Latent Feature Disclosure under Perfect Sample Privacy

Borzoo Rassouli[*]
University of Essex
Colchester CO43SQ, UK
b.rassouli@essex.ac.uk

Fernando Rosas[*]
Imperial College London
London SW72AZ, UK
f.rosas@imperial.ac.uk

Deniz Gündüz
Imperial College London
London SW72AZ, UK
d.gunduz@imperial.ac.uk

Abstract

Guaranteeing perfect data privacy seems to be incompatible with the economical and scientific opportunities provided by extensive data collection and processing. This paper tackles this challenge by studying how to disclose latent features of data sets without compromising the privacy of individual data samples. We leverage counter-intuitive properties of the multivariate statistics of data samples, and propose a technique to disclose collective properties of data sets while keeping each data sample confidential. For a given statistical description of the data set, we show how to build an optimal disclosure strategy/mapping using linear programming techniques. We provide necessary and sufficient conditions that determine when our approach is feasible, and illustrate the optimal solution in some simple scenarios. We observe that the disclosure strategy may be independent of the latent feature in some scenarios, for which explicit expressions for the performance are provided.

1. Introduction

The explosive developments in sensing, storage and networking technologies are allowing institutions to capture and exchange massive amounts of data, providing an unprecedented potential for important scientific and economic opportunities. For example, neuroimaging data can nowadays be shared effortlessly between researchers to allow parallel analyses, or consumer preferences can be extracted from online activity to aid the design of new products and services. However, recent issues related to the misuse of user data (e.g. the well-known case of Cambridge Analytica) are raising major concerns about information privacy, which is becoming a preeminent topic with important social, legal, and business consequences. Within this context, a key question that motivates this work is how to satisfy sufficient privacy requirements while still benefiting from extensive data sharing in a digital society.

The central intuition that drives our work comes from a key distinction between appropriate and inappropriate exploitation of user data: while the former just looks for statistical regularities, the latter is concerned about properties of specific entries/users. This suggests that a possible approach to preserve privacy would be to extract and share global properties of data, while keeping information about specific samples confidential. This manuscript is an attempt to formalize this intuition.

1.1. Scenario and related work

We consider a scenario where a user has a private data set, denoted by $X = (X_1, \ldots, X_n)$, which is correlated with a latent variable of interest W that the user wishes to share with an analyst. For example, X may represent measurements of a patient's vital signals, while W may be a unique health indicator, e.g., the risk of heart attack. While it would be desirable for the patient to share W with a remote assessment unit to provide alerts in case of an emergency, she may not want to share the data samples themselves as this could reveal unintended personal information.

We follow the framework for privacy against inference attacks [14], [4], which proposes to disclose a variable Y that is obtained through a mapping from the data set. In this context, the highest privacy standard is the *perfect sample privacy*; that is, we would like Y to not provide any useful information to foster statistical inference on particular samples of the data set, X_i, $\forall i$. This is equivalent to considering only those mappings whereby Y and X_i become statistically independent $\forall i$, while $W - X - Y$ form a Markov chain. To quantify the quality of Y as an estimator of W, we consider the mutual information between the two, $I(Y;W)$. This quantity is an adequate proxy (with more attractive algebraic properties) for the classification error rate [6, 8], which is a central performance metric for many machine learning tasks.

Note that the above conditions are not equivalent to imposing statistical independence between the disclosed variable Y and the whole data set X. In fact, if X and Y are independent then the data-processing inequality leads to

[*] equal contribution

$I(Y;W) \leq I(X;Y) = 0$, implying that under this condition the analyst would be unable to receive any information about W. At this point, it is useful to recall an intriguing but largely underexploited feature of multivariate statistics: variables that are pairwise independent can still be globally interdependent [9]. Said differently, although $I(Y;X) = 0$ implies $I(Y;X_i) = 0$ for $i = 1, \ldots, n$, the converse does not hold. For example, it is well-known that if X_1 and X_2 are two independent fair coins, then $Y = X_1 \oplus X_2$ (i.e., their exclusive OR) is independent of each of them, while $I(X_1, X_2; Y) > 0$ [15]. Therefore, in this case Y reveals a collective property (whether the entries of X are equal or not), while saying nothing about the individuals X_1 or X_2.

A related problem to the one considered here is the *privacy funnel*, in which the goal is to reveal the data set X within a given accuracy under some utility measure, while keeping the latent variable W as private as possible [10]. Perfect privacy in the privacy funnel setting has been studied in [5,12]. Also, various metrics for quantifying the quality of the disclosure strategy has been studied in [1,12,13].

1.2. Contributions

This paper presents an information disclosure technique that guarantees perfect sample privacy, which we call "synergistic information disclosure" as the mapping carries information about the whole data set (i.e., about X), but not about any of its constituting elements (i.e., X_i's). We derive necessary and sufficient conditions that determine when information about a latent feature can be disclosed under perfect sample privacy, and present a simple upper bound on the performance of the best-case scenario. Moreover, we show that the optimal disclosure mapping can be build via linear programming (LP). We illustrate these findings on a simple scenario where the data set X consists of two binary samples, for which we provide an explicit expression for the performance of the optimal synergistic disclosure mapping.

The rest of the paper is structured as follows. Section 2 introduces the notion of perfect sample privacy, and develops the conditions and bounds that characterize the private disclosure capacity. Subsequently, Section 3 proves that the optimal mapping can be found through an LP, and develops the case where the data set consists of two binary samples. Finally, Section 4 conveys our final remarks.

