Fundamental limits of perfect privacy

Fundamental Limits of Perfect Privacy

Flavio P. Calmon, Ali Makhdoumi, Muriel Médard

***Abstract*—"To be considered for an 2015 IEEE Jack Keil Wolf ISIT Student Paper Award." We investigate the problem of intentionally disclosing information about a set of measurement points X (useful information), while guaranteeing that little or no information is revealed about a private variable S (private information). Given that S and X are drawn from a finite set with joint distribution $p_{S,X}$, we prove that a non-trivial amount of useful information can be disclosed while not disclosing any private information if and only if the smallest principal inertia component of the joint distribution of S and X is 0. This fundamental result characterizes when useful information can be privately disclosed for any privacy metric based on statistical dependence. We derive sharp bounds for the tradeoff between disclosure of useful and private information, and provide explicit constructions of privacy-assuring mappings that achieve these bounds.**

***Index Terms*—Statistical Privacy; Privacy Funnel; Privacy-Utility Trade-off; Principal Inertia Components.**

I. Introduction

We adopt the privacy against statistical inference framework presented in [1] with the mutual information utility function. This setup, called the *Privacy Funnel*, was introduced in [2]. Consider two communicating parties, namely Alice and Bob. Alice's goal is to disclose to Bob information about a set of measurement points, represented by the random variable X. Alice discloses this information in order to receive some utility from Bob. Simultaneously, Alice wishes to limit the amount of information revealed about a private random variable S that is dependent on X. For example, X may represent Alice's movie ratings, released to Bob in order to receive movie recommendations, whereas S may represent Alice's political preference or yearly income. Bob is honest but curious, and will try to extract the maximum amount of information about S from the data disclosed by Alice.

Instead of revealing X directly to Bob, Alice releases a new random variable, denoted by Y. This random variable is produced from X through a random mapping $p_{Y|X}$, called the *privacy-assuring mapping*. We assume that $p_{S,X}$ is fixed and known by both Alice and Bob, and $S \to X \to Y$. Alice's goal is to find a mapping $p_{Y|X}$ that minimizes $I(S;Y)$, while guaranteeing that the information disclosed about X is above a certain threshold t, i.e. $I(X;Y) \geq t$. We refer to the quantity $I(S;Y)$ as the *disclosed private information*, and $I(X;Y)$ as the *disclosed useful information*. When $I(S;Y) = 0$, we say that *perfect privacy* is achieved, i.e. Y does not reveal any information about S. We consider here the non-interactive, one-shot regime, where Alice discloses information once, and no additional information is released. We also assume that Bob knows the privacy-assuring mapping $p_{Y|X}$ chosen by Alice, and no side information is available to Bob about S besides the value Y.

In this paper, we present necessary and sufficient conditions for achieving perfect privacy while disclosing a non-trivial amount of useful information when both S and X have finite support $\mathcal{S}$ and $\mathcal{X}$, respectively. We prove that the smallest principal inertia component ([3], [4]) of $p_{S,X}$ plays a central role for achieving perfect privacy: If $|\mathcal{X}| \leq |\mathcal{S}|$, then perfect privacy is achievable with $I(X;Y) > 0$ if and only if the smallest principal inertia component of $p_{S,X}$ is 0. Since $I(S;Y) = 0$ (perfect privacy) if and only if $S \perp\!\!\!\perp Y$, this fundamental result holds for any privacy metric where statistical independence implies perfect privacy. We also provide an explicit lower bound for the amount of useful information that can be released while guaranteeing perfect privacy, and demonstrate how to construct $p_{Y|X}$ in order to achieve this bound.

In addition, we derive general bounds for the minimum amount of disclosed private information $I(S;Y)$ given that, on average, at least t bits of useful information is revealed to Bob, i.e. $I(X;Y) \geq t$. These bounds are sharp, and delimit the achievable privacy-utility region for the considered setting. Adopting an analysis related to the information bottleneck [5] and for characterizing the hypercontractivity coefficient in [6], [7], we determine the smallest achievable ratio between disclosed private and useful information, i.e. $\inf_{p_{Y|X}} I(S;Y)/I(X;Y)$. We prove that this value is upper-bounded by the smallest principal inertia component, and is zero if and only if the smallest principal inertia component is zero. In this case, we present an explicit construction of a privacy-assuring mapping that discloses a non-trivial amount of useful information while guaranteeing perfect privacy.

The rest of the paper is organized as follows. Section II introduces the privacy funnel and ancillary results. Section III relates the smallest achievable ratio between disclosed private and useful information with the principal inertia components. Section IV presents a necessary and sufficient condition for achieving perfect privacy in terms of the smallest principal inertia component and the cardinality of $\mathcal{X}$. Finally, Section V presents an explicit threshold for the amount of useful information that can be disclosed with perfect privacy, and investigates the case where S and X are vectors of i.i.d. random variables.

A. Related Work

Information-theoretic formulations for privacy have appeared in [8]–[12]. For an overview, we refer the reader to [1], [9] and the references therein. The privacy against statistical inference framework considered here was further studied in [13], [14]. The results presented in this paper are closely connected to the study of hypercontractivity coefficients and strong data processing results, such as in [6], [7], [15]–[17]. The principal inertia components were studied in [3], [4], [18]–[20], and we refer the readers to the references therein for additional related work. In particular, principal inertia component-based analysis were used in the context of security in [21], [22]. Extremal properties of privacy were also investigated in [23].

F. P. Calmon, A. Makhdoumi and M. Médard are with the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, MA (e-mail: {flavio, makhdoum, medard}@mit.edu).

