Optimal Privacy-Utility Trade-off under a Rate Constraint

Sreejith Sreekumar and Deniz Gündüz
Dept. of Electrical and Electronic Engineering, Imperial College London
{s.sreekumar15,d.gunduz}@imperial.ac.uk

***Abstract*—We study the privacy-utility trade-off in data release under a rate constraint. An agent observes random variable X and reveals information U to the utility provider over a rate-constrained channel, such that $I(X;U) \leq R$, in return for utility $I(U;Y)$, where Y denotes a latent random variable correlated with X. While the objective is to maximize the utility, the agent also wants to protect a private information S, also correlated with X and Y from the utility provider. The trade-off between rate, utility and private information leakage is studied. This problem can be thought of as a generalization of both the information bottleneck and privacy funnel problems, reducing to either of the two problems in special cases. A necessary and sufficient condition for the existence of positive utility under zero private information leakage (or *perfect privacy*) is established. Subsequently, the problem of maximizing the utility subject to perfect privacy constraint is shown to be a linear program when the rate constraint is inactive. Also, the maximum value of the ratio of utility to infinitesimal private information leakage for an arbitrary rate constraint is obtained.**

I. Introduction

We consider the setup shown in Fig. 1 consisting of two parties, an agent and a remote utility provider. The agent observes a random variable (r.v.) X, and it acquires utility based on the information it shares with the utility provider. In particular, the utility acquired depends on the information the agent reveals to the utility provider about a latent r.v. Y, which may not be directly available to the agent. In particular, denoting the information the agent reveals to the utility provider by r.v. U, the utility acquired is measured by $I(U;Y)$. A rate constraint is imposed on the amount of information that can be shared with the utility provider; and therefore, the r.v. U must satisfy $I(X;U) \leq R$. We further assume that there is a latent private part of agent's information, represented by r.v. S, which the agent does not want to reveal to the utility provider. The amount of *private information leakage* (henceforth also referred to as *leakage*) to the utility provider is measured by $I(S;U)$. While X, Y and S can be arbitrarily correlated, since the revealed information U is generated observing only X, the Markov chain condition $(S,Y)-X-U$ must hold. This scenario occurs commonly in practice, e.g., a medical organization (agent) communicating patient data, e.g., test results, MRI, etc., to another laboratory (utility provider) over a rate-limited link. The goal is to keep as much information as possible about the attributes of data relevant for the requested analysis, while maintaining the privacy of some other latent features that can be inferred from this data, e.g., gender, age or identity of the patient.

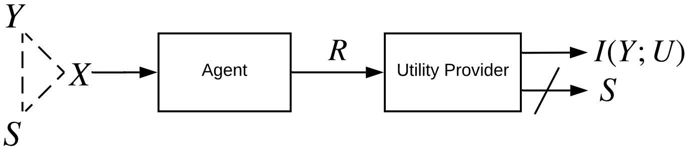

Fig. 1: System model

While it is clear that maximum privacy can be achieved by generating a r.v. U independent of X, this results in zero utility. More critically, revealing more information results in higher utility, but at the expense of leaking more private information, a dilemma of most modern information networks. We are interested in studying the fundamental trade-off between the rate, utility and leakage for a given joint distribution $\mathbf{p}_{SXY}$ of the r.v.'s S, X and Y, i.e., we are interested in characterizing the set $\mathcal{R}(\mathbf{p}_{SXY})$, defined as follows:

$$\mathcal{R}(\mathbf{p}_{SXY}) := \left\{ \begin{matrix} (R,\kappa,\Omega) \in \mathbb{R}_3^+ :\ \exists\ U \text{ s.t. } I(X;U) \leq R, \\ I(Y;U) \geq \kappa,\ I(S;U) \leq \Omega,\ (S,Y)-X-U \end{matrix} \right\}. \quad (1)$$

Note that $\mathcal{R}(\mathbf{p}_{SXY})$ depends on the joint distribution $\mathbf{p}_{SXY}$ only through the marginal distributions $\mathbf{p}_{SX}$ and $\mathbf{p}_{XY}$.

The problem of characterizing $\mathcal{R}(\mathbf{p}_{SXY})$ is equivalent to that of determining $J(\mathbf{p}_{SXY},R,\Omega)$ for all values of (R,Ω), defined as

$$J(\mathbf{p}_{SXY},R,\Omega) := \sup \left\{ \begin{matrix} I(Y;U) :\ \exists\ U \text{ s.t. } I(X;U) \leq R, \\ I(S;U) \leq \Omega \text{ and } (S,Y)-X-U \end{matrix} \right\}. \quad (2)$$

This is a generalization of two distinct problems that have received a lot of recent attention in information theory and machine learning, namely the *information bottleneck* [1] (also see [2] that studies an equivalent problem) and *privacy funnel* [3] problems. For a given joint distribution $\mathbf{p}_{XY}$ and fixed parameter R, the information bottleneck problem is defined as the optimization problem that maximizes utility $I(U;Y)$ under a given rate constraint $I(X;U) \leq R$ such that $Y-X-U$ form a Markov chain. On the other hand, for a given joint distribution $\mathbf{p}_{SX}$ and fixed parameter κ, the privacy funnel is defined as the optimization problem that minimizes the leakage

This work is supported in part by the European Research Council (ERC) through Starting Grant BEACON (agreement #677854).

