On the Privacy-Utility Trade-off With and Without Direct Access to the Private Data

Amirreza Zamani, *Member, IEEE,* Tobias J. Oechtering, *Senior Member, IEEE,* Mikael Skoglund, *Fellow, IEEE*

Division of Information Science and Engineering, KTH Royal Institute of Technology
Email: amizam@kth.se, oech@kth.se, skoglund@kth.se

***Abstract*—We study an information theoretic privacy mechanism design problem for two scenarios where the private data is either observable or hidden. In the hidden private data scenario, an agent observes useful data Y that is correlated with private data X, and generate disclosed data U which maximizes the revealed information about Y while satisfying a bounded privacy leakage constraint. Considering the other scenario, the agent has additional access to X. To design the privacy mechanism, we first extend the Functional Representation Lemma and Strong Functional Representation Lemma by relaxing the independence condition and thereby allowing a certain leakage. We then find lower and upper bounds on the privacy-utility trade-offs in both scenarios. In particular, for the case where no leakage is allowed and X is observable, our upper and lower bounds improve previous bounds. Considering bounded mutual information as privacy constraint and the observable private data scenario we show that if the common information and mutual information between X and Y are equal, then the attained upper bound is tight. Finally, the privacy-utility trade-off with prioritized private data is studied where part of X is more private than the remaining part.**

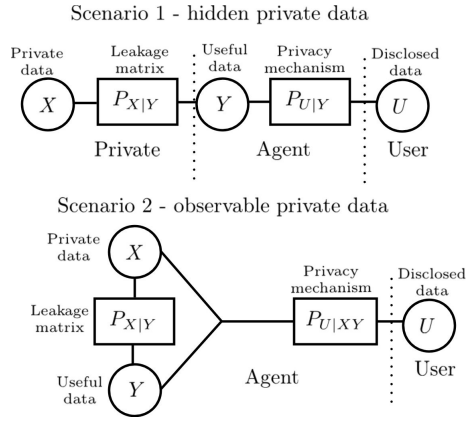

Fig. 1. In the *hidden private data* scenario the agent has only access to Y and in the *observable private data* scenario the agent has additionally access to X.

I. Introduction

The amount of data produced by robots, humans, networked sensors that record and analyze signals from physical environments, information processing and software systems is growing rapidly. Disclosing unprocessed data may lead to privacy threats through adversarial inferences. Moreover, theoretical and practical approaches to information-theoretic privacy (secrecy) can be utilized for many information processing systems. Perfect privacy (secrecy) is often not achievable in applications and we may need to relax the restriction [1]. Altogether, we need privacy mechanism designs for the disclosure of the data.

In this paper, random variable (RV) Y denotes the useful data and is correlated with the private data denoted by RV X. Furthermore, RV U describes the disclosed data. Two scenarios are considered, where in both, an agent wants to disclose the useful information to a user as shown in Fig. 1. In the *hidden private data* scenario, the agent observes Y and has no direct access to X, i.e., the private data is hidden. The goal is to design U based on Y that reveals as much information as possible about Y and satisfies a bounded privacy criterion. In *observable private data* scenario, the agent has access to both X and Y and can design U based on (X, Y) to release as much information as possible about Y while satisfying the bounded leakage constraint. In both scenarios we consider privacy constraints using different non-zero privacy measures.

The privacy mechanism design problem from an information theory perspective is recently receiving increased attention and related results can be found in [2]–[28]. In more detail, in [2], the concept of a privacy funnel is introduced, where the privacy utility trade-off has been studied considering a distortion measure for utility and the log-loss as privacy measure. The concept of maximal leakage has been introduced in [3] and used in [4] for the Shannon cipher system. Furthermore, some bounds on the privacy-utility trade-off are derived. Fundamental limits of the privacy utility trade-off measuring the leakage using estimation-theoretic guarantees are studied in [5].

In both [29] and [6], the privacy-utility trade-offs considering expected distortion and equivocation as a measures of utility and privacy are studied. In [7], the hypothesis test performance of an adversary is used to measure privacy leakage. In [8], maximal correlation either mutual information is used for measuring the privacy and properties of rate-privacy functions are studied. In [9], average total variation is used as a privacy measure and a χ^2-privacy criterion is considered in [5], where an upper bound and a lower bound on the

This work was funded in part by the Swedish research council under contract 2019-03606. This work was presented in part at the 2022 IEEE International Symposium on Information Theory and the 2022 IEEE Information Theory Workshop. A. Zamani, M. Skoglund and T. J. Oechtering are with the Division of Information Science and Engineering, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden (e-mail: amizam@kth.se; oech@kth.se; skoglund@kth.se).

privacy-utility trade-off have been derived. The problem of privacy-utility trade-off considering mutual information both as measures of utility and privacy given the Markov chain $X - Y - U$ is studied in [10]. Under the perfect privacy assumption it is shown that the privacy mechanism design problem can be reduced to a linear program. This has been extended in [11] considering the privacy utility trade-off with a rate constraint on the disclosed data. Moreover, in [10], it has been shown that information can only be revealed if the kernel (leakage matrix) between useful data and private data is not invertible. In [12], we generalize [10] by relaxing the perfect privacy assumption allowing some small bounded leakage. More specifically, we design privacy mechanisms with a per-letter privacy criterion considering an invertible kernel where a small leakage is allowed. We generalized this result to a non-invertible leakage matrix in [13]. In both [14] and [15], the optimal privacy-utility trade-offs have been studied considering two scenarios where the private data is either observable or hidden. Sufficient conditions for equality of the optimal trade-offs in the considered two scenarios have been derived where the utility is measured by a distortion metric. A multi-user data disclosure design problem with a privacy constraint is studied in [21] where upper and lower bounds on the privacy-utility trad-off have been derived. Another multi-user privacy mechanism design problem is considered in [22] where each user is equipped with a local cache. In [16], by using the Functional Representation Lemma bounds on privacy-utility trade-off for the two scenarios are derived. These results are derived under the perfect secrecy assumption, i.e., no leakages are allowed. The bounds are tight when the private data is a deterministic function of the useful data.

The concept of differential privacy is introduced in [30] and it has been used in [31] to minimize the statistical risk of identifying membership in a database. The concept of mutual information as differential privacy is introduced in [32]. A related secure source coding problem is studied in [29].

Our results in this work can be divided into three main parts as follows:

Part I (*Privacy-utility trade-off with non-zero leakage*): In the first part of the paper, we consider the problems studied in [33] and [10]. Furthermore, our setup is related to [14] and [16], where in [16] the problem of *secrecy by design* is studied. We generalize the privacy problems considered in [16] by relaxing the perfect privacy constraint and allowing some leakage. More specifically, we consider bounded mutual information, i.e., $I(U;X) \leq \epsilon$ for privacy leakage constraint. To this end, we extend the Functional Representation Lemma and the Strong Functional Representation Lemma, introduced in [34] by relaxing the independence condition to derive lower bounds for the observable private data scenario. The key idea to extend both lemmas is to use randomized response introduced in [35]. We show that if the *common information* and mutual information between X and Y are equal, then the maximum utility in two scenarios and the attained upper bound in the observable private data scenario are equal. Furthermore, in the special case of perfect privacy we find a new upper bound for the perfect privacy function by using the *excess functional information* introduced in [34]. We show that this new bound generalizes the bound in [16]. Moreover, we show that the bound is tight when $|\mathcal{Y}| = 2$. Finally, we compare our new lower and upper bounds with the bounds found in [16] when the leakage is zero. The conference version regarding this part can be found in [19].

Part II (*Privacy-utility trade-off with non-zero leakage and per-letter privacy constraints*): In the second part, for each scenario we use two different per-letter privacy constraints instead of the bounded mutual information constraint. As argued in [13], it can be more desirable to protect the private data individually and not merely on average. We first find similar results as the extended versions of the Functional Representation Lemma (FRL) and the Strong Functional Representation Lemma (SFRL) found in the previous part considering the per-letter privacy constraint rather than bounded mutual information. Using these results we find a lower bound for the privacy-utility trade-off in the observable private data scenario. Furthermore, we provide bounds for three other problems and study a special case where X is a deterministic function of Y. We show that the obtained upper and lower bounds in the hidden private data scenario are asymptotically optimal when X is a deterministic function of Y. In [13], one of the problems considered in this part has been studied. It has been shown that by using methods from Euclidean information geometry as used in [36], [37], we can simplify the design problem in the high privacy regime and the main problem can be solved approximately by a linear program. In this work, we provide upper bounds on the error of the approximation considered in [13]. Finally we compare the attained bounds in a numerical example. The conference version related to this part can be found in [19].

Part III (*Privacy-utility trade-off with non-zero leakage and prioritized private data*): Finally, we consider the problem in the observable private data scenario where the private data is divided into two parts, i.e., $X = (X_1, X_2)$. In this part we use bounded mutual information as privacy constraint. We assume that the first part is more private than the second part, i.e., the privacy leakage of X_1 is less than or equal to the privacy leakage of X_2. Furthermore, we assume that the total leakage between (X_1, X_2) and U is bounded by ϵ and we derive upper and lower bounds. Similar to the previous parts we use the extended versions of Functional Representation Lemma and the Strong Functional Representation Lemma to find lower bounds. The key idea to obtain lower bounds on the privacy-utility trade-off considering prioritized private data is to use randomization technique introduced in [35] over X_2 instead of X. We compare the obtained lower bounds on the privacy-utility trade-off with the results corresponding to Part I. We show that the lower bounds obtained in Part III can tighten the lower bounds derived in Part I.

Our contribution can be summarized as follows:

(i) We extend the Functional Representation Lemma and the Strong Functional Representation Lemma by a randomized response output that allows some controlled leakage. Various extended versions are introduced using different leakage measures.

(ii) We formulate and study various privacy mechanism design problems through the lens of information theory with

controlled leakage, demonstrating the use of the extended versions of Functional Representation Lemma and the Strong Functional Representation Lemma.

(iii) We provide discussion and comparison of the obtained results with each other and the literature.