B. Notation

We denote matrices by bold capitalized letters (e.g. $\mathbf{A}$) and vectors by bold lower-case letters (e.g. $\mathbf{x}$). The i-th component of a vector $\mathbf{x}$ is denoted by x_i, and for $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{x} = [x_1, x_2, \ldots, x_m]$. Random variables are denoted by upper-case letters (e.g. X and Y). We define $[n] \triangleq \{1, \ldots, n\}$. Sets are represented by calligraphic upper-case letters (e.g. $\mathcal{A}$).

Throughout the text we assume that S, X and Y are discrete random variables with finite support sets $\mathcal{S}$, $\mathcal{X}$ and $\mathcal{Y}$, respectively. We assume, without loss of generality, $\mathcal{S} = [|\mathcal{S}|]$, and equivalently for $\mathcal{X}$ and $\mathcal{Y}$. For random variable S and X, the joint distribution matrix of $\mathbf{P}_{S,X}$ is an $|\mathcal{S}| \times |\mathcal{X}|$ matrix with (i,j)-th entry equal to $p_{S,X}(i,j)$. For a random variable X with distribution p_X, we denote by $\mathbf{p}_X$ the vector with i-th entry equal to $p_X(i)$, $i \in \mathcal{X}$. In addition, $\mathbf{D}_X = \operatorname{diag}(\mathbf{p}_X)$ is a matrix with diagonal entries equal to $\mathbf{p}_X$, and all other entries equal to 0. The matrix $\mathbf{P}_{X|S} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{X}|}$ denotes the matrix with (i,j)-th entry equal to $p_{X|S}(j|i)$. Note that $\mathbf{P}_{S,X} = \mathbf{D}_S \mathbf{P}_{X|S}$.

II. The Privacy Funnel

We define next the privacy funnel function, which captures the smallest amount of disclosed private information for a given threshold on the amount of disclosed useful information. We then characterize properties of the privacy funnel function in the rest of this section.

Definition 1. For $0 \leq t \leq H(X)$ and a joint distribution $p_{S,X}$ over $\mathcal{S} \times \mathcal{X}$, we define the *privacy funnel function* $G_I(t, p_{S,X})$ as

$$G_I(t, p_{S,X}) \triangleq \inf \{I(S;Y) | I(X;Y) \geq t, S \to X \to Y\}, \tag{1}$$

where the infimum is over all mappings $p_{Y|X}$ such that $\mathcal{Y}$ is finite. For a fixed $p_{S,X}$ and $t \geq 0$, the set of pairs $\{(t, G_I(t, p_{S,X}))\}$ is called the *privacy region* of $p_{S,X}$.

A. Properties of the Privacy Funnel Function

We now enunciate a few useful properties of $G_I(t, p_{S,X})$ and the privacy region.

Lemma 1.

$$G_I(t, p_{S,X}) = \min_{p_{Y|X}} \{I(S;Y) | I(X;Y) \geq t, S \to X \to Y, |\mathcal{Y}| \leq |\mathcal{X}| + 2\}. \tag{2}$$

Proof: The proof is presented in the appendix. ∎

Lemma 2. *For a fixed $p_{S,X}$, the mapping $t \to \frac{G_I(t,p_{S,X})}{t}$ is non-decreasing.*

Proof: The prove is presented in the appendix. ∎

Lemma 3. *For $0 \leq t \leq H(X)$,*

$$\min\{t - H(X|S), 0\} \leq G_I(t, p_{S,X}) \leq \frac{t I(X;S)}{H(X)}. \tag{3}$$

Proof: The proof is presented in the appendix. ∎

Figure 1 illustrates the bounds from Lemma 3. The privacy region is contained withing the shaded area. The next two examples illustrate that both the upper bound (red line) and the lower bound (blue line) of the privacy region can be achieved for particular instances of $p_{S,X}$.

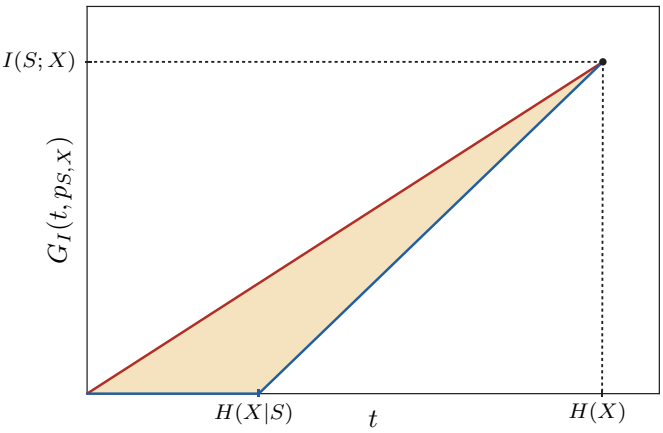

Fig. 1. For a fixed $p_{S,X}$, the privacy region is contained within the shaded area. The red and the blue lines correspond, respectively, to the upper and lower bounds presented in Lemma 3.

Example 1. Let $X = (S, W)$, where $W \perp\!\!\!\perp S$. Then by setting $Y = W$, we have $I(S;Y) = 0$ and $I(X;Y) = H(W) = H(X|S)$. Consequently, from Lemmas 2 and 3, $G_I(t, p_{S,X}) = 0$ for $t \in [0, H(X|S)]$. By letting $Y = W$ w.p. λ and $Y = (S, W)$ w.p. $1 - \lambda$ for $\lambda \in [0, 1]$, the lower-bound $G_I(t, p_{S,X}) = t - H(X|S)$ can be achieved for $H(X|S) = H(W) \leq t \leq H(X)$. Consequently, the lower bound in (3) is sharp.

Example 2. Now let $X = f(S)$. Then $I(X;S) = H(X)$ and

$$I(S;Y) = I(X;Y) - I(X;Y|S) = I(X;Y).$$

Consequently, $G_I(t, p_{S,X}) = t$, and the upper bound in (3) is sharp.