2. Definition and basic properties

2.1. Notations and Preliminaries

Random variables are denoted by capital letters, their realizations by lower case letters, and their alphabets/supports by capital letters in calligraphic font. Matrices and vectors are denoted by bold capital and bold lower case letters, respectively. For two random variables X and Y, $X \perp\!\!\!\perp Y$ indicates that they are statistically independent. For integers $m \leq n$, we define the discrete interval $[m:n] \triangleq \{m, m+1, \ldots, n\}$. For an integer $n \geq 1$, $\mathbf{1}_n$ denotes an n-dimensional all-one column vector. For a random variable $X \in \mathcal{X}$, with finite $|\mathcal{X}|$, the probability simplex $\mathcal{P}(\mathcal{X})$ is the standard $(|\mathcal{X}|-1)$-simplex given by

$$\mathcal{P}(\mathcal{X}) = \left\{ \mathbf{v} \in \mathbb{R}^{|\mathcal{X}|} \,\middle|\, \mathbf{1}_{|\mathcal{X}|}^T \cdot \mathbf{v} = 1,\ v_i \geq 0,\ \forall i \in [1:|\mathcal{X}|] \right\}.$$

To each probability mass function (pmf) on X, denoted by $p_X(\cdot)$ (or written simply as p_X), corresponds a probability vector $\mathbf{p}_X \in \mathcal{P}(\mathcal{X})$, whose i-th element is $p_X(x_i)$ ($i \in [1:|\mathcal{X}|]$). Likewise, for a pair of random variables (X,Y) with joint pmf $p_{X,Y}$, the probability vector $\mathbf{p}_{X|y}$ corresponds to the conditional pmf $p_{X|Y}(\cdot|y), \forall y \in \mathcal{Y}$, and $\mathbf{P}_{X|Y}$ is an $|\mathcal{X}| \times |\mathcal{Y}|$ matrix with columns $\mathbf{p}_{X|y}, \forall y \in \mathcal{Y}$. For matrix $\mathbf{P}_{m\times k}$, the null space, rank, and nullity are denoted by $\text{Null}(\mathbf{P})$, $\text{rank}(\mathbf{P})$, and $\text{nul}(\mathbf{P})$, respectively, where $\text{rank}(\mathbf{P}) + \text{nul}(\mathbf{P}) = k$.

The next lemma, which is used in the sequel, links statistical independence within a Markov chain with algebraic properties of matrices.

Lemma 1. *Let X, Y and Z be discrete random variables, which form a Markov chain $X - Y - Z$. In this setting, X and Z are statistically independent if and only if $\left(\mathbf{p}_Y - \mathbf{p}_{Y|z}\right) \in \text{Null}(\mathbf{P}_{X|Y})$, $\forall z \in \mathcal{Z}$.*

Proof. X and Z are independent if and only if $p_X(\cdot) = p_{X|Z}(\cdot|z)$, or equivalently, $\mathbf{p}_X = \mathbf{p}_{X|z}$, $\forall z \in \mathcal{Z}$. Furthermore, due to the Markov chain assumption, we have $\mathbf{p}_{X|z} = \mathbf{P}_{X|Y}\mathbf{p}_{Y|z}$, $\forall z \in \mathcal{Z}$, and in particular, $\mathbf{p}_X = \mathbf{P}_{X|Y}\mathbf{p}_Y$. Therefore, the condition $\mathbf{p}_X = \mathbf{p}_{X|z}, \forall z \in \mathcal{Z}$, is equivalent to

$$\mathbf{P}_{X|Y}\left(\mathbf{p}_Y - \mathbf{p}_{Y|z}\right) = \mathbf{0},\ \forall z \in \mathcal{Z},$$

or equivalently, $\left(\mathbf{p}_Y - \mathbf{p}_{Y|z}\right) \in \text{Null}(\mathbf{P}_{X|Y})$, $\forall z \in \mathcal{Z}$. □

2.2. Private disclosure capacity

Consider the random variables $W, X_1, \ldots, X_n$ distributed according to a given joint distribution $p_{W,X_1,\ldots,X_n}$. We focus on the case where $|\mathcal{W}|, |\mathcal{X}_i| < \infty$, $\forall i \in [1:n]$. Let $X \triangleq (X_1, \ldots, X_n)$, whose support is given by

$$\mathcal{X} = \left\{ (x_1, \ldots, x_n) \in \prod_{i=1}^{n} \mathcal{X}_i \,\middle|\, p_X(x_1, \ldots, x_n) > 0 \right\}.$$

Define the set of admissible stochastic mappings from the data set $\mathcal{X}$ to alphabet $\mathcal{Y}$ that satisfy perfect sample privacy as

$$\mathcal{A}_X = \left\{ p_{Y|X} \,\middle|\, Y \perp\!\!\!\perp X_i, \forall i \in [1:n] \right\}. \tag{1}$$

The *private disclosure capacity* for a latent variable W under perfect sample privacy is then defined as

$$I_s \triangleq \max_{\substack{p_{Y|X}\in\mathcal{A}_X:\\ W-X-Y}} I(W;Y). \tag{2}$$

The optimal mapping that maximizes the above expression is denoted by $p^*_{Y|X}$.