Notation: Given two jointly random variables X and Y, the entropy, conditional entropy and mutual information between X and Y are given by $H(Y)=\mathbb{E}(\log(\frac{1}{P_Y(y)}))$, $H(Y|X)=\mathbb{E}(\log(\frac{1}{P_{Y|X}(y|x)}))$, and $I(X;Y)=H(Y)-H(Y|X)$. X and Y are independent if and only if $I(X;Y)=0$. Furthermore, the Markov chain $X-Y-U$ holds if and only if $I(X;U|Y)=0$. For the binary entropy $h(\cdot)$ we have $h(p)=-\left(p\log(p)+(1-p)\log(1-p)\right)$. In this work, let matrix P_{XY} defined on $\mathbb{R}^{|\mathcal{X}|\times|\mathcal{Y}|}$ denote the joint distribution of discrete random variables X and Y defined on finite alphabets $\mathcal{X}$ and $\mathcal{Y}$. We represent marginal distributions of X and Y by vectors P_X and P_Y defined on $\mathbb{R}^{|\mathcal{X}|}$ and $\mathbb{R}^{|\mathcal{Y}|}$ given by the row and column sums of P_{XY}. We represent the leakage matrix $P_{X|Y}$ by a matrix defined on $\mathbb{R}^{|\mathcal{X}|\times|\mathcal{Y}|}$ with elements $P_{X|Y}(x|y)$ for all x and y. Furthermore, for given $u\in\mathcal{U}$, $P_{X,U}(\cdot,u)$ and $P_{X|U}(\cdot|u)$ defined on $\mathbb{R}^{|\mathcal{X}|}$ are distribution vectors with elements $P_{X,U}(x,u)$ and $P_{X|U}(x|u)$ for all $x\in\mathcal{X}$ and $u\in\mathcal{U}$. The relation between U and Y is described by the kernel $P_{U|Y}$ defined on $\mathbb{R}^{|\mathcal{U}|\times|\mathcal{Y}|}$, furthermore, the relation between U and the pair (Y,X) is described by the kernel $P_{U|Y,X}$ defined on $\mathbb{R}^{|\mathcal{U}|\times|\mathcal{Y}|\times|\mathcal{X}|}$. In this work, $\min P_X$ corresponds to minimum value inside the distribution vector P_X.

II. System Model and Problem Formulation

In this work we assume that each element in vectors P_X and P_Y is non-zero. In the second part of the results, which corresponds to *privacy-utility trade-off with non-zero leakage and per-letter privacy constraints*, we assume that for the discrete random variables X and Y defined on finite alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have that $|\mathcal{X}|<|\mathcal{Y}|$. Furthermore, we assume that the leakage matrix $P_{X|Y}$ is of full rank. In the remaining parts of the results we consider arbitrary correlated discrete random variables X and Y as private data and useful data.

In the following we introduce the main problems in three different parts. In each part, we first define the problems considered in this paper, then we motivate them and study the properties of the measures for utility and privacy leakage and compare them with previous works.

A. Privacy-utility trade-off with non-zero leakage

In this part, for both design problems we use mutual information as utility and leakage measures. The privacy mechanism design problems for the two scenarios can be stated as follows

$$g_\epsilon(P_{XY})=\sup_{\substack{P_{U|Y}:X-Y-U\\ I(U;X)\leq\epsilon,}} I(Y;U), \tag{1}$$

$$h_\epsilon(P_{XY})=\sup_{P_{U|Y,X}:I(U;X)\leq\epsilon,} I(Y;U). \tag{2}$$

The function $h_\epsilon(P_{XY})$ is used when the privacy mechanism has access to both the private data and the useful data. The function $g_\epsilon(P_{XY})$ is used when the privacy mechanism has only access to the useful data. Clearly, the relation between $h_\epsilon(P_{XY})$ and $g_\epsilon(P_{XY})$ can be stated as follows

$$g_\epsilon(P_{XY})\leq h_\epsilon(P_{XY}).$$

In the following we study the case where $0\leq\epsilon<I(X;Y)$, otherwise the optimal solution of $h_\epsilon(P_{XY})$ or $g_\epsilon(P_{XY})$ is $H(Y)$ achieved by $U=Y$.

Remark 1. For $\epsilon=0$, (1) leads to the perfect privacy problem studied in [10]. It has been shown that for a non-invertible leakage matrix $P_{X|Y}$, $g_0(P_{XY})$ can be obtained by a linear program. Furthermore, for $\epsilon=0$, (2) leads to the secret-dependent perfect privacy function $h_0(P_{XY})$, studied in [16], where upper and lower bounds on $h_0(P_{XY})$ have been derived. The bounds are tight when X is deterministic function of Y.

B. Privacy-utility trade-off with non-zero leakage and per-letter privacy constraints

The privacy mechanism design problems for the two scenarios can be stated as follows

$$g_\epsilon^{w\ell}(P_{XY})=\sup_{\substack{P_{U|Y}:X-Y-U\\ P_U(u)d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u,}} I(Y;U), \tag{3}$$

$$h_\epsilon^{w\ell}(P_{XY})=\sup_{P_{U|Y,X}:P_U(u)d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u,} I(Y;U), \tag{4}$$

$$g_\epsilon^{\ell}(P_{XY})=\sup_{\substack{P_{U|Y}:X-Y-U\\ d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u}} I(Y;U), \tag{5}$$

$$h_\epsilon^{\ell}(P_{XY})=\sup_{P_{U|Y,X}:d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u} I(Y;U), \tag{6}$$

where $d(P,Q)$ corresponds to the total variation distance between two distributions P and Q, i.e., $d(P,Q)=\sum_x|P(x)-Q(x)|$. The functions $h_\epsilon^{w\ell}(P_{XY})$ and $h_\epsilon^{\ell}(P_{XY})$ are used when the privacy mechanism has access to both the private data and the useful data. The functions $g_\epsilon^{w\ell}(P_{XY})$ and $g_\epsilon^{\ell}(P_{XY})$ are used when the privacy mechanism has only access to the useful data. In this work, the privacy constraints used in (3) and (5), i.e., $P_U(u)d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u$, and $d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u$, are called the *weighted strong ℓ_1-privacy criterion* and the *strong ℓ_1-privacy criterion*. We refer to them as strong since they are per-letter privacy constraints, i.e., they must hold for every $u\in\mathcal{U}$. The difference between the two privacy constraints in this work is the weight $P_U(u)$, therefore, we refer to $P_U(u)d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u$, as weighted. We later show that the weight $P_U(u)$ enables us to use extended versions of the Functional Representation Lemma and Strong Functional Representation Lemma to find lower bounds considering the observable private data scenario, i.e., lower bounds on $h_\epsilon^{w\ell}(P_{XY})$.

Remark 2. The *strong ℓ_1-privacy criterion*, i.e., $d(P_{X|U}(\cdot|u),P_X)\leq\epsilon,\ \forall u$, has been introduced in [13], where we have provided and utilized its properties to find an approximate solution of $g_\epsilon^{\ell}(P_{XY})$.

Remark 3. For $\epsilon = 0$, both (3) and (5) lead to the perfect privacy problem studied in [10]. This follows since by letting $\epsilon = 0$, we have $\sum_u P_U(u) d(P_{X|U}(\cdot|u), P_X) = 0$ which results in independence of X and U. It has been shown that for a non-invertible leakage matrix $P_{X|Y}$, $g_0(P_{XY})$ can be obtained by a linear program [10]. Similarly, for $\epsilon = 0$, both (4) and (6) lead to the secret-dependent perfect privacy function $h_0(P_{XY})$, studied in [16], where upper and lower bounds on $h_0(P_{XY})$ have been derived. In [19], these bounds have been strengthened.

Remark 4. The privacy problem defined in (5) has been studied in [13] where a lower bound on $g_\epsilon^\ell(P_{XY})$ has been provided using the information geometry concepts. Furthermore, it is shown that it is sufficient to assume $|\mathcal{U}| \leq |\mathcal{Y}|$ so that it is ensured that the supremum can be achieved.

Intuitively, for small ϵ, both privacy constraints mean that X and U are almost independent. As we discussed in [13], closeness of $P_{X|U}(\cdot|u)$ and P_X allows us to approximate $g_\epsilon^\ell(P_{XY})$ with a series expansion and find a lower bound. In this work we show that by using a similar methodology, we can approximate $g_\epsilon^{w\ell}(P_{XY})$ exploiting the closeness of $P_{X,U}(\cdot, u)$ and $P_X P_U(u)$. This provides us a lower bound for $g_\epsilon^{w\ell}(P_{XY})$. As we discussed earlier, an advantage of per-letter privacy measures over on average measures can be stated as follows. From a privacy perspective, it is often more desirable to protect sensitive data individually and not on average. Using an average criterion means that there can exist some data points which leak more than the average privacy threshold. However, if we choose bounded mutual information rather than the strong ℓ_1-privacy criterion and the weighted strong ℓ_1-privacy criterion, we can solve the problems explicitly under some specific assumptions. The problems considering the per-letter constraints have not been explicitly solved and only approximate solutions are derived in [13]. In the high privacy regime, the approximate solutions are close to the exact solutions.

Furthermore, as we stated earlier the only difference between the strong ℓ_1-privacy criterion and the weighted strong ℓ_1-privacy criterion is the weight $P_U(u)$ that enables us to extend FRL and SFRL. The extended versions of FRL and SFRL help us to find lower bound on $h_\epsilon^{w\ell}(P_{XY})$. On the other hand, if we use the strong ℓ_1-privacy criterion we can find upper bounds on $h_\epsilon^\ell(P_{XY})$. As a summary, the strong ℓ_1-privacy criterion lead to an upper bound on the privacy-utility trade-off considering the observable private data scenario and the weighted strong ℓ_1-privacy criterion results in lower bounds on it.

Next, we study some properties of the weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion. To this end recall that the *linkage inequality* is the property that if $\mathcal{L}$ measures the privacy leakage between two random variables and the Markov chain $X - Y - U$ holds then we have $\mathcal{L}(X;U) \leq \mathcal{L}(Y;U)$. Since the weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion are per letter constraints we define $\mathcal{L}^1(X;U=u) \triangleq d(P_{X|U}(\cdot|u), P_X)$, $\mathcal{L}^1(Y;U=u) \triangleq d(P_{Y|U}(\cdot|u), P_Y)$, $\mathcal{L}^2(X;U=u) \triangleq P_U(u) d(P_{X|U}(\cdot|u), P_X)$, $\mathcal{L}^2(Y;U=u) \triangleq P_U(u) d(P_{Y|U}(\cdot|u), P_Y)$.

Proposition 1. *The weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion satisfy the linkage inequality. Thus, for each $u \in \mathcal{U}$ we have $\mathcal{L}^1(X;U=u) \leq \mathcal{L}^1(Y;U=u)$ and $\mathcal{L}^2(X;U=u) \leq \mathcal{L}^2(Y;U=u)$.*

Proof: The proof is provided in Appendix A. ∎

As discussed in [9, page 4] and [15, page 5], one benefit of the linkage inequality is to ensure privacy in layers of private information which is discussed in the following. Assume that there are "primary" and "secondary private data" where the disclosure mechanism is designed independently based on primary private data. The secondary private data may be unforeseen or not accessible by the agent and is arbitrary correlated with the primary data. The linkage inequality ensures that the privacy leakage of the secondary private data is bounded by the guarantee on the privacy leakage of the primary data. Alternatively, assume that the Markov chain $X - Y - U$ holds and the distribution of X is not known. If we can find $\tilde{X}$ such that $X - \tilde{X} - Y - U$ holds and the distribution of $\tilde{X}$ is known then by the linkage inequality we can conclude $\mathcal{L}(X;U=u) \leq \mathcal{L}(\tilde{X};U=u)$. In other words, if the framework is designed for $\tilde{X}$, then a privacy constraint on $\tilde{X}$ leads to the constraint on X, i.e., provides an upper bound for any pre-processed RV X. To have the Markov chain $X - \tilde{X} - Y - U$ consider the scenario where $\tilde{X}$ is the private data and X is a function of private data which is not known. For instance let $\tilde{X} = (X_1, X_2, X_3)$ and $X = f(X_1)$ where the function $f(\cdot)$ is not known. Thus, the mechanism that is designed based on $\tilde{X} - Y - U$ preserves the leakage constraint on X and U. As pointed out in [9, Remark 2], among all the L^p-norms ($p \geq 1$), only the ℓ_1 norm satisfies the linkage inequality. Next, given a leakage measure $\mathcal{L}$ and let the Markov chain $X - Y - U$ hold, if we have $\mathcal{L}(X;U) \leq \mathcal{L}(X;Y)$, then we say that the *post processing inequality* holds. In this work we use $\mathcal{L}^1(X;U) = \sum_u P_U(u) \mathcal{L}^1(X;U=u)$, $\mathcal{L}^2(X;U) = \sum_u \mathcal{L}^2(X;U=u)$ and $\mathcal{L}^1(Y;U) = \sum_u P_U(u) \mathcal{L}^1(Y;U=u)$, $\mathcal{L}^2(Y;U) = \sum_u \mathcal{L}^2(Y;U=u)$.