$I(S;U)$ subject to a given constraint $I(X;U) \geq \kappa$ on the amount of information revealed about r.v. X, such that $S-X-U$ form a Markov chain. It is easy to see that the information bottleneck and privacy funnel problems are special cases of $\mathcal{R}(\mathbf{p}_{SXY})$ when $\Omega \geq H(S)$, and $Y = X$ and $R \geq H(X)$, respectively.

It is well known that the information bottleneck and privacy funnel problems are non-convex optimization problems, and hence, closed form solutions or efficient algorithms to obtain the global optimal solutions do not exist in general. In [1], an alternating minimization algorithm is proposed for the information bottleneck problem by introducing a distortion measure based on Kullback-Leibler divergence. The optimal privacy-utility coefficient, defined as

$$\nu^*(\mathbf{p}_{SX}) := \inf_{\substack{\mathbf{p}_{U|X}: \\ S-X-U}} \frac{I(S;U)}{I(X;U)} \tag{3}$$

is studied in [4]. Necessary and sufficient conditions under which $\nu^*(\mathbf{p}_{SX}) = 0$, referred to as *perfect privacy* therein, is established, based on the smallest singular value of a normalized probability distribution matrix. In [5], an equivalent form of this problem is studied. Defining utility as $I(X;U)$, the problem of utility maximization is considered with a constraint on the privacy leakage, $I(S;U) \leq \epsilon$. This can be written as

$$g_\epsilon(\mathbf{p}_{SX}) := \sup_{\substack{\mathbf{p}_{U|X}: \\ S-X-U, \\ I(S;U)\leq\epsilon}} I(X;U). \tag{4}$$

In [5], perfect privacy is said to be *feasible* if $g_0(\mathbf{p}_{SX}) > 0$, and a necessary and sufficient condition on the joint distribution $\mathbf{p}_{SX}$ is established, such that perfect privacy is feasible.

The problem of characterizing $\mathcal{R}(\mathbf{p}_{SXY})$ in (1) can be considered as a generalization of computing $g_\epsilon(\mathbf{p}_{SX})$ in (4), in which, there is an additional r.v. Y, which determines the utility to be maximized, $I(Y;U)$, such that $Y - X - U$ form a Markov chain, and an extra rate constraint is also introduced, i.e., $I(X;U) \leq R$. We also briefly mention that while mutual information leakage is the privacy measure considered here, several other measures of privacy exist in the literature, e.g., *differential privacy* [6], *k-anonymity* [7], *total variation distance* [8], etc.. In the rest of the paper, we focus on the equivalent optimization problem in the right hand side (R.H.S.) of (2). Before proceeding, we introduce the notations required to present our results.

A. Notations

Random variables, their realizations and their supports are denoted by upper case, lower case and calligraphic letters, e.g., X, x and $\mathcal{X}$, respectively. Column and row vectors are denoted by lower case bold letters, e.g., $\mathbf{a}$ and $\mathbf{a}^T$, respectively. The probability mass function (p.m.f.) of a discrete r.v. X (with finite support $\mathcal{X}$) is treated as a column vector of length $|\mathcal{X}|$ and denoted by $\mathbf{p}_X$. For r.v.'s X and Y, the conditional probability distribution $\mathbf{p}_{Y|X}$ is represented by a probability matrix of dimension $|\mathcal{Y}| \times |\mathcal{X}|$, i.e., the sum of the entries in each column is equal to 1. Singular values of a matrix A of size $m \times n$ are denoted by $\sigma_i(A)$ (or simply as σ_i), $1 \leq i \leq \min(m,n)$.The set of all p.m.f.'s defined over $\mathcal{X}$, represented as probability vectors of length $|\mathcal{X}|$, is denoted by $\mathcal{P}_\mathcal{X}$. The i^{th} component of a vector $\mathbf{x}$ is denoted by $\mathbf{x}(i)$. $\mathbf{1}_{|\mathcal{X}|}$ denotes a column vector of length $|\mathcal{X}|$ with all the components equal to 1.

II. Perfect privacy

We assume that X, Y and S have finite support, i.e., $|\mathcal{X}|$, $|\mathcal{Y}|$, $|\mathcal{S}| < \infty$. The next proposition shows that the supremum in (2) can be replaced by a maximum.

Proposition 1. *In* (2)*, it suffices to consider auxiliary r.v.'s U such that $|\mathcal{U}| \leq |\mathcal{X}| + 2$. Also, the supremum in* (2) *can be replaced by a maximum.*

The proof follows from standard arguments based on the Fenchel-Eggleston-Carathéodory's Theorem [9]. By Proposition 1, $J(\mathbf{p}_{SXY}, R, \Omega)$ can be written as

$$\begin{aligned} &J(\mathbf{p}_{SXY}, R, \Omega) \\ &\quad = \sup\{I(Y;U) \text{ s.t. } U \in \mathcal{G}(\mathbf{p}_{SXY}, R, \Omega)\}, \end{aligned} \tag{5}$$

where,

$$\begin{aligned} &\mathcal{G}(\mathbf{p}_{SXY}, R, \Omega) \\ &\quad := \begin{Bmatrix} U: \ I(X;U) \leq R, \ I(S;U) \leq \Omega, \\ (S,Y) - X - U, \ |\mathcal{U}| \leq |\mathcal{X}| + 2 \end{Bmatrix}. \end{aligned} \tag{6}$$

