On Perfect Privacy

Borzoo Rassouli and Deniz Gündüz
Department of Electrical and Electronic Engineering, Imperial College London, London, U.K.
{b.rassouli12; d.gunduz}@imperial.ac.uk.

***Abstract*—For a pair of (dependent) random variables (X, Y), the following problem is addressed: What is the maximum information that can be revealed about Y, while disclosing no information about X? Assuming that a Markov kernel maps Y to the revealed information U, it is shown that the maximum mutual information between Y and U, i.e., $I(Y;U)$, can be obtained as the solution of a standard linear program, when X and U are required to be independent, called *perfect privacy*. The resulting quantity is shown to be greater than or equal to the *non-private information about* X *carried by* Y. For jointly Gaussian (X, Y), it is shown that perfect privacy is not possible if the kernel is applied to only Y; whereas perfect privacy can be achieved if the mapping is from both X and Y; that is, if the private variables can also be observed at the encoder. Finally, it is shown that when Y is not a deterministic function of X, perfect privacy is always feasible when the mapping has access to both X and Y.**[1]

***Index Terms*—Privacy, perfect privacy, non-private information, mutual information.**

I. Introduction

Consider a situation in which Alice wants to release some *useful* information to Bob, represented by random variable Y, and she receives some utility from this disclosure of information. At the same time, she wishes to conceal from Bob some *private* information which depends on Y, represented by X. To this end, a *privacy-preserving mapping* is applied, whereby a distorted version of Y, denoted by U, is revealed to Bob. In this context, privacy and utility are competing goals: The more distorted version of Y is revealed by the privacy mapping, the less information can Bob infer about X, while the less utility can be obtained. An extreme point of this trade-off is the scenario termed as *perfect privacy*, which refers to the situation where nothing is allowed to be inferred about X by Bob through the disclosure of U. This condition is modelled by the statistical independence of X and U.

In [1], a general statistical inference framework is proposed to capture the loss of privacy in legitimate transactions of data. In [2], the privacy-utility trade-off under the self-information cost function (log-loss) is considered and called the *privacy funnel*. In [3], sharp bounds on the optimal privacy-utility trade-off for the privacy funnel are derived, and an alternative characterization of the perfect privacy condition (see [4]) is proposed. Measuring both the privacy and the utility in terms of mutual information, perfect privacy is fully characterized in [5] for the binary case.

[1]This research was supported in part by the European Research Council (ERC) through Starting Grant BEACON (agreement 677854), and by the UK Engineering and Physical Sciences Research Council (EPSRC) through the project COPES (EP/N021738/1).

We study the information-theoretic perfect privacy in this paper, and our main contributions are as follows:

- Adopting mutual information as the utility measure, i.e., $I(Y;U)$, we show that the maximum utility under perfect privacy is the solution to a standard linear program (LP) that can be efficiently solved[2].
- We show that when (X, Y) is a jointly Gaussian pair with non-zero correlation coefficient, for the privacy mapping $p_{U|Y}$, perfect privacy is not feasible. In other words, maximum privacy is obtained at the expense of zero utility. This, however, is not the case when the mapping is of the form $p_{U|X,Y}$; that is, when the encoder has access to the private latent variables as well as the data.
- Denoting the maximum $I(Y;U)$ under perfect privacy by $g_0(X, Y)$, we characterize the relationship between the *non-private information about* X *carried by* Y, $D_X(Y)$ as defined in [5], and $g_0(X, Y)$.

Notations. Random variables are denoted by capital letters, their realizations by lower case letters, and their alphabets by capital letters in calligraphic font. Matrices and vectors are denoted by bold capital and bold lower case letters, respectively. For integers $m \leq n$, we have the discrete interval $[m : n] \triangleq \{m, m+1, \ldots, n\}$, and the tuple $(a_m, a_{m+1}, \ldots, a_n)$ is written in short as $a_{[m:n]}$. 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

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

Furthermore, to each probability mass function (pmf) on X, denoted by $p_X(\cdot)$, 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 $\mathbf{p}_{X|y}, \forall y \in \mathcal{Y}$. $F_Y(\cdot)$ denotes the cumulative distribution function (CDF) of random variable Y, and if it admits a density, its probability density function (pdf) is denoted by $f_Y(\cdot)$. For $0 \leq t \leq 1$, $H_b(t) \triangleq -t \log_2 t - (1-t)\log_2(1-t)$ denotes the binary entropy function with the convention $0 \log 0 = 0$. Throughout the paper, for a random variable Y with the corresponding probability vector $\mathbf{p}_Y$, the entropies $H(Y)$ and $H(\mathbf{p}_Y)$ are written interchangeably.

[2]Similar results are obtained in the longer version in [6] when the utility is measured by the decrease in the mean-square error or the probability of error.

