On Perfect Privacy

Borzoo Rassouli[1], and Deniz Gündüz[2]
[1] School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK
[2] Department of Electrical and Electronic Engineering, Imperial College London, , London SW7 2AZ, UK
b.rassouli@essex.ac.uk, d.gunduz@imperial.ac.uk

***Abstract*—The problem of private data disclosure is studied from an information theoretic perspective. Considering a pair of dependent random variables (X,Y), where X and Y denote the private and useful data, respectively, the following problem is addressed: What is the maximum information that can be revealed about Y, measured by mutual information $I(Y;U)$, in which U denotes the revealed data, while disclosing no information about X, captured by the condition of statistical independence, i.e., $X \perp\!\!\!\perp U$, and henceforth called *perfect privacy*)? We analyze the supremization of *utility*, i.e., $I(Y;U)$ under the condition of perfect privacy for two scenarios: *output perturbation* and *full data observation* models, which correspond to the cases where a Markov kernel, called *privacy-preserving mapping*, applies to Y and the pair (X,Y), respectively. When both X and Y have a finite alphabet, the linear algebraic analysis involved in the solution provides some interesting results, such as upper/lower bounds on the size of the released alphabet and the maximum utility. Afterwards, it is shown that for the jointly Gaussian (X,Y), perfect privacy is not possible in the output perturbation model in contrast to the full data observation model. Finally, an asymptotic analysis is provided to obtain the rate of released information when a sufficiently small leakage is allowed. In particular, in the context of output perturbation model, it is shown that this rate is always finite when perfect privacy is not feasible, and two lower bounds are provided for it; When perfect privacy is feasible, it is shown that under mild conditions, this rate becomes unbounded.**

I. Introduction

With the explosion of machine learning algorithms, and their applications in many areas of science, technology, and governance, data is becoming an extremely valuable asset. However, with the growing power of machine learning algorithms in learning individual behavioral patterns from diverse data sources, privacy is becoming a major concern, calling for strict regulations on data ownership and distribution. On the other hand, many recent examples of de-anonymization attacks on publicly available anonymized data (e.g.,[2], [3]) show that regulation alone will not be sufficient to limit access to private data. An alternative approach, also considered in this paper, is to process the data at the time of release, such that no private information is leaked, called *perfect privacy*. Assuming that the joint distribution of the observed data, useful data and the private data that should be kept private is known, an information-theoretic study is carried out in this paper to characterize the fundamental limits on perfect privacy.

Consider a situation in which Alice wants to release some *useful* information about herself to Bob, represented by random variable Y, and she receives some utility from this disclosure of information. This may represent data measured and recorded by a health monitoring system [4], her smart meter measurements [5], or the sequence of a portion of her DNA to detect potential illnesses [6]. At the same time, she wishes to conceal from Bob some *private* information which depends on Y, represented by X. To this end, instead of letting Bob have a direct access to Y, 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 that result in the *utility-privacy trade-off*: The more Y is distorted by the privacy-preserving mapping, the less information can Bob infer about X, but also the less the utility that can be obtained. This trade-off is the very result of the dependencies between X and Y. 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.

The concern of privacy and the design of privacy-preserving mappings have been the focus of a broad area of research in recent years, e.g., [7]–[10], while the information-theoretic view of privacy has gained increasing attention more recently [11]. In [12], the utility-privacy trade-off under the *log-loss* cost function is considered, called as the *privacy funnel*, which is closely related to the *information bottleneck* introduced in [13]. In [14] and [15], the utility-privacy trade-off is investigated from an information theoretic perspective, and bounds on the optimal trade-off are derived. Measuring both the privacy and the utility in terms of mutual information, perfect privacy is fully characterized in [16] for the binary case. Furthermore, a new quantity is introduced to capture the amount of private information about the latent variable X carried by the useful data Y. In [17]–[19], the authors address this trade-off in a data-driven approach by setting an adversarial game between the competing neural networks.

We study the information theoretic perfect privacy in this paper, and our main contributions can be briefly summarized as follows:

A. Non-asymptotic analysis - Sections III, IV and V

1) Output perturbation model (sections III and IV):

- Denoting the supremum of $I(Y;U)$ under perfect privacy by $g_0(X,Y)$, we analyze its solution through a linear programming (LP) for finite alphabets to obtain upper and lower bounds on the cardinality of the released data, where the former is a sufficient condition, and the latter is necessary.

A conference version of this paper is provided in [1].

This work was funded by the European Research Council (ERC) through Starting Grant BEACON (no. 677854) and by the UK EPSRC (grant no. EP/N021738/1).

- From the LP solution, upper and lower bounds on $g_0(X,Y)$ are derived, which are tighter than any other known bounds in the literature in some scenarios.
- For a jointly Gaussian (X,Y), we obtain $g_\epsilon(X,Y)$ for the whole permissible range $\epsilon \in [0, I(X;Y)]$. Furthermore, we generalize this result for $\epsilon > 0$ to any joint distribution that satisfies smoothness, and for $\epsilon = 0$ to the additive case, i.e., $X = Y + N$, with $N(\perp\!\!\!\perp Y)$ being Gaussian.
- In the same setting, in the case of a finite release alphabet, say of cardinality M, we show that the utility reaches its maximum of $\log_2 M$ for a vanishingly small leakage. This is shown by using two types of practical filters: equiprobable and uniform quantizers.
- We show that in the case of finite release alphabet, the supremum in the definition of $g_\epsilon(X,Y)$ is actually a maximum, in spite of the non-compactness of the search space.
- We establish the relationship between $g_0(X,Y)$ and *non-private information about X carried by Y*, $D_X(Y)$, as defined in [16], and provide the necessary and sufficient conditions when the two aforementioned quantities are equal.

2) Full data observation model (section V):

- We provide the necessary and sufficient condition for the feasibility of perfect privacy. In this context, the maximum utility is denoted by $G_0(X,Y)$.
- We provide a lower bound on $G_0(X,Y)$, which can become relevant to the *maximal leakage* defined in [20].
- We show that for a jointly Gaussian (X,Y), we have $G_0(X,Y) = \infty$, which is the direct opposite of $g_0(X,Y) = 0$. We actually state this result for the broader class of additive noise, i.e., $Y = X + N$, in which, N does not need to be independent of X, but it needs to admit a density for each realization x of X.

B. Asymptotic analysis in the context of output pertaurbation model - Section VI

- We show that when perfect privacy is not feasible, the slope of $g_\epsilon(X,Y)$ at the origin, i.e., $\epsilon = 0$, is always finite, and provide two lower bounds on this slope, which are tighter than the previously known bounds in the literature.
- We show that when perfect privacy is feasible, for a broad range of cases, this slope at the origin is infinite.
- We provide a general lower bound on this slope when perfect privacy is feasible.

Notation. 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 a matrix $\mathbf{A}_{m\times k}$, the null space, rank, and nullity are denoted by $\text{Null}(\mathbf{A})$, $\text{rank}(\mathbf{A})$, and $\text{nul}(\mathbf{A})$, respectively, with $\text{rank}(\mathbf{A}) + \text{nul}(\mathbf{A}) = k$. For integers $m \le n$, 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 $a_{[m:n]}$. The set $[1:n]$ is written in short as $[n]$. For an integer $n \ge 1$, the notation $\mathbf{1}_n$, and $\mathbf{0}_n$ denote the n-dimensional all-one, and all-zero column vectors, respectively. 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

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

whose interior is denoted by $\text{int}(\mathcal{P}(\mathcal{X}))$. Furthermore, to each probability mass function (pmf) on $\mathcal{X}$, denoted by $p_X(\cdot)$, corresponds a matrix $\mathbf{P}_X = \text{diag}(\mathbf{p}_X)$, where $\mathbf{p}_X$ is a probability vector in $\mathcal{P}(\mathcal{X})$, whose i-th element is $p_X(x_i)$ $(i \in [|\mathcal{X}|])$. For a pair of random variables (X,Y) with joint pmf $p_{X,Y}$, $\mathbf{P}_{X,Y}$ is an $|\mathcal{X}| \times |\mathcal{Y}|$ matrix with (i,j)-th entry equal to $p_{X,Y}(i,j)$. Likewise, $\mathbf{P}_{X|Y}$ is an $|\mathcal{X}| \times |\mathcal{Y}|$ matrix with (i,j)-th entry equal to $p_{X|Y}(i|j)$. $F_Y(\cdot)$ denotes the cumulative distribution function (CDF) of random variable Y, and if it admits a density, its probability density function (pdf) is denoted by $f_Y(\cdot)$. Throughout the paper, for a random variable Y with the corresponding probability vector $\mathbf{p}_Y$, $H(Y)$ and $H(\mathbf{p}_Y)$ are written interchangeably, and so are the quantities $D(p_Y(\cdot)||q_Y(\cdot))$ and $D(\mathbf{p}_Y||\mathbf{q}_Y)$. All logarithms in this paper are in base 2. Given two positive integers a, b, a modulo b is abbreviated as a mod b. Finally, d_{TV}, $\lfloor\cdot\rfloor$, and $\lceil\cdot\rceil$ denote the total variation distance, the floor, and the ceiling operators, respectively. [1]

II. System model and preliminaries

Consider a triplet of random variables $(X,Y,W) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{W}$, distributed according to the joint distribution $p_{X,Y,W}$. Let X denote the *private/sensitive data* that the user/curator wants to conceal, Y denote the *useful data* the user wishes to disclose, and W denote the *observable data* that the curator observes, which can be regarded as a noisy version of (X,Y). Assume that the *privacy-preserving mapping/data release mechanism* takes W as input, and maps it to the *released data*, denoted by U. In this scenario, $(X,Y) - W - U$ form a Markov chain, and the privacy-preserving mapping is captured by the conditional distribution $p_{U|W}$.

Definition 1. *Perfect privacy is feasible if there exists a privacy-preserving mapping $p_{U|W}$ whose output U is statistically dependent on the useful data Y, while being statistically independent of the private data X; that is, $Y \not\perp\!\!\!\perp U$ and $X \perp\!\!\!\perp U$.*

Unless otherwise stated explicitly, we assume that all the alphabets/supports $\mathcal{X}, \mathcal{Y}, \mathcal{W}$ are finite. In this context, we assume that $p_X(x), p_Y(y), p_W(w) > 0, \forall (x,y,w) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{W}$, since otherwise the alphabets could have been modified accordingly. This equivalently means that the corresponding probability vectors $\mathbf{p}_X, \mathbf{p}_Y, \mathbf{p}_W$ are in the interior of their corresponding probability simplices, i.e., $\mathcal{P}(\mathcal{X}), \mathcal{P}(\mathcal{Y}), \mathcal{P}(\mathcal{W})$, respectively.

The following proposition states the necessary and sufficient condition for the feasibility of perfect privacy.

[1] Due to space constraints, some of the proofs are provided in the extended online version [21].

Proposition 1. *Perfect privacy is feasible for* $(X,Y,W) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{W}$ *if and only if*

$$\dim\Big(\text{Null}(\mathbf{P}_{X|W})\backslash\text{Null}(\mathbf{P}_{Y|W})\Big) \neq 0. \tag{1}$$

Proof. The proof is a simple generalization of [22, Theorem 4], by noting that both $X-W-U$ and $Y-W-U$ form Markov chains. In other words, we have $X \perp\!\!\!\perp U$ if and only if for all $u\in\mathcal{U}$, $\mathbf{p}_X = \mathbf{P}_{X|W}\mathbf{p}_{W|u}$. On the other hand, we have $Y \not\perp\!\!\!\perp U$ if and only if there exists a $u'\in\mathcal{U}$, such that $\mathbf{p}_Y \neq \mathbf{P}_{Y|W}\mathbf{p}_{W|u'}$. Equivalently, there exists a vector $\mathbf{v}' \triangleq \mathbf{p}_W - \mathbf{p}_{W|u'}$ in $\text{Null}(\mathbf{P}_{X|W})$ ($\mathbf{v}'\neq\mathbf{0}$) that does not belong to $\text{Null}(\mathbf{P}_{Y|W})$, which is equivalent to (1). □

The special cases of *full data observation* and *output perturbation* ([23]) refer to the scenarios in which the privacy-preserving mapping has direct access to both the private and useful data (i.e., $W=(X,Y)$) and only to the useful data (i.e., $W=Y$), respectively. The whole paper is devoted to these two models.

By adopting mutual information as the measure of both *utility* and *privacy* (i.e., $I(Y;U)$, and $I(X;U)$, respectively), the optimal utility-privacy trade-off in the output perturbation model is defined as[2]

$$g_\epsilon(X,Y) \triangleq \sup_{\substack{p_{U|Y}:\\ X-Y-U\\ I(X;U)\leq\epsilon}} I(Y;U), \tag{2}$$

and in the full data observation model, the trade-off can be formulated as

$$G_\epsilon(X,Y) \triangleq \sup_{\substack{p_{U|X,Y}:\\ I(X;U)\leq\epsilon}} I(Y;U), \tag{3}$$

where the effective range of ϵ is $[0, I(X;Y)]$.

Finally, we can say that perfect privacy being feasible in the output perturbation and full data observation models is equivalent to having $g_0(X,Y)>0$ and $G_0(X,Y)>0$, respectively.

III. Output perturbation model

In this model, we have $X-Y-U$ form a Markov chain, and in order to derive $g_0(X,Y)$, we proceed as follows. 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 matrix of right eigenvectors is $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_{|\mathcal{Y}|}\end{bmatrix}$. By assuming (without loss of generality) that the singular values are arranged in a descending order, only the first $\text{rank}(\mathbf{P}_{X|Y})$ singular values are non-zero. Therefore, the null space of $\mathbf{P}_{X|Y}$ can be written as $\text{Null}(\mathbf{P}_{X|Y}) = \text{Span}\{\mathbf{v}_{\text{rank}(\mathbf{P}_{X|Y})+1}, \mathbf{v}_{\text{rank}(\mathbf{P}_{X|Y})+2}, \dots, \mathbf{v}_{|\mathcal{Y}|}\}$.