III. The Optimal Privacy-Utility Coefficient and the Principal Inertia Components

We now study the smallest possible ratio between disclosed private and useful information, defined next.

Definition 2. The *optimal privacy-utility coefficient* for a given distribution $p_{S,X}$ is given by

$$v^*(p_{S,X}) \triangleq \inf_{p_{Y|X}} \frac{I(S;Y)}{I(X;Y)}. \tag{4}$$

It follows directly from Lemma 2 that

$$v^*(p_{S,X}) = \lim_{t \to 0} \frac{G_I(t, p_{S,X})}{t}. \tag{5}$$

We show in Section IV that the value of $v^*(p_{S,X})$ is related to the smallest principal inertia component of $p_{S,X}$ (i.e. the smallest eigenvalue of the spectrum of the conditional expectation operator, defined below). We also prove that $v^*(p_{S,X}) = 0$ is a necessary and sufficient condition for achieving perfect privacy while disclosing a non-trivial amount of useful information. Before introducing these results, we present an alternative characterization of $v^*(p_{S,X})$ (Lemma 4), and introduce the principal inertia components (Definition 3) and an auxiliary result (Lemma 5).

Remark 1. The proof of Lemma 4 and Theorem 1 in this paper are closely related to [7]. We acknowledge that their proof techniques inspired some of the results presented here.

A. Divergence characterization of v^*

Lemma 4. *Let q_S denote the distribution of S when $p_{S|X}$ is fixed and $X \sim q_X$. Then*

$$v^*(p_{S,X}) = \inf_{q_X \neq p_X} \frac{D(q_S||p_S)}{D(q_X||p_X)}. \tag{6}$$

Proof: The proof is presented in the appendix. ∎

B. The Smallest Principal Inertia Component

The principal inertia components, defined below, provide a fine-grained decomposition of the dependency between two random variables. As shown in the next section, the smallest principal inertia component is of particular interest for privacy, and upper bounds the value of $v^*(p_{S,X})$. For an overview of the topic, we refer the reader to [3], [4]. We also encourage the reader to review briefly the notation section before perusing the next definition.

Definition 3. Let $\mathbf{Q} \triangleq \mathbf{D}_S^{-1/2}\mathbf{P}_{S,X}\mathbf{D}_X^{-1/2}$, and $m \triangleq \min\{|\mathcal{S}|, |\mathcal{X}|\}$. The largest singular value of $\mathbf{Q}$ is 1, and let $\sigma_1(\mathbf{Q}) \geq \sigma_2(\mathbf{Q}) \geq \cdots \geq \sigma_{m-1}(\mathbf{Q})$ denote the remaining $m-1$ singular values of $\mathbf{Q}$. Then $\sigma_1^2(\mathbf{Q}), \ldots, \sigma_{m-1}^2(\mathbf{Q})$ are the *principal inertia components* of $p_{S,X}$. In particular, $\sigma_{m-1}^2(\mathbf{Q})$ is the *smallest principal inertia component* of $p_{S,X}$. We define

$$\delta(p_{S,X}) \triangleq \begin{cases} \sigma_{m-1}^2(\mathbf{Q}) & \text{if } |\mathcal{X}| \leq |\mathcal{S}|, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Observe that $\delta(p_{S,X})$ is the smallest eigenvalue of $\mathbf{Q}^T\mathbf{Q}$.

The following lemma provides a useful characterization of $\delta(p_{S,X})$, related to the interpretation of the principal inertia components as the spectrum of the conditional expectation operator, discussed in [18] and [19]. A similar characterization can be obtained for the other principal inertia components by imposing appropriate orthogonality constraints on $f(X)$ below. For further details, we refer the reader to [3].

Lemma 5. *For a given $p_{S,X}$,*

$$\delta(p_{S,X}) = \min\left\{ \|\mathbb{E}\left[f(X)|S\right]\|_2^2 \middle| f : \mathcal{X} \to \mathbb{R}, \mathbb{E}\left[f(X)\right] = 0, \|f(X)\|_2 = 1 \right\}. \tag{8}$$

Proof: The proof is included in the appendix. ∎

IV. Information Disclosure with Perfect Privacy

If $v^*(p_{S,X}) = 0$, then it may be possible to disclose some information about X without revealing any information about S. However, since $G_I(0, p_{X,S}) = 0$, it is not immediately clear that $v^*(p_{S,X}) = 0$ implies that there exists t strictly bounded away from 0 such $G_I(t, p_{X,S}) = 0$. This would represent the ideal privacy setting, since, from Lemma 1, there would exist a privacy-assuring mapping that allows the disclosure of some non-negligible amount of useful information for $I(S;Y) = 0$. This, in turn, would mean that perfect privacy is achievable with non-negligible utility *regardless of the specific privacy metric used*, since S and Y would be independent.

In this section, we prove that if the optimal privacy-utility coefficient is 0, then there indeed exists a privacy-assuring mapping that allows the disclosure of a non-trivial amount of useful information while guaranteeing perfect privacy. We also show that the value of $\delta(p_{S,X})$ is closely related to $v^*(p_{S,X})$. This relationship is analogous to the one between the hypercontractivity coefficient s^*, defined in [6] and [24], and the maximal correlation ρ_m. In particular, as shown in the next two theorems, $v^*(p_{S,X}) \leq \delta(p_{S,X})$ and $v^*(p_{S,X}) = 0 \iff \delta(p_{S,X}) = 0$.