Using the same proof as [9, Theorem 3] which is based on the convexity of the ℓ_1-norm, the average of the weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion with weights equal to one and $P_U(u)$, respectively, satisfy the post-processing inequality. In other words, under the Markov chain $X - Y - U$ we have $\mathcal{L}^1(X;U) \leq \mathcal{L}^1(Y;U)$ and $\mathcal{L}^2(X;U) \leq \mathcal{L}^2(Y;U)$. Before stating the next result we present the inference threat model as introduced in [27]. Let $C(\cdot,\cdot) : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ be an inference cost function, where $\mathcal{P}(\mathcal{X})$ denotes all possible distribution vectors for X. The adversarial attacker chooses a belief distribution for X as q_X which is the solution of $c_0^* = \min_{q_X \in \mathcal{P}(\mathcal{X})} \mathbb{E}_X C(X, q_X)$. After observing $U = u$, he updates the belief distribution as the solution of $c_u^* = \min_{q_X \in \mathcal{P}(\mathcal{X})} \mathbb{E}_{X|U}[C(X, q_X)|U = u]$. Let $\Delta C = c_0^* - \mathbb{E}_U[c_U^*]$ denote the on average gain cost attained by the attacker. Note that ΔC specifies how much improvement the attacker can obtain in his inference. Next result asserts that by using ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion, ΔC can be bounded by a constant.

Proposition 2. *The weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion result in bounded inference*

threat that is modeled in [27].

Proof: The weighted strong ℓ_1-privacy criterion and the strong ℓ_1-privacy criterion lead to a bounded on average constraint $\sum_u P_U(u) \left\| P_{X|U=u} - P_X \right\|_1 \leq \epsilon$. Thus, using [9, Theorem 4], we conclude that adversarial inference performance is bounded as follows

$$\Delta C \leq 2L \sum_u P_U(u) \left\| P_{X|U=u} - P_X \right\|_1 \leq 2L\epsilon, \tag{7}$$

where $L = \sup_{x \in \mathcal{X}, q_X \in \mathcal{P}(\mathcal{X})} |C(x, q_X)|$. ∎

Another property of the ℓ_1 distance is the relation between the ℓ_1-norm and probability of error in a hypothesis test. As argued in [38, Remark 6.5], for the binary hypothesis test with $H_0 : X \sim P$ and $H_1 : X \sim Q$, the expression $1 - \frac{1}{2}d(P, Q)$ is the sum of false alarm and missed detection probabilities. Thus, we have $\frac{1}{2}d(P, Q) = 1 - 2P_e$, where P_e is the error probability (the probability that we can not decide the right distribution for X with equal prior probabilities for H_0 and H_1). For instance, consider the scenario where we want to decide whether X and U are independent or correlated. To this end, let $P = P_{X,U}$, $Q = P_X P_U$, $H_0 : X, U \sim P$ and $H_1 : X, U \sim Q$. We have

$$\frac{1}{2}d(P, Q) = \frac{1}{2} \sum_u P_U(u) d(P_{X|U}(\cdot|u), P_X) \leq \frac{1}{2}\epsilon.$$

Thus, by increasing the leakage, which means that $d(P, Q)$ increases, then the error of probability decreases.

Another property is the relation between the strong ℓ_1-privacy criterion and MMSE$(X|U)$ which has been used in [39, Corrolary 2]. An interesting result is that the strong ℓ_1-privacy criterion leads to a lower bound on MMSE$(X|U)$ [13, Proposition 9]. Practical examples based on the MNIST data set and medical experiment with real data have been provided in [13, Experiment 2] and [13, Experiment 3] which show the applicability of the strong ℓ_1-privacy criterion. Meaningful interpretations of the solution to $g_\epsilon^\ell(P_{XY})$ have been provided in [13] for both experiments. Geometrical interpretation of the solution to $g_\epsilon^\ell(P_{XY})$ have been derived in [13, Section IV]. An interesting result is that all candidates of optimizers for $g_\epsilon^\ell(P_{XY})$ are inside an ℓ_1-ball with a bounded radius and certain centers [13, Section IV]. The role of the strong ℓ_1-privacy criterion can be evaluated by using different metrics such as probability of error and MMSE for measuring the utility and the privacy leakage in $g_\epsilon^\ell(P_{XY})$ and comparing the results with the previous works. More detail can be found in [13, Section V-B].

Finally, if we use ℓ_1 distance as privacy leakage, after approximating $g_\epsilon^{w\ell}(P_{XY})$ and $g_\epsilon^\ell(P_{XY})$, we face a linear program problem in the end, which are much easier to handle.

C. Privacy-utility trade-off with non-zero leakage and prioritized private data

In this part, we assume that the private data X is divided into two parts X_1 and X_2, where the first part is more private than the other part, i.e., the privacy leakage of X_1 is less than or equal to the privacy leakage of X_2. We use mutual information for measuring both privacy leakage and utility and we only consider the observable private data scenario where the privacy mechanism has access to both X and Y. Hence, the problem can be stated as follows

$$h_\epsilon^p(P_{X_1X_2Y}) = \sup_{\substack{P_{U|YX_1X_2} : I(U;X_1,X_2) \leq \epsilon, \\ I(U;X_1) \leq I(U;X_2)}} I(Y;U). \tag{8}$$

The constraint $I(U; X_1, X_2) \leq \epsilon$ ensures that the total leakage is bounded by ϵ and the constraint $I(U; X_1) \leq I(U; X_2)$ corresponds to the priority of X_1. In practice, we usually have different levels of privacy leakage for the private data and in this work we consider two levels.

Remark 5. For $\epsilon = 0$, (8) leads to the secret-dependent perfect privacy function $h_0(P_{XY})$.

In certain scenarios, the privacy problem containing prioritized private data leads to the non-prioritized problem, i.e., $h_\epsilon^p(P_{X_1X_2Y}) = h_\epsilon(P_{X_1X_2Y})$. For instance, consider the scenario where X_1 is a deterministic function of X_2, i.e., $X_1 = f(X_2)$. In this case, the inequality corresponding to the priority $I(U; X_1) = I(U; f(X_2)) \leq I(U; X_2)$ holds, hence, $h_\epsilon^p(P_{X_1X_2Y}) = h_\epsilon(P_{X_1X_2Y})$. In a more general setting, if the Markov chain $X_1 - X_2 - Y - U$ holds, due to the data processing inequality we have $h_\epsilon^p(P_{X_1X_2Y}) = h_\epsilon(P_{X_1X_2Y})$. In other words, under certain assumptions, the prioritized privacy problem reduces to the non-prioritized version.

III. Overview and Relation between the Privacy Measures and Problems

In this section we first present an overview of the privacy problems outlined in (1) and (2). We then provide essential lemmas and definitions corresponding to the functional representation lemma, strong functional representation lemma, *excess functional information*, and *common information between two RVs*. Finally, we study the relation between the privacy measures and problems considered in this work.

As mentioned earlier, (1) and (2) have been studied in previous works, e.g., [10], and [33]. In [33, Lemma 1], lower and upper bounds on $g_\epsilon(P_{XY})$ have been derived, where we have

$$\frac{H(Y)}{I(X;Y)}\epsilon \leq g_\epsilon(P_{XY}) \leq H(Y|X) + \epsilon. \tag{9}$$

The result in [33, Lemma 2] asserts that the mapping $\epsilon \to g_\epsilon(P_{XY})$ is concave for any $\epsilon \geq 0$. Concavity of $g_\epsilon(P_{XY})$ can be used to show the lower bound in (9), since in this case, when $\epsilon = I(X;Y)$, $U = Y$ is feasible and the utility $H(Y)$ is achieved. Hence, the line $\frac{H(Y)}{I(X;Y)}\epsilon$ is achievable. As stated in [33, Remark 1], by using the concavity of $g_\epsilon(P_{XY})$, the lower bound (9) can be improved. In this case, we have

$$\epsilon \frac{H(Y)}{I(X;Y)} + g_0(P_{XY}) \left(1 - \frac{\epsilon}{I(X;Y)}\right) \leq g_\epsilon(P_{XY}). \tag{10}$$

One benefit of the lower bound in (10) is that we can use it as a lower bound on $h_\epsilon(P_{XY})$, c.f., see Theorem 2. Considering perfect privacy, i.e., $\epsilon = 0$, $g_0(P_{XY})$ can be obtained by solving a linear program [10, Theorem 1]. This means the optimal mapping is the solution to a linear program.

Moreover, it has been shown that to find the optimal privacy-preserving mapping, it is sufficient to consider U such that $|\mathcal{U}| \leq \text{null}(P_{X|Y})+1$. Solving the linear program proposed in [10] can become challenging when the size of $\mathcal{Y}$ or $\mathcal{X}$ grow. Hence, simple lower and upper bounds are derived in [10, Corrolary 2] as follows

$$(H(Y) - \log(\text{rank}(P_{X|Y})))^+ \leq g_0(P_{XY}) \leq \min\{H(Y|X), \log(\text{null}(P_{X|Y})+1)\},$$

where $a^+ = \begin{cases} 0, & a<0, \\ a, & a \geq 0 \end{cases}$. The upper bound $\log(\text{null}(P_{X|Y})+1)$ can be obtained using the sufficiency condition $|\mathcal{U}| \leq \text{null}(P_{X|Y})+1$, since in this case we have $H(U) \leq \log(\text{null}(P_{X|Y})+1)$. Necessary and sufficient conditions for attaining non-zero utility when the private data is hidden, i.e., $g_0(P_{XY})>0$, have been derived in [17] and [40]. It has been shown that $g_0(P_{XY})>0$ if and only if rows of $P_{X|Y}$ are linearly dependent. This result has been also shown and generalized in [10], e.g., see [10, Proposition 1]. Moreover, considering observable private data scenario, necessary and sufficient conditions for attaining non-zero utility, i.e., $h_0(P_{XY})>0$, have been derived in [16, Theorem 5]. It has been shown that $h_0(P_{XY})>0$ if and only if Y is not a deterministic function of X, i.e., $H(Y|X)>0$. Next, we recall the Functional Representation Lemma (FRL) and the Strong Functional Representation Lemma (SFRL).