In the Markov chain $X-Y-U$, the random variables X and U are independent if and only if $\mathbf{P}_{X|Y}(\mathbf{p}_Y - \mathbf{p}_{Y|u}) = \mathbf{0}$, $\forall u\in\mathcal{U}$, which is equivalent to $(\mathbf{p}_Y - \mathbf{p}_{Y|u}) \in \text{Null}(\mathbf{P}_{X|Y})$, $\forall u\in\mathcal{U}$. Let $\mathbf{A}$ be defined as $\mathbf{A}\triangleq\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_{\text{rank}(\mathbf{P}_{X|Y})}\end{bmatrix}^T$. Therefore, we have $X \perp\!\!\!\perp U$ in $X-Y-U$ if and only if $\mathbf{A}(\mathbf{p}_Y - \mathbf{p}_{Y|u}) = \mathbf{0}$, $\forall u\in\mathcal{U}$. Let $\mathbb{S}_{X,Y}$ be defined as

$$\mathbb{S}_{X,Y} \triangleq \left\{\mathbf{t}\in\mathbb{R}^{|\mathcal{Y}|}\,\middle|\,\mathbf{A}\mathbf{t} = \mathbf{A}\mathbf{p}_Y\ ,\ \mathbf{t}\geq 0\right\}, \tag{4}$$

which is a convex polytope in $\mathcal{P}(\mathcal{Y})$, since it can be written as the intersection of a finite number of half-spaces in $\mathcal{P}(\mathcal{Y})$. With this definition, we have that having $X \perp\!\!\!\perp U$ in $X-Y-U$ results in $\mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y}$, $\forall u\in\mathcal{U}$. On the other hand, for any pair (Y,U), for which $\mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y}$, $\forall u\in\mathcal{U}$, we can simply have $X-Y-U$ and $X\perp\!\!\!\perp U$. Therefore, we can write

$$X-Y-U,\ X\perp\!\!\!\perp U \iff \mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y},\ \forall u\in\mathcal{U}. \tag{5}$$

Theorem 1. *The supremum in (2) is attained, and hence, it is a maximum. Furthermore, in the evaluation of* $g_0(X,Y)$*, the optimal privacy-preserving mapping is the solution to a standard linear program (LP), and it is sufficient to have* $|\mathcal{U}| \leq \text{nul}(\mathbf{P}_{X|Y}) + 1$*. Finally, if* $p^*_{Y,U}$ *corresponds to a maximizer* $p^*_{U|Y}$*, for any* $u\in\mathcal{U}$*, we have*

$$|\{y\in\mathcal{Y}|p^*(y|u)>0\}| \leq \text{rank}(\mathbf{P}_{X|Y}). \tag{6}$$

Proof. The proof of the attainability of the supremum, and the upper bound $|\mathcal{U}|\leq\text{nul}(\mathbf{P}_{X|Y})+1$[3] are provided in Appendix A. We have

$$g_0(X,Y) = H(Y) - \min_{\substack{p_U(\cdot),\ \mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y},\ \forall u\in\mathcal{U}:\\ \sum_u p_U(u)\mathbf{p}_{Y|u}=\mathbf{p}_Y}} H(Y|U), \tag{7}$$

where in (7), since the minimization is over $\mathbf{p}_{Y|u}$ rather than $p_{U|Y}$, a constraint was added to preserve the marginal distribution $\mathbf{p}_Y$. The minimization of the concave functional in (7) simplifies to an LP as stated in [24].

In order to prove the final claim in the statement of this Theorem, we need to address the solution to this LP, whose linear algebraic analysis (i.e., characterizations of the null space, extreme points, etc.) is the basis for some of the main results obtained in this paper. We address this solution as follows.

Lemma 1. *In minimizing* $H(Y|U)$ *over* $\{\mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y}|\sum_u p_U(u)\mathbf{p}_{Y|u}=\mathbf{p}_Y\}$*, it is sufficient to consider only* $\text{nul}(\mathbf{P}_{X|Y})+1$ *extreme points of* $\mathbb{S}_{X,Y}$*.*

Proof. The proof is provided in [21, Appendix B]. □

From lemma 1, the solution to the minimization in (7) can be obtained in two phases: in phase one, the extreme points of set $\mathbb{S}_{X,Y}$ are identified, while in phase two, proper weights over these extreme points are obtained to minimize the objective function.

For the first phase, we proceed as follows. The extreme points of $\mathbb{S}_{X,Y}$ are the basic feasible solutions (see [25], [26]) of $\{\mathbf{x}\in\mathbb{R}^{|\mathcal{Y}|}|\mathbf{A}\mathbf{x}=\mathbf{b}\ ,\ \mathbf{x}\geq 0\}$, where $\mathbf{b}=\mathbf{A}\mathbf{p}_Y$. The procedure of finding the extreme points of $\mathbb{S}_{X,Y}$ is as follows. Pick a set $\mathcal{B}\subset[|\mathcal{Y}|]$ of indices that correspond to $\text{rank}(\mathbf{P}_{X|Y})$ linearly independent columns of matrix $\mathbf{A}$ defined prior to (4). Let $\mathbf{A}_\mathcal{B}$ be a $\text{rank}(\mathbf{P}_{X|Y})\times\text{rank}(\mathbf{P}_{X|Y})$ matrix whose columns are the columns of $\mathbf{A}$ indexed by the indices in $\mathcal{B}$. Also, for any $\mathbf{x}\in\mathbb{S}_{X,Y}$, define a corresponding

[2] This is the same notation as in [16].

[3] The proof of this upper bound follows the application of cardinality bounding technique and taking into account the convex polytope $\mathbb{S}_{X,Y}$ in (4). Although we are considering perfect privacy here, i.e, $g_0(X,Y)$, in the evaluation of $g_\epsilon(X,Y), \forall\epsilon>0$, it is sufficient to have $|\mathcal{U}|\leq|\mathcal{Y}|+1$ as in [15].

vector $\tilde{\mathbf{x}} \triangleq \begin{bmatrix}\mathbf{x}_{\mathcal{B}}^T & \mathbf{x}_{\mathcal{N}}^T\end{bmatrix}^T$, where $\mathbf{x}_{\mathcal{B}}$ and $\mathbf{x}_{\mathcal{N}}$ are $\text{rank}(\mathbf{P}_{X|Y})$-dimensional and $\text{nul}(\mathbf{P}_{X|Y})$-dimensional vectors whose elements are the elements of $\mathbf{x}$ indexed by the indices in $\mathcal{B}$ and $[|\mathcal{Y}|]\backslash\mathcal{B}$, respectively.

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

$$\mathbf{x}_{\mathcal{N}}^* = \mathbf{0}, \quad \mathbf{x}_{\mathcal{B}}^* = \mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b}, \quad \mathbf{x}_{\mathcal{B}}^* \geq 0, \tag{8}$$

where the inequality is element-wise. On the other hand, for any set $\mathcal{B} \subset [|\mathcal{Y}|]$ of indices that correspond to a set of $\text{rank}(\mathbf{P}_{X|Y})$ linearly independent columns of $\mathbf{A}$, if $\mathbf{A}_{\mathcal{B}}^{-1}\mathbf{b} \geq 0$, then $\begin{bmatrix}\mathbf{b}^T\mathbf{A}_{\mathcal{B}}^{-T} & \mathbf{0}^T\end{bmatrix}^T$ is the corresponding vector of a basic feasible solution. Hence, the extreme points of $\mathbb{S}_{X,Y}$ are obtained as mentioned above, and their number is at most $\binom{|\mathcal{Y}|}{\text{rank}(\mathbf{P}_{X|Y})}$, which is justified as follows. Since an extreme point is identified if and only if A) the $\text{rank}(\mathbf{P}_{X|Y})$ selected columns are linearly independent, B) the corresponding $\mathbf{x}_{\mathcal{B}}$ has all non-negative elements, it is concluded that the total number of extreme points is upper bounded by the total number of ways to choose $\text{rank}(\mathbf{P}_{X|Y})$ linearly independent columns out of $|\mathcal{Y}|$ columns. The latter is also upper bounded by the total number of ways to choose $\text{rank}(\mathbf{P}_{X|Y})$ columns out of $|\mathcal{Y}|$ columns, which is $\binom{|\mathcal{Y}|}{\text{rank}(\mathbf{P}_{X|Y})}$. Furthermore, each extreme point has at most $\text{rank}(\mathbf{P}_{X|Y})$ non-zero elements corresponding to $\mathbf{x}_{\mathcal{B}}$, which is equivalent to $|\{y \in \mathcal{Y} | p^*(y|u) > 0\}| \leq \text{rank}(\mathbf{P}_{X|Y})$ for any $u \in \mathcal{U}$.

For the second phase, we proceed as follows. Assume that $\mathbb{S}_{X,Y}$ has K (a positive integer) extreme points, denoted by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_K$, which were identified in the first phase. Then, (7) is equivalent to

$$\begin{aligned} g_0(X,Y) = H(Y) - \min_{\mathbf{w}\geq 0} \; & \begin{bmatrix}H(\mathbf{p}_1) & H(\mathbf{p}_2) & \ldots & H(\mathbf{p}_K)\end{bmatrix} \cdot \mathbf{w} \\ \text{s.t.} \; & \begin{bmatrix}\mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_K\end{bmatrix} \mathbf{w} = \mathbf{p}_Y, \end{aligned} \tag{9}$$

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 (9) is met. The problem in (9) is a standard LP. □

Corollary 1.1. *In the evaluation of $g_0(X,Y)$, it is necessary to have $|\mathcal{U}| \geq \left\lceil \frac{|\mathcal{Y}|}{\text{rank}(\mathbf{P}_{X|Y})} \right\rceil$.*

Proof. In the proof of Theorem 1, in order to write the $|\mathcal{Y}|$-dimensional probability vector $\mathbf{p}_Y$ as a convex combination of the extreme points of $\mathbb{S}_{X,Y}$, that have at most $\text{rank}(\mathbf{P}_{X|Y})$ non-zero elements, at least $\left\lceil \frac{|\mathcal{Y}|}{\text{rank}(\mathbf{P}_{X|Y})} \right\rceil$ points are needed, which results in $|\mathcal{U}| \geq \left\lceil \frac{|\mathcal{Y}|}{\text{rank}(\mathbf{P}_{X|Y})} \right\rceil$. □

Corollary 1.2. *We have the following bounds on $g_0(X,Y)$.*

$$\begin{aligned} \left(H(Y) - \log \text{rank}(\mathbf{P}_{X|Y})\right)^+ &\leq g_0(X,Y), \\ g_0(X,Y) &\leq \min\{\log\left(\text{nul}(\mathbf{P}_{X|Y}) + 1\right), H(Y|X)\}. \end{aligned}$$

Proof. The first term in the upper bound is immediate from $I(Y;U) \leq H(U) \leq \log|\mathcal{U}| \leq \log\left(\text{nul}(\mathbf{P}_{X|Y}) + 1\right)$, where the last inequality follows from Theorem 1. The second term in the upper bound follows from [14]. The lower bound is proved as follows. As mentioned in the proof of Theorem 1, each extreme point of $\mathbb{S}_{X,Y}$ has at most $\text{rank}(\mathbf{P}_{X|Y})$ non-zero elements, which means that the entropy of each extreme point is upper bounded by $\log(\text{rank}(\mathbf{P}_{X|Y}))$. Hence,

$$\min_{\substack{p_U(\cdot),\ \mathbf{p}_{Y|u}\in\mathbb{S}_{X,Y},\ \forall u\in\mathcal{U}: \\ \sum_u p_U(u)\mathbf{p}_{Y|u}=\mathbf{p}_Y}} H(Y|U) \leq \log(\text{rank}(\mathbf{P}_{X|Y})),$$

which results in the lower bound on $g_0(X,Y)$. □

An example of the LP solution in Theorem 1 is provided in [21].

Thus far, we have investigated perfect privacy when $|\mathcal{X}|, |\mathcal{Y}| < \infty$. In what follows, i.e., Theorem 2, it is shown that perfect privacy is not feasible for the (correlated) jointly Gaussian pair. Part of the proof of Theorem 2 relies on using a privacy-preserving mapping $p_{U|Y}$ that quantizes the useful data Y with infinitely small quantization intervals, which in turn, is based on the following lemma.

Lemma 2. *Let Z be an r.v. distributed over an interval $[a,b]$, in which $a, b \in \mathbb{R}$ $(a < b)$, with a bounded smooth pdf denoted by $f_Z(\cdot)$[4]. For positive integers M, n, define a partition $a = a_0 < a_1 < a_2 < \ldots < a_{Mn-1} < a_{Mn} = b$. Let $\mathcal{I}_i \triangleq [a_{i-1}, a_i)$, $\forall i \in [Mn-1]$, and $\mathcal{I}_{Mn} \triangleq [a_{Mn-1}, b]$. Let U be a function of Z defined as*

$$u(z) \triangleq (i-1) \bmod M, \text{ if } z \in \mathcal{I}_i, \text{ for some } i \in [Mn]. \tag{10}$$

If for all $i \in [Mn]$, we have $(a_i - a_{i-1}) \to 0$ with $n \to \infty$, then

$$\lim_{n\to\infty} H(U) = \log M. \tag{11}$$

Proof. Let p_U denote the pmf of U, whose realizations are given in (10). Also, let $\mathcal{J}_i \triangleq \cup_{k=(i-1)M+1}^{iM} \mathcal{I}_k$, $\forall i \in [n]$. Let $\hat{Z}$ be an r.v. whose pdf is a piecewise uniform approximation of f_Z over the intervals $\mathcal{J}_i, i \in [n]$. Hence, we have

$$f_{\hat{Z}}(\hat{z}) = \frac{1}{l(\mathcal{J}_i)} \int_{\mathcal{J}_i} f_Z(z)dz, \; \forall \hat{z} \in \mathcal{J}_i, \; \forall i \in [n], \tag{12}$$

where $l(\mathcal{J}_i) = a_{iM} - a_{(i-1)M}$ denotes the length of the segment $\mathcal{J}_i$, $i \in [n]$.