Theorem 1. *For any $p_{S,X}$ with finite support $\mathcal{S} \times \mathcal{X}$,*

$$v^*(p_{S,X}) \leq \delta(p_{S,X}). \tag{9}$$

Proof: Let $p_{S|X}$ be fixed, and define

$$g_\lambda(p_X) \triangleq H(S) - \lambda H(X),$$

where $H(S)$ and $H(X)$ are the entropy of S and X, respectively, when $(S, X) \sim p_{S|X}p_X$. For $0 < \epsilon \ll 1$, let

$$p_\epsilon(i) \triangleq p_X(i)(1 + \epsilon f(i))$$

be a perturbed version of p_X, where $\mathbb{E}\left[f(X)\right] = 0$ and, w.l.o.g., $\|f(X)\|_2 = 1$. The second derivative of $g_\lambda(p_\epsilon)$ at $\epsilon = 0$ is[1]

$$\begin{aligned} \left.\frac{\partial^2 g_\lambda(p_\epsilon)}{\partial \epsilon^2}\right|_{\epsilon=0} &= \log_2(e)\left(-\|\mathbb{E}\left[f(X)|S\right]\|_2^2 + \lambda\|f(X)\|_2^2\right) \\ &= \log_2(e)\left(-\|\mathbb{E}\left[f(X)|S\right]\|_2^2 + \lambda\right). \end{aligned} \tag{10}$$

Thus, from Lemma 5, if $\lambda \leq \delta(p_{S,X})$ then for any sufficiently small perturbation of p_X, (10) will be non-positive. Conversely, if $\lambda > \delta(p_{S,X})$, then we can find a perturbation $f(X)$ such that (10) is positive. Therefore, $g_\lambda(p_X)$ has a negative semi-definite Hessian if and only if $0 \leq \lambda \leq \delta(p_{S,X})$.

For any $S \to X \to Y$, we have $I(S;Y)/I(X;Y) \geq v^*(p_{S,X})$, and, consequently, for $0 \leq \lambda^\dagger \leq v^*(p_{S,X})$,

$$g_{\lambda^\dagger}(p_X) \geq H(S|Y) - \lambda^\dagger H(X|Y),$$

and $g_{\lambda^\dagger}(p_X)$ touches the upper-concave envelope of $g_{\lambda^\dagger}$ at p_X. Consequently, $g_{\lambda^\dagger}$ has a negative semi-definite Hessian at p_X and, from (10), $\lambda^\dagger \leq \delta(p_{S,X})$. Since this holds for any $0 \leq \lambda^\dagger \leq v^*(p_{S,X})$, we find $v^*(p_{S,X}) \leq \delta(p_{S,X})$. ∎

Remark 2. For a fixed $p_{S|X}$, the function $g_\lambda(p_X)$ is concave when $\lambda = 0$ and convex when $\lambda = 1$. A consequence of Theorem 1 is that the maximum λ for which $g_\lambda(p_X)$ has a negative Hessian at p_X is $\delta(p_{S,X})$. Furthermore, Lemma 4 implies that the value of λ for which $g_{\lambda_1}(p_X)$ touches it's lower concave envelope at p_X for all $\lambda_1 \geq \lambda$ is $v^*(p_{S,X})$. Therefore, both $\inf_{p_X} v^*(p_{S,X})$ and $\inf_{p_X} \delta(p_{S,X})$ equal the maximum value of λ such that the function $g_\lambda(p_X)$ is concave at all values of p_X. Therefore, we established that for a given $p_{S|X}$,

$$\inf_{p_X} v^*(p_{S,X}) = \inf_{p_X} \delta(p_{S,X}).$$

[1] This was observed in [7] and [24], and follows directly from $-\frac{\partial^2}{\partial \epsilon^2} a(1 + b\epsilon)\log_2 a(1 + b\epsilon) = -b^2 a \log_2(e)$.

The next theorem proves that $\delta(p_{S,X})$ can serve as a proxy for perfect privacy, since the optimal privacy-utility coefficient is 0 if and only if $\delta(p_{S,X})$ is also 0.

Theorem 2. *Let $p_{S,X}$ be such that $H(X) > 0$ and $\mathcal{S}$ and $\mathcal{X}$ are finite. Then*

$$v^*(p_{S,X}) = 0 \iff \delta(p_{S,X}) = 0. \tag{11}$$

Proof: Theorem 1 immediately gives $\delta(p_{S,X}) = 0 \Rightarrow v^*(p_{S,X}) = 0$. Let $v^*(p_{S,X}) = 0$. Then, since $D(q_X||p_X) \leq -\min_{i\in\mathcal{X}} \log_2 p_X(i)$ and $\mathcal{X}$ is finite, Lemma 4 implies that for any $\epsilon > 0$ there exists q_X and $0 < \delta \leq -\min_{i\in\mathcal{X}} \log_2 p_X(i)$ such that

$$D(q_X||p_X) \geq \delta > 0$$

and

$$D(q_S||p_S) < \epsilon.$$

We can then construct a sequence $q_X^1, q_X^2, q_X^3, \dots$ such that $q_X^i \neq p_X$, $D(q_S^k||p_S) \leq \epsilon_k$ and

$$\lim_{k\to\infty} \epsilon_k = 0.$$

Let $\mathbf{q}_S^k$ be a vector whose entries are $q_S^k(\cdot)$. Then, from Pinsker's inequality,

$$\epsilon_k \geq \frac{1}{2}\|\mathbf{q}_S^k - \mathbf{p}_S\|_1^2 \geq \frac{1}{2}\|\mathbf{q}_S^k - \mathbf{p}_S\|_2^2. \tag{12}$$