Lemma 1. *(Functional Representation Lemma [16, Lemma 1]): For any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite, there exists a RV U defined on $\mathcal{U}$ such that X and U are independent, i.e., we have*

$$I(U;X)=0, \tag{11}$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X)=0, \tag{12}$$

and

$$|\mathcal{U}| \leq |\mathcal{X}|(|\mathcal{Y}|-1)+1. \tag{13}$$

Lemma 2. *(Strong Functional Representation Lemma [34, Theorem 1]): For any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite with $I(X,Y)<\infty$, there exists a RV U defined on $\mathcal{U}$ such that X and U are independent, i.e., we have*

$$I(U;X)=0,$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X)=0,$$

$I(X;U|Y)$ can be upper bounded as follows

$$I(X;U|Y) \leq \log(I(X;Y)+1)+4,$$

and $|\mathcal{U}| \leq |\mathcal{X}|(|\mathcal{Y}|-1)+2$.

Remark 6. *By checking the proof in [34, Th. 1], the term $e^{-1}\log(e)+2+\log(I(X;Y)+e^{-1}\log(e)+2)$ can be used instead of $\log(I(X;Y)+1)+4$.*

Both FRL and SFRL have constructive proofs that can be useful to find lower bounds on the privacy-utility trade-offs considered in this work. Lower and upper bounds on $h_0(P_{XY})$ have been derived in [16, Theorem 6, 7] and we have

$$H(Y)-H(X)=H(Y|X)-H(X|Y) \leq h_0(P_{XY}) \leq H(Y|X). \tag{14}$$

The lower bound in (14) is attained by Lemma 1 and is tight if and only if X is a deterministic function of Y, i.e., $H(X|Y)=0$. In [16, Theorem 7] it is claimed that the necessary and sufficient conditions for equalities $h_0(P_{XY})=g_0(P_{XY})=H(Y|X)$ are fulfilled when $H(X|Y)=0$. In this paper, we show that $H(X|Y)=0$ can be improved using the concept of *common information*. In the following we state the definition of *excess functional information* defined in [34] as

$$\psi(X \to Y) = \inf_{P_{U|YX}: I(U;X)=0,\ H(Y|X,U)=0} I(X;U|Y).$$

The lower bound on $\psi(X \to Y)$ derived in [34, Prop. 1] is given in the next lemma. Since this lemma is useful for deriving the upper bound on $h_\epsilon(P_{XY})$ we state it here.

Lemma 3. *[34, Prop. 1] For discrete Y we have*

$$\psi(X \to Y) \geq - \sum_{y \in \mathcal{Y}} \int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt - I(X;Y), \tag{15}$$

where for $|\mathcal{Y}|=2$ the equality holds and it is attained by the Poisson functional representation [34].

Remark 7. *The lower bound in* (15) *can be negative. For instance, let Y be a deterministic function of X, i.e., $H(Y|X)=0$. In this case we have $-\sum_{y\in\mathcal{Y}}\int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\})dt - I(X;Y) = -I(X;Y) = -H(Y)$.*

More properties of the excess functional information $\psi(X \to Y)$ can be found in [34, Prop. 3]. For instance, using [34, Prop. 3] an alternative characterization for $\psi(X \to Y)$ can be stated as follows

$$\psi(X \to Y) = \inf_{P_{U|XY}: I(U;X)=0} H(Y|U) - I(X;Y).$$

Next we recall the definition of the common information between X and Y using [41]. For any pair of RVs (X,Y) defined on discrete alphabets $\mathcal{X}$ and $\mathcal{Y}$, the common information between X and Y can be defined as follows

$$C(X;Y) = \inf_{P_{W|XY}: X-W-Y} I(X,Y;W). \tag{16}$$

As shown in [41, Remark A] we have

$$I(X;Y) \leq C(X;Y) \leq \min\{H(X), H(Y)\}. \tag{17}$$

One simple observation is that when $H(X|Y) = 0$ or $H(Y|X) = 0$ we have $I(X;Y) = C(X;Y)$. This follows since when $H(X|Y) = 0$ we have

$$I(X,Y;W) = I(Y;W)$$

and W can be chosen as X, hence $X - W - Y$ holds. However, these are not the only cases where we have $I(X;Y) = C(X;Y)$ [42]. In this paper, we show that the equalities $h_0(P_{XY}) = g_0(P_{XY}) = H(Y|X)$ are fulfilled when $I(X;Y) = C(X;Y)$. Due to the observation the constraint $I(X;Y) = C(X;Y)$ generalizes the constraint $H(X|Y) = 0$.

In the following we recall the definition of the common information using [43]. The Gács-Körner common information between two RVs X and Y is defined as the entropy of the common part between X and Y, i.e., $C(X;Y) = H(U)$, where U is the common part between X and Y [43]. The common part U is defined based on the graphical representation of X and Y, which is shown to be a deterministic function of only X and only Y. Moreover, the Gács-Körner common information satisfies $C(X;Y) \leq I(X;Y)$ with equality if and only if the Markov chain $X - U - Y$ holds [42]. For more information about the Gács-Körner common information see [43] and [15, Appendix B].

In this work, we can use both notions of common information defined in [41] or [43], since our focus is on scenarios where the common information and mutual information between X and Y are equal, i.e., $C(X;Y) = I(X;Y)$. In other words, our results hold for both Wyner and Gács-Körner notions of common information.

In the following, we present the relation between the weighted strong ℓ_1-privacy criterion and bounded mutual information.

Using Pinsker's inequality [44], we show that for any $\epsilon \geq 0$ and pair (X,U), bounded mutual information implies the weighted strong ℓ_1-privacy criterion with a different leakage. In more detail, we have

$$I(X;U) \leq \epsilon \Rightarrow P_U(u) d(P_{X|U}(\cdot,u), P_X) \leq \sqrt{2\epsilon},\ \forall u. \quad (18)$$

The proof for (18) is provided in Appendix A. Note that (18) can help us to find relations between the privacy problems considering bounded mutual information and the weighted strong ℓ_1-privacy criterion in both scenarios where the private data is hidden or observable. By using (18) we have

$$h_\epsilon(P_{XY}) \leq h^{w\ell}_{\bar{\epsilon}}(P_{XY}), \quad (19)$$
$$g_\epsilon(P_{XY}) \leq g^{w\ell}_{\bar{\epsilon}}(P_{XY}), \quad (20)$$

where $\bar{\epsilon} = \sqrt{2\epsilon}$. Next, we present the relation between the strong ℓ_1-privacy criterion and bounded mutual information.

Using reverse Pinsker's inequality [44], we show that for any $\epsilon \geq 0$ and pair (X,U), the strong ℓ_1-privacy criterion implies the bounded mutual information with a different leakage. In more detail we have

$$d(P_{X|U}(\cdot|u), P_X) \leq \epsilon,\ \forall u \Rightarrow I(X;U) \leq \frac{\epsilon^2}{\min P_X}. \quad (21)$$

The proof for (21) is provided in Appendix A. Similarly, (21) enables us to find relations between the privacy problems considering bounded mutual information and the strong ℓ_1-privacy criterion in both scenarios where the private data is hidden or observable. Using (21) we have

$$h^\ell_\epsilon(P_{XY}) \leq h_{\epsilon'}(P_{XY}), \quad (22)$$
$$g^\ell_\epsilon(P_{XY}) \leq g_{\epsilon'}(P_{XY}), \quad (23)$$

where $\epsilon' = \frac{\epsilon^2}{\min P_X}$. Finally, using (19), (20), (22), and (23) we have

$$h^\ell_{\tilde{\epsilon}}(P_{XY}) \leq h_\epsilon(P_{XY}) \leq h^{w\ell}_{\bar{\epsilon}}(P_{XY}),$$
$$g^\ell_{\tilde{\epsilon}}(P_{XY}) \leq g_\epsilon(P_{XY}) \leq g^{w\ell}_{\bar{\epsilon}}(P_{XY}),$$

where $\tilde{\epsilon} = \sqrt{\epsilon \min P_X}$ and $\bar{\epsilon} = \sqrt{2\epsilon}$. Another simple relation can be established considering the strong ℓ_1-privacy criterion and the weighted strong ℓ_1-privacy criterion. Clearly, the strong ℓ_1-privacy criterion implies the weighted strong ℓ_1-privacy criterion. In other words we have

$$d(P_{X|U}(\cdot|u), P_X) \leq \epsilon, \forall u \Rightarrow P_U(u) d(P_{X|U}(\cdot|u), P_X) \leq \epsilon, \forall u, \quad (24)$$

which lead to following results

$$h^\ell_\epsilon(P_{XY}) \leq h^{w\ell}_\epsilon(P_{XY}), \quad (25)$$
$$g^\ell_\epsilon(P_{XY}) \leq g^{w\ell}_\epsilon(P_{XY}). \quad (26)$$

Finally, we present the relation between the privacy problem with prioritized private data with other problems. Let $X = (X_1, X_2)$, in this case using (19) we have

$$h^p_\epsilon(P_{X_1X_2Y}) \leq h_\epsilon(P_{X_1X_2Y}) \leq h^{w\ell}_{\bar{\epsilon}}(P_{X_1X_2Y}),$$

where $\bar{\epsilon} = \sqrt{2\epsilon}$. Consequently, lower bounds on $h^p_\epsilon(P_{X_1X_2Y})$ can be used as lower bounds on $h_\epsilon(P_{X_1X_2Y})$.

IV. Main results

In this part, we provide lower and upper bounds for the privacy problems defined in (1), (2), (3), (4), (5), (6) and (8). We study the tightness of the bounds in special cases and compare them in examples. In more detail, in the first part of the results, which corresponds to *privacy-utility trade-off with non-zero leakage*, we show that the upper bound on $h_\epsilon(P_{XY})$ is achieved when the common information and mutual information between X and Y are equal. We provide necessary and sufficient conditions for the achievability of the obtained upper bound in general. Moreover, in cases where no leakage is allowed, i.e., $\epsilon = 0$, we provide new bounds that generalize the previous bounds. In the second part of the results in this section corresponding to *privacy-utility trade-off with non-zero leakage and per-letter privacy criterions* we use concepts from information geometry to find lower bounds on $g^{w\ell}_\epsilon(P_{XY})$ and $g^\ell_\epsilon(P_{XY})$. In the remaining parts of the main results, we provide lower and upper bounds for $h^p_\epsilon(P_{X_1X_2Y})$ and study them for special cases.

A. Privacy-utility trade-off with non-zero leakage

In this section, we extend FRL and SFRL for correlated random variables X and U, i.e., $0 \leq I(U;X) = \epsilon$. We refer to them as Extended Functional Representation Lemma (EFRL) and Extended Strong Functional Representation Lemma (ESFRL). We show that the extended lemmas, i.e., EFRL and ESFRL, enable us to find lower bounds on $h_\epsilon(P_{XY})$.