Let $\hat{U}$ be a function of $\hat{Z}$ in exactly the same way that U is defined as a function of Z, i.e., as in (10). Since $f_{\hat{Z}}$ is flat over $\mathcal{J}_i$, $i \in [n]$, $\hat{U}$ is uniform over $[0 : M-1]$. Since $(a_i - a_{i-1}) \to 0$ as $n \to \infty$, $\forall i \in [Mn]$, we conclude that $l(\mathcal{J}_i) \to 0$ as $n \to \infty$, $\forall i \in [n]$. As a result, $f_{\hat{Z}}(\cdot)$ converges pointwise to $f_Z(\cdot)$, due to the smoothness of the latter. Therefore, we have $d_{\text{TV}}(f_{\hat{Z}}, f_Z) = \int |f_{\hat{Z}} - f_Z| dz \to 0$ as $n \to \infty$, which is a direct consequence of Lebesgue's Dominated Convergence Theorem. Hence, by viewing Z and $\hat{Z}$ as the inputs to a (deterministic) channel in (10), with the corresponding outputs U

[4]Note that since Z admits a density, its support being a segment, as (a,b), or an interval, as $[a,b]$, or a mixture does not change the result in this lemma. Hence, with a slight abuse of notation, segment and interval are used interchangeably.

and $\hat{U}$, respectively, we get $\lim_{n\to\infty} d_{\text{TV}}(p_U, p_{\hat{U}}) = 0$, which follows from the data processing inequality of f-divergences, i.e., $d_{\text{TV}}(p_U, p_{\hat{U}}) \leq d_{\text{TV}}(f_{\hat{Z}}, f_Z)$. Finally, from the fact that $H(\hat{U}) = \log M$, and by the continuity of entropy, (11) is proved. □

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

$$\mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}, \tag{13}$$

in which $\rho \neq 0$, since otherwise $X \perp\!\!\!\perp Y$. We have

$$g_\epsilon(X, Y) = \begin{cases} 0 & \epsilon = 0 \\ \infty & \text{o.w.} \end{cases} \tag{14}$$

Proof. First, it is shown that $g_0(X, Y) = 0$. 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 has a density. Equivalently, we must have

$$f_X(\cdot) = \int f_{X|Y}(\cdot|y) dF_{Y|U}(y|u),\ \forall u \in \mathcal{U}. \tag{15}$$

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

$$F_{Y|U}(\cdot|u_1) \neq F_{Y|U}(\cdot|u_2). \tag{16}$$

In what follows we show that if (15) holds, (16) 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

$$X|\{Y = y\} \sim \mathcal{N}\Big(\underbrace{\frac{\rho\sigma_X}{\sigma_Y}(y - \mu_Y) + \mu_X}_{\alpha y + \beta}, \underbrace{(1 - \rho^2)\sigma_X^2}_{\sigma^2}\Big). \tag{17}$$

From (15), (17), and for $u_1, u_2 \in \mathcal{U}$, we have

$$\begin{aligned} f_X(x) &= \int \frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dF_{Y|U}(y|u_1) \\ &= \int \frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dF_{Y|U}(y|u_2),\ \forall x \in \mathbb{R}, \end{aligned}$$

or, equivalently,

$$\int \frac{e^{-\frac{(x-\alpha y-\beta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big) = 0,\ \forall x \in \mathbb{R}. \tag{18}$$

Multiplying both sides of (18) by $e^{j\omega x}$, and taking the integral with respect to x, we obtain

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

By Fubini's theorem[5], we can write

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

and after some manipulations, we get

$$\int e^{j\omega\alpha y} d\Big(F_{Y|U}(y|u_1) - F_{Y|U}(y|u_2)\Big) = 0. \tag{19}$$

Since $\rho \neq 0$, from (17), we have $\alpha \neq 0$. Hence, the LHS of (98) 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, (16) does not hold and perfect privacy is not feasible for the (correlated) jointly Gaussian pair (X, Y).

In order to show $g_\epsilon(X, Y) = \infty,\ \forall \epsilon > 0$, two proofs/methods are provided. Both of them aim to construct a privacy-preserving mapping $p_{U|Y}$ as an M-level quantizer (for an arbitrary integer $M > 0$), which satisfies the privacy constraint, and results in a utility that grows with M. Hence, the proof is completed by letting $M \to \infty$. In the first method, this is done by quantizing the support of Y into equiprobable intervals, while in the second method, a uniform quantizer is employed, which partitions the support of Y into intervals of the same length (denoted by Δ). The advantages/disadvantages of these two methods are elaborated further in the remarks that follow the Theorem.

In what follows, without loss of optimality, we consider that both X and Y have the standard Normal distribution[6].

A. First method for showing $g_\epsilon(X, Y) = \infty,\ \forall \epsilon > 0$: Equiprobable quantizer

Fix $\epsilon > 0$, and a positive integer M. For each integer $n > 1$, define

$$\mathcal{B}_n \triangleq \Big[\Phi^{-1}(\frac{1}{n}), \Phi^{-1}(1 - \frac{1}{n})\Big]^2, \tag{20}$$

$$p_n \triangleq \Pr\Big\{(X, Y) \notin \mathcal{B}_n\Big\}, \tag{21}$$

where $\Phi^{-1}(\cdot)$ is the inverse function of standard Normal CDF $\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$.

As $n \to \infty$, we have $p_n \to 0$, and hence, $(p_n \log M + H_b(p_n)) \to 0$. Therefore, there exists a positive integer N_0 such that $p_n \log M + H_b(p_n) \leq \frac{\epsilon}{2}$, for all $n \geq N_0$. Let $N_1(\epsilon)$ denote the minimum N_0 for which the previous statement holds.

Let $E \triangleq \mathbb{1}_{(X,Y)\in\mathcal{B}_{N_1(\epsilon)}}$ be a binary indicator which is 1 when $(X, Y) \in \mathcal{B}_{N_1(\epsilon)}$, and 0 otherwise. For a positive integer n, let $\{\Phi^{-1}(\frac{i}{MN_1(\epsilon)n})\}_{i=1}^{MN_1(\epsilon)n-1}$ be a set of points that divide the support of a standard Normal into $MN_1(\epsilon)n$ equiprobable intervals, which are denoted by $\mathcal{I}_1 = \Big(-\infty, \Phi^{-1}(\frac{1}{MN_1(\epsilon)n})\Big)$, and $\mathcal{I}_i = \Big[\Phi^{-1}(\frac{i-1}{MN_1(\epsilon)n}), \Phi^{-1}(\frac{i}{MN_1(\epsilon)n})\Big)$ for $i \in [2 : MN_1(\epsilon)n)$, with the convention $\Phi^{-1}(1) = \infty$. Define U as a function of Y according to

$$u(y) \triangleq (i - 1) \bmod M,\ \text{if } y \in \mathcal{I}_i,\ \text{for some } i \in [MN_1(\epsilon)n]. \tag{22}$$

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

[6]This is due to the fact that for the tuple (X, Y, U), $I(aX + b; U) = I(X; U)$, and $I(cY + d; U) = I(Y; U)$ for constants a, b, c and d. Furthermore, the set of mappings from Y has a one-to-one correspondence with the set of mappings from $cY + d$.

From the construction in (22), we have that U is uniform over $[0:M-1]$, and $I(Y;U)=\log M$ for any positive integer n. In the sequel, we show that there exists a positive integer N, such that $I(X;U)\leq\epsilon,\ \forall n\geq N$, which results in $g_\epsilon(X,Y)\geq\log M$. Since M is arbitrary, we get $g_\epsilon(X,Y)=\infty$, which concludes the proof.

Conditioned on the event $\{E=1\}$, we have $H(U|E=1)=\log M$, since $Y|\{E=1\}$ with pdf $f_{Y|E}(\cdot|1)$, which is a scaled version of $f_Y(\cdot)$, is distributed over $[\Phi^{-1}(\frac{1}{N_1(\epsilon)}),\Phi^{-1}(1-\frac{1}{N_1(\epsilon)})]$, which has been divided into $Mn(N_1(\epsilon)-2)$ equiprobable intervals, i.e., $\{\mathcal{I}_i\}_{i=Mn+1}^{Mn(N_1(\epsilon)-1)}$, and from (22), $U|\{E=1\}$ becomes uniform over $[0:M-1]$.

For each realization $x\in[\Phi^{-1}(\frac{1}{N_1(\epsilon)}),\Phi^{-1}(1-\frac{1}{N_1(\epsilon)})]$ of X, the conditional pdf $f_{Y|X,E}(\cdot|x,1)$ is a bounded smooth density. Hence, from lemma 2, there exists a positive integer $N(x,\epsilon)$, such that

$$H(U|X=x,E=1)\geq\log M-\frac{\epsilon}{2},\ \forall n\geq N(x,\epsilon). \quad (23)$$

Furthermore, since $[\Phi^{-1}(\frac{1}{N_1(\epsilon)}),\Phi^{-1}(1-\frac{1}{N_1(\epsilon)})]$ is a compact subset of the real line, we can define

$$N_2(\epsilon)\triangleq\max_{x\in[\Phi^{-1}(\frac{1}{N_1(\epsilon)}),\Phi^{-1}(1-\frac{1}{N_1(\epsilon)})]}N(x,\epsilon).$$

Therefore, for all $n\geq N_2(\epsilon)$, we have

$$\begin{aligned}I(X;U|E=1)&=H(U|E=1)-H(U|X,E=1)\\&=\log M-H(U|X,E=1) \quad (24)\\&\leq\frac{\epsilon}{2}, \quad (25)\end{aligned}$$

where (24) and (25) follow, respectively, from $U|\{E=1\}$ being uniform, and (23).

Finally, we have that for $n\geq\max\{N_1(\epsilon),N_2(\epsilon)\}$,

$$\begin{aligned}I(X;U)&\leq I(X;U,E)\\&\leq H(E)+I(X;U|E) \quad (26)\\&=H(\Pr\{E=0\})+\Pr\{E=0\}I(X;U|E=0)\\&\quad+\Pr\{E=1\}I(X;U|E=1)\\&<H(\Pr\{E=0\})+\Pr\{E=0\}\log M\\&\quad+I(X;U|E=1) \quad (27)\\&\leq\frac{\epsilon}{2}+\frac{\epsilon}{2} \quad (28)\\&=\epsilon,\end{aligned}$$

where (27) follows from the trivial upper bound $I(X;U|E=0)\leq\log M$; (28) results from (25) and the fact that $p_n\log M+H_b(p_n)\leq\frac{\epsilon}{2}$ for $n\geq N_1(\epsilon)$, in which $p_n=\Pr\{E=0\}$. Therefore, for any $\epsilon>0$, we have constructed a privacy-preserving mapping $p_{U|Y}$ for which $I(Y;U)=\log M$, and $I(X;U)\leq\epsilon$. Finally, letting $M\to\infty$ completes the first proof.

B. Second method for showing $g_\epsilon(X,Y)=\infty,\ \forall\epsilon>0$: Uniform quantizer

Fix $\epsilon>0$. Fix a positive integer M, and set $U_M^\Delta\triangleq\lfloor\frac{MY}{\Delta}\rfloor \mod M,\ \forall\Delta>0$. From lemma 2, we have

$$\lim_{\Delta\to0}I(Y;U_M^\Delta)=\lim_{\Delta\to0}H(U_M^\Delta)=\log M, \quad (29)$$

which follows from $H(U_M^\Delta|Y)=0$. Therefore, we have that for any $\delta>0$, there exists $\Delta_0>0$, such that $I(Y;U_M^\Delta)\geq\log M-\delta$, for all $\Delta\leq\Delta_0$. Since the conditional distribution of Y given $\{X=x\}$ still satisfies the conditions of lemma 2 (i.e., replace f_Z with $f_{Y|X}$), we obtain that for any $x\in\mathbb{R}$, there exists $\Delta_x>0$, such that $H(U_M^\Delta|X=x)\geq\log M-\frac{\epsilon}{2}$, for all $\Delta\leq\Delta_x$.

Let $\mathcal{I}_0\triangleq[-\Phi^{-1}(1-\frac{\epsilon}{4\log M}),\Phi^{-1}(1-\frac{\epsilon}{4\log M})]$, and define $E_0\triangleq\mathbb{1}_{\{X\in\mathcal{I}_0\}}$ as an r.v. indicating if X belongs to $\mathcal{I}_0$. Finally, set $\Delta_1\triangleq\min_{x\in\mathcal{I}_0}\Delta_x$, where Δ_x is given in the previous paragraph. Note that since $\Delta_x>0$, and the minimization is over a compact set, i.e., $\mathcal{I}_0$, we have $\Delta_1>0$. For any $\Delta\leq\Delta_1$, we can write

$$\begin{aligned}I(X;U_M^\Delta)&=I(X,E_0;U_M^\Delta) \quad (30)\\&=H(U_M^\Delta)-\Pr\{E_0=1\}H(U_M^\Delta|X,E_0=1)\\&\quad-\Pr\{E_0=0\}H(U_M^\Delta|X,E_0=0)\\&\leq\log M-(1-\frac{\epsilon}{2\log M})(\log M-\frac{\epsilon}{2}) \quad (31)\\&=\frac{\epsilon}{2}(2-\frac{\epsilon}{2\log M})\\&\leq\epsilon,\end{aligned}$$

where (30) follows from having E_0 as a deterministic function of X; (31) results from $H(U_M^\Delta)\leq\log M$, $\Pr\{E_0=1\}=1-\frac{\epsilon}{2\log M}$, and the non-negativity of entropy. Hence, it is shown that for any $\epsilon,\delta>0$ and integer $M>0$, there exists $\Delta_2\triangleq\min\{\Delta_0,\Delta_1\}$, such that

$$\begin{aligned}&I\left(X;U_M^\Delta\right)\leq\epsilon,\\&I\left(Y;U_M^\Delta\right)\geq\log M-\delta,\ \forall\Delta\in(0,\Delta_2).\end{aligned}$$

Finally, by letting $M\to\infty$, the proof is completed. [7] □

Remark 1. ***(Generalization of $g_0(X,Y)=0$.)*** *It is important to note that in the first claim of (14), i.e., $g_0(X,Y)=0$, Gaussianity of Y is not used, and the proof relies solely on the characteristics of the conditional pdf $f_{X|Y}(\cdot|\cdot)$. Therefore, for an arbitrary pair of random variables (X,Y), in which $X=Y+N$, with N being Gaussian and independent of Y, we have $g_0(X,Y)=0$. This means that perfect privacy for an additive Gaussian noisy version of any random variable comes at the cost of zero utility.*

Remark 2. *It is important to note that in the second claim of (14), i.e., $g_\epsilon(X,Y)=\infty,\ \forall\epsilon>0$, Gaussianity of the pair (X,Y) is not necessary. Hence, for an arbitrary pair (X,Y), $g_\epsilon(X,Y)=\infty,\ \forall\epsilon>0$, if Y conditioned on $\{X=x\}$, i.e., $Y|\{X=x\}$, admits a bounded smooth pdf for any $x\in\mathcal{X}$.*

[7] As it can be observed, in the first method of showing $g_\epsilon(X,Y)=\infty,\ \forall\epsilon>0$ (i.e., equiprobable quantization), which is more complicated than the second one, we have $I(Y;U)=\log M$, while in the second method (i.e., uniform quantization), $I(Y;U)$ approaches $\log M$, since $I\left(Y;\lfloor\frac{MY}{\Delta}\rfloor \mod M\right)<\log M,\ \forall\Delta>0$ in general. This advantage of the first method is used in the proof of Remark 4, i.e., replacing the supremum with maximum when $\mathcal{U}$ is a finite set. However, the second method is simpler and has a fixed quantization interval.