Defining $\mathbf{x}^k = \mathbf{q}_X^k - \mathbf{p}_X$, observe that $0 < \|\mathbf{x}^k\|_2^2 \leq 2$ and, from (12), $\|\mathbf{P}_{S|X}\mathbf{x}^k\|_2 \leq \sqrt{2\epsilon_k}$. Hence,

$$\lim_{k\to\infty} \frac{\|\mathbf{P}_{S|X}\mathbf{x}^k\|_2^2}{\|\mathbf{x}^k\|_2^2} = 0. \tag{13}$$

In addition, denoting $s_m \triangleq \min_{s\in\mathcal{S}} p_S(s)$ and $x_M \triangleq \min_{x\in\mathcal{X}} p_X(x)$, for each k we have

$$\begin{aligned}
\frac{\|\mathbf{P}_{S|X}\mathbf{x}^k\|_2^2}{\|\mathbf{x}^k\|_2^2} &\geq \min_{\|\mathbf{y}\|_2^2>0} \frac{\|\mathbf{P}_{S|X}\mathbf{y}\|_2^2}{\|\mathbf{y}\|_2^2} \\
&= \min_{\|\mathbf{y}\|_2^2>0} \frac{\|\mathbf{P}_{S,X}\mathbf{D}_X^{-1/2}\mathbf{y}\|_2^2}{\|\mathbf{D}_X^{1/2}\mathbf{y}\|_2^2} && (14)\\
&\geq \min_{\|\mathbf{y}\|_2^2>0} \frac{s_m\|\mathbf{D}_S^{-1/2}\mathbf{P}_{S,X}\mathbf{D}_X^{-1/2}\mathbf{y}\|_2^2}{x_M\|\mathbf{y}\|_2^2} && (15)\\
&= \frac{s_m}{x_M} \min_{\|\mathbf{y}\|_2^2>0} \frac{\|\mathbf{Q}\mathbf{y}\|_2^2}{\|\mathbf{y}\|_2^2} && (16)\\
&= \frac{s_m\delta(p_{S,X})}{x_M}. && (17)
\end{aligned}$$

In the derivation above, (14) follows from $\mathbf{D}_X$ being invertible (by definition), (15) is a direct consequence of $\|\mathbf{D}_S^{-1/2}\mathbf{y}\|_2^2 \leq s_m^{-1}\|\mathbf{y}\|_2^2$ and $\|\mathbf{D}_X^{11/2}\mathbf{y}\|_2^2 \leq x_M\|\mathbf{y}\|_2^2$ for any $\mathbf{y}$, and (16) and (17) follow from the definition of $\mathbf{Q}$ and $\delta(p_{S,X})$, respectively. Combining (17) with (13), it follows that $\delta(p_{S,X}) = 0$, proving the desired result. ∎

We are now ready to prove that a non-trivial amount of useful information can be disclosed without revealing any private information if and only if $v^*(p_{S,X}) = 0$ (or equivalently, $\delta(p_{S,X}) = 0$). This result follows naturally from Theorem 2, since $v^*(p_{S,X}) = 0$ implies that $\delta(p_{S,X}) = 0$, which means that the matrix $\mathbf{Q}$ and, consequently, $\mathbf{P}_{S|X}$, is either not full rank or has more columns than rows (i.e. $|\mathcal{X}| > |\mathcal{S}|$). This, in turn, can be exploited in order to find a mapping $p_{Y|X}$ such that Y reveals some information about X, but no information about S. This argument is made precise in the following theorem.

Theorem 3. *For a given $p_{S,X}$, there exists a privacy-assuring mapping $p_{Y|X}$ such that $S \to X \to Y$, $I(X;Y) > 0$ and $I(S;Y) = 0$ if and only if $\delta(p_{S,X}) = 0$ (equivalently $v^*(p_{S,X}) = 0$). In particular,*

$$\exists t_0 > 0 : G_I(t_0, p_{S,X}) = 0 \iff \delta(p_{S,X}) = 0. \tag{18}$$

Proof: The direct part of the theorem follows directly from the definition of $v^*(p_{S,X})$ and Theorem 2. Assume that $\delta(p_{S,X}) = 0$. Then, from Lemma 5, there exists $f : \mathcal{X} \to \mathbb{R}$ such that $\|f(X)\|_2 = 1$, $\mathbb{E}[f(X) = 0]$, and $\|\mathbb{E}[f(X)|S]\|_2 = 0$. Consequently, $\mathbb{E}[f(X)|S = s] = 0$ for all $s \in \mathcal{S}$.

Fix $\mathcal{Y} = [2]$, and, for $\epsilon > 0$ and ϵ appropriately small,

$$p_{Y|X}(y|x) = \begin{cases} \frac{1}{2} - \epsilon f(x), & y = 1, \\ \frac{1}{2} + \epsilon f(x), & y = 2. \end{cases}$$

Note that it is sufficient to choose $\epsilon = (2\max_{x\in\mathcal{X}}|f(X)|)^{-1}$, so ϵ is strictly bounded away from 0. In addition, $p_Y(1) = 1/2$. Therefore,

$$I(X;Y) = 1 - \sum_{x\in\mathcal{X}} p_X(x) h_b\left(\frac{1}{2} + \epsilon f(x)\right) > 0 \tag{19}$$

where $h_b(x) \triangleq -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function. Since $S \to X \to Y$,

$$\begin{aligned}
p_{Y|S}(y|s) &= \sum_{x\in\mathcal{X}} p_{Y|X}(y|x) p_{X|S}(x|s) \\
&= \sum_{x\in\mathcal{X}} \left(\frac{1}{2} + (-1)^y \epsilon f(x)\right) p_{X|S}(x|s) \\
&= \frac{1}{2} + (-1)^y \epsilon \mathbb{E}[f(X)|S = s] \\
&= 1/2,
\end{aligned}$$

and, consequently, S and Y are independent. Then $I(S;Y) = 0$, and the result follows. ∎

The previous result proves that if either $|\mathcal{X}| > |\mathcal{S}|$ or the smallest principal inertia component of $p_{S,X}$ is 0 (i.e. $\delta(p_{S,X}) = 0$), then it is possible to achieve perfect privacy while disclosing some useful information. In particular, the value of t_0 in (10) is lower-bounded by the expression in (19). We note that this result would not necessarily hold if $\mathcal{S}$ and $\mathcal{X}$ are not finite sets.