Remark 8. *The idea of extending Functional Representation Lemma and Strong Functional Representation Lemma is basically adding a randomized response argument to the random variable U found by Lemma 1 and Lemma 2. The idea is simple, when it can be connected with an old principle it gets creditable. The randomized response technique has been introduced in [35] and has been used in many related works, e.g., [45, Theorem 2] and [46, Theorem 2].*

Lemma 4. *(Extended Functional Representation Lemma): For any $0 \leq \epsilon < I(X;Y)$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite, there exists a RV U defined on $\mathcal{U}$ such that the leakage between X and U is equal to ϵ, i.e., we have*

$$I(U;X) = \epsilon,$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X) = 0,$$

and $|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+1]\,[|\mathcal{X}|+1]$.

Proof: The proof is provided in Appendix B. ∎

Lemma 5. *(Extended Strong Functional Representation Lemma): For any $0 \leq \epsilon < I(X;Y)$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite with $I(X,Y) < \infty$, there exists a RV U defined on $\mathcal{U}$ such that the leakage between X and U is equal to ϵ, i.e., we have*

$$I(U;X) = \epsilon,$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X) = 0,$$

$I(X;U|Y)$ can be upper bounded as follows

$$I(X;U|Y) \leq \alpha H(X|Y) + (1-\alpha)\left[\log(I(X;Y)+1)+4\right],$$

and $|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+2]\,[|\mathcal{X}|+1]$, where $\alpha = \frac{\epsilon}{H(X)}$.

Proof: The proof is provided in Appendix B. ∎

In Lemma 8, which is proved in Appendix B, we show that there exists a RV U that satisfies (11), (12) and has bounded entropy. The lemma is a generalization of [16, Lemma 2] for dependent X and U.

Before stating the next theorem we derive an expression for $I(Y;U)$. We have

$$\begin{aligned} I(Y;U) &= I(X,Y;U) - I(X;U|Y), \\ &= I(X;U) + I(Y;U|X) - I(X;U|Y), \\ &= I(X;U) + H(Y|X) - H(Y|U,X) - I(X;U|Y). \end{aligned} \tag{27}$$

As argued in [16], (27) is an important observation to find lower and upper bounds for $h_\epsilon(P_{XY})$ and $g_\epsilon(P_{XY})$.

In next theorem we generalize [16, Theorem 5] considering non-zero leakage. Necessary and sufficient conditions that the utility in the observable private data scenario is larger than ϵ, i.e., $h_\epsilon(P_{XY}) > \epsilon$, are characterized.

Theorem 1. *For any $0 \leq \epsilon < I(X;Y)$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$, we have*

$$h_\epsilon(P_{XY}) > \epsilon \tag{28}$$

if and only if

$$H(Y|X) > 0. \tag{29}$$

Proof: The proof is provided in Appendix B. ∎

In the next theorem we present lower bounds on $h_\epsilon(P_{XY})$ and find the conditions under which the bounds are tight. The following theorem is a generalization of [16, Th. 6] for correlated X and U, i.e., $I(X;U) \leq \epsilon$.

Theorem 2. *For any $0 \leq \epsilon < I(X;Y)$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$, we have*

$$h_\epsilon(P_{XY}) \geq \max\{L_h^1(\epsilon), L_h^2(\epsilon), L_h^3(\epsilon)\}, \tag{30}$$

where

$$\begin{aligned} L_h^1(\epsilon) &= H(Y|X) - H(X|Y) + \epsilon = H(Y) - H(X) + \epsilon, \\ L_h^2(\epsilon) &= H(Y|X) - \alpha H(X|Y) + \epsilon \\ &\quad - (1-\alpha)\left(\log(I(X;Y)+1)+4\right), \\ L_h^3(\epsilon) &= \epsilon\frac{H(Y)}{I(X;Y)} + g_0(P_{XY})\left(1 - \frac{\epsilon}{I(X;Y)}\right), \end{aligned}$$

and $\alpha = \frac{\epsilon}{H(X)}$. The lower bound in (30) *is tight if $H(X|Y) = 0$, i.e., X is a deterministic function of Y. Furthermore, if the lower bound $L_h^1(\epsilon)$ is tight then we have $H(X|Y) = 0$.*

Proof: The proof is provided in Appendix B. The lower bounds $L_h^1(\epsilon)$ and $L_h^2(\epsilon)$ are derived by using Lemma 4 and Lemma 5. ∎

Next, we argue that when non-zero leakage is allowed, EFRL and ESFRL can strictly improve the utility compared with FRL and SFRL, respectively. Using (27) utility achieved by FRL is $H(Y|X) - H(X|Y)$, which is less than or equal to utility achieved by EFRL, i.e., $H(Y|X) - H(X|Y) + \epsilon = L_1^\epsilon$. Furthermore, utility achieved by SFRL is $H(Y|X) - (\log(I(X;Y)+1)+4)$, which is less than or equal to utility attained by ESFRL, i.e., $H(Y|X) + \epsilon - \alpha H(X|Y) - (1-\alpha)\left(\log(I(X;Y)+1)+4\right) = L_h^2(\epsilon)$, since we have

$$\begin{aligned} &L_h^2(\epsilon) - \left(H(Y|X) - (\log(I(X;Y)+1)+4)\right) = \\ &\epsilon + \frac{\epsilon}{H(X)}\left(\log(I(X;Y)+1)+4\right) - \frac{\epsilon}{H(X)}H(X|Y) \\ &\geq 0. \end{aligned}$$

The latter holds since $H(X|Y) \leq H(X)$. Equality holds if and only if $\epsilon = 0$. For the only if part, noting that $\epsilon + \frac{\epsilon}{H(X)}\left(\log(I(X;Y)+1)+4\right) \geq \epsilon$ and $\epsilon \geq \frac{\epsilon}{H(X)}H(X|Y)$.

Hence, for non-zero leakage EFRL and ESFRL strictly improve the bounds attained by the FRL and SFRL.

In the following we let $\epsilon = 0$ in Theorem 2 and derive lower bounds on $h_0(P_{XY})$. In this case, for any pair of RVs (X, Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have

$$h_0(P_{XY}) \geq \max\{L_1^0, L_2^0\}, \tag{31}$$

where

$$\begin{aligned} L_1^0 &= H(Y|X) - H(X|Y) = H(Y) - H(X), \\ L_2^0 &= H(Y|X) - (\log(I(X;Y)+1)+4). \end{aligned}$$

Note that the lower bound L_1^0 has been derived in [16, Th. 6], while the lower bound L_2^0 is new. Hence, the lower bound derived in (31) generalizes the bound found in [16, Th. 6].

In the next two examples we compare the bounds L_1^ϵ, L_2^ϵ and L_3^ϵ in special cases where $I(X;Y) = 0$ and $H(X|Y) = 0$.

Example 1. *Let X and Y be independent. Then, we have*

$$\begin{aligned} L_h^1(\epsilon) &= H(Y) - H(X) + \epsilon, \\ L_h^2(\epsilon) &= H(Y) - \frac{\epsilon}{H(X)} H(X) + \epsilon - 4(1 - \frac{\epsilon}{H(X)}), \\ &= H(Y) - 4(1 - \frac{\epsilon}{H(X)}). \end{aligned}$$

Thus,

$$\begin{aligned} L_h^2(\epsilon) - L_h^1(\epsilon) &= H(X) - 4 + \epsilon(\frac{4}{H(X)} - 1), \\ &= (H(X) - 4)(1 - \frac{\epsilon}{H(X)}). \end{aligned}$$

Consequently, for independent X and Y if $H(X) > 4$, then $L_h^2(\epsilon) > L_h^1(\epsilon)$, i.e., the second lower bound is dominant and $h_\epsilon(P_X P_Y) \geq L_h^2(\epsilon)$.

Example 2. *Let X be a deterministic function of Y. As we have shown in Theorem 2, if $H(X|Y) = 0$, then*

$$\begin{aligned} L_h^1(\epsilon) = L_h^3(\epsilon) &= H(Y|X) + \epsilon \\ &\geq H(Y|X) + \epsilon - (1 - \frac{\epsilon}{H(X)})(\log(H(X)+1)+4) \\ &= L_h^2(\epsilon). \end{aligned}$$

Therefore, L_1^ϵ and L_3^ϵ become dominants.

In Lemma 9 which is provided in Appendix B, we find a lower bound for $\sup_U H(U)$ where U satisfies the leakage constraint $I(X;U) \leq \epsilon$, the bounded cardinality stated in Lemma 4 and $H(Y|U,X) = 0$.

In the next result, using (27) we derive an upper bound on $h_\epsilon(P_{XY})$.

Lemma 6. *For any $0 \leq \epsilon < I(X;Y)$ and pair of RVs (X, Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have*

$$g_\epsilon(P_{XY}) \leq h_\epsilon(P_{XY}) \leq H(Y|X) + \epsilon = U_h^1(\epsilon).$$

Proof: By using (27) we have

$$h_\epsilon(P_{XY}) \leq H(Y|X) + \sup I(U;X) \leq H(Y|X) + \epsilon.$$

■

Corollary 1. *If X is a deterministic function of Y, then by using Theorem 2 and Lemma 6 we have*

$$g_\epsilon(P_{XY}) = h_\epsilon(P_{XY}) = H(Y|X) + \epsilon,$$

since in this case the Markov chain $X - Y - U$ holds.

In the next result we find a larger set of distributions P_{XY} compared to Corollary 1 for which we have $g_\epsilon(P_{XY}) = h_\epsilon(P_{XY})$, where common information corresponds to the Wyner [41] or Gács-Körner [43] notions of common information. One advantage of having $g_\epsilon(P_{XY}) = h_\epsilon(P_{XY})$ is discussed after Theorem 3, where we show that under the assumption of equality between common information and mutual information the upper bound in Lemma 6 is tight.

Proposition 3. *If the common information and the mutual information between X and Y are equal, then we have*

$$g_\epsilon(P_{XY}) = h_\epsilon(P_{XY}).$$

Proof: The proof follows similar arguments as the proof of [14, Th. 2]. Let U^* be an optimizer of $h_\epsilon(P_{XY})$, then by using the proof of [14, Th. 2] we can construct U' satisfying the Markov chain $X - Y - U'$, $I(U^*;Y) = I(U';Y)$ and $I(U';X) \leq I(U^*;X)$ which completes the proof. ■

Noting that if X is a deterministic function of Y, then the common information and mutual information between X and Y are equal, i.e., we have

$$H(Y|X) = 0 \Rightarrow C(X;Y) = I(X;Y). \tag{32}$$

The proof for (32) is provided in Appendix B. Thus, the constraint in Proposition 3 contains a larger set of joint distributions P_{XY} compared to the constraint used in Corollary 1. In the next lemma we provide an important property of an optimizer of $h_\epsilon(P_{XY})$ which is used to derive equivalencies in Theorem 3.