Definition 2. *For a positive integer M, the utility-privacy trade-off with an M-ary release alphabet is defined as*

$$g_\epsilon^M(X,Y) \triangleq \sup_{\substack{p_{U|Y}:|\mathcal{U}|\leq M \\ I(X;U)\leq\epsilon \\ X-Y-U}} I(Y;U). \tag{32}$$

Remark 3. *For any pair (X,Y) that satisfies the condition in Remark 2, from the proof of Theorem 2, we get*

$$g_\epsilon^M(X,Y) = \log M,\ \forall \epsilon > 0. \tag{33}$$

Therefore, the fact that mutual information may not be a suitable measure of utility (or privacy[8]) for the continuous alphabet scenarios is not only because the utility can be unbounded. Even when constrained to a finite alphabet $\mathcal{U}$, it can reach its supremum of $\log|\mathcal{U}|$ for arbitrarily small leakage, rendering the term "trade-off" pointless.[9] Therefore, in order to fully capture the utility-privacy trade-off for continuous alphabets, we can either use mutual information but impose additional constraints (such as the restrictions on the set of permissible $p_{U|Y}$ in (2) as in [27]), or adopt a different measure .

Remark 4. *For an arbitrary pair (X,Y), where $Y|\{X=x\}$ admits a bounded, and positive smooth pdf for any $x\in\mathcal{X}$, the supremum in (32) can be replaced by maximum[10] for $\epsilon>0$ as*

$$g_\epsilon^M(X,Y) = \max_{\substack{p_{U|Y}:|\mathcal{U}|\leq M \\ I(X;U)\leq\epsilon \\ X-Y-U}} I(Y;U) = \log M,\ \forall\epsilon>0.$$

This follows similar steps as in the first method in the prrof of Theorem 2 with the modification of replacing $\Phi^{-1}(\cdot)$ with the inverse function of $F_Y(\cdot)$.

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

For a pair of random variables $(X,Y)\in\mathcal{X}\times\mathcal{Y}$, the *private information about X carried by Y* is defined in [16] as

$$C_X(Y) \triangleq \min_{\substack{W:X-W-Y, \\ H(W|Y)=0}} H(W). \tag{34}$$

Since $H(W|Y)=0$ implies that W is a deterministic function of Y, (34) 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. It can be verified that $I(X;Y)\leq C_X(Y)\leq H(Y)$, where the first inequality is due to the data processing inequality applied on the Markov chain $X-W-Y$, i.e., $I(W;Y)\geq I(X;Y)$, and the second inequality is a direct result of the fact that $W=Y$ satisfies the constraints in (34).

The *non-private information about X carried by Y* is defined in [16] as

$$D_X(Y) \triangleq H(Y) - C_X(Y). \tag{35}$$

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 [16, Theorem 3] that the minimizer in (34) is $W^*=T^{\mathcal{X}}(Y)$; and hence,

$$D_X(Y) = H(Y) - H(T^{\mathcal{X}}(Y)). \tag{36}$$

Furthermore, it was proved in [16, lemma 5] that $C_X(Y)=H(Y)$, i.e., $D_X(Y)=0$, 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 [16], 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. In Theorem 3, we characterize the relation between $D_X(Y)$ and $g_0(X,Y)$. To this end, some preliminaries and two lemmas are needed, as explained in the sequel.

If $\mathbf{P}_{X|Y}$ has at least two identical columns, we define $\hat{\mathbf{P}}_{X|Y}$ as follows[11]. Let $\mathcal{E}_m\subset[|\mathcal{Y}|],\forall m\in[B]$, for some integer $B\geq 1$, be a set of indices corresponding to the columns in $\mathbf{P}_{X|Y}$ that are equal, i.e., $\mathbf{p}_{X|y_i}=\mathbf{p}_{X|y_j},\ \forall i,j\in\mathcal{E}_m,\forall m\in[B]$, and $\mathbf{p}_{X|y_i}\neq\mathbf{p}_{X|y_k},\ \forall i\in\mathcal{E}_m,\forall k\in[|\mathcal{Y}|]\backslash\mathcal{E}_m,\ \forall m\in[B]$. 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

$$\mathbf{P}_{X|Y}=\begin{bmatrix}0.3 & 0.3 & 0.4 & 0.5 & 0.4\\ 0.2 & 0.2 & 0.5 & 0.5 & 0.5\\ 0.5 & 0.5 & 0.1 & 0 & 0.1\end{bmatrix},\hat{\mathbf{P}}_{X|Y}=\begin{bmatrix}0.3 & 0.4 & 0.5\\ 0.2 & 0.5 & 0.5\\ 0.5 & 0.1 & 0\end{bmatrix}, \tag{37}$$

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

Since $\mathbf{p}_{X|y_i}=\mathbf{p}_{X|y_j},\forall i,j\in\mathcal{E}_m,\forall m\in[B]$, we have $T^{\mathcal{X}}(y_i)=T^{\mathcal{X}}(y_j),\forall i,j\in\mathcal{E}_m,\forall m\in[B]$. Hence, $T^{\mathcal{X}}(Y)$ is a random variable whose support has the cardinality $|\mathcal{Y}|-G+B$ and whose mass probabilities are the elements of the following set

$$\left\{\sum_{i\in\mathcal{E}_1}p_Y(y_i),\ldots,\sum_{i\in\mathcal{E}_B}p_Y(y_i)\right\}\cup\left\{p_Y(y_i)\middle| i\notin\cup_{m=1}^B\mathcal{E}_m\right\}. \tag{38}$$

Let $\mathbb{S}'_{X,Y}$ be a set of $\prod_{i=1}^B|\mathcal{E}_i|$ probability vectors in the simplex $\mathcal{P}(\mathcal{Y})$ given by

$$\mathbb{S}'_{X,Y}=\left\{\mathbf{s}_{m_{[B]}}\middle|\forall m_{[B]}\in\prod_{i=1}^B\mathcal{E}_i\right\}, \tag{39}$$

where the tuple $(m_1,m_2,\ldots,m_B)$ is written in short as $m_{[B]}$ and the probability vectors $\mathbf{s}_{m_{[B]}}$ are defined element-wise as

$$s_{m_{[B]}}(k)=\begin{cases}\sum_{t\in\mathcal{E}_i}p_Y(y_t) & k=m_i,i\in[B]\\ p_Y(y_k) & k\notin\cup_{i=1}^B\mathcal{E}_i\\ 0 & \text{otherwise}\end{cases}, \tag{40}$$

[8]For example, if mutual information is used as the privacy measure, while the total variation (TV) distance is used as the utility measure, i.e., $T(Y;U)\triangleq d_{\text{TV}}(F_{Y,U},F_Y\cdot F_U)$, the maximum utility (which is 2 for TV distance) can be achieved for any $\epsilon>0$ for jointly Gaussian (X,Y). Hence, this phenomenon is not restricted mutual information as a utility measure. This, however, is not the case when MMSE or the probability of error is used as the utility measure.

[9]In general, $g_\epsilon^M(X,Y)$ is not a continuous function of ϵ, which results from two facts i) relative entropy is lower semi-continuous, and so is mutual information, ii) supremum of continuous functions is itself lower semi-continuous. Hence, having $g_0^M(X,Y)=0$, while $g_\epsilon^M(X,Y)=\log M$ for $\epsilon>0$ is permissible. However, for the finite alphabet scenarios, all the probability simplices are compact, and it can be shown that g_ϵ is continuous.

[10]This holds in spite of the non-compactness of the search space.

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

for all $k \in [|\mathcal{Y}|]$, and $\forall m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i$.

Lemma 3. *For the set $\mathbb{S}'_{X,Y}$ in (39) and the set $\mathbb{S}_{X,Y}$ in (4), we have $\mathbb{S}'_{X,Y} \subseteq \mathbb{S}_{X,Y}$ and $H(\mathbf{s}) = H(T^{\mathcal{X}}(Y)),\ \forall \mathbf{s} \in \mathbb{S}'_{X,Y}$. Furthermore, the probability vector $\mathbf{p}_Y$ can be written as a convex combination of the points in $\mathbb{S}'_{X,Y}$, i.e.*

$$\mathbf{p}_Y = \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} \mathbf{s}_{m_{[B]}}, \tag{41}$$

where $\alpha_{m_{[B]}} \geq 0, \forall m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i$ and $\sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} = 1$.

Proof. The proof is provided in [21]. □

For example, assume that in the example in (37), we have $\mathbf{p}_Y = \begin{bmatrix}0.1 & 0.2 & 0.15 & 0.25 & 0.3\end{bmatrix}^T$. We can write $\mathbf{p}_Y = \frac{1}{9}\mathbf{s}_{1,3} + \frac{2}{9}\mathbf{s}_{1,5} + \frac{2}{9}\mathbf{s}_{2,3} + \frac{4}{9}\mathbf{s}_{2,5}$, where $\mathbf{s}_{1,3} = \begin{bmatrix}0.3 & 0 & 0.45 & 0.25 & 0\end{bmatrix}^T$, $\mathbf{s}_{1,5} = \begin{bmatrix}0.3 & 0 & 0 & 0.25 & 0.45\end{bmatrix}^T$, $\mathbf{s}_{2,3} = \begin{bmatrix}0 & 0.3 & 0.45 & 0.25 & 0\end{bmatrix}^T$, and $\mathbf{s}_{2,5} = \begin{bmatrix}0 & 0.3 & 0 & 0.25 & 0.45\end{bmatrix}^T$.

Lemma 4. *If $\text{nul}(\hat{\mathbf{P}}_{X|Y}) = 0$, we have $\text{ext}(\mathbb{S}_{X,Y}) = \mathbb{S}'_{X,Y}$. Otherwise, none of the elements in $\mathbb{S}'_{X,Y}$ belongs to $\text{ext}(\mathbb{S}_{X,Y})$, where $\text{ext}(\mathbb{S}_{X,Y})$ denotes the set of extreme points of $\mathbb{S}_{X,Y}$.*

Proof. The proof is provided in [21]. □

Theorem 3. *For a pair of random variables $(X, Y) \in \mathcal{X} \times \mathcal{Y}$, we have*

$$g_0(X,Y) \geq D_X(Y), \tag{42}$$

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

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

Proof. The proof of the inequality in (42) is as follows. It is obvious that when there exist no $y_1, y_2 \in \mathcal{Y}$ such that $\mathbf{p}_{X|Y}(\cdot|y_1) = \mathbf{p}_{X|Y}(\cdot|y_2)$, we have $D_X(Y) = 0$, and (42) holds from the non-negativity of $g_0(X,Y)$. Assume that there exist index sets $\mathcal{E}_m, \forall m \in [B]$, corresponding to equal columns of $\mathbf{P}_{X|Y}$, as defined before. We can write

$$\begin{aligned}
g_0(X,Y) &= H(Y) - \min_{\substack{F_U(\cdot),\ \mathbf{p}_{Y|u} \in \mathbb{S}_{X,Y},\ \forall u \in \mathcal{U}: \\ \int_{\mathcal{U}} \mathbf{p}_{Y|u} dF(u) = \mathbf{p}_Y}} H(Y|U) && (43)\\
&\geq H(Y) - \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} H(\mathbf{s}_{m_{[B]}}) && (44)\\
&= H(Y) - \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} H(T^{\mathcal{X}}(Y)) && (45)\\
&= H(Y) - H(T^{\mathcal{X}}(Y))\\
&= D_X(Y), && (46)
\end{aligned}$$

where (43) is from (7); (44) is justified as follows. According to lemma 3, $\mathbb{S}'_{X,Y} \subseteq \mathbb{S}_{X,Y}$, and $\mathbf{p}_Y$ is preserved from (41). Hence, the vectors in $\mathbb{S}'_{X,Y}$ belong to the constraint of the minimization in (43), and the inequality follows. (45) is from lemma 3, and (46) is due to (36). This proves the inequality (42).

The proof of the sufficient conditions for the equality in (42) is as follows. If $\text{nul}(\hat{\mathbf{P}}_{X|Y}) = 0$, from lemma 4, we can say that for any vector $\mathbf{s}$ that is an extreme point of $\mathbb{S}_{X,Y}$, we have $H(\mathbf{s}) = H(T^{\mathcal{X}}(Y))$, which means that

$$\min_{\substack{F_U(\cdot),\ \mathbf{p}_{Y|u} \in \mathbb{S}_{X,Y},\ \forall u \in \mathcal{U}: \\ \int_{\mathcal{U}} \mathbf{p}_{Y|u} dF(u) = \mathbf{p}_Y}} H(Y|U) = H(T^{\mathcal{X}}(Y)).$$

This is equivalent to $g_0(X,Y) = D_X(Y)$, from (7) and (36).

The proof of the necessary conditions for the equality in (42) is as follows. Assume that $g_0(X,Y) = D_X(Y)$. If $g_0(X,Y) = 0$, we have that perfect privacy is not feasible and the proof is complete. However, if $g_0(X,Y) > 0$, we must have $D_X(Y) > 0$, according to our assumption of $g_0(X,Y) = D_X(Y)$. In this case, as in [16], there must exist index sets $\mathcal{E}_m, \forall m \in [B]$, corresponding to equal columns of $\mathbf{P}_{X|Y}$. We prove that $\text{nul}(\hat{\mathbf{P}}_{X|Y}) = 0$ by contradiction. Assume that $\text{nul}(\hat{\mathbf{P}}_{X|Y}) \neq 0$. From Proposition 4, we conclude that none of the elements in $\mathbb{S}'_{X,Y}$ is an extreme point of $\mathbb{S}_{X,Y}$. In other words, for any $\mathbf{s}$ in $\mathbb{S}'_{X,Y}$, which is also a member of $\mathbb{S}_{X,Y}$ according to lemma 3, we can find the triplet $(\mathbf{s}', \mathbf{s}'', \beta)$, such that $\mathbf{s} = \beta\mathbf{s}' + (1-\beta)\mathbf{s}''$, where $\mathbf{s}', \mathbf{s}'' \in \mathbb{S}_{X,Y}$ ($\mathbf{s}' \neq \mathbf{s}''$) and $\beta \in (0,1)$. Therefore,

$$\begin{aligned}
H(T^{\mathcal{X}}(Y)) &= \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} H(\mathbf{s}_{m_{[B]}})\\
&= \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \alpha_{m_{[B]}} H\left(\beta_{m_{[B]}} \mathbf{s}'_{m_{[B]}} + (1-\beta_{m_{[B]}}) \mathbf{s}''_{m_{[B]}}\right)\\
&> \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} \beta_{m_{[B]}} \alpha_{m_{[B]}} H(\mathbf{s}'_{m_{[B]}})\\
&\quad + \sum_{m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i} (1-\beta_{m_{[B]}}) \alpha_{m_{[B]}} H(\mathbf{s}''_{m_{[B]}}) && (47)\\
&\geq \min_{\substack{F_U(\cdot),\ \mathbf{p}_{Y|u} \in \mathbb{S}_{X,Y},\ \forall u \in \mathcal{U}: \\ \int_{\mathcal{U}} \mathbf{p}_{Y|u} dF(u) = \mathbf{p}_Y}} H(Y|U), && (48)
\end{aligned}$$

where (47) is due to the strict concavity of the entropy; (48) comes from the fact that $\mathbf{s}'_{m_{[B]}}$ and $\mathbf{s}''_{m_{[B]}}$ with corresponding mass probabilities $\beta_{m_{[B]}}\alpha_{m_{[B]}}$ and $(1-\beta_{m_{[B]}})\alpha_{m_{[B]}}$, $\forall m_{[B]} \in \prod_{i=1}^{B} \mathcal{E}_i$, belong to the constraints of minimization in (48). This results in $g_0(X,Y) > D_X(Y)$, which is a contradiction. Hence, we must have $\text{nul}(\hat{\mathbf{P}}_{X|Y}) = 0$.