Following the literature on information-theoretic privacy, this approach assumes that an adequate statistical description of the database and the latent feature is available, and that adversaries do not posses additional information that could aid an inference attack [16]. Statistical descriptions of data are commonly estimated using analytic or numerical methods from the Bayesian or machine learning literature [3, 7]. However, one should keep in mind that estimation errors (e.g. due to insufficient training data) could potentially compromise the guarantees of perfect privacy.

2.3. Fundamental properties

We first investigate the conditions under which employing a synergistic disclosure strategy is feasible. Proposition 1 answers this by characterizing the necessary and sufficient conditions for having $I_\text{s} > 0$. Define matrix $\mathbf{P}$ as

$$\mathbf{P} \triangleq \begin{bmatrix} \mathbf{P}_{X_1|X} \\ \vdots \\ \mathbf{P}_{X_n|X} \end{bmatrix}_{G\times|\mathcal{X}|}, \tag{3}$$

where $G \triangleq \sum_{i=1}^n |\mathcal{X}_i|$ (note that in general $|\mathcal{X}| \neq \Pi_{i=1}^n |\mathcal{X}_i|$). Note that $\mathbf{P}$ is a binary matrix, as X_i's are deterministic functions of X. Examples of matrix $\mathbf{P}$ are presented in Section 3.

Proposition 1. *We have $I_s > 0$ if and only if* $\text{nul}(\mathbf{P}) \neq 0$ *and* $\text{Null}(\mathbf{P}) \not\subset \text{Null}(\mathbf{P}_{W|X})$.

Proof. As a preliminary remark, note that $X_i - X - Y$ form a Markov chain for any index $i \in [1:n]$. Therefore, from Lemma 1, X_i and Y are independent if and only if $(\mathbf{p}_X - \mathbf{p}_{X|y}) \in \text{Null}(\mathbf{P}_{X_i|X}), \forall y \in \mathcal{Y}$. This results in the following equivalence:

$$X_i \perp\!\!\!\perp Y,\ \forall i \in [1:n] \iff (\mathbf{p}_X - \mathbf{p}_{X|y}) \in \text{Null}(\mathbf{P}), \forall y \in \mathcal{Y}, \tag{4}$$

where matrix $\mathbf{P}$ is defined in (3).

For the first direction in the statement of the proposition, we proceed as follows. If $I_s > 0$, we have $W \not\perp\!\!\!\perp Y$. Therefore, there exist $y_1, y_2 \in \mathcal{Y}$, where $y_1 \neq y_2$, such that $\mathbf{p}_{W|y_1} \neq \mathbf{p}_{W|y_2}$, and hence, $\mathbf{p}_{X|y_1} \neq \mathbf{p}_{X|y_2}$. Since $X_i \perp\!\!\!\perp Y$, $\forall i \in [1:n]$, (4) implies that $(\mathbf{p}_X - \mathbf{p}_{X|y_1}), (\mathbf{p}_X - \mathbf{p}_{X|y_2}) \in \text{Null}(\mathbf{P})$, which results in $\text{nul}(\mathbf{P}) \neq 0$. Also, $\text{Null}(\mathbf{P}) \not\subset \text{Null}(\mathbf{P}_{W|X})$, since otherwise $\mathbf{P}_{W|X}(\mathbf{p}_X - \mathbf{p}_{X|y_1}) = \mathbf{P}_{W|X}(\mathbf{p}_X - \mathbf{p}_{X|y_2}) = \mathbf{0}$, which implies $\mathbf{p}_{W|y_1} = \mathbf{p}_{W|y_2}$ leading to a contradiction.

The second direction in the statement of the Proposition is proved as follows. If $\text{nul}(\mathbf{P}) \neq 0$ and $\text{Null}(\mathbf{P}) \not\subset \text{Null}(\mathbf{P}_{W|X})$, there exists a non-zero vector $\mathbf{v} \in \text{Null}(\mathbf{P})$, such that $\mathbf{v} \notin \text{Null}(\mathbf{P}_{W|X})$. Let $\mathcal{Y} = \{y_1, y_2\}$, $Y \sim \text{Bern}(\frac{1}{2})$, and for sufficiently small $\epsilon > 0$, let $\mathbf{p}_{X|y_1} = \mathbf{p}_X + \epsilon\mathbf{v}$ and $\mathbf{p}_{X|y_2} = \mathbf{p}_X - \epsilon\mathbf{v}$. This construction is possible as $\mathbf{p}_X$ lies in the interior of $\mathcal{P}(\mathcal{X})$, and $\mathbf{1}^T_{|\mathcal{X}|}\cdot\mathbf{v} = \mathbf{1}^T_G\cdot\mathbf{P}\mathbf{v} = 0$, which follows from $\mathbf{1}^T_{|\mathcal{X}|} = \mathbf{1}^T_G\cdot\mathbf{P}$, and $\mathbf{v} \in \text{Null}(\mathbf{P})$. Accordingly, since $\mathbf{p}_X - \mathbf{p}_{X|y_i} \in \text{Null}(\mathbf{P})$, $i = 1, 2$, from (4), we have $X_i \perp\!\!\!\perp Y$, $\forall i \in [1:n]$. Also, in the construction of the pair (X, Y), $\mathbf{p}_X$ is preserved, as specified in $p_{W,X}$, therefore, we have $W - X - Y$. Finally, since $\mathbf{v} \notin \text{Null}(\mathbf{P}_{W|X})$, from $\mathbf{p}_{W|y} = \mathbf{P}_{W|X}\mathbf{p}_{X|y}$, we get $\mathbf{p}_{W|y_1} \neq \mathbf{p}_{W|y_2}$, or equivalently, $I_s > 0$. □

In what follows, we propose an upper bound on I_s, which is tight as shown in Example 2.