II. System model and preliminaries

Consider a pair of random variables $(X,Y) \in \mathcal{X} \times \mathcal{Y}$ ($|\mathcal{X}|, |\mathcal{Y}| < \infty$) distributed according to the joint distribution $p_{X,Y}$. We assume that $p_Y(y) > 0, \forall y \in \mathcal{Y}$ and $p_X(x) > 0, \forall x \in \mathcal{X}$, since otherwise the supports $\mathcal{Y}$ or/and $\mathcal{X}$ could have been modified accordingly. Let X denote the *private/sensitive data* that the user wants to conceal and Y denote the *useful data* the user wishes to disclose. Assume that the *privacy mapping* takes Y as input and maps it to the *released data* denoted by U. In this scenario, $X-Y-U$ form a Markov chain, and the privacy mapping is captured by the conditional distribution $p_{U|Y}$.

Let $g_\epsilon(X,Y)$ be defined as [5]

$$g_\epsilon(X,Y) \triangleq \sup_{\substack{p_{U|Y}: \\ X-Y-U \\ I(X;U)\leq\epsilon}} I(Y;U). \tag{1}$$

In other words, when mutual information is adopted as a measure of both *utility* and *privacy*, (1) gives the optimal utility-privacy trade-off.

Proposition 1. In the evaluation of $g_0(X,Y)$, it is sufficient to restrict our attention to $|\mathcal{U}| \leq |\mathcal{Y}|$.

Proof. The proof is provided in [6, Appendix B]. □

III. Perfect Privacy

Definition. For a pair of random variables (X,Y), we say that *perfect privacy* is feasible if there exists a random variable U, such that $X-Y-U$ form a Markov chain, $X \perp\!\!\!\perp U$, i.e., X and U are independent, and $Y \not\perp\!\!\!\perp U$.

From the above definition, we can say that perfect privacy being feasible is equivalent to having $g_0(X,Y) > 0$.

Proposition 2. Perfect privacy is feasible if and only if

$$\dim\Big(\text{Null}(\mathbf{P}_{X|Y})\Big) \neq 0. \tag{2}$$

Proof. In [4, Theorem 4], the authors showed that for a given pair of random variables $(X,Y) \in \mathcal{X} \times \mathcal{Y}$ ($|\mathcal{X}|, |\mathcal{Y}| < \infty$), there exists a random variable U satisfying the conditions of perfect privacy if and only if the columns of $\mathbf{P}_{X|Y}$ are linearly dependent. Equivalently, there must exist a non-zero vector $\mathbf{v}$, such that $\mathbf{P}_{X|Y}\mathbf{v} = \mathbf{0}$, which is equivalent to (2). □

Proposition 3. For the null space of $\mathbf{P}_{X|Y}$, we have

$$\mathbf{z} \in \text{Null}(\mathbf{P}_{X|Y}) \Longrightarrow \mathbf{1}_{|\mathcal{Y}|}^T \cdot \mathbf{z} = 0.$$

Therefore, for any $\mathbf{z} \in \text{Null}(\mathbf{P}_{X|Y})$, there exists a positive real number α, such that $\mathbf{p}_Y + \alpha\mathbf{z} \in \mathcal{P}(\mathcal{Y})$.

Proof. For any $\mathbf{z}$ in the null space of $\mathbf{P}_{X|Y}$,

$$\mathbf{1}_{|\mathcal{Y}|}^T \cdot \mathbf{z} = \mathbf{1}_{|\mathcal{X}|}^T \mathbf{P}_{X|Y}\mathbf{z} = 0. \tag{3}$$

The last claim of the proposition is due to the fact that $\mathbf{p}_Y$ is in the interior of $\mathcal{P}(\mathcal{Y})$, i.e., $p_Y(y) > 0, \forall y \in \mathcal{Y}$. □

Theorem 1. For a pair of random variables $(X,Y) \in \mathcal{X} \times \mathcal{Y}$ ($|\mathcal{X}|, |\mathcal{Y}| < \infty$), $g_0(X,Y)$ is the solution to a standard linear program (LP) given in (10).