□

V. Full data observation model

In this section, we assume that the curator has access to both X and Y, and investigate $G_\epsilon(X,Y)$, as defined in (3), at $\epsilon = 0$.

Define the support of a given pair (X,Y) as

$$\text{supp}(X,Y) \triangleq \left\{(x,y) \in \mathcal{X} \times \mathcal{Y} \middle| p_{X,Y}(x,y) > 0\right\}.$$

Proposition 2. *In the evaluation of $G_0(X,Y)$, we must have*

$$\max_x |\{y \in \mathcal{Y} | p(y|x) > 0\}| \leq |\mathcal{U}| \leq |\text{supp}(X,Y)| - |\mathcal{X}| + 1, \tag{49}$$

where the first and second inequalities are necessary and sufficient conditions, respectively.

Proof. The proof is provided in Appendix B. □

Theorem 4. *Perfect privacy is feasible in the full data observation model, i.e., $G_0(X,Y) > 0$, if and only if Y is not a deterministic function of X.*

Proof. If Y is a deterministic function of X, we have $Y - X - U$ form a Markov chain. From data processing inequality, $I(X;U) = 0$ results in $I(Y;U) = 0$. This proves the first direction of the theorem.

For the second direction, we proceed as follows. 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

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

$$p_{X,Y|U}(x,y|u_2) = 2p_{X,Y}(x,y) - p_{X,Y|U}(x,y|u_1),\ \forall (x,y).$$

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 $Y \not\perp\!\!\!\perp U$.

In the light of Proposition 2, an alternative proof for this Theorem is provided as follows. If Y is a deterministic function of X, we have $|\text{supp}(X,Y)| = |\mathcal{X}|$, which results in $|\mathcal{U}| \leq 1$ according to Proposition 2, which in turn results in $G_0(X,Y) = 0$. If Y is not a deterministic function of X, we have $|\{y \in \mathcal{Y} | p(y|x) > 0\}| \geq 2$ for some $x \in \mathcal{X}$, which results in the necessity of having $|\mathcal{U}| \geq 2$ according to Proposition 2, which in turn results in $G_0(X,Y) > 0$ (since otherwise, $|\mathcal{U}| = 1$ is sufficient, and $|\mathcal{U}| \geq 2$ is not necessary, which is a contradiction). □

In Theorem 5, a lower bound is provided on the utility of the full data observation model. Prior to that, a quantity, which is used in the sequel, needs to be defined and investigated.

Definition 3. *For a given pair (X,Y), define the mapping $J : \mathcal{X} \times \text{int}(\mathcal{P}(\mathcal{Y})) \longrightarrow (0,1]$ as*

$$J(x, q_Y) \triangleq \frac{1}{\max_{y \in \mathcal{Y}} \frac{p_{Y|X}(y|x)}{q_Y(y)}}. \tag{50}$$

Therefore, we have that $J(x,q_Y) = q_Y(x)$, if $X = Y$, and $J(x,q_Y) = \min_y \frac{q_Y(y)}{p_Y(y)}$, if $X \perp\!\!\!\perp Y$. The following Proposition relates the above quantity to the *maximal leakage* from Y to X defined as ([20])

$$\mathcal{L}(Y \to X) \triangleq \log \sum_{x \in \mathcal{X}} \max_{\substack{y \in \mathcal{Y} \\ p_Y(y) > 0}} p_{X|Y}(x|y). \tag{51}$$

Proposition 3. *For a given pair (X,Y), we have*

$$\mathbb{E}_X[J(X,p_Y)] \geq 2^{-\mathcal{L}(Y \to X)}, \tag{52}$$

where equality holds if and only if $\max_{y \in \mathcal{Y}} \frac{p_{Y|X}(y|x)}{p_Y(y)}$ does not vary with $x \in \mathcal{X}$, which includes the special cases of i) $X \perp\!\!\!\perp Y$, and ii) $X = Y$ and uniformly distributed.

Proof. The proof is provided in Appendix C. □

Theorem 5. *For a given pair $(X,Y) \in \mathcal{X} \times \mathcal{Y}$, we have*

$$G_0(X,Y) \geq \left(\mathbb{E}_X[J(X,q^*)] \log |\mathcal{Y}| - 1\right)^+, \tag{53}$$

where q^* denotes the uniform pmf over $\mathcal{Y}$.

Proof. Fix an arbitrary pmf in the interior of $\mathcal{P}(\mathcal{Y})$ and denote it by q^*. A privacy-preserving mapping $p_{U|X,Y}$ is designed such that the conditional pmf of U conditioned on $\{X = x\}$ is the same as q^* for any $x \in \mathcal{X}$. Therefore, for this privacy-preserving mapping, we have that $X \perp\!\!\!\perp U$, and the resulting $I(Y;U)$ serves as a lower bound on $G_0(X,Y)$. The only reason for selecting q^* as the uniform pmf over $\mathcal{Y}$ is that its corresponding $I(Y;U)$ can be further lower bounded in a closed form way.

Let $\Theta_X \in \{0,1\}$ be a Bernoulli r.v., parametrized by X, with $\Pr\{\Theta_X = 1 | X = x\} = J(x,q^*),\ \forall x \in \mathcal{X}$. The privacy-preserving mapping is designed as $U \triangleq \Theta_X Y + (1 - \Theta_X)\tilde{Y}_X$, where for each $x \in \mathcal{X}$, $\tilde{Y}_x$ is an r.v. over $\mathcal{Y}$, which is distributed according to $\frac{q^*(\cdot) - J(x,q^*)p_{Y|X}(\cdot|x)}{1 - J(x,q^*)}$, when $J(x,q^*) < 1$, and arbitrarily distributed when $J(x,q^*) = 1$.[12] Conditioned on each realization $x \in \mathcal{X}$, we have that $p_{U|X}(\cdot|x)$ is a convex combination of $p_{Y|X}(\cdot|x)$ and $\frac{q^*(\cdot) - J(x,q^*)p_{Y|X}(\cdot|x)}{1 - J(x,q^*)}$ with weights $J(x,q^*)$, and $1 - J(x,q^*)$, respectively. Hence, U conditioned on $\{X = x\}$, i.e., $U|\{X = x\}$, is distributed according to q^* for each $x \in \mathcal{X}$, and therefore, $X \perp\!\!\!\perp U$. Since q^* is an arbitrary point in $\text{int}(\mathcal{P}(\mathcal{Y}))$, we have

$$\begin{aligned}
G_0(X,Y) &\geq \max_{q^*} I(Y;U) \\
&= \max_{q^*} \{H(U) - H(U|Y)\} \\
&\geq \max_{q^*} \{H(q^*) - H(U|Y,\Theta_X) - H(\Theta_X|Y)\} \\
&\geq \max_{q^*} \{H(q^*) - H(U|Y,\Theta_X)\} - 1 \qquad (54) \\
&= \max_{q^*} \{H(q^*) - H(\Theta_X Y + (1-\Theta_X)\tilde{Y}_X | Y, \Theta_X)\} - 1 \\
&= \max_{q^*} \{H(q^*) - \Pr\{\Theta_X = 1\} H(\tilde{Y}_X | Y, \Theta_X = 0)\} - 1 \qquad (55) \\
&\geq \Pr\{\Theta_X = 1\} \log |\mathcal{Y}| - 1 \qquad (56) \\
&= \mathbb{E}_X[J(X,q^*)] \log |\mathcal{Y}| - 1,
\end{aligned}$$

where (54) follows from Θ_X being binary, and (55) results from $H(U|Y,\Theta_X = 1) = 0$. In (56), we pick the uniform q^*, and use the fact that $H(\tilde{Y}_X|Y,\Theta_X = 0) \leq \log |\mathcal{Y}|$. [13]

□

Considering the output perturbation model, Theorem 2 proved that perfect privacy is not feasible for the (correlated) jointly Gaussian pair. This is not the case in the full data observation model. By using the following Theorem, which includes a broad range of joint distributions on $\mathcal{X} \times \mathcal{Y}$, we show that $G_0(X,Y)$ is actually unbounded, which is stated in corollary 6.1.

[12] Note that when $J(x,q^*) = 1$, we have $\Theta_x = 1$, and therefore, the coefficient of $\tilde{Y}_x$ in U becomes zero.

[13] Note that if $\tilde{Y}_x$ and Y are not independent, the bound $H(\tilde{Y}_X|Y,\Theta_X = 0) \leq \log |\mathcal{Y}|$ can be further tightened. This, in turn, calls for an algorithmic approach to this problem that aims to maximize $I(Y;\tilde{Y}_x)$ over the joint distribution for fixed marginals.

Theorem 6. *For the class of additive noise, i.e., when $Y = X+N$ (where N is not necessarily independent of X), if there exists $\alpha, \beta \in \mathbb{R}$, and $\Delta > 0$ such that $\mathcal{I} \triangleq (\mathcal{X} \cap [\alpha, \alpha+\Delta]) \times [\beta, \beta+\Delta] \subset \mathcal{X} \times \mathcal{N}$, and $N|\{X = x\}$ admits a bounded smooth density over $[\beta, \beta+\Delta]$ for each $x \in \mathcal{X} \cap [\alpha, \alpha+\Delta]$, we have*

$$G_0(X,Y) = \infty. \tag{57}$$

Proof. The sketch of the proof is as follows. A privacy-preserving mapping $p_{U|X,Y}$ is designed whose output, i.e., U is independent of X. Therefore, we conclude that $G_0(X,Y)$ is lower bounded by the utility, i.e., $I(Y;U)$, of this privacy-preserving mapping. By showing that the latter can grow unboundedly, the proof of (57) is complete.

Define the Bernoulli r.v. E, which is 1 when $(X,N) \in \mathcal{I}$, and 0 elsewhere. Define

$$f_x^* \triangleq \max_{n \in [\beta, \beta+\Delta]} f_{N|X,E}(n|x,1), \ \forall x \in \mathcal{X} \cap [\alpha, \alpha+\Delta]. \tag{58}$$

which is defined by the assumption of having a bounded $f_{N|X,E}(\cdot|x,1)$ over $[\beta, \beta+\Delta]$ for each $x \in \mathcal{X} \cap [\alpha, \alpha+\Delta]$. Moreover, we have $f_x^* \geq \frac{1}{\Delta}$, since otherwise $f_{N|X,E}(\cdot|x,1)$ will not integrate to 1 over its support, i.e., $[\beta, \beta+\Delta]$. Also, we have $f_x^* = \frac{1}{\Delta}$ if and only if $f_{N|X,E}(\cdot|x,1)$ is uniform.

Let $\tilde{N}_x$ be a continuous r.v., independent of (X,N), with the following density

$$f_{\tilde{N}_x}(t) = \frac{f_x^* - f_{N|X,E}(t|x,1)}{\Delta f_x^* - 1},$$

for all $(x,t) \in (\mathcal{X} \cap [\alpha, \alpha+\Delta]) \times [\beta, \beta+\Delta]$, when $f_x^* > \frac{1}{\Delta}$, and arbitrarily distributed when $f_x^* = \frac{1}{\Delta}$.

Define the Bernoulli r.v. $\Theta_X \in \{0,1\}$, with $\Pr\{\Theta_X = 1|X = x\} = \frac{1}{\Delta f_x^*}$, and set $R \triangleq \Theta_X N + (1-\Theta_X)\tilde{N}_X$. We have $R \sim \text{Uniform}[\beta, \beta+\Delta]$, since for each $x \in \mathcal{X} \cap [\alpha, \alpha+\Delta]$, the pdf of R, which is a convex combination of $f_{N|X,E}(\cdot|x,1)$ and $f_{\tilde{N}_x}(\cdot)$, with the corresponding weights of $\frac{1}{\Delta f_x^*}$ and $1 - \frac{1}{\Delta f_x^*}$, is equal to $\frac{1}{\Delta}$ over $[\beta, \beta+\Delta]$.

Let M be an arbitrary positive integer and set

$$U \triangleq E\left(\left\lfloor \frac{M(X+R)}{\Delta} \right\rfloor \bmod M\right) + (1-E)\tilde{U}, \tag{59}$$

where $\tilde{U}$ is a uniform pmf over $[0 : M-1]$.

With some simple calculations, it can be verified that the conditional distribution of U conditioned on $\{X = x\}$ remains uniform over $[0 : M-1]$ for any realization $x \in \mathcal{X}$.[14] Hence, $U \perp\!\!\!\perp X$.

We can write

$$\begin{aligned} I(X+N;U) &= I(X+N, \Theta_X; U|E) + I(E;U) \\ &\quad I(E;U|X+N,\Theta_X) - I(\Theta_X;U|X+N) \\ &\geq I(X+N,\Theta_X;U|E) - 2 \qquad (60) \\ &\geq p_E(1) I(X+N,\Theta_X;U|E=1) - 2 \qquad (61) \\ &= p_E(1)\Big(H(U|E=1) - H(U|X+N,\Theta_X,E=1)\Big) \\ &\quad - 2 \\ &\geq p_E(1)\Pr\{\Theta_X = 1\}\log M - 2 \\ &\quad - p_E(1)\Pr\{\Theta_X = 1\} H(U|X+N,\Theta_X=1,E=1) \qquad (62) \\ &\geq p_E(1)\Pr\{\Theta_X = 1\}\log M - 2, \qquad (63) \end{aligned}$$

where (60), and (61) follow, respectively, from the facts that E, Θ_X are binary (having a maximum entropy of 1), and mutual information is non-negative; (62) and (63) result, respectively, from having $H(U|X+N,\Theta_X = 0, E=1) \leq \log M$, and $H(U|X+N,\Theta_X=1,E=1) = 0$.