Proposition 2. *The following upper bound holds for I_s:*

$$I_s \leq \min_{j\in\{1,\ldots,n\}} I(W;X_{-j}|X_j), \tag{5}$$

where $X_{-j} \triangleq \{X_1, \ldots, X_n\}\backslash X_j$.

Proof. Let $j \in [1:n]$ be an arbitrary index. Then,

$$\begin{aligned} I(W;Y) &= I(W;X) - I(W;X|Y) & (6)\\ &= I(W;X_{-j}|X_j) + I(W;X_j) \\ &\quad - I(W;X_j|Y) - I(W;X_{-j}|X_j,Y) \\ &= I(W;X_{-j}|X_j) + I(W;X_j) \\ &\quad - I(W,Y;X_j) - I(W;X_{-j}|X_j,Y) & (7)\\ &= I(W;X_{-j}|X_j) - I(Y;X_j|W) \\ &\quad - I(W;X_{-j}|X_j,Y) \\ &\leq I(W;X_{-j}|X_j), & (8) \end{aligned}$$

where (6) follows from the Markov chain $W - X - Y$, and (7) from the independence of X_j and Y. Since j is chosen arbitrarily, (8) holds for all $j \in [1:n]$, resulting in (5). □

3. Finding the optimal mapping

3.1. General solution

This section presents the main result of this work, which provides a practical method for computing the optimal latent feature disclosure strategy/mapping under perfect sample privacy.

Theorem 1. *The maximizer in (2), i.e., the optimal mapping $p^*_{Y|X}$, can be obtained as the solution to a standard LP.*

Proof. In what follows, we assume that $\text{nul}(\mathbf{P}) \neq 0$, since otherwise we have from Proposition 1 that $I_s = 0$, making the result trivial.

The singular value decomposition (SVD) of $\mathbf{P}$ gives $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where the matrix of right eigenvectors is

$$\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_{|\mathcal{X}|}\end{bmatrix}_{|\mathcal{X}|\times|\mathcal{X}|}. \tag{9}$$

By assuming (without loss of generality) that the singular values are arranged in a descending order, only the first $\text{rank}(\mathbf{P})$ singular values are strictly positive. Therefore, it is direct to check that the null space of $\mathbf{P}$ is given by

$$\text{Null}(\mathbf{P}) = \text{Span}\{\mathbf{v}_{\text{rank}(\mathbf{P})+1}, \dots, \mathbf{v}_{|\mathcal{X}|}\}. \tag{10}$$

Let $\mathbf{A} \triangleq \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_{\text{rank}(\mathbf{P})}\end{bmatrix}^T$, which has the useful property $\text{Null}(\mathbf{P}) = \text{Null}(\mathbf{A})$. Hence, from (4), having $X_i \perp\!\!\!\perp Y$, $\forall i \in [1:n]$ is equivalent to

$$\mathbf{A}(\mathbf{p}_X - \mathbf{p}_{X|y}) = \mathbf{0},\ \forall y \in \mathcal{Y}. \tag{11}$$

Let $\mathbb{S}$ be defined as

$$\mathbb{S} \triangleq \left\{\mathbf{t} \in \mathbb{R}^{|\mathcal{X}|} \middle| \mathbf{A}\mathbf{t} = \mathbf{A}\mathbf{p}_X\ ,\ \mathbf{t} \geq 0\right\}, \tag{12}$$

which is a convex polytope in $\mathcal{P}(\mathcal{X})$. If $W - X - Y$ with $p_{Y|X} \in \mathcal{A}_X$, from (11), one can see that $\mathbf{p}_{X|y} \in \mathbb{S}$, $\forall y \in \mathcal{Y}$. On the other hand, for any $p_{X,Y}$ for which $\mathbf{p}_{X|y} \in \mathbb{S}$, $\forall y \in \mathcal{Y}$, it is guaranteed that if one uses the corresponding mapping $p_{Y|X}$ to build a Markov chain $W - X - Y$, then the condition $X_i \perp\!\!\!\perp Y, \forall i \in [1:n]$ holds. Hence, we have proven the following equivalence:

$$W - X - Y,\ p_{Y|X} \in \mathcal{A}_X \iff \mathbf{p}_{X|y} \in \mathbb{S},\ \forall y \in \mathcal{Y}. \tag{13}$$

This leads us to

$$\begin{aligned} I_s =& H(W) - \min_{\substack{p_{Y|X}\in\mathcal{A}_X:\\ W-X-Y}} H(W|Y) \qquad (14)\\ =& H(W) - \min_{\substack{p_Y(\cdot),\mathbf{p}_{X|y}\in\mathbb{S},\ \forall y\in\mathcal{Y}:\\ \sum_y p_Y(y)\mathbf{p}_{X|y}=\mathbf{p}_X}} \sum_y p_Y(y) H\left(\mathbf{P}_{W|X}\mathbf{p}_{X|y}\right), \\ & \qquad (15)\end{aligned}$$

where, since the minimization is over $\mathbf{p}_{X|y}$ rather than $p_{Y|X}$, the constraint $\sum_y p_Y(y)\mathbf{p}_{X|y} = \mathbf{p}_X$ has been added to preserve the distribution $\mathbf{p}_X$ specified in $p_{W,X}$.