Lemma 7. *Let $\bar{U}$ be an optimizer of $h_\epsilon(P_{XY})$. We have*

$$H(Y|X,\bar{U}) = 0.$$

Proof: The detailed proof is provided in Appendix B and is similar to the proof of [16, Lemma 5]. The proof is by contradiction and we show that if for an optimizer of $h_\epsilon(P_{XY})$, denoted by $\bar{U}$, we have $H(Y|X,\bar{U}) > 0$, then we can build U such that it satisfies $I(U;X) \leq \epsilon$ and achieves strictly greater utility than $\bar{U}$, which contradicts the assumption. ■

In the next theorem we generalize the equivalent statements in [16, Th. 7] for bounded leakage between X and U, i.e., $I(X;U) \leq \epsilon$.

Theorem 3. *For any $\epsilon < I(X;Y)$, we have the following equivalencies*

i. $g_\epsilon(P_{XY}) = H(Y|X) + \epsilon$,
ii. $g_\epsilon(P_{XY}) = h_\epsilon(P_{XY})$,
iii. $h_\epsilon(P_{XY}) = H(Y|X) + \epsilon$.

Proof: The statements i $\Rightarrow$ ii and iii $\Rightarrow$ i can be shown by using Lemma 6 and Lemma 7, respectively. For proving ii $\Rightarrow$ iii, let $\bar{U}$ be an optimizer of $g_\epsilon(P_{XY})$, by using Lemma 7,

(27) and Markov chain $X-Y-U$ we show that $I(\bar{U};Y)=I(X;\bar{U})+H(Y|X)$. Furthermore, we show that $I(X;\bar{U})=\epsilon$, thus, $h_\epsilon(P_{XY})=H(Y|X)+\epsilon$. A detailed proof is provided in Appendix B. ∎

By using Theorem 3 and Proposition 3, Corollary 1 can be strengthened as follows.

Corollary 2. *If the common information and mutual information between X and Y are equal then we have*

$$g_\epsilon(P_{XY})=h_\epsilon(P_{XY})=H(Y|X)+\epsilon.$$

Lemma 7 and Theorem 3 generalize [16, Th. 7] for non-zero leakage. Now let $\epsilon=0$ in Lemma 7 and Theorem 3. As argued in [16] when X is a deterministic function of Y, the necessary and sufficient conditions for having equality in Lemma 7 are fulfilled. Furthermore, this result holds when X and Y are independent or Y is a deterministic function of X. However, in this work we have shown that for any $0\leq\epsilon<I(X;Y)$, this statement can be generalized and we can substitute the condition that X is a deterministic function of Y by the condition that the common information and mutual information between X and Y are equal.

Special case: $H(Y|X)=0$

In this section we study the bounds derived in Theorem 2 considering the scenario where $H(Y|X)=0$. Considering zero leakage scenario, $\epsilon=0$, the constraint $H(Y|X)=0$ results in zero utility, i.e., $h_0(P_{XY})=0$, [16, Theorem 5]. This result can be verified by using (27). We have

$$\begin{aligned}I(U;Y)&=I(U;X)+H(Y|X)-H(Y|X,U)-I(U;X|Y)\\&=-I(U;X|Y).\end{aligned}$$

Using $0\leq I(U;Y)=-I(U;X|Y)\leq 0$ we must have $I(U;Y)=0$. Here, we show that when non-zero leakage is allowed we can achieve non-zero utility, however, Theorem 1 ensures that utilities larger than ϵ can not be attained, i.e., $h_\epsilon(P_{XY})\leq\epsilon$. Next, we show that if $H(Y|X)=0$ we have $h_\epsilon(P_{XY})=\epsilon$, we then propose a RV U that attains it. Using a similar proof as (32) we conclude that if Y is a deterministic function of X the common information and mutual information between X and Y are equal. As a result, by using Corollary 2 we have $h_\epsilon(P_{XY})=\epsilon$. To achieve ϵ let $U=\begin{cases}Y, & \text{w.p. } \alpha\\ c, & \text{w.p. } 1-\alpha\end{cases}$, where c is a constant which does not belong to the support of X and $\alpha=\frac{\epsilon}{I(X;Y)}$. We emphasize that since we only consider the range $\epsilon<I(X;Y)$, we have $\alpha<1$. To verify the privacy constraint we have

$$\begin{aligned}I(U;X)&=H(X)-H(X|U)\\&=H(X)-\alpha H(X|Y)-(1-\alpha)H(X)\\&=\alpha I(X;Y)\\&=\epsilon.\end{aligned}$$

Using (27) we have

$$\begin{aligned}I(U;Y)&\overset{(a)}{=}\epsilon-H(X|Y)+H(X|Y,U)\\&=\epsilon-H(X|Y)+H(X|Y)\\&=\epsilon,\end{aligned}$$

where in (a) we use $I(U;X)=\epsilon$ and $H(Y|X)=H(Y|X,U)=0$.

Special case: $\epsilon=0$ (Independent X and U)

In this section we derive new lower and upper bounds for $h_0(P_{XY})$ and compare them with the previous bounds found in [16]. In the next theorem lower and upper bounds on $h_0(P_{XY})$ are provided.

Theorem 4. *For any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have*

$$\max\{L_1^0,L_2^0\}\leq h_0(P_{XY})\leq\min\{U_1^0,U_2^0\},$$

where L_1^0 and L_2^0 are defined in (31) *and*

$$\begin{aligned}&U_1^0=H(Y|X),\\&U_2^0=H(Y|X)+\\&\sum_{y\in\mathcal{Y}}\int_0^1\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt+\\&I(X;Y).\end{aligned}$$

Furthermore, if $|\mathcal{Y}|=2$, then we have

$$h_0(P_{XY})=U_2^0.$$

Proof: The proof is provided in Appendix B. ∎

As mentioned before the upper bound U_1^0 has been derived in [16, Th. 7]. The upper bound U_2^0 is a new upper bound. Thus, the lower and upper bounds on $h_0(P_{XY})$ stated in Theorem 4 generalize the bounds in [16]. Furthermore, in case of binary Y the exact expression for $h_0(P_{XY})$ has been derived.

To study the bounds U_2^0 and U_1^0 let us consider a case where X is a deterministic function of Y. Using Corollary 1 we know that the upper bound U_1^0 is tight. If X is a deterministic function of Y, i.e., $H(X|Y)=0$, we have

$$\begin{aligned}&\sum_{y\in\mathcal{Y}}\int_0^1\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt+\\&I(X;Y)=0.\end{aligned}\tag{33}$$

The proof for (33) is provided in Appendix B. According to (33), if X is a deterministic function of Y, then we have $U_2^0=U_1^0$. In other words, in this case U_2^0 is tight as well as U_1^0. In the next example we compare the bounds U_1^0 and U_2^0 for a $BSC(\theta)$ and we show that U_2^0 can improve U_1^0.

Example 3. *(Binary Symmetric Channel) Let the binary RVs $X\in\{0,1\}$ and $Y\in\{0,1\}$ have the following joint distribution*

$$P_{XY}(x,y)=\begin{cases}\frac{1-\theta}{2}, & x=y\\ \frac{\theta}{2}, & x\neq y\end{cases},$$

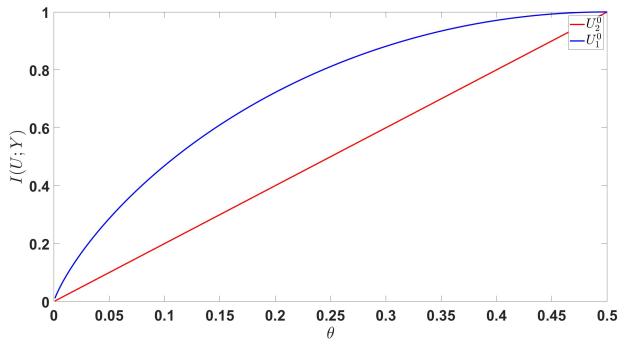

Fig. 2. Comparing the upper bounds U_1^0 and U_2^0 for $BSC(\theta)$. The blue curve illustrates the upper bound found in [16] and the red line shows the upper bound found in this work. In fact by using Theorem 4 the red curve corresponding to U_2^0 can be achieved and it presents the solution for $h_0(P_{XY})$.

where $\theta < \frac{1}{2}$. We obtain

$$\sum_{y\in\mathcal{Y}}\int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt =$$
$$\int_\theta^{1-\theta} \mathbb{P}_X\{P_{Y|X}(0|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(0|X)\geq t\})dt +$$
$$\int_\theta^{1-\theta} \mathbb{P}_X\{(P_{Y|X}(1|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(1|X)\geq t\})dt =$$
$$(1-2\theta)\left(P_X(0)\log(P_X(0)) + P_X(1)\log(P_X(1))\right) =$$
$$-(1-2\theta)H(X) =$$
$$-(1-2\theta).$$

Thus,

$$U_2^0 = H(Y|X) +$$
$$\sum_{y\in\mathcal{Y}}\int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt +$$
$$I(X;Y) = h(\theta) - (1-2\theta) + (1-h(\theta)) = 2\theta,$$
$$U_1^0 = H(Y|X) = h(\theta),$$

where $h(\cdot)$ corresponds to the binary entropy function. As shown in Fig. 2, we have

$$h_0(P_{XY}) \leq U_2^0 \leq U_1^0.$$

However by using Theorem 4, since $|\mathcal{Y}| = 2$ the upper bound U_2^0 is achieved and we have

$$h_0(P_{XY}) = U_2^0.$$

Example 4. *(Erasure Channel) Let the RVs $X \in \{0,1\}$ and $Y \in \{0,e,1\}$ have the following joint distribution*

$$P_{XY}(x,y) = \begin{cases} \frac{1-\theta}{2}, & x = y \\ \frac{\theta}{2}, & y = e \\ 0, & else \end{cases},$$

where $\theta < \frac{1}{2}$. We have

$$\sum_{y\in\mathcal{Y}}\int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt$$
$$= \int_0^{1-\theta} \mathbb{P}_X\{P_{Y|X}(0|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(0|X)\geq t\})dt$$
$$+ \int_0^{1-\theta} \mathbb{P}_X\{P_{Y|X}(1|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(1|X)\geq t\})dt$$
$$= -(1-\theta)H(X) = -(1-\theta).$$

Thus,

$$U_2^0 = H(Y|X) +$$
$$\sum_{y\in\mathcal{Y}}\int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X)\geq t\}\log(\mathbb{P}_X\{P_{Y|X}(y|X)\geq t\})dt +$$
$$I(X;Y) = h(\theta) - (1-\theta) + h(\theta) + 1 - \theta - h(\theta) = h(\theta),$$
$$U_1^0 = H(Y|X) = h(\theta).$$

Hence, in this case, $U_1^0 = U_2^0 = h(\theta)$. Furthermore, in [16, Example 8], it has been shown that for this pair of (X,Y) we have $g_0(P_{XY}) = h_0(P_{XY}) = h(\theta)$.