Finally, as $\Pr\{\Theta_X = 1\} = \int_{x \in \mathcal{X} \cap [\alpha, \alpha_\Delta]} \frac{1}{\Delta f_x^*} dF_X(x) > 0$, by letting $M \to \infty$ in (63), (57) is proved [15].

□

Corollary 6.1. *For the jointly Gaussian pair (X,Y), $G_0(X,Y)$ is unbounded.*

VI. Asymptotic analysis

In the previous sections, we have mainly focused on one extreme point of the utility-privacy trade-off curve, corresponding to perfect privacy either in the output perturbation or full data observation models. In general, characterizing the whole of this trade-off curve is analytically challenging. Therefore, to better understand the fundamental trade-off between utility and privacy, we will next consider the output perturbation model, and study the slope of $g_\epsilon(X,Y)$ as $\epsilon \to 0$. This will reveal us how much utility we can gain at the expense of a small amount of privacy leakage. The analysis depends on whether perfect privacy is feasible or not.

Consider a pair of random variables $(X,Y) \in \mathcal{X} \times \mathcal{Y}$ distributed according to $\mathbf{P}_{X,Y}$, with the marginals $\mathbf{p}_X$ and $\mathbf{p}_Y$. The matrix $\mathbf{P}_{X|Y}$ can be viewed as a channel with input Y and output X. When the input of this channel is distributed according to $\mathbf{q}_Y$, the output is distributed according to $\mathbf{q}_X = \mathbf{P}_{X|Y}\mathbf{q}_Y$.

Define $r : \mathcal{P}(\mathcal{Y}) \backslash \{\mathbf{p}_Y\} \to [0,1]$ as

$$r(\mathbf{q}_Y) \triangleq \frac{D(\mathbf{q}_X||\mathbf{p}_X)}{D(\mathbf{q}_Y||\mathbf{p}_Y)}. \tag{64}$$

Let $V^* \in [1, +\infty]$ be defined as

$$V^* \triangleq \sup_{\substack{\mathbf{q}_Y: \\ \mathbf{q}_Y \neq \mathbf{p}_Y}} \frac{1}{r(\mathbf{q}_Y)} = \sup_{\substack{\mathbf{q}_Y: \\ \mathbf{q}_Y \neq \mathbf{p}_Y}} \frac{D(\mathbf{q}_Y||\mathbf{p}_Y)}{D(\mathbf{q}_X||\mathbf{p}_X)}, \tag{65}$$

with the convention that if for some $\mathbf{q}_Y (\neq \mathbf{p}_Y)$, we have $\mathbf{q}_X = \mathbf{p}_X$, then $V^* = +\infty$.

Proposition 4. *We have $g_0(X,Y) = 0$ if and only if $V^* < +\infty$.*[16]

Proof. The proof is provided in [21]. □

[14] The fact that $U|\{X = x, E = 1\}$ is uniform over $[0 : M-1]$ is immediate from noting that $\left(\lfloor \frac{MA}{\Delta} \rfloor \bmod M\right)$ is uniform if A is uniform. The uniform distribution of $U|\{X = x, E = 0\}$ is immediate from construction.

[15] Note that the set of all mappings $p_{U|\{X,N\}}$ can be put into a one-to-one correspondence with the set of all mappings $q_{U|\{X,X+N\}}$.

[16] A claim, similar to this proposition, is provided in [14]; however, the proof is incomplete as elaborated in [21].

A. Perfect privacy is not feasible.

If perfect privacy is not feasible, i.e., $g_0(X,Y)=0$, then the slope of $g_\epsilon(X,Y)$ at $\epsilon=0$ is equal to V^* as shown in [14]. However, V^* itself is written as a supremization, and hence, practical approximations of the this slope based on the properties of the joint distribution $p_{X,Y}$ is of interest. The following Theorem provides a lower bound on this slope.

Let $\hat{\mathcal{Y}}$ denote the set of all the subsets of $\mathcal{Y}$ excluding the empty set and $\mathcal{Y}$, i.e., $\hat{\mathcal{Y}} \triangleq \{\mathcal{V}|\mathcal{V}\subset\mathcal{Y}\}-\{\mathcal{Y},\emptyset\}$.

Theorem 7. *We have*

$$\lim_{\epsilon\to 0}\frac{g_\epsilon(X,Y)}{\epsilon} \geq \max\{A(X,Y),B_0(X,Y)\} \tag{66}$$

$$\geq \frac{H(Y)}{I(X;Y)}, \tag{67}$$

where

$$A(X,Y) \triangleq \max_{\mathcal{B}:\mathcal{B}\in\hat{\mathcal{Y}}} \frac{-\log\left(\sum_{y\in\mathcal{B}}p_Y(y)\right)}{D\left(\frac{\sum_{y\in\mathcal{B}}p_Y(y)p_{X|Y}(\cdot|y)}{\sum_{y\in\mathcal{B}}p_Y(y)}\middle\|p_X(\cdot)\right)}, \tag{68}$$

$$B_\alpha(X,Y) \triangleq \frac{(H(Y)-1)^+-\alpha}{I(X;Y)-\max_{x\in\mathcal{X}}p(x)D(p_{Y|X}(\cdot|x)||p_Y(\cdot))}, \tag{69}$$

for $\alpha\geq 0$.

Proof. It is known that the LHS of (66) equals V^* (see [14]). Fix an arbitrary $\mathcal{B}\in\hat{\mathcal{Y}}$, and define the pmf q'_Y as

$$q'_Y(y) \triangleq \frac{p_Y(y)}{\sum_{t\in\mathcal{B}}p_Y(t)}\cdot\mathbb{1}_{\{y\in\mathcal{B}\}}. \tag{70}$$

From (65), we have

$$V^* \geq \frac{D(\mathbf{q}'_Y||\mathbf{p}_Y)}{D(\mathbf{P}_{X|Y}\mathbf{q}'_Y||\mathbf{p}_X)} \tag{71}$$

$$= \frac{-\log\left(\sum_{y\in\mathcal{B}}p_Y(y)\right)}{D\left(\frac{\sum_{y\in\mathcal{B}}p_Y(y)p_{X|Y}(\cdot|y)}{\sum_{y\in\mathcal{B}}p_Y(y)}\middle\|p_X(\cdot)\right)}, \tag{72}$$

where (71) follows from the definition in (65) and the fact that $q'_Y(\cdot)\neq p_Y(\cdot)$, since $\mathcal{Y}\notin\hat{\mathcal{Y}}$. Since (72) is valid for any $\mathcal{B}\in\hat{\mathcal{Y}}$, by taking its maximum over $\mathcal{B}\in\hat{\mathcal{Y}}$, we have

$$\lim_{\epsilon\to 0}\frac{g_\epsilon(X,Y)}{\epsilon}\geq A(X,Y). \tag{73}$$

Let $x^*\triangleq\arg\max_{x\in\mathcal{X}}p(x)D(p_{Y|X}(\cdot|x)||p_Y(\cdot))$. Define the binary r.v. $\hat{X}$ as a deterministic function of X given by $\hat{x}(x)\triangleq\mathbb{1}_{\{x=x^*\}}$. Hence, we have $\hat{X}-X-Y$ form a Markov chain. From Corollary 1.2, we have $g_0(\hat{X},Y)\geq(H(Y)-\log\text{rank}(\mathbf{P}_{\hat{X}|Y}))^+=(H(Y)-1)^+$. Therefore, there exists a privacy-preserving mapping $p_{U|Y}$, such that $\hat{X}-X-Y-U$ form a Markov chain, $I(Y;U)\geq(H(Y)-1)^+$, and $I(\hat{X};U)=0$. We have

$$I(X;U)=\sum_{x\in\mathcal{X}}p(x)D\left(p_{U|X}(\cdot|x)||p_U(\cdot)\right)$$

$$=p(x^*)D\left(p_{U|X}(\cdot|x^*)||p_U(\cdot)\right)+\sum_{x\in\mathcal{X}\setminus\{x^*\}}p(x)D\left(p_{U|X}(\cdot|x)||p_U(\cdot)\right)$$

$$=p_{\hat{X}}(1)D\left(p_{U|\hat{X}}(\cdot|1)||p_U(\cdot)\right)+\sum_{x\in\mathcal{X}\setminus\{x^*\}}p(x)D\left(p_{U|X}(\cdot|x)||p_U(\cdot)\right) \tag{74}$$

$$=\sum_{x\in\mathcal{X}\setminus\{x^*\}}p(x)D\left(p_{U|X}(\cdot|x)||p_U(\cdot)\right) \tag{75}$$

$$\leq\sum_{x\in\mathcal{X}\setminus\{x^*\}}p(x)D\left(p_{Y|X}(\cdot|x)||p_Y(\cdot)\right) \tag{76}$$

$$=I(X;Y)-p(x^*)D\left(p_{Y|X}(\cdot|x^*)||p_Y(\cdot)\right), \tag{77}$$

where (74) follows from $\hat{x}(x)\triangleq\mathbb{1}_{x=x^*}$, which results in $p_{\hat{X}}(1)=p_X(x^*)$, and $p_{U|\hat{X}}(\cdot|1)=p_{U|X}(\cdot|x^*)$; (75) follows from having $U\perp\!\!\!\perp\hat{X}$, and hence, $p_{U|\hat{X}}(\cdot|1)=p_U(\cdot)$, and (76) results from the data processing inequality by viewing two pmfs $p_{Y|X}(\cdot|x)$ and $p_Y(\cdot)$ entering the channel $p_{U|Y}$.

In this construction, we have $X-Y-U$ form a Markov chain and a point with utility of at least $(H(Y)-1)^+$, and privacy leakage of at most $I(X;Y)-p(x^*)D\left(p_{Y|X}(\cdot|x^*)||p_Y(\cdot)\right)$ is achievable in the utility-privacy trade-off curve. By noting the concavity of $g_\epsilon(X,Y)$ in ϵ (see [15, lemma 2]), the slope at $(0,0)$ is lower bounded by the slope of the straight line connecting this point to the origin. Hence,

$$\lim_{\epsilon\to 0}\frac{g_\epsilon(X,Y)}{\epsilon}\geq B_0(X,Y). \tag{78}$$

From (73), (78), the lower bound in (66) is proved. As a special case, if in (68), $\mathcal{B}$ is restricted to the space of singletons, i.e., subsets of $\mathcal{Y}$ with only one element, we get a lower bound on the slope at origin as

$$\max_{y\in\mathcal{Y}}\frac{-\log\left(p_Y(y)\right)}{D\left(p_{X|Y}(\cdot|y)||p_X(\cdot)\right)},$$

which is proved differently in [15, lemma 19], and shown to satisfy the inequality in (67). □

Thus far, we have observed that when perfect privacy is not feasible, the slope of the trade-off curve is finite. In other words, for a vanishingly small privacy leakage, only a linearly proportional vanishingly small utility can be attained. This is not necessarily the case when perfect privacy is feasible, which is discussed next.

B. Perfect privacy is feasible.

For a given pair (X,Y), assume that $g_0(X,Y)$, obtained through the LP formulation in Theorem 1, is achieved by

$$U^*\in\mathcal{U}^*=\{u_1^*,u_2^*,\ldots,u_{|\mathcal{U}^*|}^*\},\ \mathbf{p}_{Y|u^*},\forall u^*\in\mathcal{U}^*, \tag{79}$$

where the vectors $\mathbf{p}_{Y|u^*},\ \forall u^*\in\mathcal{U}^*$ belong to the extreme points of the set $\mathbb{S}_{X,Y}$, as in (4).

Definition 4. *Define*

$$\psi(u^*)\triangleq\sup_{\substack{\mathbf{q}_Y:\\ 0<D(\mathbf{q}_Y||\mathbf{p}_{Y|u^*})<+\infty}}\frac{D(\mathbf{q}_Y||\mathbf{p}_{Y|u^*})}{D(\mathbf{q}_X||\mathbf{p}_X)},\ \forall u^*\in\mathcal{U}^*, \tag{80}$$

and if for some u^*, there is no $\mathbf{q}_Y$ for which $0 < D(\mathbf{q}_Y||\mathbf{p}_{Y|u^*}) < +\infty$ (which happens exactly when $\mathbf{p}_{Y|u^*}$ is a corner point of the probability simplex), then let $\psi(u^*) \triangleq 0$. Therefore, in order to evaluate $\psi(u^*)$ for some $u^* \in \mathcal{U}^*$, the search space in (80) includes the set of all probability vectors in $\mathcal{P}(\mathcal{Y})$ such that i) they are not equal to the extreme point $\mathbf{p}_{\mathbf{Y}|\mathbf{u}^*}$ (equivalently, $0 < D(\mathbf{q}_Y||\mathbf{p}_{Y|u^*})$), ii) if $\mathbf{p}_{\mathbf{Y}|\mathbf{u}^*}$ has a zero entry, they will also have a zero in the same entry (equivalently, $D(\mathbf{q}_Y||\mathbf{p}_{Y|u^*}) < +\infty$).

The following lemma is needed in the sequel.