Next, we present a result that allows us to further simplify the optimization domain in (15).

Proposition 3. *For minimizing $H(W|Y)$ over $\mathbf{p}_{X|y} \in \mathbb{S}$ in (15), it is sufficient to consider only the extreme points of $\mathbb{S}$.*

Proof. Let $\mathbf{p}$ be an arbitrary point in $\mathbb{S}$. $\mathbf{p}$ can be written as[1] $\mathbf{p} = \sum_{i=1}^{|\mathcal{X}|} \alpha_i \mathbf{p}_i$, where $\alpha_i \geq 0$ $(\forall i \in [1:|\mathcal{X}|])$ and $\sum_{i=1}^{|\mathcal{X}|} \alpha_i = 1$; points $\mathbf{p}_i$ $(\forall i \in [1:|\mathcal{X}|])$ belong to the extreme points of $\mathbb{S}$ and $\mathbf{p}_i \neq \mathbf{p}_j$ $(i \neq j)$. From the concavity of entropy, we have

$$H(\mathbf{P}_{W|X}\mathbf{p}) \geq \sum_{i=1}^{|\mathcal{X}|} \alpha_i H(\mathbf{P}_{W|X}\mathbf{p}_i). \tag{16}$$

Therefore, from (16), it is sufficient to consider only the extreme points of $\mathbb{S}$ in the minimization. □

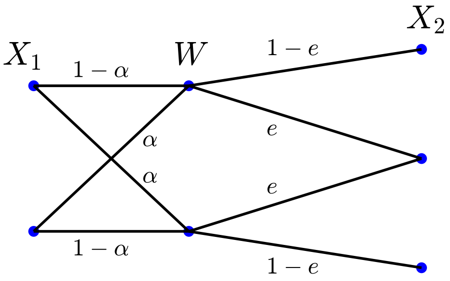

Figure 1: Example 1, where X_1 and X_2 are observations of W through a $BSC(\alpha)$ and a $BEC(e)$, respectively.

Using Proposition 3, the problem in (15) can be solved in two steps: a first step in which the extreme points of set $\mathbb{S}$ are identified, followed by a second step where proper weights over these extreme points are obtained to minimize the objective function.

For the first step, we first note that the extreme points of $\mathbb{S}$ are the basic feasible solutions (see [2], [11]) of it, i.e., the basic feasible solutions of the set

$$\left\{\mathbf{t} \in \mathbb{R}^{|\mathcal{X}|} \middle| \mathbf{A}\mathbf{t} = \mathbf{b}\ ,\ \mathbf{t} \geq 0\right\},$$

where $\mathbf{b} = \mathbf{A}\mathbf{p}_X$. The procedure of finding the extreme points of $\mathbb{S}$ is as follows. Pick a set $\mathcal{B} \subset [1:|\mathcal{X}|]$ of indices that correspond to $\text{rank}(\mathbf{P})$ linearly independent columns of matrix $\mathbf{A}$. Let $\mathbf{A}_\mathcal{B}$ be a $\text{rank}(\mathbf{P}) \times \text{rank}(\mathbf{P})$ matrix whose columns are the columns of $\mathbf{A}$ indexed by the indices in $\mathcal{B}$. Also, for any $\mathbf{x} \in \mathbb{S}$, let $\tilde{\mathbf{x}} = \begin{bmatrix}\mathbf{x}_\mathcal{B}^T & \mathbf{x}_\mathcal{N}^T\end{bmatrix}^T$, where $\mathbf{x}_\mathcal{B}$ and $\mathbf{x}_\mathcal{N}$ are $\text{rank}(\mathbf{P})$-dimensional and $\text{nul}(\mathbf{P})$-dimensional vectors whose elements are the elements of $\mathbf{x}$ indexed by the indices in $\mathcal{B}$ and $[1:|\mathcal{X}|]\backslash\mathcal{B}$, respectively.

For any basic feasible solution $\mathbf{x}^*$, there exists a set $\mathcal{B} \subset [1:|\mathcal{X}|]$ of indices that correspond to a set of linearly independent columns of $\mathbf{A}$, such that the corresponding vector of $\mathbf{x}^*$, i.e. $\tilde{\mathbf{x}}^* = \begin{bmatrix}\mathbf{x}_\mathcal{B}^{*T} & \mathbf{x}_\mathcal{N}^{*T}\end{bmatrix}^T$, satisfies the following

$$\mathbf{x}_\mathcal{N}^* = \mathbf{0},\quad \mathbf{x}_\mathcal{B}^* = \mathbf{A}_\mathcal{B}^{-1}\mathbf{b},\quad \mathbf{x}_\mathcal{B}^* \geq 0.$$

On the other hand, for any set $\mathcal{B} \subset [1:|\mathcal{X}|]$ of indices that correspond to a set of linearly independent columns of $\mathbf{A}$,

[1] The set $\mathbb{S}$ is an at most $(|\mathcal{X}|-1)$-dimensional convex subset of $\mathbb{R}^{|\mathcal{X}|}$. Therefore, any point in $\mathbb{S}$ can be written as a convex combination of at most $|\mathcal{X}|$ extreme points of $\mathbb{S}$.

if $\mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b} \geq 0$, then $\begin{bmatrix}\mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b} \\ \mathbf{0}\end{bmatrix}$ is the corresponding vector of a basic feasible solution. Hence, the extreme points of $\mathbb{S}$ are obtained as mentioned above, and their number is upper bounded by $\binom{|\mathcal{X}|}{\mathrm{rank}(\mathbf{P})}$.