In [34, Prop. 2] it has been shown that for every $\alpha \geq 0$, there exist a pair (X,Y) such that $I(X;Y) \geq \alpha$ and

$$\psi(X \to Y) \geq \log(I(X;Y)+1) - 1. \tag{34}$$

Let (X,Y) be as in [34, Prop. 2], i.e. (X,Y) satisfies (34). Then for such pair we have

$$H(Y|X) - \log(I(X;Y)+1) - 4 \leq h_0(P_{XY}) \tag{35}$$
$$\leq H(Y|X) - \log(I(X;Y)+1) + 1. \tag{36}$$

The lower bound in (35) follows from (31). For the upper bound, we use (27) and (34) so that

$$I(U;Y) \leq H(Y|X) - \psi(X \to Y)$$
$$\leq H(Y|X) - \log(I(X;Y)+1) + 1.$$

For such pair using (35) and (31) we can conclude that the lower bound $L_2^0 = H(Y|X) - (\log(I(X;Y)+1)+4)$ is tight within 5 bits.

B. Privacy-utility trade-off with non-zero leakage and per-letter privacy constraints

In this section, we provide lower and upper bounds on the privacy problems defined in (3), (4), (5), and (6). To do so, we first introduce similar lemmas as Lemma 4 and Lemma 5 in Appendix C, where we have replaced the mutual information constraint, i.e., $I(U;X) = \epsilon$, with the weighted strong privacy criterion 1 defined in (3) and (4). In the remaining part of this work $d(\cdot,\cdot)$ corresponds to the total variation distance, i.e., $d(P,Q) = \sum_x |P(x) - Q(x)|$.
In the next proposition we find a lower bound on $h_\epsilon^{w\ell}(P_{XY})$ using Lemma 10 and Lemma 11 which are provided in Appendix C.

Proposition 4. *For any $0 \leq \epsilon < \sqrt{2I(X;Y)}$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have*

$$h_\epsilon^{w\ell}(P_{XY}) \geq \max\{L^1_{h^{w\ell}}(\epsilon), L^2_{h^{w\ell}}(\epsilon)\}, \tag{37}$$

where

$$\begin{aligned} L^1_{h^{w\ell}}(\epsilon) =& H(Y|X) - H(X|Y) + \frac{\epsilon^2}{2}, \\ L^2_{h^{w\ell}}(\epsilon) =& H(Y|X) - \alpha H(X|Y) + \frac{\epsilon^2}{2} \\ & -(1-\alpha)\left(\log(I(X;Y)+1)+4\right), \end{aligned}$$

with $\alpha = \frac{\epsilon^2}{2H(X)}$.

Proof: The proof is provided in Appendix C. ∎

Considering the lower bounds on $h_\epsilon^{w\ell}(P_{XY})$ and $h_\epsilon(P_{XY})$, i.e., $L^1_{h^{w\ell}}(\epsilon)$ and $L^1_h(\epsilon)$, derived in Proposition 4 and Theorem 2, we can conclude that in high privacy regimes the utility attained by $L^1_{h^{w\ell}}(\epsilon)$ is lees than the utility achieved by $L^1_h(\epsilon)$. On the other hand, by letting $\bar{\epsilon} = \sqrt{2}\epsilon$, we have $L^1_{h^{w\ell}}(\bar{\epsilon}) = L^1_h(\epsilon)$ and $L^2_{h^{w\ell}}(\bar{\epsilon}) = L^2_h(\epsilon)$. This means that the lower bounds $h_\epsilon^{w\ell}(P_{XY})$ and $h_\epsilon(P_{XY})$, i.e., $L^1_{h^{w\ell}}(\epsilon)$, $L^2_{h^{w\ell}}(\epsilon)$, $L^1_h(\epsilon)$, and $L^2_h(\epsilon)$, follow the same inequality as (19). In the next section, we provide a lower bound on $g_\epsilon^{w\ell}(P_{XY})$ by following the same approach as in [13]. For more details about the proofs and steps of approximation see [13, Section III].

Lower bound on $g_\epsilon^{w\ell}(P_{XY})$

In [13], we show that $g_\epsilon^{\ell}(P_{XY})$ can be approximated by a linear program using information geometry concepts. Using this result we can derive a lower bound for $g_\epsilon^{\ell}(P_{XY})$. In this part, we follow a similar approach to approximate $g_\epsilon^{w\ell}(P_{XY})$, which results in a lower bound. We emphasize that in contrast with [13] $P_{X,U}(\cdot,u)$ and $P_X P_U(u)$ are not distribution vectors since the sum of elements in both vectors equal to $P_U(u)$, however, the approach to approximate $g_\epsilon^{w\ell}(P_{XY})$ does not change. Similar to [13], for sufficiently small ϵ, by using the leakage constraint in $g_\epsilon^{w\ell}(P_{XY})$, i.e., the weighted strong ℓ_1-privacy criterion, we can rewrite $P_{X,U}(\cdot,u)$ as a perturbation of $P_X P_U(u)$. Thus, for any u we can write $P_{X,U}(\cdot,u) = P_X P_U(u) + \epsilon J_u$, where $J_u \in \mathbb{R}^{|\mathcal{X}|}$ is a perturbation vector and satisfies the following properties:

$$\mathbf{1}^T \cdot J_u = 0, \ \forall u, \tag{38}$$

$$\sum_u J_u = \mathbf{0} \in \mathbb{R}^{|\mathcal{X}|}, \tag{39}$$

$$\mathbf{1}^T \cdot |J_u| \leq 1, \ \forall u, \tag{40}$$

where $|\cdot|$ corresponds to the absolute value of the vector. We define matrix $M \in \mathbb{R}^{|\mathcal{X}|\times|\mathcal{Y}|}$, which is used in the remaining part, as follows: Let V be the matrix of right eigenvectors of $P_{X|Y}$, i.e., $P_{X|Y} = U\Sigma V^T$ and $V = [v_1,\ v_2,\ ...,\ v_{|\mathcal{Y}|}]$, then M is defined as

$$M \triangleq \left[v_1,\ v_2,\ ...,\ v_{|\mathcal{X}|}\right]^T.$$

Similar to [13, Proposition 2], we have the following result.

Proposition 5. *In (1), it suffices to consider U such that $|\mathcal{U}| \leq |\mathcal{Y}|$. Since the supremum in (1) is achieved, we can replace the supremum by the maximum.*

Proof: The proof follows the similar lines as the proof of [13, Proposition 2]. The only difference is that the new convex and compact set is as follows

$$\Psi = \left\{ y \in \mathbb{R}_+^{|\mathcal{Y}|} | My = MP_Y + \frac{\epsilon}{P_U(u)} M \begin{bmatrix} P^{-1}_{X|Y_1} J_u \\ 0 \end{bmatrix}, J_u \in \mathcal{J} \right\},$$

where $\mathcal{J} = \{J \in \mathbb{R}_+^{|\mathcal{X}|} | \ \|J\|_1 \leq 1,\ \mathbf{1}^T \cdot J = 0\}$ and $\mathbb{R}_+$ corresponds to non-negative real numbers. Only non-zero weights $P_U(u)$ are considered since in the other case the corresponding $P_{Y|U}(\cdot|u)$ does not appear in $H(Y|U)$. ∎

Next, we show that $P_{Y|U}(\cdot|u)$ lies in a convex polytope. If the Markov chain $X - Y - U$ holds, for sufficiently small ϵ and every $u \in \mathcal{U}$, the vector $P_{Y|U}(\cdot|u)$ lies in the following convex polytope

$$\mathbb{S}_u = \left\{ y \in \mathbb{R}_+^{|\mathcal{Y}|} | My = MP_Y + \frac{\epsilon}{P_U(u)} M \begin{bmatrix} P^{-1}_{X|Y_1} J_u \\ 0 \end{bmatrix} \right\}, \tag{41}$$

where J_u satisfies (38), (39), and (40). Furthermore, $P_U(u) > 0$, otherwise $P_{Y|U}(\cdot|u)$ does not appear in $I(Y;U)$. To prove (41) using the Markov chain $X - Y - U$, we have

$$P_{X|U=u} - P_X = P_{X|Y}[P_{Y|U=u} - P_Y] = \epsilon \frac{J_u}{P_U(u)}.$$

Thus, by following the similar lines as [13, Lemma 2] and using the properties of Null(M) as [13, Lemma 1], we have

$$MP_{Y|U}(\cdot|u) = MP_Y + \frac{\epsilon}{P_U(u)} M \begin{bmatrix} P^{-1}_{X|Y_1} J_u \\ 0 \end{bmatrix}.$$

By using the same arguments as [13, Lemma 3], it can be shown that any vector inside $\mathbb{S}_u$ is a standard probability vector. Thus, by using [13, Lemma 3] and (41) we have following result.

Theorem 5. *We have the following equivalency*

$$\min_{\substack{P_{U|Y}: X-Y-U \\ d(P_{X,U}(\cdot,u), P_X P_U(u)) \leq \epsilon,\ \forall u \in \mathcal{U}}} H(Y|U) = \min_{\substack{P_U,\ P_{Y|U=u} \in \mathbb{S}_u,\ \forall u \in \mathcal{U}, \\ \sum_u P_U(u) P_{Y|U=u} = P_Y, \\ J_u \text{ satisfies (38), (39), and (40)}}} H(Y|U). \tag{42}$$

Furthermore, similar to [13, Proposition 3], it can be shown that the minimum of $H(Y|U)$ occurs at the extreme points of the sets $\mathbb{S}_u$, i.e., for each $u \in \mathcal{U}$, $P^*_{Y|U}(\cdot|u)$ that minmizes $H(Y|U)$ must belong to the extreme points of $\mathbb{S}_u$. To find the extreme points of $\mathbb{S}_u$ let Ω be the set of indices which correspond to $|\mathcal{X}|$ linearly independent columns of M, i.e., $|\Omega| = |\mathcal{X}|$ and $\Omega \subset \{1,..,|\mathcal{Y}|\}$. Let $M_\Omega \in \mathbb{R}^{|\mathcal{X}|\times|\mathcal{X}|}$ be the submatrix of M with columns indexed by the set Ω. Assume that $\Omega = \{\omega_1,..,\omega_{|\mathcal{X}|}\}$, where $\omega_i \in \{1,..,|\mathcal{Y}|\}$ and all elements are arranged in an increasing order. The ω_i-th element of the extreme point V^*_Ω can be found as i-th element of $M_\Omega^{-1}(MP_Y + \frac{\epsilon}{P_U(u)} M \begin{bmatrix} P^{-1}_{X|Y_1} J_u \\ 0 \end{bmatrix})$, i.e., for $1 \leq i \leq |\mathcal{X}|$

we have

$$V^*_\Omega(\omega_i) = \left(M_\Omega^{-1} M P_Y + \frac{\epsilon}{P_U(u)} M_\Omega^{-1} M \begin{bmatrix} P^{-1}_{X|Y_1} J_u \\ 0 \end{bmatrix} \right)(i). \tag{43}$$

Other elements of V^*_Ω are set to be zero. Now we approximate the entropy of V^*_Ω.