Let

$$\begin{aligned} J(\mathbf{p}_{SXY}, R) &:= J(\mathbf{p}_{SXY}, R, 0) \\ &:= \max\{I(Y;U): \ U \in \mathcal{G}(\mathbf{p}_{SXY}, R, 0)\} \end{aligned} \tag{7}$$

denote the maximum utility achievable under the given rate constraint R and perfect privacy constraint $\Omega = 0$. In the sequel, we will establish the conditions on the joint probability distribution $\mathbf{p}_{SXY}$ under which $J(\mathbf{p}_{SXY}, R) > 0$. Note that $J(\mathbf{p}_{SXY}, 0) = 0$ since $0 \leq I(Y;U) \leq I(U;X) = 0$; and hence, we may assume $R > 0$. Since $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{S}$ are finite discrete sets by assumption, without loss of generality (w.l.o.g.), we may assume that $\mathbf{p}_X(x) > 0$, $\mathbf{p}_Y(y) > 0$ and $\mathbf{p}_S(s) > 0$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and $s \in \mathcal{S}$, respectively (since otherwise, we may modify the support by discarding those points with zero probability without affecting anything). Due to Proposition 1, we may assume w.l.o.g. that $\mathcal{U} = \{1, \ldots, |\mathcal{U}|\}$, $|\mathcal{U}| \leq |\mathcal{X}| + 2$, and using the same arguments as above, that, in the supremum in (7), $\mathbf{p}_U(u) > 0, \ \forall \ u \in \mathcal{U}$. Let $\mathcal{P}_U^+ := \{\mathbf{p}_U \in \mathcal{P}_U : \mathbf{p}_U(u) > 0, \ \forall u \in \mathcal{U}\}$ and

$$\begin{aligned} &\mathcal{G}'(\mathbf{p}_{SXY}, R, \mathbf{p}_U) \\ &:= \Big\{\mathbf{p}_{X|U} : \textstyle\sum_{u\in\mathcal{U}} \mathbf{p}_U(u)\mathbf{p}_{X|U}(x|u) = \mathbf{p}_X(x), \ \forall \ x \in \mathcal{X}, \\ &\qquad (S,Y) - X - U, \ I(S;U) = 0, \ I(X;U) \leq R\Big\}. \end{aligned} \tag{8}$$

Note that we can write

$$J(\mathbf{p}_{SXY}, R) = \max_{\substack{\mathbf{p}_U \in \mathcal{P}_U^+, \\ |\mathcal{U}| \leq |\mathcal{X}|+2}} \ \max_{\mathbf{p}_{X|U} \in \mathcal{G}'(\mathbf{p}_{SXY}, R, \mathbf{p}_U)} I(Y;U). \tag{9}$$

Let $\mathcal{N}(A)$ and $\mathcal{N}(A)^{\perp}$ denote the null space of a matrix A and its orthogonal complement, respectively. We define

$$\begin{aligned}&\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)\\ &:= \Big\{\mathbf{p}_{X|U}=[\mathbf{x}_1\cdots\mathbf{x}_{|\mathcal{U}|}] : \mathbf{x}_i\geq \mathbf{0},\ \mathbf{x}_i=\mathbf{p}_X+\mathbf{a}_i,\\ &\mathbf{a}_i\in\mathcal{N}(\mathbf{p}_{S|X}),\ 1\leq i\leq|\mathcal{U}|, \sum\nolimits_{u\in\mathcal{U}}\mathbf{p}_U(u)\mathbf{p}_{X|U}(x|u)\\ &=\mathbf{p}_X(x),\ \forall\ x\in\mathcal{X},\ I(X;U)\leq R,\ (S,Y)-X-U\Big\}.\end{aligned}$$

The next result, which is similar in spirit to Proposition 2 in [5], provides a necessary and sufficient condition under which $J(\mathbf{p}_{SXY},R)>0$.

Proposition 2. *(i)* *$J(\mathbf{p}_{SXY},R)>0$ if and only if $\dim\left(\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp}\right)\geq 1$. Moreover, it is sufficient to consider U such that $|\mathcal{U}|\leq|\mathcal{X}|+1$ and $\mathcal{G}'(\mathbf{p}_{SXY},R,\mathbf{p}_U)=\mathcal{J}^*(\mathbf{p}_{SXY},R,\mathbf{p}_U)$ for the maximization in* (9)*, where $\mathcal{J}^*(\mathbf{p}_{SXY},R,\mathbf{p}_U)$ denotes the extreme points of the convex set $\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)$.*

(ii) *$J(\mathbf{p}_{SXY},R)$ is a non-decreasing function of R.*

Proof: To begin with, note that $\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp}$ is a linear space since it is the intersection of two linear spaces, and hence, its dimension is well-defined. We first prove that $J(\mathbf{p}_{SXY},R)>0$ if $\dim\left(\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp}\right)\geq 1$. By definition, $I(S;U)=0$, or equivalently, $S\perp U$, if and only if

$$\mathbf{p}_{S|U}(s|u)=\mathbf{p}_S(s),\forall\ u\in\mathcal{U},s\in\mathcal{S}. \tag{10}$$