For the second step, assume that the extreme points of $\mathbb{S}$, found in the first step, are denoted by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_K$. Then (15) is equivalent to

$$H(W) - \min_{\mathbf{u}\geq 0} \begin{bmatrix} H(\mathbf{P}_{W|X}\mathbf{p}_1) & \ldots & H(\mathbf{P}_{W|X}\mathbf{p}_K)\end{bmatrix} \cdot \mathbf{u}$$
$$\text{s.t. } \begin{bmatrix}\mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_K\end{bmatrix} \mathbf{u} = \mathbf{p}_X, \tag{17}$$

where $\mathbf{u}$ is a K-dimensional weight vector, and it can be verified that the constraint $\mathbf{1}_K^T \cdot \mathbf{u} = 1$ is satisfied if the constraint in (17) is met. The problem in (17) is a standard LP. □

The following example clarifies the optimization procedure in the proof of Theorem 1.

Example 1. *Let $W \sim$ Bern($\frac{1}{2}$) be the random variable that the user wishes to share with an analyst, and assume that the user has data samples denoted by X_1 and X_2, which are, respectively, the observations of W through a binary symmetric channel with crossover probability α, i.e., BSC(α), and a binary erasure channel with erasure probability e, i.e., BEC(e). Figure 1 provides an illustrative representation of this setting. Set $\alpha = \frac{2}{3}$, and $e = \frac{1}{2}$, which results in $\mathbf{p}_X = \frac{1}{12}\begin{bmatrix}1 & 3 & 2 & 2 & 3 & 1\end{bmatrix}^T$, and*

$$\mathbf{P}_{W|X} = \begin{bmatrix} 1 & \frac{2}{3} & 0 & 1 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 1 & 0 & \frac{2}{3} & 1 \end{bmatrix}. \tag{18}$$

Matrix $\mathbf{P}$ in (3) is given by

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix},$$

and by obtaining an SVD of $\mathbf{P}$, we obtain matrix $\mathbf{A}$ as

$$\mathbf{A} = \frac{1}{10^4}\begin{bmatrix} 4082 & 4082 & 4082 & 4082 & 4082 & 4082 \\ 4082 & 4082 & 4082 & -4082 & -4082 & -4082 \\ -4677 & 5270 & -593 & -4677 & 5270 & -593 \\ -3385 & -2385 & 5743 & -3385 & -2358 & 5743 \end{bmatrix}.$$

There are at most 15 ways of choosing 4 linearly independent columns of $\mathbf{A}$. From $\mathbf{x}_{\mathcal{B}} = \mathbf{A}_{\mathcal{B}}^{-1}\mathbf{A}\mathbf{p}_X$, and the condition $\mathbf{x}_{\mathcal{B}} \geq 0$, we obtain the extreme points of $\mathbb{S}$ as

$$\mathbf{p}_1 = \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{1}{4} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ \frac{1}{4} \\ 0 \\ \frac{1}{4} \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ 0 \\ 0 \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}, \mathbf{p}_4 = \begin{bmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ 0 \end{bmatrix}.$$

Finally, the LP is given by

$$\min_{\mathbf{u}\geq 0} \begin{bmatrix} H(\mathbf{P}_{W|X}\mathbf{p}_1) & \ldots & H(\mathbf{P}_{W|X}\mathbf{p}_4)\end{bmatrix} \cdot \mathbf{u} = 0.9866 \textit{ bits}$$
$$\textit{s.t. } \begin{bmatrix}\mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \mathbf{p}_4\end{bmatrix} \mathbf{u} = \mathbf{p}_X, \tag{19}$$

where $\mathbf{u}^* = \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3}\end{bmatrix}^T$. Therefore, the maximum information that can be shared with an analyst about W, while preserving the privacy of the observations, is $I_s = 0.0134$ bits, which is achieved by the following synergistic disclosure strategy

$$\mathbf{P}^*_{Y|X} = \begin{bmatrix} 1 & 0 & \frac{1}{2} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{2} & 0 & 1 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}. \tag{20}$$

3.2. Two binary observations

To illustrate the above results, in what follows, we consider the case where two binary (noisy) observations X_1, X_2 of an underlying phenomenon W are available. As before, the goal is to maximally inform an analyst about W, while preserving the privacy of both observations.

Consider the tuple (W, X_1, X_2) distributed according to a given joint distribution $p_{W,X_1,X_2} = p_{X_1,X_2}p_{W|X_1,X_2}$. In this setting, no condition is imposed on the conditional $p_{W|X_1,X_2}$. Without loss of generality, p_{X_1,X_2} is parametrized as

$$\mathbf{p}_X = \begin{bmatrix}\alpha - r & r & (\beta-\alpha)+r & (1-\beta)-r\end{bmatrix}^T, \tag{21}$$

where $\alpha, \beta \in (0,1)$ are degrees of freedom that determine the marginals , i.e., $X_1 \sim$ Bern(α) and X_2Bern(β), while $r \in [0, R]$ with $R \triangleq \min\{\alpha, 1-\beta\}$ determines the interdependency between X_1 and X_2. In particular, $X_1 \perp\!\!\!\perp X_2$, if and only if $r = \alpha(1-\beta)$.