Proof. From the singular value decomposition of $\mathbf{P}_{X|Y}$, we have $\mathbf{P}_{X|Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where the right-singular vectors are denoted by $\mathbf{v}_i, i \in [1 : |\mathcal{Y}|]$. The condition in (2) is equivalent to having the null space of $\mathbf{P}_{X|Y}$ written as[3]

$$\text{Null}(\mathbf{P}_{X|Y}) = \text{Span}\{\mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{|\mathcal{Y}|}\}, \text{ for some } m \leq |\mathcal{Y}|. \tag{4}$$

For any pair (Y,U), for which $X-Y-U$ form a Markov chain, the independence of X and U, i.e., $\mathbf{p}_X = \mathbf{p}_{X|u}, \forall u \in \mathcal{U}$, is equivalent to the following for any $u \in \mathcal{U}$:

$$\mathbf{P}_{X|Y}(\mathbf{p}_Y - \mathbf{p}_{Y|u}) = \mathbf{0} \Longleftrightarrow (\mathbf{p}_Y - \mathbf{p}_{Y|u}) \in \text{Null}(\mathbf{P}_{X|Y}).$$

For the index m given in (4), construct the matrix $\mathbf{A}$ as

$$\mathbf{A} \triangleq \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_{m-1}\end{bmatrix}^T. \tag{5}$$

From (4) and (5), we can write

$$(\mathbf{p}_Y - \mathbf{p}_{Y|u}) \in \text{Null}(\mathbf{P}_{X|Y}) \Longleftrightarrow \mathbf{A}(\mathbf{p}_Y - \mathbf{p}_{Y|u}) = \mathbf{0}, \ \forall u \in \mathcal{U}.$$

Therefore, for any pair (Y,U), if $X-Y-U$ form a Markov chain and $X \perp\!\!\!\perp U$, we must have $\mathbf{p}_{Y|u} \in \mathbb{S}, \ \forall u \in \mathcal{U}$, where $\mathbb{S}$ is a convex polytope defined as

$$\mathbb{S} \triangleq \left\{\mathbf{x} \in \mathbb{R}^{|\mathcal{Y}|} \,\middle|\, \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{p}_Y \ , \ \mathbf{x} \geq 0\right\}. \tag{6}$$

It can be verified that any element of $\mathbb{S}$ is a probability vector.[4]

On the other hand, for any pair (Y,U), for which $\mathbf{p}_{Y|u} \in \mathbb{S}, \ \forall u \in \mathcal{U}$, we can simply make $X-Y-U$, where $X \perp\!\!\!\perp U$. Therefore, we can write

$$X-Y-U, \ X \perp\!\!\!\perp U \Longleftrightarrow \mathbf{p}_{Y|u} \in \mathbb{S}, \ \forall u \in \mathcal{U}. \tag{7}$$

This leads us to write $g_0(X,Y)$ as

$$\begin{aligned}\max_{\substack{p_{U|Y}: \\ X-Y-U \\ I(X;U)=0}} I(Y;U) &= \max_{\substack{p_{U|Y}: \\ \mathbf{p}_{Y|u}\in\mathbb{S}, \ \forall u\in\mathcal{U}}} I(Y;U) \\ &= H(Y) - \min_{\substack{p_U(\cdot), \mathbf{p}_{Y|u}\in\mathbb{S}, \ \forall u\in\mathcal{U}: \\ \sum_u p_U(u)\mathbf{p}_{Y|u}=\mathbf{p}_Y}} H(Y|U),\end{aligned} \tag{8}$$

where in (8), since the minimization is over $p_U(\cdot)$ and $\mathbf{p}_{Y|u}$ rather than $p_{U|Y}$, a constraint was added to preserve the marginal distribution $\mathbf{p}_Y$.

Proposition 4. In minimizing $H(Y|U)$ over $\mathbf{p}_{Y|u} \in \mathbb{S}$, it is sufficient to consider only $|\mathcal{Y}|$ extreme points of $\mathbb{S}$.

Proof. Assume that the minimum in (8) is achieved by $N(\leq |\mathcal{Y}|)$ points in $\mathbb{S}$, which follows from Proposition 1. Let $\mathbf{p}$ be an arbitrary point among these N points. $\mathbf{p}$ can be written as[5] $\mathbf{p} = \sum_{i=1}^{|\mathcal{Y}|} \alpha_i \mathbf{p}_i$, where $\alpha_i \geq 0$ ($\forall i \in [1 : |\mathcal{Y}|]$) and $\sum_{i=1}^{|\mathcal{Y}|} \alpha_i = 1$; points $\mathbf{p}_i$ ($\forall i \in [1 : |\mathcal{Y}|]$) belong to the extreme

[3] We assume, without loss of generality, that the singular values are arranged in a descending order.

[4] We have $\forall \mathbf{x} \in \mathbb{S}, (\mathbf{x} - \mathbf{p}_Y) \in \text{Null}(\mathbf{P}_{X|Y})$, and from Proposition 3, $\mathbf{1}_{|\mathcal{Y}|}^T \cdot \mathbf{x} = 1$.