The proof of Theorem 3 holds for *any* measure of information J that satisfies $J(X;Y) = 0$ if and only if X and Y are independent, since it depends solely on the properties of $p_{S,X}$. Examples of J are maximal correlation or information metrics based on f-divergences [25]. This leads to the following result.

Corollary 1. *Let $p_{S,X}$ be given, and J be a non-negative measure of information (e.g. total variation or maximal correlation) such that for any two random variable A and B $J(A;B) = 0 \iff A \perp\!\!\!\perp B$. Then there exists $p_{Y|X}$ such that $S \to X \to Y$, $I(X;Y) > 0$ and $I(S;Y) = 0$ if and only if $\delta(p_{S,X}) = 0$.*

Proof: This is a direct consequence of Theorem 3, since, by assumption, $J(X;Y) = 0 \iff I(X;Y) = 0$ and $J(S;Y) = 0 \iff I(S;Y) = 0$. ∎

Remark 3. As long as privacy is measured in terms of statistical dependence (with perfect privacy implying statistical independence) and some utility can be derived when Y is not independent of X, then $\delta(p_{S,X})$ fully characterizes when perfect privacy is achievable with non-trivial utility.

V. Further Results

We present next an explicit lower bound for the largest amount of useful information that can be disclosed while guaranteeing perfect privacy. The result follows directly from the construction used in the proof of Theorem 3.

Corollary 2. *For fixed $p_{S,X}$, let*

$$\mathcal{F}_0 \triangleq \{f : \mathcal{X} \to \mathbb{R} | \mathbb{E}[f(X)] = 0, \|f(X)\|_2 = 1, \|\mathbb{E}[f(X)|S]\|_2 = 0\} \cup f_0,$$

where f_0 is the trivial function that maps $\mathcal{X}$ to $\{0\}$. Then $G_I(t, p_{S,X}) = 0$ for $t \in [0, t^*]$, where

$$t^* \geq 1 - \max_{f \in \mathcal{F}_0} \mathbb{E}\left[h_b\left(\frac{1}{2} + \frac{f(X)}{2\|f\|_\infty}\right)\right]. \tag{20}$$

Furthermore, the lower bound for t^* is sharp when $\delta(p_{S,X}) = 0$, i.e. there exists a $p_{S,X}$ such that $t^* > 0$ and $G_I(t, p_{S,X}) = 0$ if and only if $t \in [0, t^*]$.

Proof: The proof is included in the appendix. ∎

The previous bound for t^* can be loose, especially if $|\mathcal{X}|$ is large. In addition, the right-hand side of (20) can be made arbitrarily small by decreasing $\min_{x \in \mathcal{X}} p_X(x)$. Nevertheless, (20) is an explicit estimate of the amount of useful information that can be disclosed with perfect privacy.

When $S^n = (S_1, \ldots, S_n)$ and $X^n = (X_1, \ldots, X_n)$, where $(S_i, X_i) \sim p_{S,X}$ are i.i.d. random variables, the next proposition states that $\delta(p_{S^n,X^n}) = \delta(p_{S,X})^n$. Consequently, as long as $\delta(p_{S,X}) < 1$, it is possible to disclose a non-trivial amount of useful information while disclosing an arbitrarily small amount of private information by making n sufficiently large.

Proposition 1. *Let $S^n = (S_1, \ldots, S_n)$ and $X^n = (X_1, \ldots, X_n)$, where $(S_i, X_i) \sim p_{S,X}$ are i.i.d. random variables. Then*

$$v^*(p_{S^n,X^n}) \leq \delta(p_{S^n,X^n}) = \delta(p_{S,X})^n. \tag{21}$$

Proof: The result is a direct consequence of the tensorization property of the principal inertia components, presented in [3], [16], [20]. ∎

References

[1] F. P. Calmon and N. Fawaz, "Privacy against statistical inference," in *Proc. 50th Ann. Allerton Conf. Commun., Contr., and Comput.* IEEE, 2012, pp. 1401–1408.

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

[3] F. P. Calmon, M. Varia, M. Médard, M. Christiansen, K. Duffy, and S. Tessaro, "Bounds on inference," in *Proc. 51st Ann. Allerton Conf. Commun., Contr., and Comput.* IEEE, 2013, pp. 567–574.

[4] F. P. Calmon, M. Varia, and M. Médard, "An exploration of the role of principal inertia components in information theory," in *Information Theory Workshop (ITW), 2014 IEEE*, 2014, pp. 252–256.

[5] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proc. 37th Ann. Allerton Conf. Commun., Contr., and Comput.* IEEE, 1999, pp. 368–377.

[6] R. Ahlswede and P. Gács, "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, pp. 925–939, 1976.

[7] V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On maximal correlation, hypercontractivity, and the data processing inequality studied by erkip and cover," *arXiv:1304.6133[cs.IT]*, 2013.

[8] D. Rebollo-Monedero, J. Forné, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 11, pp. 1623 –1636, Nov. 2010.

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

[10] R. Tandon, L. Sankar, and H. Poor, "Discriminatory lossy source coding: Side information privacy," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5665–5677, Sep. 2013.

[11] I. S. Reed, "Information Theory and Privacy in Data Banks," in *Proceedings of the June 4-8, 1973, national computer conference and exposition*, ser. AFIPS '73. ACM, 1973, pp. 581–587.