From the Markov chain $S-X-U$, we have $\mathbf{p}_{S|U}=\mathbf{p}_{S|X}\mathbf{p}_{X|U}$. Writing $\mathbf{p}_S=\mathbf{p}_{S|X}\mathbf{p}_X$, it follows that (10) holds if and only if,

$$\mathbf{p}_{S|X}(\mathbf{p}_{X|U=u}-\mathbf{p}_X)=0,\ \forall\ u\in\mathcal{U}. \tag{11}$$

Similarly, from $Y-X-U$, it follows that $I(V;W)>0$ if and only if

$$\mathbf{p}_{Y|X}(\mathbf{p}_{X|U=u}-\mathbf{p}_X)\neq 0 \text{ for some } u\in\mathcal{U}. \tag{12}$$

Also, note that for $\mathbf{a}\in\mathcal{N}(\mathbf{p}_{S|X})$,

$$\mathbf{1}^T_{|\mathcal{X}|}\mathbf{a}=\mathbf{1}^T_{|\mathcal{S}|}\mathbf{p}_{S|X}\mathbf{a}=0. \tag{13}$$

Hence, for sufficiently small $\theta\in\mathbb{R}$ and $\mathbf{a}\in\mathcal{N}(\mathbf{p}_{S|X})$, $\mathbf{p}_X+\theta\mathbf{a}$ is a probability vector. Now, suppose that $\dim(\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp})\neq\varnothing$. This implies that there exists $\mathbf{a}\neq\mathbf{0}$, $\mathbf{a}\in\mathcal{N}(\mathbf{p}_{S|X})\cap\ \mathcal{N}(\mathbf{p}_{Y|X})^{\perp}$. We will now show that there exists an auxiliary r.v. U^* and conditional probability distribution $\mathbf{p}_{X|U^*}$ such that $I(Y;U^*)>0$, $I(S;U^*)=0$, $I(X;U^*)\leq R$, and $(S,Y)-X-U^*$ form Markov chains. Let $U^*\sim$ Bernoulli (0.5). For arbitrary $\theta>0$, let

$$\begin{aligned}&\mathbf{p}_{X|U^*=0}:=\mathbf{p}_X+\theta\mathbf{a},\\ &\mathbf{p}_{X|U^*=1}:=\mathbf{p}_X-\theta\mathbf{a},\\ &\mathbf{p}_{S|U^*=u}:=\sum\nolimits_{x\in\mathcal{X}}\mathbf{p}_{X|U^*=u}(x|u)\mathbf{p}_{S|X}(s|x),\ u\in\{0,1\},\\ &\mathbf{p}_{Y|U^*=u}:=\sum\nolimits_{x\in\mathcal{X}}\mathbf{p}_{X|U^*=u}(x|u)\mathbf{p}_{Y|X}(y|x),\ u\in\{0,1\}.\end{aligned}$$

Due to (13), $\mathbf{p}_{X|U^*=0}$ and $\mathbf{p}_{X|U^*=1}$ are probability vectors for $\theta>0$ sufficiently small. Also, note that

$$\sum\nolimits_{u\in\mathcal{X}^*}\mathbf{p}_{U^*}(u)\mathbf{p}_{X|U^*}(x|u)=\mathbf{p}_X(x),\ \forall\ x\in\mathcal{X}.$$

Thus, the marginal distribution of X is preserved. From (11) and (12), it follows that $I(S;U^*)=0$ and $I(Y;U^*)>0$. Also,

$$I(X;U^*)=\frac{1}{2}\left(D(\mathbf{p}_X+\theta\mathbf{a}||\mathbf{p}_X)+D(\mathbf{p}_X-\theta\mathbf{a}||\mathbf{p}_X)\right).$$

From the continuity of $D(\mathbf{p}_X||\mathbf{q}_X)$ in $\mathbf{p}_X$ (for $\mathbf{p}_X<<\mathbf{q}_X$) for a fixed $\mathbf{q}_X$, it follows that there exists $\theta^*>0$ (sufficiently small) such that

$$D(\mathbf{p}_X+\theta^*\mathbf{a}||\mathbf{p}_X)\leq R, \tag{14}$$

$$D(\mathbf{p}_X-\theta^*\mathbf{a}||\mathbf{p}_X)\leq R. \tag{15}$$

For such a θ^*, it follows that $I(X;U^*)\leq R$. Hence, $J(\mathbf{p}_{SXY},R)>0$.