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

points of $\mathbb{S}$ and $\mathbf{p}_i \neq \mathbf{p}_j$ $(i \neq j)$. From the concavity of entropy, we have

$$H(\mathbf{p}) \geq \sum_{i=1}^{|\mathcal{Y}|} \alpha_i H(\mathbf{p}_i), \tag{9}$$

where the equality holds if and only if all of the α_is but one are zero. From the definition of an extreme point, if $\mathbf{p}$ is not an extreme point of $\mathbb{S}$, it can be written with at least two non-zero α_is, which makes the inequality in (9) strict. However, this violates the assumption that the N points achieve the minimum. Hence, all of the N points of the minimizer must belong to the set of extreme points of $\mathbb{S}$. □

Therefore, the problem in (8) has two phases: in phase one, the extreme points of set $\mathbb{S}$ are identified, while in phase two, proper weights over these extreme points are obtained to minimize the objective function, i.e., $H(Y|U)$.

For the first phase, we proceed as follows. The extreme points of $\mathbb{S}$ are the basic feasible solutions (see [7], [8]) of it, i.e., the basic feasible solutions of the set

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

where $\mathbf{b} = \mathbf{A}\mathbf{p}_Y$. For any basic feasible solution $\mathbf{x}^*$, there exists a set $\mathcal{B} \subset [1 : |\mathcal{Y}|]$ 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: $\mathbf{x}_{\mathcal{N}}^* = \mathbf{0}, \mathbf{x}_{\mathcal{B}}^* = \mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b}, \mathbf{x}_{\mathcal{B}}^* \geq 0$. On the other hand, for any set $\mathcal{B} \subset [1 : |\mathcal{Y}|]$ of indices that correspond to a set of linearly independent columns of $\mathbf{A}$, 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 at most $\binom{|\mathcal{Y}|}{m-1}$.

For the second phase, we proceed as follows. Assume that the extreme points of $\mathbb{S}$, found in the previous phase, are denoted by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_K$. Then (8) is equivalent to

$$\begin{aligned} g_0(X,Y) = H(Y) - \min_{\mathbf{w} \geq 0} \ & \begin{bmatrix} H(\mathbf{p}_1) & H(\mathbf{p}_2) & \ldots & H(\mathbf{p}_K) \end{bmatrix} \cdot \mathbf{w} \\ \text{s.t.} \ & \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_K \end{bmatrix} \mathbf{w} = \mathbf{p}_Y, \end{aligned} \tag{10}$$

where $\mathbf{w}$ is a K-dimensional weight vector, and it can be verified that the constraint $\sum_{i=1}^{K} w_i = 1$ is satisfied if the constraint in (10) is met. The problem in (10) is a standard linear program (LP), which can be efficiently solved. □

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

Example. Consider the pair $(X, Y) \in [1:2] \times [1:4]$ whose joint distribution is specified by the following matrix:

$$\mathbf{P}_{X,Y} = \begin{bmatrix} 0.15 & 0.2 & 0.0625 & 0.05 \\ 0.35 & 0.05 & 0.0625 & 0.075 \end{bmatrix}.$$

Since $|\mathcal{Y}| > |\mathcal{X}|$, we have $\dim\Big(\text{Null}(\mathbf{P}_{X|Y})\Big) \neq 0$; and therefore, $g_0(X,Y) > 0$. From the singular value decomposition of $\mathbf{P}_{X|Y}$, we have $\sigma_1 = 1.41, \sigma_2 = 0.53, \sigma_3 = \sigma_4 = 0$, and

$$\mathbf{V} = \begin{bmatrix} -0.5 & 0.5345 & -0.4163 & -0.5394 \\ -0.5 & -0.8018 & -0.3154 & -0.0876 \\ -0.5 & 0 & 0.8452 & -0.1889 \\ -0.5 & 0.2673 & -0.1135 & 0.8159 \end{bmatrix},$$

where columns 3 and 4 of $\mathbf{V}$ span the null space of $\mathbf{P}_{X|Y}$, i.e., $m = 3$. Hence, the matrix $\mathbf{A}$ in (5) is given by

$$\mathbf{A} = \begin{bmatrix} -0.5 & -0.5 & -0.5 & -0.5 \\ 0.5345 & -0.8018 & 0 & 0.2673 \end{bmatrix}.$$