If $r \in (0, R)$[2], we have $\mathcal{X} = \{(0,0),(0,1),(1,0),(1,1)\}$, and correspondingly one finds that

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

A direct calculation shows that the Null($\mathbf{P}$) is spanned by the single vector $\mathbf{n} = \begin{bmatrix}1 & -1 & -1 & 1\end{bmatrix}^T$. As the null space of $\mathbf{P}$ is one-dimensional, one can check that $\mathbb{S}$ has only two extreme points given by $\mathbf{a}_1 = \mathbf{p}_X - (R-r)\mathbf{n}$ and $\mathbf{a}_2 = \mathbf{p}_X + r\mathbf{n}$ (see Figure 2). Note that the original distribution can be recovered as a convex combination of these two extreme points, i.e.,

$$\mathbf{p}_X = \frac{r}{R}\mathbf{a}_1 + \frac{R-r}{R}\mathbf{a}_2. \tag{22}$$

[2] For the uninteresting cases where $r \in \{0, R\}$, we have $|\mathcal{X}| < 4$ and nul($\mathbf{P}$) $= 0$. Consequently, from Proposition 1, we get $I_s = 0$.

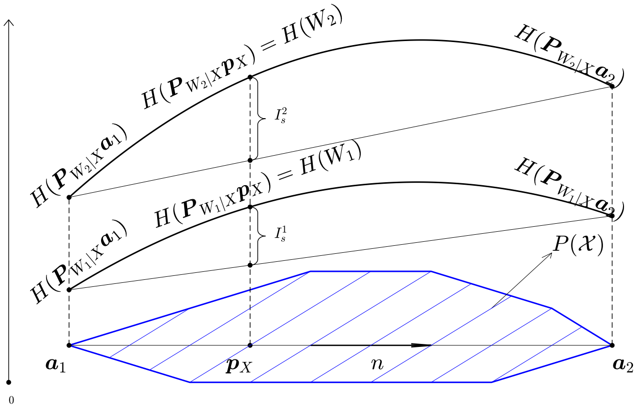
$H(\boldsymbol{P}_{W_2|X}\boldsymbol{p}_X) = H(W_2)$
$H(\boldsymbol{P}_{W_2|X}\boldsymbol{a}_2)$
$H(\boldsymbol{P}_{W_2|X}\boldsymbol{a}_1)$
I_s^2
$H(\boldsymbol{P}_{W_1|X}\boldsymbol{p}_X) = H(W_1)$
$H(\boldsymbol{P}_{W_1|X}\boldsymbol{a}_2)$
$H(\boldsymbol{P}_{W_1|X}\boldsymbol{a}_1)$
I_s^1
$P(\mathcal{X})$
$\boldsymbol{a}_1$
$\boldsymbol{p}_X$
n
$\boldsymbol{a}_2$
0

Figure 2: Diagram of private information disclosure for two tuples (W_1, X_1, X_2) and (W_2, X_1, X_2), where (X_1, X_2) are binary and distributed according to $\mathbf{p}_X$ as given in (21), and $p_{W_1|X} \neq p_{W_1|X}$. While their private disclosure capacities, i.e., I_s^i, $i = 1, 2$, are different, their optimal synergistic disclosure strategies are the same, as regardless of the tuples, we have $\mathbf{p}_X = \frac{r}{R}\mathbf{a}_1 + \frac{R-r}{R}\mathbf{a}_2$.

Therefore, using (15), I_s can be computed as

$$\begin{aligned} I_\mathrm{s} &= H(W) - \frac{r}{R} H(\mathbf{P}_{W|X}\mathbf{a}_1) - \frac{R-r}{R} H(\mathbf{P}_{W|X}\mathbf{a}_2) \\ &= H(\mathbf{p}_W) - \frac{r}{R} H\left(\mathbf{p}_W - (R-r)\mathbf{P}_{W|X}\mathbf{n}\right) \\ &\quad - \frac{R-r}{R} H\left(\mathbf{p}_W + r\mathbf{P}_{W|X}\mathbf{n}\right). \end{aligned} \tag{23}$$

From the last expression, it is direct to verify that, $I_\mathrm{s} > 0$ if and only if $\mathbf{n} \notin \mathrm{Null}(\mathbf{P}_{W|X})$. Finally, the optimal mapping $\mathbf{P}^*_{Y|X}$ is derived as follows. Considering (22), let $\mathcal{Y} \triangleq \{y_1, y_2\}$, and fix $p_Y(y_1) = \frac{r}{R}$ and $\mathbf{p}_{X|y_i} = \mathbf{a}_i, i = 1, 2$. Using these, a direct calculation results in the following optimal mapping

$$\mathbf{P}^*_{Y|X} = \begin{bmatrix} \frac{r(\alpha-R)}{R(\alpha-r)} & 1 & \frac{r(\beta-\alpha+R)}{R(\beta-\alpha+r)} & \frac{r(1-\beta-R)}{R(1-\beta-r)} \\ \frac{\alpha(R-r)}{R(\alpha-r)} & 0 & \frac{(\beta-\alpha)(R-r)}{R(\beta-\alpha+r)} & \frac{(1-\beta)(R-r)}{R(1-\beta-r)} \end{bmatrix}. \tag{24}$$

Although the private disclosure capacity in (23) depends on the choice of $\mathbf{P}_{W|X}$, the optimal synergistic disclosure strategy in (24) is only a functional of $\mathbf{p}_X$ (or equivalently, α, β, r), and does not depend on $\mathbf{P}_{W|X}$. This observation is formalized in the following proposition.