To prove the opposite implication, notice from (11) and (12) that $J(\mathbf{p}_{SXY},R)>0$ only if there exists an auxiliary r.v. U and $\mathbf{a}\neq\mathbf{0}$, $\mathbf{a}\in\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp}$ such that for some $u\in\mathcal{U}$, $\mathbf{p}_{X|U=u}=\mathbf{p}_X+\mathbf{a}$. This implies that $\dim(\mathcal{N}(\mathbf{p}_{S|X})\cap\ \mathcal{N}(\mathbf{p}_{Y|X})^{\perp})\geq 1$. Thus, $\dim\left(\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp}\right)\geq 1$ is both necessary and sufficient condition for $J(\mathbf{p}_{SXY},R)>0$, and we can write

$$\begin{aligned}&J(\mathbf{p}_{SXY},R)\\ &\quad=\max_{\substack{\mathbf{p}_U\in\mathcal{P}_U^+,\\ |\mathcal{U}|\leq|\mathcal{X}|+2}}\ \max_{\mathbf{p}_{X|U}\in\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)} I(Y;U).\end{aligned} \tag{16}$$

Now, we prove the second statement of (i). From Proposition 1, it follows that $|\mathcal{U}|\leq|\mathcal{X}|+1$ is sufficient since for $\mathbf{p}_{X|U}\in\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)$, $I(S;U)=0$ is automatically satisfied. Also, note that $\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)$ is a compact and convex set since $|\mathcal{U}|$ is bounded, and that $I(Y;U)$ is a convex function of $\mathbf{p}_{U|X}$. It is well known that the maximum of a convex function over a compact convex set is achieved at the extreme points of the set. Thus, it is sufficient to consider $\mathcal{J}^*(\mathbf{p}_{SXY},R,\mathbf{p}_U)$ in place of $\mathcal{J}(\mathbf{p}_{SXY},R,\mathbf{p}_U)$ in (16). This completes the proof of (i). Part (ii) follows from the facts that $\mathcal{J}^*(\mathbf{p}_{SXY},R,\mathbf{p}_U)\subseteq\mathcal{J}^*(\mathbf{p}_{SXY},R',\mathbf{p}_U)$ for $R\leq R'$ and the supremum is non-decreasing with respect to set inclusion. This completes the proof. ■

While Proposition 2 provides a necessary and sufficient condition under which $J(\mathbf{p}_{SXY},R)>0$, it would be interesting to explicitly compute $J(\mathbf{p}_{SXY},R)$. However, this is a non-convex optimization problem and hence, a closed form solution for the global optima does not exist in general. In [5], it was shown that computing $g_0(\mathbf{p}_{SX})$ is a linear program. It can be shown similarly that computing $J(\mathbf{p}_{SXY},R)$ for $R\geq H(X)$ is also a linear program. Note that when $R\geq H(X)$, the constraint $I(X;U)\leq R$ is inactive. Hence, $J(\mathbf{p}_{SXY},R)=J(\mathbf{p}_{SXY},H(X))$ for all $R\geq H(X)$. We will assume that $\dim(\mathcal{N}(\mathbf{p}_{S|X})\cap\mathcal{N}(\mathbf{p}_{Y|X})^{\perp})\geq 1$, since otherwise $J(\mathbf{p}_{SXY},R)=0$. Let the singular value decomposition of $\mathbf{p}_{S|X}$ be given by $\mathbf{p}_{S|X}=B\Sigma C^T$, where B (resp. C) is the orthogonal matrix of dimension $|\mathcal{S}| \times |\mathcal{S}|$ (resp. $|\mathcal{X}| \times |\mathcal{X}|$), whose columns consist of the left (resp. right) singular vectors of $\mathbf{p}_{S|X}$ and Σ is a diagonal matrix whose entries are the singular values of $\mathbf{p}_{S|X}$. Assuming w.l.o.g. that the diagonal entries of Σ are arranged in decreasing order of magnitude, the null space of $\mathbf{p}_{S|X}$ is given by

$$\mathcal{N}(\mathbf{p}_{S|X}) = \text{Span}(\mathbf{c}_l, \cdots, \mathbf{c}_{|\mathcal{X}|}),$$

for some $l \leq |\mathcal{X}|$, where $\mathbf{c}_i, 1 \leq i \leq |\mathcal{X}|$ denote the column vectors of C. Let

$$E := E(\mathbf{p}_{S|X}) := [\mathbf{c}_1 \cdots \mathbf{c}_{l-1}]^T.$$

Then the constraint $\{\mathbf{x}_i \geq \mathbf{0},\ \mathbf{x}_i = \mathbf{p}_X + \mathbf{a}_i,\ \mathbf{a}_i \in \mathcal{N}(\mathbf{p}_{S|X}),\ 1 \leq i \leq |\mathcal{U}|\}$ is equivalent to $\{\mathbf{x}_i \geq \mathbf{0},\ E\mathbf{x}_i = E\mathbf{p}_X,\ 1 \leq i \leq |\mathcal{U}|\}$. Let

$$\mathcal{E}(E) := \{\mathbf{x} :\ \mathbf{x} \geq \mathbf{0},\ E\mathbf{x} = E\mathbf{p}_X\}.$$

$J(\mathbf{p}_{SXY}, H(X))$ can be written alternatively as

$$\begin{aligned} &J(\mathbf{p}_{SXY}, H(X)) \\ &\quad= \max_{\substack{\mathbf{p}_U \in \mathcal{P}_U^+,\\ |\mathcal{U}| \leq |\mathcal{X}|+1}} \ \max_{\mathbf{p}_{X|U} \in \mathcal{J}(\mathbf{p}_{SXY}, H(X), \mathbf{p}_U)} I(Y;U), \end{aligned} \tag{17}$$

where,

$$\begin{aligned} &\mathcal{J}(\mathbf{p}_{SXY}, H(X), \mathbf{p}_U) \\ &:= \Big\{ \mathbf{p}_{X|U} = [\mathbf{x}_1 \cdots \mathbf{x}_{|\mathcal{U}|}] : \mathbf{x}_i \in \mathcal{E}(E),\ 1 \leq i \leq |\mathcal{U}|, \\ &\qquad \sum_{u \in \mathcal{U}} \mathbf{p}_U(u)\mathbf{p}_{X|U}(x|u) = \mathbf{p}_X(x),\ \forall\ x \in \mathcal{X}, \\ &\qquad (S,Y) - X - U \Big\}. \end{aligned} \tag{18}$$

Let $\mathcal{E}^*(E)$ denote the extreme points of the convex set $\mathcal{E}(E)$.