For the first phase, i.e., finding the extreme points of $\mathbb{S}$, it is clear that there are $\binom{4}{2}$ possible ways of choosing 2 linearly independent columns of $\mathbf{A}$. Hence, the index set $\mathcal{B}$ can be $\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}$ or $\{3,4\}$. From $\mathbf{x}_{\mathcal{B}} = \mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b}$, where $\mathbf{b} = \mathbf{A}\mathbf{p}_Y$, we get

$$\mathbf{x}_{\{1,2\}} = \begin{bmatrix} 0.675 \\ 0.325 \end{bmatrix}, \mathbf{x}_{\{1,3\}} = \begin{bmatrix} 0.1875 \\ 0.8125 \end{bmatrix}, \mathbf{x}_{\{1,4\}} = \begin{bmatrix} -0.625 \\ 1.625 \end{bmatrix}$$
$$\mathbf{x}_{\{2,3\}} = \begin{bmatrix} -0.125 \\ 1.125 \end{bmatrix}, \mathbf{x}_{\{2,4\}} = \begin{bmatrix} 0.1563 \\ 0.8437 \end{bmatrix}, \mathbf{x}_{\{3,4\}} = \begin{bmatrix} 0.625 \\ 0.375 \end{bmatrix}.$$

It is obvious that $\mathbf{x}_{\{1,4\}}$ and $\mathbf{x}_{\{2,3\}}$ are not feasible, since they do not satisfy $\mathbf{x}_{\mathcal{B}} \geq 0$. Therefore, the extreme points of $\mathbb{S}$ are $\mathbf{p}_1$ to $\mathbf{p}_4$ as below

$$\begin{bmatrix} 0.675 \\ 0.325 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.1875 \\ 0 \\ 0.8125 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0.1563 \\ 0 \\ 0.8437 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0.625 \\ 0.375 \end{bmatrix}.$$

For the second phase, the standard LP in (10) is written as

$$\begin{aligned} \min_{\mathbf{w} \geq 0} \ & \begin{bmatrix} 0.9097 & 0.6962 & 0.6254 & 0.9544 \end{bmatrix} . \mathbf{w} \\ \text{S.t.} \ & \begin{bmatrix} 0.675 & 0.1875 & 0 & 0 \\ 0.325 & 0 & 0.1563 & 0 \\ 0 & 0.8125 & 0 & 0.625 \\ 0 & 0 & 0.8437 & 0.375 \end{bmatrix} \mathbf{w} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{8} \\ \frac{1}{8} \end{bmatrix}, \end{aligned}$$

where the minimum value is 0.8437 bits, which is achieved by $\mathbf{w}^* = \begin{bmatrix} 0.698 & 0.1538 & 0.1481 & 0 \end{bmatrix}^T$. Therefore, $g_0(X,Y) = H(Y) - 0.8437 = 0.9063, \mathcal{U} = \{u_1, u_2, u_3\}$, $\mathbf{p}_U = \begin{bmatrix} 0.698 & 0.1538 & 0.1481 \end{bmatrix}^T$ and $\mathbf{p}_{Y|u_i} = \mathbf{p}_i, \ \forall i \in [1:3]$. Finally, $p_{U|Y}^*$ can be derived from the above.

Thus far, we have investigated the constraint of perfect privacy when $|\mathcal{X}|, |\mathcal{Y}| < \infty$. The following theorem shows that perfect privacy is not feasible for the (correlated) jointly Gaussian pair.

Theorem 2. Let $(X,Y) \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ be a pair of jointly Gaussian random variables, where

$$\mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}, \tag{11}$$

in which $\rho \neq 0$, since otherwise $X \perp\!\!\!\perp Y$. We have $g_0(X,Y) = 0$ for the above pair.

Proof. If there exists a random variable U such that $X-Y-U$ form a Markov chain and $X \perp\!\!\!\perp U$, we must have $F_X(\cdot) = F_{X|U}(\cdot|u),\ \forall u \in \mathcal{U}$, and hence, $f_X(\cdot) = f_{X|U}(\cdot|u),\ \forall u \in \mathcal{U}$, since X admits a density. Equivalently, we must have

$$f_X(\cdot) = \int f_{X|Y}(\cdot|y) dF_{Y|U}(y|u),\ \forall u \in \mathcal{U}. \tag{12}$$

Also, to have $g_0(X,Y) > 0$, there must exists at least $u_1, u_2 \in \mathcal{U}$, such that

$$F_{Y|U}(\cdot|u_1) \neq F_{Y|U}(\cdot|u_2). \tag{13}$$

In what follows we show that if (12) holds, (13) cannot be satisfied; and therefore, perfect privacy is not feasible for a jointly Gaussian (X,Y) pair.

