," in *Proc. Ann. IEEE Symp. Found. Comput. Sci. (FOCS)*, Las Vegas, NV, 2010, pp. 81–90.

[25] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy and mutual-information privacy," *arXiv:1402.3757 [cs.CR]*, Feb. 2014.

[26] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proc. Ann. IEEE Symp. Found. Comput. Sci. (FOCS)*, Providence, RI, 2007, pp. 94–103.

[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2006.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge Univ. Press, 2004.