Proposition 3. *It is sufficient to restrict to $\mathcal{E}^*(E)$ in place of $\mathcal{E}(E)$ in* (18).

Thanks to Proposition 3, we can write

$$\begin{aligned} &J(\mathbf{p}_{SXY}, H(X)) \\ &\quad= \max_{\substack{\mathbf{p}_U \in \mathcal{P}_U^+,\\ |\mathcal{U}| \leq |\mathcal{X}|+1}} \ \max_{\mathbf{p}_{X|U} \in \mathcal{J}^*(\mathbf{p}_{SXY}, H(X), \mathbf{p}_U)} I(Y;U), \end{aligned} \tag{19}$$

where,

$$\begin{aligned} &\mathcal{J}^*(\mathbf{p}_{SXY}, H(X), \mathbf{p}_U) \\ &:= \Big\{ \mathbf{p}_{X|U} = [\mathbf{x}_1 \cdots \mathbf{x}_{|\mathcal{U}|}] : \mathbf{x}_i \in \mathcal{E}^*(E),\ 1 \leq i \leq |\mathcal{U}|, \\ &\qquad \sum_{u \in \mathcal{U}} \mathbf{p}_U(u)\mathbf{p}_{X|U}(x|u) = \mathbf{p}_X(x),\ \forall\ x \in \mathcal{X}, \\ &\qquad (S,Y) - X - U \Big\}. \end{aligned}$$

Since $\mathcal{E}(E)$ is a set defined by linear constraints, it is well known that $\mathcal{E}^*(E)$ consists of the basic feasible solutions of the system of linear equations defining $\mathcal{E}(E)$ [10]. Since the rank of E is $l-1$, the number of elements in $\mathcal{E}^*(E)$ is atmost $\binom{|\mathcal{X}|}{l-1}$. The problem in (19) can be solved in two steps as follows.

1) First, the matrix $\mathbf{p}_{X|U}$ is constructed by choosing columns from the set $\mathcal{E}^*(E)$, i.e., for some $\mathbf{x}_i \in \mathcal{E}^*(E)$, $1 \leq i \leq |\mathcal{U}|$, $\mathbf{p}_{X|U} = [\mathbf{x}_1 \cdots \mathbf{x}_{|\mathcal{U}|}]$.
2) Note that given $\mathbf{p}_{X|U}$ as constructed in step 1, $\sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i)\mathbf{p}_{X|U}(x|i)$, $H(Y|U)$ and $H(X|U)$ are all linear functions of the weights $\mathbf{p}_U(i), i \in [1 : |\mathcal{U}|]$ that satisfies $\sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i) = 1$. More specifically, we have

$$H(Y|U) = \sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i) f_1(\mathbf{p}_{Y|X}\mathbf{x}_i),$$

$$H(X|U) = \sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i) f_1(\mathbf{x}_i),$$

where

$$f_1(\mathbf{x}) := -\sum_{j=1}^{|\mathcal{X}|} \mathbf{x}(j)\log(\mathbf{x}(j)).$$

Hence, finding the optimal $\mathbf{p}_U$ that maximizes (17) is equivalent to finding the weights $\mathbf{p}_U(i)$, $1 \leq i \leq |\mathcal{U}|$, that maximizes $H(Y) - \sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i) f_1(\mathbf{x}_i)$ subject to

$$\sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i) = 1 \text{ and } \sum_{i=1}^{|\mathcal{U}|} \mathbf{p}_U(i)\mathbf{x}_i = \mathbf{p}_X.$$

This is a simple linear program and can be solved efficiently. To illustrate the algorithm above, consider the following example.

Example 1. *Let $\mathcal{S} = \mathcal{X} = \{0,1,2,3\}$, $\mathcal{Y} = \{0,1\}$, and let $\mathbf{p}_{SXY}$ be defined by*

$$\mathbf{p}_{SX} = 0.125 \cdot \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{p}_{Y|X} = \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \\ 0.3 & 0.7 \\ 0.1 & 0.9 \end{bmatrix},$$

and $S - X - Y$. The singular value decomposition of $\mathbf{p}_{S|X}$ yields $\mathbf{p}_{S|X} = B\Sigma C^T$, where $\sigma_1 = \sigma_2 = 1$, $\sigma_3 = \sigma_4 = 0$, $B = [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3\ \mathbf{b}_4]$, $C = [\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3\ \mathbf{c}_4]$, $\mathbf{b}_1 = \mathbf{c}_1 = (-0.7071\ -0.7071\ 0\ 0)^T$, $\mathbf{b}_2 = \mathbf{c}_2 = (0\ 0\ -0.7071 - 0.7071)^T$, $\mathbf{b}_3 = -\mathbf{c}_3 = (-0.7071\ 0.7071\ 0\ 0)^T$ and $\mathbf{b}_4 = -\mathbf{c}_4 = (0\ 0\ -0.7071\ 0.7071)^T$. Then,

$$\mathcal{N}(\mathbf{p}_{S|X}) = \textit{Span}\,(\mathbf{c}_3, \mathbf{c}_4)$$

and

$$E = [\mathbf{c}_1\ \mathbf{c}_2]^T = \begin{bmatrix} -0.7071 & -0.7071 & 0 & 0 \\ 0 & 0 & -0.7071 & -0.7071 \end{bmatrix}.$$