It is known that X conditioned on $\{Y = y\}$ is also Gaussian, given by

$$X|\{Y=y\} \sim \mathcal{N}\Big(\underbrace{\frac{\rho\sigma_X}{\sigma_Y}(y-\mu_Y)+\mu_X}_{\alpha y+\beta}, \underbrace{(1-\rho^2)\sigma_X^2}_{\sigma^2}\Big). \tag{14}$$

From (12), (14), and for $u_1, u_2 \in \mathcal{U}$, we have

$$\int \frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big) = 0,\ \forall x \in \mathbb{R}. \tag{15}$$

Multiplying both sides of (15) by $e^{j\omega x}$, and taking the integral with respect to x, we obtain

$$\int e^{j\omega x}\left[\int \frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big)\right] dx = 0.$$

By Fubini's theorem[6], we can write

$$\int\left[\int e^{j\omega x}\frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx\right] d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big) = 0.$$

After some manipulations, we get

$$\int e^{j\omega\alpha y} d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big) = 0. \tag{16}$$

Since $\rho \neq 0$, from (14), we have $\alpha \neq 0$. Hence, the LHS of (16) is a Fourier transform. Due to the invertiblity of the Fourier transform, i.e. $\int e^{j\omega t} dg(t) = 0 \iff dg(t) = 0$, we must have $F_{Y|U}(\cdot|u_1) = F_{Y|U}(\cdot|u_2)$. Therefore, (13) does not hold and perfect privacy is not feasible for the (correlated) jointly Gaussian pair (X,Y). □

IV. Non-private information vs. $g_0(X,Y)$

For a pair of random variables $(X,Y) \in \mathcal{X}\times\mathcal{Y}$ ($|\mathcal{X}|, |\mathcal{Y}| < \infty$), the *private information about X carried by Y* is defined in [5] as

$$C_X(Y) \triangleq \min_{\substack{W: X-W-Y,\\ H(W|Y)=0}} H(W). \tag{17}$$

Since $H(W|Y) = 0$ implies that W is a deterministic function of Y, (17) means that among all the functions of Y that make X and Y conditionally independent, we want to find the one with the lowest entropy.

The *non-private information about X carried by Y* is defined in [5] as

$$D_X(Y) \triangleq H(Y) - C_X(Y). \tag{18}$$

Let $T^{\mathcal{X}} : \mathcal{Y} \to \mathcal{P}(\mathcal{X})$ be a mapping from $\mathcal{Y}$ to the probability simplex on $\mathcal{X}$ defined by $y \to p_{X|Y}(\cdot|y)$. It was shown in [5, Theorem 3] that the minimizer in (17) is $W^* = T^{\mathcal{X}}(Y)$; and hence, $D_X(Y) = H(Y) - H(T^{\mathcal{X}}(Y))$. Furthermore, it was proved in [5, Lemma 5] that $D_X(Y) = 0$ if and only if there do not exist $y_1, y_2 \in \mathcal{Y}$ such that $p_{X|Y}(\cdot|y_1) = p_{X|Y}(\cdot|y_2)$.

In [5], $g_0(X,Y)$ and $D_X(Y)$ were loosely connected to each other, as the latter represents roughly the amount of information contained in Y and not correlated with X. Three examples were provided, where in two of them $g_0(X,Y) = D_X(Y)$, while in the last one $g_0(X,Y) > D_X(Y)$. Finally a question was raised regarding the condition on the joint distribution $p_{X,Y}$ under which $g_0(X,Y) = D_X(Y)$ holds. The next theorem, whose proof is provided in [6], characterizes the relation between $D_X(Y)$ and $g_0(X,Y)$.

If $\mathbf{P}_{X|Y}$ has at least two identical columns, we define $\hat{\mathbf{P}}_{X|Y}$ as follows[7]. Let $\mathcal{E}_m \subset [1:|\mathcal{Y}|], \forall m \in [1:B]$, for some $B \geq 1$, be a set of indices corresponding to the columns in $\mathbf{P}_{X|Y}$ that are equal, i.e., $\mathbf{p}_{X|y_i} = \mathbf{p}_{X|y_j},\ \forall i,j \in \mathcal{E}_m, \forall m \in [1:B]$, and $\mathbf{p}_{X|y_i} \neq \mathbf{p}_{X|y_k},\ \forall i \in \mathcal{E}_m, \forall k \in [1:|\mathcal{Y}|]\backslash\mathcal{E}_m,\ \forall m \in [1:B]$. Let $G \triangleq \sum_{i=1}^{B} |\mathcal{E}_i|$. We construct a corresponding $|\mathcal{X}| \times (|\mathcal{Y}| - G + B)$-dimensional matrix $\hat{\mathbf{P}}_{X|Y}$ from $\mathbf{P}_{X|Y}$ by eliminating all the columns in each $\mathcal{E}_m$, except one. For example, we have the following pair for $\mathbf{P}_{X|Y}$ and $\hat{\mathbf{P}}_{X|Y}$

$$\begin{bmatrix} 0.3 & 0.3 & 0.4 & 0.5 & 0.4 \\ 0.2 & 0.2 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.1 & 0 & 0.1 \end{bmatrix}, \begin{bmatrix} 0.3 & 0.4 & 0.5 \\ 0.2 & 0.5 & 0.5 \\ 0.5 & 0.1 & 0 \end{bmatrix},$$

where $B = 2$, $G = 4$, $\mathcal{E}_1 = \{1,2\}$, and $\mathcal{E}_2 = \{3,5\}$.