Lemma 5. *We have $\psi(u^*) < +\infty,\ \forall u^* \in \mathcal{U}^*$.*

Proof. The proof is provided in [21]. □

Theorem 8. *For a given pair (X,Y), if perfect privacy is feasible, we have*

$$\lim_{\epsilon\to 0} \frac{g_\epsilon(X,Y) - g_0(X,Y)}{\epsilon} \geq \max\left\{L(X,Y), B_{g_0(X,Y)}(X,Y), \frac{H(Y)-g_0(X,Y)}{I(X;Y)}\right\}, \tag{81}$$

where $B_\alpha(X,Y)$ is defined in (69), and

$$L(X,Y) \triangleq \max_{u^*\in\mathcal{U}^*} \psi(u^*). \tag{82}$$

Proof. First, we note that from lemma 5, $L(X,Y)$ is well defined. Denote a/the maximizer of (82) by u_j^* for some $j \in [|\mathcal{U}^*|]$. From (80) and (82), for an arbitrary fixed $\delta > 0$, there exists $\mathbf{q}_Y \neq \mathbf{p}_{Y|u_j^*}$, such that $D(\mathbf{q}_Y||\mathbf{p}_{Y|u_j^*}) < +\infty$, and

$$L(X,Y) - \delta < \frac{D(\mathbf{q}_Y||\mathbf{p}_{Y|u_j^*})}{D(\mathbf{q}_X||\mathbf{p}_X)} \leq L(X,Y). \tag{83}$$

Construct the pair (Y,U) as follows. Let $\mathcal{U} = \{u_1, u_2, \ldots, u_{|\mathcal{U}^*|}, \hat{u}_j\}$, and for sufficiently small $\gamma > 0$, let $p_U(u_i) = p_{U^*}(u_i^*)$, $\mathbf{p}_{Y|u_i} = \mathbf{p}_{Y|u_i^*},\ \forall i \in [|\mathcal{U}^*|], i \neq j$, $p_U(u_j) = \gamma p_{U^*}(u_j^*)$, $p_U(\hat{u}_j) = (1-\gamma)p_{U^*}(u_j^*)$, and $\mathbf{p}_{Y|u_j} = \mathbf{q}_Y$, $\mathbf{p}_{Y|\hat{u}_j} = \frac{1}{1-\gamma}(\mathbf{p}_{Y|u_j^*} - \gamma\mathbf{q}_Y)$. Note that for sufficiently small $\gamma > 0$, $\mathbf{p}_{Y|\hat{u}_j}$ is a probability vector, as we have $D(\mathbf{q}_Y||\mathbf{p}_{Y|u_j^*}) < +\infty$. In other words, for any entry of the vector $\mathbf{p}_{Y|u_j^*}$ that is zero (note that it is an extreme point of $\mathbb{S}_{X,Y}$), the corresponding entry in $\mathbf{q}_Y$ is also zero. Finally, it can be verified that the marginal probability vector $\mathbf{p}_Y$ is also preserved.

With $I_\gamma(Y;U)$, and $I_\gamma(X;U)$ denoting the corresponding mutual information terms in this construction, and from the concavity of $g_\epsilon(X,Y)$ in ϵ (see [15]), the LHS of (81) is lower bounded by

$$\frac{I_\gamma(Y;U) - g_0(X,Y)}{I_\gamma(X;U)} = \frac{\sum_{u\in\mathcal{U}} p_U(u) D(\mathbf{p}_{Y|u}||\mathbf{p}_Y) - \sum_{u^*\in\mathcal{U}^*} p_{U^*}(u^*) D(\mathbf{p}_{Y|u^*}||\mathbf{p}_Y)}{\sum_{u\in\mathcal{U}} p_U(u) D(\mathbf{p}_{X|u}||\mathbf{p}_X)} = \frac{\sum_{u\in\{u_j,\hat{u}_j\}} p_U(u) D(\mathbf{p}_{Y|u}||\mathbf{p}_Y) - p_{U^*}(u_j^*) D(\mathbf{p}_{Y|u_j^*}||\mathbf{p}_Y)}{\sum_{u\in\{u_j,\hat{u}_j\}} p_U(u) D(\mathbf{p}_{X|u}||\mathbf{p}_X)} \tag{84}$$

$$= \frac{\gamma D(\mathbf{q}_Y||\mathbf{p}_Y) + (1-\gamma) D\left(\frac{1}{1-\gamma}(\mathbf{p}_{Y|u_j^*} - \gamma\mathbf{q}_Y)\middle|\middle|\mathbf{p}_Y\right) - D(\mathbf{p}_{Y|u_j^*}||\mathbf{p}_Y)}{\gamma D(\mathbf{q}_X||\mathbf{p}_X) + (1-\gamma) D\left(\frac{1}{1-\gamma}(\mathbf{p}_X - \gamma\mathbf{q}_X)\middle|\middle|\mathbf{p}_X\right)}, \tag{85}$$

where the numerator in (84) follows from the construction of (Y,U); the denominator in (84) is from the fact that $\mathbf{p}_{X|u_i} = \mathbf{p}_{X|u_i^*}, \forall i \in [|\mathcal{U}^*|], i \neq j$ and $\mathbf{p}_{X|u^*} = \mathbf{P}_{X|Y}\mathbf{p}_{Y|u^*} = \mathbf{p}_X, \forall u^* \in \mathcal{U}^*$; (85) follows from the construction of (Y,U). For three generic pmfs p on $\mathcal{Y}$, and q, r on $\tilde{Y} \subset \mathcal{Y}$, when $\gamma \to 0$, we can write

$$\begin{aligned} &D\left(\frac{1}{1-\gamma}(\mathbf{r} - \gamma\mathbf{q})\middle|\middle|\mathbf{p}\right) \\ &= \sum_{y\in\tilde{\mathcal{Y}}} \frac{r(y)-\gamma q(y)}{1-\gamma}\left[\log\frac{1}{1-\gamma} + \log\frac{r(y)-\gamma q(y)}{p(y)}\right] \\ &= -\log(1-\gamma) + \sum_{y\in\tilde{\mathcal{Y}}} \frac{r(y)-\gamma q(y)}{1-\gamma}\log\frac{r(y)}{p(y)}\left(1 - \gamma\frac{q(y)}{r(y)}\right) \\ &\approx -\log(1-\gamma) + \sum_{y\in\tilde{\mathcal{Y}}} \frac{r(y)-\gamma q(y)}{1-\gamma}\left[\log\frac{r(y)}{p(y)} - \gamma\frac{q(y)}{r(y)}\right] \end{aligned} \tag{86}$$

$$= \frac{1}{1-\gamma}\left(D(r||p) - \gamma\sum_{y\in\tilde{\mathcal{Y}}} q(y)\log\frac{r(y)}{p(y)} + O(\gamma^2)\right), \tag{87}$$

where in (86), the first order approximation, i.e., $\log(1+x) \approx x$ for $x \to 0$, is used[17].

Using the approximation in (87) for both of the second terms in the numerator and denominator of (85), after some manipulation, we get

$$\begin{aligned} \lim_{\epsilon\to 0}\frac{g_\epsilon(X,Y)-g_0(X,Y)}{\epsilon} &\geq \lim_{\gamma\to 0}\frac{I_\gamma(Y;U)-g_0(X,Y)}{I_\gamma(X;U)}, \\ &= \lim_{\gamma\to 0}\frac{\gamma D(\mathbf{q}_Y||\mathbf{p}_{Y|u_j^*}) + O(\gamma^2)}{\gamma D(\mathbf{q}_X||\mathbf{p}_X) + O(\gamma^2)}, \\ &= \frac{D(\mathbf{q}_Y||\mathbf{p}_{Y|u_j^*})}{D(\mathbf{q}_X||\mathbf{p}_X)}, \\ &> L(X,Y) - \delta, \end{aligned} \tag{88}$$

where (88) follows from (83). Since $\delta > 0$ was chosen arbitrarily, we have

$$\lim_{\epsilon\to 0}\frac{g_\epsilon(X,Y)-g_0(X,Y)}{\epsilon} \geq L(X,Y).$$

The second term in the RHS of (81) follows similarly to the discussion in the proof of Theorem 7. In other words, it is a lower bound for the slope of a straight line that connects $(0, g_0(X,Y))$ to a point with utility of at least $(H(Y)-1)^+$, and privacy leakage of at most $I(X;Y) - p(x^*)D\left(p_{Y|X}(\cdot|x^*)||p_Y(\cdot)\right)$. Finally, the second term in the RHS of (81) is the slope of the straight line connecting the

[17] Note that this is true only if the logarithm is natural, i.e., to the base of the mathematical constant e; however, since in this section, we are dealing with the ratios of mutual information terms, and hence the ratios of logarithms, this has no effect on the results, as we can multiply both the numerator and denominator by $\log_e 2$.

end points of the curve of $g_\epsilon(X,Y)$ vs. ϵ, i.e., $(0, g_0(X,Y))$ and $(I(X;Y), H(Y))$. The fact that these are lower bounds on the slope at the origin follow from the concavity of $g_\epsilon(X,Y)$ in ϵ.

□

Assume that the maximizer in $g_0(X,Y)$, i.e., (79), induces the joint distribution $p^*_{Y,U}(\cdot,\cdot)$. In what follows, it is shown that under certain conditions, when perfect privacy is feasible, the slope at origin is infinite. The following lemmas are needed in the sequel.

Lemma 6. *Let p, q denote two pmfs on $\mathcal{Y}$, and assume that $p(y) > 0,\ \forall y \in \mathcal{Y}$. We have*

$$\sum_y \frac{q^2(y)}{p(y)} \geq 1, \tag{89}$$

with equality if and only if $q = p$.

Proof. We have $\sum_y \frac{q^2(y)}{p(y)} = 1 + \sum_y \frac{(q(y)-p(y))^2}{p(y)} \geq 1$, with equality if and only if $p = q$.[18] □

Lemma 7. *For a given pair (X,Y), if there exists $y_0 \in \mathcal{Y}$, for which $p_{X|Y}(\cdot|y_0) = p_X(\cdot)$, there must exist $u_0^* \in \mathcal{U}^*$, defined in (79), such that*

$$p_{Y|U^*}(y_0|u_0^*) = 1,\ p_{Y|U^*}(y_0|u^*) = 0,\ \forall u^* \in \mathcal{U}^* \backslash \{u_0^*\}.$$

Proof. The proof is provided in Appendix D. □

Theorem 9. *If there exist $y_0 \in \mathcal{Y}$, and $u_0, u_1 \in \mathcal{U}^*$, such that $p^*(y_0,u_0), p^*(y_0,u_1) > 0$, and $p^*(y_0|u_0) \neq p^*(y_0|u_1)$, then*

$$\lim_{\epsilon \to 0} \frac{g_\epsilon(X,Y) - g_0(X,Y)}{\epsilon} = \infty. \tag{90}$$

Proof. Without loss of generality, assume that $p^*(y_0|u_0) > p^*(y_0|u_1)$. Consider the tuple (X,Y,U) distributed according to $p_{X|Y} \cdot p'_{Y,U}$, where $p'(y,u) = p^*(y,u) + \eta \cdot i(y,u)$, in which $i(y,u)$ is non-zero only for two cases: $i(y_0,u_0) = -i(y_0,u_1) = 1$. The value of $\eta > 0$ is chosen arbitrarily small such that $p'(y,u)$ is a pmf[19]. Therefore, we have the marginal pmf $p'(u) = p^*(u) + \eta \cdot i(y_0,u),\ \forall u \in \mathcal{U}$. It can be verified that with this construction, the marginal pmf $p_{X,Y}$ is preserved in the tuple (X,Y,U).

Let $I_\eta(Y;U)$ and $I_\eta(X;U)$ denote the mutual information terms induced by $p_{X|Y} \cdot p'_{Y,U}$. When $\eta \to 0$, we have

$$\begin{aligned}
&I_\eta(Y;U) = D(p'_{Y,U}||p_Y \cdot p'_U)\\
&= D\bigg(p^*(y,u) + \eta \cdot i(y,u)\bigg|\bigg|p(y)\left(p^*(u) + \eta \cdot i(y_0,u)\right)\bigg),\\
&= \sum_{y,u} (p^*(y,u) + \eta \cdot i(y,u)) \log \frac{p^*(y,u) + \eta \cdot i(y,u)}{p(y)\left(p^*(u) + \eta \cdot i(y_0,u)\right)},\\
&= \sum_{y,u} (p^*(y,u) + \eta \cdot i(y,u)) \log \frac{p^*(y,u)\left(1 + \eta \frac{i(y,u)}{p^*(y,u)}\right)}{p(y)p^*(u)\left(1 + \eta \frac{i(y_0,u)}{p^*(u)}\right)},\\
&\approx \sum_{y,u} (p^*(y,u) + \eta \cdot i(y,u)) \left[\log \frac{p^*(y,u)}{p(y)p^*(u)} + \eta \frac{i(y,u)}{p^*(y,u)}\right.\\
&\qquad \left. - \eta \frac{i(y_0,u)}{p^*(u)}\right], \qquad (91)\\
&= D(p^*(y,u)||p(y)p^*(u)) + \eta\bigg(\sum_{y,u} i(y,u) - i(y_0,u)p^*(y|u)\bigg)\\
&\quad + \eta \sum_{y,u} i(y,u) \log \frac{p^*(y,u)}{p(y)p^*(u)} + O(\eta^2),\\
&= g_0(X,Y) + 0 + \eta(\log p^*(y_0|u_0) - \log p^*(y_0|u_1)) + O(\eta^2), \qquad (92)
\end{aligned}$$

where all the summations above are over the support of (Y,U), i.e., supp(Y,U); in (91), the first order approximation, i.e., $\log(1+x) \approx x$ for $x \to 0$, has been used. Also, (92) follows from having $g_0(X,Y) = D(p^*(y,u)||p(y)p^*(u))$, and the properties of $i(\cdot,\cdot)$.

Similarly, when $\eta \to 0$, we can write

$$\begin{aligned}
&I_\eta(X;U) = D(p'_{X,U}||p_X \cdot p'_U)\\
&= D\bigg(p^*(x,u) + \eta \cdot p(x|y_0)i(y_0,u)\bigg|\bigg|p(x)(p^*(u) + \eta \cdot i(y_0,u))\bigg),\\
&= \sum_{x,u} (p^*(x,u) + \eta \cdot p(x|y_0)i(y_0,u)) \log \frac{p^*(x,u) + \eta \cdot p(x|y_0)i(y_0,u)}{p(x)(p^*(u) + \eta \cdot i(y_0,u))},\\
&= \sum_{x,u} (p^*(x,u) + \eta \cdot p(x|y_0)i(y_0,u)) \log \frac{p^*(x,u)\left(1 + \eta \frac{p(x|y_0)i(y_0,u)}{p^*(x,u)}\right)}{p(x)p^*(u)\left(1 + \eta \frac{i(y_0,u)}{p^*(u)}\right)},\\
&\approx \sum_{x,u} (p^*(x,u) + \eta \cdot p(x|y_0)i(y_0,u)) \left[\log \frac{p^*(x,u)}{p(x)p^*(u)} - \eta \frac{i(y_0,u)}{p^*(u)}\right.\\
&\quad \left. - \eta^2 \frac{p^2(x|y_0)i^2(y_0,u)}{2p^{*2}(x,u)} + \eta \frac{p(x|y_0)i(y_0,u)}{p^*(x,u)} + \eta^2 \frac{i^2(y_0,u)}{2p^{*2}(u)}\right], \qquad (93)\\
&= \eta^2 \sum_{x,u} \left(\frac{p^2(x|y_0)i^2(y_0,u)}{2p^*(x,u)} - \frac{p(x|y_0)i^2(y_0,u)}{p^*(u)} + \frac{p(x)i^2(y_0,u)}{2p^*(u)}\right)\\
&\quad + O(\eta^3), \qquad (94)\\
&= \eta^2 \sum_u \frac{i^2(y_0,u)}{2p^*(u)} \sum_x \left(\frac{p^2(x|y_0)}{p(x)} - 2p(x|y_0) + p(x)\right) + O(\eta^3),\\
&= \eta^2 \cdot \underbrace{\sum_{u \in \{u_0,u_1\}} \frac{1}{2p^*(u)} \left(\sum_x \frac{p^2(x|y_0)}{p(x)} - 1\right)}_{\triangleq A} + O(\eta^3), \qquad (95)
\end{aligned}$$

where (93) follows from the second order approximation $\log(1+x) \approx x - \frac{x^2}{2}$ for $x \to 0$; (94) follows from having $p^*(x,u) = p(x)p^*(u)$, due to the condition of perfect privacy, and the equality $\sum_{x,u} i(y_0,u)\left(p(x|y_0) - p(x)\right) = 0$. In (95), we make use of the fact that $i^2(y_0,u) = 1$ for $u = u_0, u_1$, and zero otherwise. Hence, from the construction of the tuple (X,Y,U), we obtain a lower bound for the LHS of (90) as

$$\lim_{\epsilon \to 0} \frac{g_\epsilon(X,Y) - g_0(X,Y)}{\epsilon} \geq \lim_{\eta \to 0} \frac{I_\eta(Y;U) - g_0(X,Y)}{I_\eta(X;U)}$$

[18] Alternatively, the LHS of (89) is $2^{D_2(q||p)}$, where $D_\alpha(\cdot||\cdot)$ is the Rényi divergence of order α. By noting that $D_\alpha(p||q) \geq 0$, with equality if and only if $p = q$, the proof is complete.