It is easy to see that the rank of E is 2, and there are 4 possible ways (at most $\binom{4}{2} = 6$ in general) of choosing two linearly independent columns of E, which correspond to the basic feasible solutions of the system of linear equations that define $\mathcal{E}(E)$. Denoting by $\mathbf{e}_j$, the j^{th} column of E, these are $\{\mathbf{e}_1, \mathbf{e}_3\}$, $\{\mathbf{e}_1, \mathbf{e}_4\}$, $\{\mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}_2, \mathbf{e}_4\}$, and lead to the basic feasible solutions $[0.5\ 0\ 0.5\ 0]^T$, $[0.5\ 0\ 0\ 0.5]^T$,

$[0\ 0.5\ 0.5\ 0]^T$ *and* $[0\ 0.5\ 0\ 0.5]^T$*, respectively. Thus, finding the maximum utility under perfect privacy amounts to solving the following optimization problem:*

$$\max_{\mathbf{p}_U \in \mathcal{P}_U^+} \quad 0.8113 - [0.9341\ 0.8113\ 0.8113\ 0.6098]\ \mathbf{p}_U \qquad (20)$$

$$\textit{s.t.} \quad \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 1 & 1 & 1 & 1 \end{bmatrix} \mathbf{p}_U = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \\ 1 \end{bmatrix}. \qquad (21)$$

Solving (20) *yields the solution* $\mathbf{p}_U^* = [0.5\ 0\ 0\ 0.5]^T$ *with the maximum utility* $I(Y;U^*) = 0.0393$*.*

III. Privacy-Utility Trade-off

In Proposition 2, we characterized the conditions under which a positive utility is achievable along with perfect privacy. Next, we study the behaviour of the privacy-utility trade-off when a small amount of leakage is allowed. More specifically, we provide a characterization of the maximum value of the slope of the utility-leakage trade-off $\frac{J(\mathbf{p}_{SXY}, R, \Omega)}{\Omega}$, when $\Omega > 0$ is small, i.e.,

$$\Lambda(\mathbf{p}_{SXY}, R) := \lim_{\Omega \to 0^+} \frac{J(\mathbf{p}_{SXY}, R, \Omega)}{\Omega}. \qquad (22)$$

When $J(\mathbf{p}_{SXY}, R, 0)$ can be efficiently computed, as discussed above, $\Lambda(\mathbf{p}_{SXY}, R)$ together with $J(\mathbf{p}_{SXY}, R, 0)$ (which is the vertical intercept in the utility-leakage graph) provides a good linear approximation to $J(\mathbf{p}_{SXY}, R, \Omega)$ for small positive values of Ω.

When $I(Y;X) = 0$, then $I(Y;U) = 0$ for all $Y - X - U$, hence $J(\mathbf{p}_{SXY}, R, \Omega) = 0$, and consequently, $\Lambda(\mathbf{p}_{SXY}, R) = 0$. On the other hand, when $I(Y;X) > 0$ and $I(S;X) = 0$ (which implies $I(S;U) = 0$ for all $S - X - U$), then taking U such that $I(X;U) \leq R$ and $I(Y;U) > 0$ (such a U always exists by time-sharing between $U = X$ and U equal to a constant), we obtain $\Lambda(\mathbf{p}_{SXY}, R) = \infty$. Hence, in the sequel, we assume that $I(Y;X)$ and $I(S;X)$ are both positive.

Let $\mathbf{p}_X$ and $\mathbf{q}_X$ denote two p.m.f.'s on $\mathcal{X}$. Let $\mathbf{p}_S$ (resp. $\mathbf{q}_S$) and $\mathbf{p}_Y$ (resp. $\mathbf{q}_Y$) denote the output of the channel $\mathbf{p}_{S|X}$ and $\mathbf{p}_{Y|X}$ with input $\mathbf{p}_X$ (resp. $\mathbf{q}_X$), i.e., $\mathbf{p}_S := \mathbf{p}_{S|X}\mathbf{p}_X$, $\mathbf{q}_S := \mathbf{p}_{S|X}\mathbf{q}_X$, $\mathbf{p}_Y := \mathbf{p}_{Y|X}\mathbf{p}_X$ and $\mathbf{q}_Y := \mathbf{p}_{Y|X}\mathbf{q}_X$. Define

$$\Lambda^*(\mathbf{p}_{SX}, \mathbf{p}_{YX}) := \sup_{\substack{\mathbf{q}_X: \\ \mathbf{q}_Y \neq \mathbf{p}_Y}} \frac{D(\mathbf{q}_Y || \mathbf{p}_Y)}{D(\mathbf{q}_S || \mathbf{p}_S)},$$

with the convention that if there exists $\mathbf{q}_X$ such that $\mathbf{q}_Y \neq \mathbf{p}_Y$ and $\mathbf{q}_S = \mathbf{p}_S$, then $\Lambda^*(\mathbf{p}_{SX}, \mathbf{p}_{YX}) = \infty$. Note that $\Lambda^*(\mathbf{p}_{SX}, \mathbf{p}_{YX}) \in [0, \infty]$, and is well-defined.