[12] A. Evfimievski, J. Gehrke, and R. Srikant, "Limiting privacy breaches in privacy preserving data mining," in *Proceedings of the twenty-second ACM Symposium on Principles of Database Systems*, New York, NY, USA, 2003, pp. 211–222.

[13] S. Salamatian, A. Zhang, F. Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data," *IEEE GlobalSIP*, 2013.

[14] S. Salamatian, A. Zhang, F. P. Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "Managing your private and public data: Bringing down inference attacks against your privacy," *arXiv:1408.3698[cs.CR]*, 2014.

[15] Y. Polyanskiy and Y. Wu, "Dissipation of information in channels with input constraints," *arXiv:1405.3629 [cs, math]*, May 2014.

[16] Y. Polyanskiy, "Hypothesis testing via a comparator," in *Proc. 2012 IEEE Int. Symp. on Inf. Theory*, Jul. 2012, pp. 2206–2210.

[17] M. Raginsky, "Logarithmic Sobolev inequalities and strong data processing theorems for discrete channels," in *Proc. 2013 IEEE Int. Symp. on Inf. Theory*, Jul. 2013, pp. 419–423.

[18] A. Rényi, "On measures of dependence," *Acta mathematica hungarica*, vol. 10, no. 3, pp. 441–451, 1959.

[19] O. Sarmanov, "Maximum correlation coefficient (nonsymmetric case)," *Selected Translations in Mathematical Statistics and Probability*, vol. 2, pp. 207–210, 1962.

[20] W. Kang and S. Ulukus, "A new data processing inequality and its applications in distributed source and channel coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 56–69, 2011.

[21] C. T. Li and A. E. Gamal, "Maximal correlation secrecy," *arXiv:1412.5374 [cs, math]*, Dec. 2014.

[22] F. P. Calmon, M. Varia, and M. Médard, "On information-theoretic metrics for symmetric-key encryption and privacy," in *Proc. 52nd Annual Allerton Conference on Communication, Control, and Computing*, 2014.

[23] S. Chakraborty, N. Bitouze, M. Srivastava, and L. Dolecek, "Protecting data against unwanted inferences," in *2013 IEEE Information Theory Workshop (ITW)*, Sep. 2013, pp. 1–5.

[24] S. Kamath and V. Anantharam, "Non-interactive simulation of joint distributions: The hirschfeld-gebelein-rényi maximal correlation and the hypercontractivity ribbon," in *Proc. 50th Ann. Allerton Conf. Commun., Contr., and Comput.* IEEE, 2012, pp. 1057–1064.

[25] I. Csiszár and P. C. Shields, "Information theory and statistics: A tutorial," *Communications and Information Theory*, vol. 1, no. 4, pp. 417–528, 2004.

APPENDIX
ADDITIONAL PROOFS

A. Proof of Lemma 1

Let $p_{S,X}$ and $p_{Y|X}$ be given, with $S \to X \to Y$ and $|\mathcal{Y}| > |\mathcal{X}| + 2$. Denote by $\mathbf{w}_i$ the vector in the $|\mathcal{X}|$-simplex with entries $p_{X|Y}(\cdot|i)$. Furthermore, let $a_i \triangleq H(X) - H(X|Y=i)$, and $b_i \triangleq H(S) - H(S|Y=i)$. Therefore

$$\sum_{i=1}^{\mathcal{Y}} p_Y(i) \left[\mathbf{w}_i, a_i, b_i\right] = \left[\mathbf{p}_X, I(X;Y), I(S;Y)\right].$$

Since $\mathbf{w}_i$ belongs to the $|\mathcal{X}|$-simplex, the vector $[\mathbf{w}_i, a_i, b_i]$ is taken from a $|\mathcal{X}|+1$ dimensional space. Then, from Carathéodory's theorem, the point $[\mathbf{p}_X, I(X;Y), I(S;Y)]$ can also be achieved by at most $|\mathcal{X}|+2$ non-zero values of $p_Y(i)$. It follows directly that it is sufficient to consider $|\mathcal{Y}| \leq |\mathcal{X}|+2$ for the infimum (1).

The set of all mappings $p_{Y|X}$ for $|\mathcal{Y}| \leq |\mathcal{X}| + 2$ is compact, and both $p_{Y|X} \to I(S;Y)$ and $p_{Y|X} \to I(X;Y)$ are continuous and bounded when S, X and Y have finite support. Consequently, the infimum in (1) is attainable. ∎

B. Proof of Lemma 2

For $0 < t \leq H(X)$ and $p_{S,X}$ fixed, let $G_I(t, p_{S,X}) = \alpha$. From Lemma 1, there exists $p_{Y|X}$ that achieves $I(W;Y) = \alpha$ for $I(X;Y) \geq t$. Now consider $p_{\widetilde{Y}|X}$ where $\widetilde{\mathcal{Y}} = [|\mathcal{Y}| + 1]$ and, for $0 < \lambda \leq 1$,

$$p_{\widetilde{Y}|X}(y|x) = (1-\lambda)\mathbf{1}_{\{y=|\mathcal{Y}|+1\}} + \lambda \mathbf{1}_{\{y\neq|\mathcal{Y}|+1\}} p_{Y|X}(y|x).$$

Note that $\widetilde{Y}$ is an "erased" version of Y, with the erasure symbol being $|\mathcal{Y}|+1$. It follows directly that $I(S;\widetilde{Y}) = \lambda I(S;Y) = \lambda\alpha$, $I(X;\widetilde{Y}) = \lambda I(X;Y) \geq \lambda t$, and

$$\frac{G_I(\lambda t, P_{S,X})}{\lambda t} \leq \frac{\lambda I(S;Y)}{\lambda t} = \frac{G_I(t, p_{S,X})}{t}.$$

Since this holds for any $0 < \lambda \leq 1$, the result follows. ∎

C. Proof of Lemma 3

Observe that $G_I(H(X), p_{S,X}) = I(X;S)$, since $I(X;Y) = H(X)$ implies that $p_{Y|X}$ is a one-to-one mapping of X. The upper bound then follows directly from Lemma 2.