[19] To this end, a necessary condition is to have $\eta \leq \min\{p^*(y_0,u_1), 1 - p^*(y_0,u_0)\}$

$$= \lim_{\eta \to 0} \frac{\eta(\log(p^*(y_0|u_0)) - \log(p^*(y_0|u_1))) + O(\eta^2)}{A\eta^2 + O(\eta^3)}$$
$$= +\infty, \tag{96}$$

which is valid if it can be shown that A is a positive real number. From lemma 6, we have that $A \geq 0$. However, since $p_{X|Y}(\cdot|y_0) \neq p_X(\cdot)$[20], we have $A > 0$. This proves (90). □

VII. Conclusions

This paper addresses the problem of perfect privacy, where the goal is to find the maximum $I(Y;U)$, while guaranteeing $I(X;U) = 0$. This problem boils down to a standard linear program when the utility is measured by the mutual information between Y and U, as well as other utility measures such as mean-square error and probability of error. By solving this LP, upper and lower bounds for the cardinality of the disclosed data and the maximum utility are obtained. It is shown that when the private variable and the useful data form a jointly Gaussian pair, utility can be obtained only at the expense of privacy; that is, perfect privacy is not feasible. On the other hand, when the privacy-preserving mapping has direct access to both the useful data Y and the latent variable X, perfect privacy is feasible and the utility is actually unbounded. Finally, we have investigated the slope of the optimal utility-privacy trade-off curve as we approach to the perfect privacy point, i.e., $I(X;U) = 0$. We observe that if perfect privacy is not feasible, this slope is finite, and provide two lower bounds on it. However, when perfect privacy is feasible, under mild conditions, this slope is infinite, i.e., the rate of disclosing information is infinite for a vanishingly small privacy leakage.

Appendix A

Let $\mathcal{U}$ be an arbitrary set. Let $\mathcal{Q}$ denote an index set of rank$(\mathbf{P}_{X|Y})$ linearly independent columns of $\mathbf{P}_{X|Y}$. Hence, $\mathcal{Q}^c = [|\mathcal{Y}|] \backslash \mathcal{Q}$. Let $\pi : [\text{nul}(\mathbf{P}_{X|Y})] \to \mathcal{Q}^c$ such that $\pi(i) < \pi(j)$ for $i < j, \forall i, j \in [\text{nul}(\mathbf{P}_{X|Y})]$. Let $\mathbf{r} : \mathbb{S}_{X,Y} \to \mathbb{R}^{\text{nul}(\mathbf{P}_{X|Y})+1}$ be a vector-valued mapping defined element-wise as $r_i(\mathbf{p}) = \mathbf{p}(\pi(i)),\ \forall i \in [\text{nul}(\mathbf{P}_{X|Y})]$, $r_{\text{nul}(\mathbf{P}_{X|Y})+1}(\mathbf{p}) = H(\mathbf{p})$, where $\mathbf{p}(\pi(i))$ denotes the $\pi(i)$-th element of the probability vector $\mathbf{p}$. Since $\mathbb{S}_{X,Y}$ is a closed and bounded subset of $\mathcal{P}(\mathcal{Y})$, it is compact. Also, $\mathbf{r}$ is a continuous mapping from $\mathbb{S}_{X,Y}$ to $\mathbb{R}^{\text{nul}(\mathbf{P}_{X|Y})+1}$. Therefore, from the support lemma [28], for every $U \sim F(u)$ defined on $\mathcal{U}$, there exists a random variable $U' \sim p(u')$ with $|\mathcal{U}'| \leq \text{nul}(\mathbf{P}_{X|Y}) + 1$ and a collection of conditional probability vectors $\mathbf{p}_{Y|u'} \in \mathbb{S}_{X,Y}$ indexed by $u' \in \mathcal{U}'$, such that

$$\int_{\mathcal{U}} r_i(\mathbf{p}_{Y|u}) dF(u) = \sum_{u' \in \mathcal{U}'} r_i(\mathbf{p}_{Y|u'}) p(u'),\ i \in [\text{nul}(\mathbf{P}_{X|Y})+1].$$

It can be verified that by knowing the marginal $\mathbf{p}_X$, and the nul$(\mathbf{P}_{X|Y})$ elements of $\mathbf{p}_Y$ corresponding to index set $\mathcal{Q}^c$, the remaining rank$(\mathbf{P}_{X|Y})$ elements of $\mathbf{p}_Y$ can be uniquely identified by solving $\mathbf{p}_X = \mathbf{P}_{X|Y}\mathbf{p}_Y$. Therefore, for an arbitrary U in $X - Y - U$, that satisfies $X \perp\!\!\!\perp U$, the terms $p_Y(\cdot)$, and $I(Y;U)$ are preserved if U is replaced with U'. So are the condition of independence $X \perp\!\!\!\perp U'$ as $\mathbf{p}_{Y|u'} \in \mathbb{S}_{X,Y}, \forall u' \in \mathcal{U}^*$. Since we can simply construct the Markov chain $X - Y - U'$, there is no loss of optimality in considering $|\mathcal{U}| \leq \text{nul}(\mathbf{P}_{X|Y}) + 1$. The attainability of the supremum follows from the continuity of $I(Y;U)$, and the compactness of $\mathbb{S}_{X,Y}$, since $\mathcal{X}, \mathcal{Y}$ are finite.

[20]Since otherwise, from lemma 7, this will contradict the assumption in the statement of Theorem 9. In other words, if $p_{X|Y}(\cdot|y_0) = p_X(\cdot)$, no $u_0, u_1 \in \mathcal{U}$ exist such that $p^*(y_0, u_0), p^*(y_0, u_1) > 0$.

Appendix B

Let $W \triangleq (X, Y)$. For the binary matrix $\mathbf{P}_{X|W}$ (all elements being 0 or 1), which has $|\mathcal{X}|$ rows and $|\text{supp}(X, Y)|$ columns, we have rank$(\mathbf{P}_{X|W}) = |\mathcal{X}|$. Hence, the upper bound in (49) follows similarly to the analysis in Appendix A. Fix an arbitrary $x_0 \in \mathcal{X}$. We have that for each $y' \in \{y \in \mathcal{Y} | p(x_0, y) > 0\}$, there must exist a corresponding $u' \in \mathcal{U}$ such that $p(x_0, y'|u') > 0$, since otherwise, we get $p(x_0, y') = 0$, which is a contradiction. Moreover, from [21, lemma 5], for this u', we have $p(x_0, y|u') = 0,\ \forall y \neq y'$, which results in $|\mathcal{U}| \geq |\{y \in \mathcal{Y} | p(x_0, y) > 0\}|$. Finally, by noting that x_0 is chosen arbitrarily, and $p(x_0) > 0,\ \forall x_0 \in \mathcal{X}$, the lower bound in (49) is obtained.

Appendix C

We have

$$\mathbb{E}_X[J(X, p_Y)] = \mathbb{E}_X\left[\frac{1}{\max_{y \in \mathcal{Y}} \frac{p_{Y|X}(y|X)}{p_Y(y)}}\right]$$
$$\geq \frac{1}{\mathbb{E}_X\left[\max_{y \in \mathcal{Y}} \frac{p_{Y|X}(y|X)}{p_Y(y)}\right]} \tag{97}$$
$$= \frac{1}{\mathbb{E}_X\left[\max_{y \in \mathcal{Y}} \frac{p_{X|Y}(X|y)}{p_X(X)}\right]} \tag{98}$$
$$= \frac{1}{\sum_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} p_{X|Y}(x|y)}$$
$$= 2^{-\mathcal{L}(Y \to X)}, \tag{99}$$

where (97) follows from Jensen's inequality and the convexity of $f(t) = \frac{1}{t}$ for $t > 0$; (98) follows from Bayes's rule, and $p_X(x) > 0,\ \forall x \in \mathcal{X}$; in (99), we note that $p_Y(y) > 0,\ \forall y \in \mathcal{Y}$. Since the function $f(t) = \frac{1}{t}$ for $t > 0$ is strictly convex, the inequality in (97) is tight if and only if the term inside the expectation does not vary with x.

Appendix D

Without loss of generality, assume that y_0 is the first element of $\mathcal{Y}$. Hence, $\mathbf{e}_1 \triangleq \begin{bmatrix} 1 & \mathbf{0}_{|\mathcal{Y}|-1}^T \end{bmatrix}^T$ is an extreme point of $\mathbb{S}_{X,Y}$, since we have $\mathbf{p}_X = \mathbf{P}_{X|Y}\mathbf{e}_1$, and $\mathbf{e}_1$ cannot be written as a convex combination of two distinct probability vectors in $\mathbb{S}_{X,Y}$. Furthermore, the first element of all the other extreme points of $\mathbb{S}_{X,Y}$ is zero as proved by contradiction in what follows. Let $\mathbf{r} = \begin{bmatrix} \alpha & \mathbf{v}^T \end{bmatrix}^T$ ($\alpha \neq 0, \neq \mathbf{e}_1$) be an extreme point of $\mathbb{S}_{X,Y}$, where $\mathbf{v}$ is a vector of probability masses that sum to $1 - \alpha$. Since $\mathbf{r} \in \mathbb{S}_{X,Y}$, we must have $\mathbf{p}_X = \mathbf{P}_{X|Y}\mathbf{r}$,

which, from $p_{X|Y}(\cdot|y_0) = p_X(\cdot)$, results in $\mathbf{p}_X = \mathbf{P}_{X|Y}\mathbf{r}_0$, where $\mathbf{r}_0 = \left[0 \ \frac{1}{1-\alpha}\mathbf{v}^T\right]^T$ is a probability vector. Therefore, $\mathbf{r}_0 \in \mathbb{S}_{X,Y}$. However, since $\mathbf{r}$ can be written as a convex combination of $\mathbf{e}_1$ and $\mathbf{r}_0$, i.e., $\mathbf{r} = \alpha\mathbf{e}_1 + (1-\alpha)\mathbf{r}_0$, it is concluded that $\mathbf{r}$ cannot be an extreme point of $\mathbb{S}_{X,Y}$. Finally, by noting that in the evaluation of $g_0(X,Y)$, only the extreme points of $\mathbb{S}_{X,Y}$ are involved, the proof is complete.

REFERENCES

[1] B. Rassouli and D. Gunduz, "On perfect privacy," in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2551–2555.

[2] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *IEEE Symp. on Security and Privacy (SP)*, 2008, pp. 111–125.

[3] X. Ding, L. Zhang, and W. Zhiguo, "A brief survey on de-anonymization attacks in online social networks," in *International Conf. on Computational Aspects of Social Networks (CASoN)*, 2010, pp. 611–615.

[4] S. Kumar, W. Nilsen, M. Pavel, and M. Srivastava, "Mobile health: Revolutionizing healthcare through transdisciplinary research," *Computer*, vol. 46, pp. 28–35, 2013.

[5] J. Gomez-Vilardebo and D. Gündüz, "Smart meter privacy for multiple users in the presence of an alternative energy source," *IEEE Trans. on Information Forensics and Security*, pp. 132–141, 2015.

[6] A. Motahari, G. Bresler, and D. Tse, "Information theory of DNA shotgun sequencing," *IEEE Trans. on Information Theory*, vol. 59, no. 10, pp. 6273–6289, Oct. 2013.

[7] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," *Theory of Cryptography, Springer*, pp. 265–284, 2006.

[8] L. Sweeney, "k-anonymity: A model for protecting privacy," *Intl. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[9] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," *ACM Trans. on Knowledge Discovery from Data*, vol. 1.

[10] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," *IEEE Intl. Conf. on Data Eng.*, 2007.

[11] S. Sreekumar and D. Gündüz, "Optimal privacy-utility trade-off under a rate constraint," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2159–2163.

[12] 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.

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

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

[15] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, 2016. [Online]. Available: https://www.mdpi.com/2078-2489/7/1/15

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

[17] C. Huang, P. Kairouz, and L. Sankar, "Generative adversarial privacy: A data-driven approach to information-theoretic privacy," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 2162–2166.

[18] A. Tripathy, Y. Wang, and P. Ishwar, "Privacy-preserving adversarial networks," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2019, pp. 495–505.

[19] M. Bertran, N. Martinez, A. Papadaki, Q. Qiu, M. Rodrigues, G. Reeves, and G. Sapiro, "Adversarially learned representations for information obfuscation and inference," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. Long Beach, California, USA: PMLR, 09–15 Jun 2019, pp. 614–623. [Online]. Available: http://proceedings.mlr.press/v97/bertran19a.html

[20] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1625–1657, 2020.

[21] B. Rassouli and D. Gunduz, "On perfect privacy," *https://arxiv.org/pdf/1712.08500v8.pdf*.

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

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

[24] H. Witsenhausen and A. Wyner, "A conditional entropy bound for a pair of discrete random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 493–501, 1975.

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

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

[27] S. Asoodeh, F. Alajaji, and T. Linder, "Almost perfect privacy for additive Gaussian privacy filters," in *Information Theoretic Security*, A. C. Nascimento and P. Barreto, Eds. Cham: Springer International Publishing, 2016, pp. 259–278.

[28] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.