Theorem 4. *For any* $R > 0$ *and given distribution* $\mathbf{p}_{SXY}$ *such that* $I(Y;X) > 0$ *and* $I(S;X) > 0$*,*

$$\Lambda(\mathbf{p}_{SXY}, R) = \Lambda^*(\mathbf{p}_{SX}, \mathbf{p}_{YX}). \qquad (23)$$

Note that the R.H.S. of (23) does not depend on the rate constraint R, as intuitively expected.

While S and Y are assumed to be latent variables in our setting, the results stated above easily extend to the scenario when S or Y, or both S and Y are directly observed by the agent. In fact, these scenarios are special cases of our setting in which $X = (S, W)$, $X = (Y, W)$, or $X = (S, Y, W)$, respectively, for some r.v. W with finite support. We may assume w.l.o.g. that $|\mathcal{W}| \geq 1$, since we may take W to be a constant if $W = \varnothing$. We next show that when $X = (S, Y, W)$, it is always possible to obtain a positive utility under perfect privacy, provided Y is not a deterministic function of S.

Proposition 5. *If* $X = (S, Y, W)$ *for some r.v.* W *(*$|\mathcal{W}| < \infty$*), then* $J(\mathbf{p}_{SXY}, R) > 0$ *if and only if* Y *is not a deterministic function of* S*.*

Proof: If Y is not a deterministic function of S, there exists some $s_1 \in \mathcal{S}$, $w_1, w_2 \in \mathcal{W}$ and $y_1, y_2 \in \mathcal{Y}$ such that $y_1 \neq y_2$, $\mathbf{p}_{SYW}(s_1, y_1, w_1) > 0$ and $\mathbf{p}_{SYW}(s_1, y_2, w_2) > 0$. Let U denote a Bernoulli(0.5) r.v. and

$$\mathbf{p}_{SYW|U}(s, y, w|0)$$
$$= \begin{cases} \mathbf{p}_{SYW}(s, y, w) + \epsilon, & \text{if } (s, y, w) = (s_1, y_1, w_1), \\ \mathbf{p}_{SYW}(s, y, w) - \epsilon, & \text{if } (s, y, w) = (s_1, y_2, w_2), \\ \mathbf{p}_{SYW}(s, y, w), & \text{otherwise}, \end{cases}$$

$$\mathbf{p}_{SYW|U}(s, y, w|1)$$
$$= 2\mathbf{p}_{SYW}(s, y, w) - \mathbf{p}_{SYW|U}(s, y, w|0), \forall\, (s, y, w) \in \mathcal{S} \times \mathcal{Y} \times \mathcal{W},$$

where $\epsilon > 0$ is chosen sufficiently small such that $\mathbf{p}_{SYW|U}(s_1, y_2, w_2|0) > 0$ and $0 < I(S, Y, W; U) < R$. It is easy to see that $\mathbf{p}_{SYW|U=0}$ and $\mathbf{p}_{SYW|U=1}$ are valid probability vectors, $\mathbf{p}_{SYW}(s, y, w)$ is preserved in $\mathbf{p}_{SYWU}$, $I(S;U) = 0$ and $I(Y;U) > 0$. Hence, $J(\mathbf{p}_{SXY}, R) > 0$. On the other hand, if $Y = f(S)$ for some $f : \mathcal{S} \to \mathcal{Y}$, then $Y - S - U$ holds. Hence, $I(S;U) = 0$ implies $I(Y;U) = 0$. This completes the proof. ■

References

[1] N. Tishby, F. Pereira, and W. Bialek, "The information bottleneck method," *arXiv preprint physics/0004057*, 2000.

[2] H. S. Witsenhausen and A. D. Wyner, "A conditional entropy bound for a pair of discrete random variables," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, Sep. 1975.

[3] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Medard, "From the information bottleneck to the privacy funnel," in *IEEE Inf. Theory Workshop*, Hobart, Australia, Nov. 2014.

[4] F. Calmon, A. Makhdoumi, and M. Medard, "Fundamental limits of perfect privacy," in *IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015.

[5] B. Rassouli and D. Gündüz, "On perfect privacy and maximal correlation," *arXiv:1712.08500 [cs.IT]*.

[6] C. Dwork, "Differential privacy," *Automata, Languages and Programming. Springer*, vol. 4052, pp. 1–12, 2006.

[7] L. Sweeney, "K-anonymity: a model for protecting privacy," *Int. Journ. on Uncertainty, Fuzziness and Knowledge based Systems*, 2002.

[8] B. Rassouli and D. Gunduz, "Optimal utility-privacy trade-off with total variation distance as a privacy measure," *IEEE Trans. on Information Forensics and Security, To appear*.

[9] A. E. Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[10] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.