Theorem 3. For a pair of random variables $(X,Y) \in \mathcal{X}\times\mathcal{Y}$ ($|\mathcal{X}|, |\mathcal{Y}| < \infty$) distributed according to $p_{X,Y}$, we have

$$g_0(X,Y) \geq D_X(Y), \tag{19}$$

where the equality holds if and only if either of the following holds:

1) Perfect privacy is not feasible, i.e. $\dim(\text{Null}(\mathbf{P}_{X|Y})) = 0$,
2) Perfect privacy is feasible, and $\dim(\text{Null}(\hat{\mathbf{P}}_{X|Y})) = 0$.

V. Full data observation vs. output perturbation

Thus far, we have assumed that the *privacy mapping* takes Y as input and maps it to the *released data* denoted by U. In a more general scenario, the privacy mapping can take a noisy version W of (X,Y) as input, as in [9]. In this case, the privacy mapping is denoted by $p_{U|W}$, and $(X,Y) - W - U$ form a Markov chain, where the triplet $(X,Y,W) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{W}$ ($|\mathcal{X}|, |\mathcal{Y}|, |\mathcal{W}| < \infty$) is distributed according to some given joint distribution $p_{X,Y,W}$. In this model, perfect privacy

[6]Note that $\int |f_{X|U}(x|u_1) - f_{X|U}(x|u_2)|dx \leq \int [|f_{X|U}(x|u_1)| + |f_{X|U}(x|u_2)|]dx = 2 < +\infty$.

[7]If this is not the case, let $\hat{\mathbf{P}}_{X|Y} \triangleq \mathbf{P}_{X|Y}$.

is feasible for the triplet (X, Y, W) if there exists a privacy mapping $p_{U|W}$ whose output (U) depends on the useful data (Y), while being independent of the private data (X); that is, $I(Y;U) > 0$ and $I(X;U) = 0$ as before.

Proposition 5. Perfect privacy is feasible for (X, Y, W) if and only if

$$\dim\Big(\text{Null}(\mathbf{P}_{X|W})\backslash\text{Null}(\mathbf{P}_{Y|W})\Big) \neq 0. \tag{20}$$

Proof. The proof follows similarly to that of Proposition 2 by noting that both $X-(X,Y)-W-U$ and $Y-(X,Y)-W-U$ form Markov chains. Hence, there must exist a vector in $\mathcal{P}(\mathcal{W})$, such that a change in $\mathbf{p}_W$ along that vector changes $\mathbf{p}_Y$ ($Y \not\perp\!\!\!\perp U$), while keeps $\mathbf{p}_X$ unchanged ($X \perp\!\!\!\perp U$). □

It can be verified that for the general scenario of $(X, Y) - W - U$, where the mapping is denoted by $p_{U|W}$, perfect privacy can be obtained through a similar LP as in Theorem 1 with the following modifications: It is sufficient to restrict our attention to $|\mathcal{U}| \leq |\mathcal{W}| + 1$; The convex polytope $\mathbb{S}$ is modified as the set of probability vectors $\mathbf{x}$ in $\mathcal{P}(\mathcal{W})$, such that $(\mathbf{p}_W - \mathbf{x}) \in \text{Null}(\mathbf{P}_{X|W})$; denoting the extreme points of $\mathbb{S}$ by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_K$, (10) changes to

$$\max_{\substack{p_{U|W}:\\(X,Y)-W-U\\X\perp\!\!\!\perp U}} I(Y;U) = H(Y) - \min_{\mathbf{w}\geq 0} \begin{bmatrix} H(\mathbf{q}_1) & \ldots & H(\mathbf{q}_K) \end{bmatrix} \cdot \mathbf{w},$$
$$\text{s.t.} \begin{bmatrix} \mathbf{p}_1 & \ldots & \mathbf{p}_K \end{bmatrix} \mathbf{w} = \mathbf{p}_W$$

where $\mathbf{q}_i = \mathbf{P}_{Y|W}\mathbf{p}_i, i \in [1:K]$.

The special cases of *full data observation* and *output perturbation* ([9]) refer to scenarios where the privacy mapping has direct access to both the private and useful data ($W = (X, Y)$) and only the useful data ($W = Y$), respectively. With these definitions, Sections II to IV consider the particular case of output perturbation. In what follows, we consider the full data observation scenario briefly.

Proposition 6. If Y is not a deterministic function of X, perfect privacy is always feasible in the full data observation model.