Proposition 4. *For the tuple (W, X_1, X_2), in which X_1 and X_2 are binary latent variables, the optimal synergistic disclosure strategy, i.e., $\mathbf{P}^*_{Y|X}$, does not depend on $p_{W|X_1,X_2}$.*

Proof. This follows from the fact that, in this setting, $\mathbb{S}$ has only two extreme points, and the condition of preserving $\mathbf{p}_X$ suffices to define the probability masses of these extreme points. Hence, the LP is solved already by its constraint. This is an example of the general case where there is only one way of writing an interior point of a set as a convex combination of its extreme points. □

This result implies that the same strategy can provide an optimal service in addressing any possible query over the data, as given by a specific $p_{W|X}$. In other words, optimal processing of the data can be done in the absence of knowledge about the query. However, $p_{W|X}$ plays a role in determining the effectiveness of the disclosure strategy (i.e. the magnitude of I_s), as illustrated in Figure 2.

We finish this section with an example.

Example 2. *Let us consider $\mathbf{p}_X$ as given by (21) with $\alpha = \beta = \frac{1}{2}$ and $r = \alpha(1-\beta) = \frac{1}{4}$, which corresponds to the case where X_1 and X_2 are two independent fair coins. From (24), one finds that*

$$\mathbf{P}^*_{Y|X} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \tag{25}$$

which proves that the optimal mapping for this case corresponds to $Y = X_1 \oplus X_2$. This, combined with Proposition 4, implies that $I_s \leq H(Y) = 1$.

As the optimal Y is independent of $p_{W|X}$, one can see that the case $W = X_1 \oplus X_2$, attains the maximal value $I_s = 1$. Moreover, a direct calculation shows that for this case $I(W; X_1|X_2) = I(W; X_2|X_1) = 1$, showing that the upper bound provided by Proposition 3 is attained.

4. Conclusions

This paper explored the idea of disclosing collective properties of a data set while ensuring element-wise confidentiality, which can be achieved by processing the data set with an adequate synergistic disclosure mapping. For the case of discrete variables that follow a known distribution, we have provided a method to build an optimal mapping that maximizes the mutual information with respect to a latent variable of interest. Moreover, we presented a tight upper bound for the optimal performance, and provided necessary and sufficient conditions that determine when this approach is effective. We have also illustrated our ideas on simple scenarios, and left the study of large data sets for a future extension of this work.

Acknowledgements

This work was carried out when B.R. was with the *Intelligent Systems and Networks* group at Imperial College London. This research was supported by the European Research Council (ERC) through the Starting Grant BEACON (agreement 677854), and by the UK Engineering and Physical Sciences Research Council (EPSRC) through the project COPES (EP/N021738/1). F.R. was supported by the EU's H2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No. 702981.

References

[1] Y. O. Basciftci, Y. Wang, and P. Ishwar. On privacy-utility tradeoffs for constrained data release mechanisms. In *2016 Information Theory and Applications Workshop (ITA)*, pages 1–6, Jan 2016. 2

[2] D. Bertsimas and J. N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientic, 1997. 4

[3] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006. 3

[4] F. Calmon and N. Fawaz. Privacy against statistical inference. In *50th Annual Allerton Conference*, pages 1401–1407, Illinois, USA, Oct. 2012. 1

[5] F. Calmon, A. Makhdoumi, and M. Médard. Fundamental limits of perfect privacy. In *IEEE Int. Symp. Inf. Theory (ISIT)*, pages 1796–1800, 2015. 2

[6] M. Feder and N. Merhav. Relations between entropy and error probability. *IEEE Transactions on Information Theory*, 40(1):259–266, 1994. 1

[7] A. Gelman, H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 2013. 3

[8] M. Hellman and J. Raviv. Probability of error, equivocation, and the chernoff bound. *IEEE Transactions on Information Theory*, 16(4):368–372, 1970. 1

[9] J. Jacod and P. Protter. *Probability essentials*. Springer Science & Business Media, 2012. 2

[10] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard. From the information bottleneck to the privacy funnel. In *IEEE Information Theory Workshop (ITW)*, pages 501–505, 2014. 2

[11] K. G. Murty. *Linear Programming*. John Wiley and Sons, 1983. 4

[12] B. Rassouli and D. Gündüz. On perfect privacy. In *IEEE Int. Symp. Inf. Theory (ISIT)*, 2018. 2

[13] B. Rassouli and D. Gündüz. Optimal utility-privacy trade-off with the total variation distance as the privacy measure. In *IEEE Information Theory Workshop (ITW), 2018*, accepted for publication. 2

[14] D. Rebollo-Monedero, J. Forne, and J. Domingo-Ferrer. From t-closeness-like privacy to postrandomization via information theory. *IEEE Trans. Knowl., Data Eng.*, 22(11):1623–1636, Nov. 2010. 1

[15] F. Rosas, V. Ntranos, C. J. Ellison, S. Pollin, and M. Verhelst. Understanding interdependency through complex information sharing. *Entropy*, 18(2):38, 2016. 2

[16] L. Sankar, S. R. Rajagopalan, and H. V. Poor. Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Transactions on Information Forensics and Security*, 8(6):838–852, 2013. 3