Clearly $G_I(t, p_{S,X}) \geq 0$. In addition, for any $p_{Y|X}$,

$$\begin{aligned} I(S;Y) &= I(X;Y) - I(X;Y|S) \\ &\geq I(X;Y) - H(X|S) \\ &\geq t - H(X|S), \end{aligned}$$

proving the lower bound. ∎

D. Proof of Lemma 4

For fixed $p_{Y|X}$ and $p_{S,X}$

$$\begin{aligned} \frac{I(S;Y)}{I(X;Y)} &= \frac{\sum_{y\in\mathcal{Y}} p_Y(y) D(p_{S|Y=y}||p_S)}{\sum_{y\in\mathcal{Y}} p_Y(y) D(p_{X|Y=y}||p_X)} \\ &\geq \min_{\substack{y\in\mathcal{Y}: \\ D(p_{X|Y=y}||p_X)>0}} \frac{D(p_{S|Y=y}||p_S)}{D(p_{X|Y=y}||p_X)} \\ &\geq \inf_{q_X \neq p_X} \frac{D(q_S||p_S)}{D(q_X||p_X)}. \end{aligned}$$

Now let d^* be the infimum in the right-hand side of (6), and q_X satisfy

$$\frac{D(q_Y||p_Y)}{D(q_X||p_X)} = d^* + \delta,$$

where $\delta > 0$. For $\epsilon > 0$ and sufficiently small, let $p_{Y|X}$ be such that $\mathcal{Y} = [2]$, $p_Y(1) = \epsilon$, $p_{X|Y}(x|1) = q_X(x)$ and

$$p_{X|Y}(x|2) = \frac{1}{1-\epsilon} p_X(x) - \frac{\epsilon}{1-\epsilon} q_X(x).$$

Since for any distribution r_X with support $\mathcal{X}$ we have $D\left((1-\epsilon)p_X + \epsilon r_X || p_X\right) = o(\epsilon)$, we find

$$\begin{aligned} I(S;Y) &= \epsilon D(p_{S|Y=1}||p_S) + (1-\epsilon) D(p_{S|Y=0}||p_S) \\ &= \epsilon D(q_S||p_S) + o(\epsilon), \end{aligned}$$

and equivalently, $I(X;Y) = \epsilon D(q_X||p_X) + o(\epsilon)$. Consequently,

$$\frac{I(S;Y)}{I(X;Y)} = \frac{\epsilon D(q_S||p_S) + o(\epsilon)}{\epsilon D(q_X||p_X) + o(\epsilon)} \to d^* + \delta,$$

where the limit is taken as $\epsilon \to 0$. Since this holds for any $\delta > 0$, then $v^*(p_{S,X}) \leq d^*$, proving the result. ∎

E. Proof of Lemma 5

Let $f : \mathcal{X} \to \mathbb{R}$, $\mathbb{E}\left[f(X)\right] = 0$ and $\|f(X)\|_2^2$=1, and $\mathbf{f} \in \mathbb{R}^{|\mathcal{X}|}$ be a vector with entries $f_i = f(i)$ for $i \in \mathcal{X}$. Observe that

$$\begin{aligned} \|\mathbb{E}\left[f(X)|S\right]\|_2^2 &= \sum_{s\in\mathcal{S}} p_S(s) \mathbb{E}\left[f(X)|S=s\right]^2 \\ &= \mathbf{f}^T \mathbf{P}_{X|S}^T \mathbf{D}_S \mathbf{P}_{X|S} \mathbf{f}^T \\ &= \mathbf{f}^T \mathbf{D}_X^{1/2} \mathbf{Q}^T \mathbf{Q} \mathbf{D}_X^{1/2} \mathbf{f} \\ &\geq \delta(p_{S,X}), \end{aligned}$$

where the last inequality follows by noting that $\mathbf{x} \triangleq \mathbf{f}^T \mathbf{D}_X^{1/2}$ satisfies $\|\mathbf{x}\|_2 = 1$ and that $\delta(p_{S,X})$ is the smallest eigenvalue of the positive semi-definite matrix $\mathbf{Q}^T\mathbf{Q}$, where $\mathbf{Q}$ is given in Definition 3. ∎

F. Proof of Corollary 2

If $\delta(p_{S,X}) = 0$, then the lower bound for t^* follows directly from the proof of Theorem 3 and, in particular, (18). If $\delta(p_{S,X}) > 0$, then $\mathcal{F}_0 = \{f_0\}$, and the lower bound (20) reduces to the trivial bound $t^* \geq 0$.

In order to prove that the lower bound is sharp, consider S being an unbiased bit, drawn from $\{1,2\}$, and X the result of sending S through an erasure channel with erasure probability $1/2$ and $\mathcal{X} = \{1,2,3\}$, with 3 playing the role of the erasure symbol. Let

$$f(x) \triangleq \begin{cases} 1, & x \in \{1,2\}, \\ -1 & x = 3. \end{cases}$$

Then $f \in \mathcal{F}_0$, $h_b\left(\frac{1}{2} + \frac{f(x)}{2\|f\|_\infty}\right) = 0$ for $x \in \mathcal{X}$ and $t^* = 1$. But, from Lemma 3, $t^* \leq H(X|S) = 1$. The result follows. ∎