Proof. If Y is not a deterministic function of X, there must exist $x_1 \in \mathcal{X}$ and $y_1, y_2 \in \mathcal{Y}$ ($y_1 \neq y_2$) such that $p_{X,Y}(x_1, y_1) > 0$ and $p_{X,Y}(x_1, y_2) > 0$. Let $\mathcal{U} = \{u_1, u_2\}$ and $p_U(u_1) = \frac{1}{2}$. Choose a sufficiently small $\epsilon > 0$ and let

$$p_{X,Y|U}(x, y|u_1) = \begin{cases} p_{X,Y}(x_1, y_1) + \epsilon & (x, y) = (x_1, y_1) \\ p_{X,Y}(x_1, y_2) - \epsilon & (x, y) = (x_1, y_2) \\ p_{X,Y}(x, y) & \text{otherwise} \end{cases},$$

$$p_{X,Y|U}(\cdot, \cdot|u_2) = 2p_{X,Y}(\cdot, \cdot) - p_{X,Y|U}(\cdot, \cdot|u_1). \tag{21}$$

It can be verified that $p_{X,Y}$ is preserved in $p_{X,Y,U}$. Also, $p_{X|U}(\cdot|u) = p_X(\cdot)$, $\forall u \in \mathcal{U}$, and $p_{Y|U}(y_1|u_1) \neq p_Y(y_1)$, where the former indicates that $X \perp\!\!\!\perp U$, and the latter shows that $Y \not\perp\!\!\!\perp U$. □

Considering the output perturbation model, Theorem 2 proved that perfect privacy is not feasible for the (correlated) jointly Gaussian pair. The following theorem states the opposite for the full data model.

Theorem 4. For a jointly Gaussian pair (X, Y) with correlation coefficient ρ ($\neq 0$), perfect privacy is feasible for the full data observation model, and we have

$$\sup_{F_{U|X,Y}:X\perp\!\!\!\perp U} I(Y;U) \geq -\log\rho. \tag{22}$$

Proof. Denoting the variances of X and Y by σ_X^2 and σ_Y^2, respectively, it is already known that we can write $Y = \frac{\rho\sigma_Y}{\sigma_X}X + \sigma_Y\sqrt{1-\rho^2}N$, where $N \sim \mathcal{N}(0,1)$ is independent of X. By letting $U = \frac{1}{\sigma_Y\sqrt{1-\rho^2}}(Y - \frac{\rho\sigma_Y}{\sigma_X}X)$, i.e., $U = N$, we have $X \perp\!\!\!\perp U$, and

$$\begin{aligned} I(Y;U) &= h(Y) - h(Y|U) \\ &= \frac{1}{2}\log 2\pi e\sigma_Y^2 - h\left(\frac{\rho\sigma_Y}{\sigma_X}X + \sigma_Y\sqrt{1-\rho^2}N \middle| N\right) && (23) \\ &= \frac{1}{2}\log 2\pi e\sigma_Y^2 - h(\frac{\rho\sigma_Y}{\sigma_X}X) && (24) \\ &= \frac{1}{2}\log 2\pi e\sigma_Y^2 - \frac{1}{2}\log 2\pi e\rho^2\sigma_Y^2 \\ &= -\log\rho, && (25) \end{aligned}$$

where (24) follows from the fact that $X \perp\!\!\!\perp N$. □

VI. Conclusions

Adopting mutual information as the utility measure, it is shown that the maximum utility under perfect privacy is the solution to a standard linear program (LP). This solution is shown to be greater than or equal to the *non-private information about X carried by Y*, $D_X(Y)$ as defined in [5]. It is shown that when (X, Y) is a jointly Gaussian pair with non-zero correlation coefficient, for the privacy mapping $p_{U|Y}$, perfect privacy is not feasible. This, however, is not the case when the mapping is of the form $p_{U|X,Y}$. Finally, it is shown that when Y is not a deterministic function of X, perfect privacy is always feasible when the mapping has direct access to both X and Y.

References

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

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

[3] F. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2015, pp. 1796–1800.

[4] T. Berger and R. Yeung, "Multiterminal source encoding with encoder breakdown," *IEEE Trans. Inf. Theory*, pp. 237–244, 1989.

[5] S. Asoodeh, F. Alajaji, and T. Linder, "Notes on information-theoretic privacy," in *52nd Annual Allerton Conference*, Illinois, USA, Oct. 2014, pp. 1272–1278.

[6] B. Rassouli and D. Gündüz, "On perfect privacy and maximal correlation," *https://arxiv.org/pdf/1712.08500.pdf*, Dec. 2017.

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

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

[9] Y. Wang, Y. Basciftci, and P. Ishwar, "Privacy-utility tradeoffs under constrained data release mechanisms," *https://arxiv.org/pdf/1710.09295.pdf*, Oct. 2017.