Let $V^*_{\Omega_u}$ be an extreme point of the set $\mathbb{S}_u$, then we have

$$\begin{aligned} H(P_{Y|U=u}) &= \sum_{y=1}^{|\mathcal{Y}|} -P_{Y|U=u}(y)\log(P_{Y|U=u}(y)) \\ &= -(b_u + \frac{\epsilon}{P_U(u)} a_u J_u) + o(\epsilon), \end{aligned} \tag{44}$$

with $b_u = l_u\left(M^{-1}_{\Omega_u} M P_Y\right)$, $a_u = l_u\left(M^{-1}_{\Omega_u} M(1{:}|\mathcal{X}|) P^{-1}_{X|Y_1}\right) \in \mathbb{R}^{1\times|\mathcal{X}|}$, $l_u = \left[\log\left(M^{-1}_{\Omega_u} M P_Y(i)\right)\right]_{i=1:|\mathcal{X}|} \in \mathbb{R}^{1\times|\mathcal{X}|}$, and $M^{-1}_{\Omega_u} M P_Y(i)$ stands for i-th ($1 \le i \le |\mathcal{X}|$) element of the vector $M^{-1}_{\Omega_u} M P_Y$. Furthermore, $M(1{:}|\mathcal{X}|)$ stands for submatrix of M with first $|\mathcal{X}|$ columns. The proof of (44) follows similar lines as [13, Lemma 4] and is based on first order Taylor expansion of $\log(1+x)$. By using (44) we can approximate (3) as follows.

For sufficiently small ϵ, the minimization problem in (42) can be approximated as follows

$$\min_{P_U(.),\{J_u, u\in\mathcal{U}\}} -\left(\sum_{u=1}^{|\mathcal{Y}|} P_U(u) b_u + \epsilon a_u J_u\right) \tag{45}$$

subject to:

$$\sum_{u=1}^{|\mathcal{Y}|} P_U(u) V^*_{\Omega_u} = P_Y,\ \sum_{u=1}^{|\mathcal{Y}|} J_u = 0,\ P_U \in \mathbb{R}_+^{|\mathcal{Y}|}.$$
$$\mathbf{1}^T |J_u| \le 1,\ \mathbf{1}^T \cdot J_u = 0,\ \forall u \in \mathcal{U},$$

The proof of (45) follows directly from (44) and the fact that the minimum of $H(Y|U)$ occurs at the extreme points of the sets $\mathbb{S}_u$. Thus, for $P_{Y|U=u} = V^*_{\Omega_u},\ u \in \{1,..,|\mathcal{Y}|\}$, where $V^*_{\Omega_u}$ is defined in (43), $H(Y|U)$ can be approximated as follows

$$H(Y|U) = \sum_u P_U(u) H(P_{Y|U=u}) \cong \sum_{u=1}^{|\mathcal{Y}|} P_U(u) b_u + \epsilon a_u J_u.$$

By using the vector $\eta_u = P_U(u)\left(M^{-1}_{\Omega_u} M P_Y\right) + \epsilon\left(M^{-1}_{\Omega_u} M(1:|\mathcal{X}|) P^{-1}_{X|Y_1}\right)(J_u)$ for all $u \in \mathcal{U}$, where $\eta_u \in \mathbb{R}^{|\mathcal{X}|}$, we can write (45) as a linear program. The vector η_u corresponds to multiple of non-zero elements of the extreme point $V^*_{\Omega_u}$, furthermore, $P_U(u)$ and J_u can be uniquely found as

$$P_U(u) = \mathbf{1}^T \cdot \eta_u,$$
$$J_u = \frac{P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u}[\eta_u - (\mathbf{1}^T \eta_u) M^{-1}_{\Omega_u} M P_Y]}{\epsilon}.$$

By solving the linear program we obtain P_U and J_u for all u, thus, $P_{Y|U}(\cdot|u)$ can be computed using (43). Let $P^*_{U|Y}$ be found by the linear program, which solves (45), and let $I(U^*;Y)$ be evaluated by this kernel. Then we have

$$g^{w\ell}_\epsilon(P_{XY}) \ge I(U^*;Y) = L^1_{g^{w\ell}}(\epsilon). \tag{46}$$

The proof directly follows since the kernel $P^*_{U|Y}$ that achieves the approximate solution satisfies the constraints in (1). In the next result we present lower and upper bounds of $g^{w\ell}_\epsilon(P_{XY})$ and $h^{w\ell}_\epsilon(P_{XY})$.

Theorem 6. *For sufficiently small $\epsilon \ge 0$ and any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have*

$$L^1_{g^{w\ell}}(\epsilon) \le g^{w\ell}_\epsilon(P_{XY}),$$

and for any $\epsilon \ge 0$ we obtain

$$g^{w\ell}_\epsilon(P_{XY}) \le \frac{\epsilon|\mathcal{Y}||\mathcal{X}|}{\min P_X} + H(Y|X) = U_{g^{w\ell}}(\epsilon),$$
$$g^{w\ell}_\epsilon(P_{XY}) \le h^{w\ell}_\epsilon(P_{XY}).$$

Furthermore, for any $0 \le \epsilon \le \sqrt{2I(X;Y)}$ we have

$$\max\{L^1_{h^{w\ell}}(\epsilon), L^2_{h^{w\ell}}(\epsilon)\} \le h^{w\ell}_\epsilon(P_{XY}),$$

where $L^1_{h^{w\ell}}(\epsilon)$ and $L^2_{h^{w\ell}}(\epsilon)$ are defined in Proposition 4.

Proof: The proof is provided in Appendix C. ∎

In the next section we provide upper and lower bounds for $g^\ell_\epsilon(P_{XY})$ and $h^\ell_\epsilon(P_{XY})$.

Lower and Upper bounds on $g^\ell_\epsilon(P_{XY})$ and $h^\ell_\epsilon(P_{XY})$

As we mentioned earlier in [13], we have provided an approximate solution for $g^\ell_\epsilon(P_{XY})$ using a local approximation of $H(Y|U)$ for sufficiently small ϵ. Furthermore, in [13, Proposition 8] we specified permissible leakages. By using [13, Proposition 8], we can write

$$\begin{aligned} g^\ell_\epsilon(P_{XY}) &= \sup_{\substack{P_{U|Y}: X-Y-U \\ d(P_{X|U}(\cdot|u), P_X)\le\epsilon,\ \forall u}} I(Y;U) \\ &= \max_{\substack{P_{U|Y}: X-Y-U \\ d(P_{X|U}(\cdot|u), P_X)\le\epsilon,\ \forall u \\ |\mathcal{U}|\le|\mathcal{Y}|}} I(Y;U). \end{aligned} \tag{47}$$

In the next lemma we find a lower bound for $g^\ell_\epsilon(P_{XY})$, where we use the approximate problem for (5). Let the kernel $P_{U^*|Y}$ achieve the optimum solution in [13, Theorem 2]. Thus, $I(U^*;Y)$ evaluated by this kernel is a lower bound for $g^\ell_\epsilon(P_{XY})$. In other words, we have

$$g^\ell_\epsilon(P_{XY}) \ge I(U^*;Y) = L^1_{g^\ell}(\epsilon). \tag{48}$$

The proof follows since the kernel $P^*_{U|Y}$ that achieves the approximate solution satisfies the constraints in (5). Next we provide upper bounds for $g^\ell_\epsilon(P_{XY})$. To do so, we first bound the approximation error in [13, Theorem 2]. Let Ω^1 be the set of all $\Omega_i \subset \{1,..,|\mathcal{Y}|\},\ |\Omega_i| = |\mathcal{X}|$, such that each Ω_i produces a valid standard distribution vector $M^{-1}_{\Omega_i} M P_Y$, i.e., all elements in the vector $M^{-1}_{\Omega_i} M P_Y$ are positive.

Proposition 6. *Let the approximation error be the distance between $H(Y|U)$ and the approximation derived in [13, Theorem 2]. Then, for all $\epsilon < \frac{1}{2}\epsilon_2$, we have*

$$|\textit{Approximation error}| < \frac{3}{4}.$$

Furthermore, for all $\epsilon < \frac{1}{2}\frac{\epsilon_2}{\sqrt{|\mathcal{X}|}}$ the upper bound can be strengthened as follows

$$|\textit{Approximation error}| < \frac{1}{2(2\sqrt{|\mathcal{X}|}-1)^2} + \frac{1}{4|\mathcal{X}|}.$$

where $\epsilon_2 = \frac{\min_{y,\Omega\in\Omega^1} M_\Omega^{-1} M P_Y(y)}{\max_{\Omega\in\Omega^1}|\sigma_{\max}(H_\Omega)|}$, $H_\Omega = M_\Omega^{-1} M(1:|\mathcal{X}|)P_{X|Y_1}^{-1}$ and $\sigma_{\max}$ is the largest right singular value.

Proof: The proof is provided in Appendix C. ∎

As a result we can find an upper bound on $g_\epsilon^\ell(P_{XY})$. To do so let approx(g_ϵ^ℓ) be the value that the kernel $P_{U^*|Y}$ in (48) achieves, i.e., the approximate value in [13, (7)].

Corollary 3. *For any $0 \le \epsilon < \frac{1}{2}\epsilon_2$ we have*

$$g_\epsilon^\ell(P_{XY}) \le \textit{approx}(g_\epsilon^\ell) + \frac{3}{4} = U_{g^\ell}^1(\epsilon),$$

furthermore, for any $0 \le \epsilon < \frac{1}{2}\frac{\epsilon_2}{\sqrt{|\mathcal{X}|}}$ the upper bound can be strengthened as

$$g_\epsilon^\ell(P_{XY}) \le \textit{approx}(g_\epsilon^\ell) + \frac{1}{2(2\sqrt{|\mathcal{X}|}-1)^2} + \frac{1}{4|\mathcal{X}|} = U_{g^\ell}^2(\epsilon).$$

In the next theorem we summarize the bounds for $g_\epsilon^\ell(P_{XY})$ and $h_\epsilon^\ell(P_{XY})$, furthermore, a new upper bound for $h_\epsilon^\ell(P_{XY})$ is derived.

Theorem 7. *For any $0 \le \epsilon < \frac{1}{2}\epsilon_2$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ we have*

$$L_{g^\ell}^1(\epsilon) \le g_\epsilon^\ell(P_{XY}) \le U_{g^\ell}^1(\epsilon),$$

where $L_{g^\ell}^1(\epsilon)$ is defined in (48) *and for any $0 \le \epsilon < \frac{1}{2}\frac{\epsilon_2}{\sqrt{|\mathcal{X}|}}$ we get*

$$L_{g^\ell}^1(\epsilon) \le g_\epsilon^\ell(P_{XY}) \le U_{g^\ell}^2(\epsilon),$$

furthermore, for any $0 \le \epsilon$

$$g_\epsilon^\ell(P_{XY}) \le h_\epsilon^\ell(P_{XY}) \le \frac{\epsilon^2}{\min P_X} + H(Y|X) = U_{h^\ell}(\epsilon).$$

Proof: It is sufficient to show that the upper bound on $h_\epsilon^\ell(P_{XY})$ holds, i.e., $U_{h^\ell}(\epsilon)$. To do so, let U satisfy $d(P_{X|U}(\cdot|u), P_X) \le \epsilon$, then we have

$$\begin{aligned} I(U;Y) &= I(X;U) + H(Y|X) - I(X;U|Y) - H(Y|X,U) \\ &\le I(X;U) + H(Y|X) \\ &\overset{(a)}{\le} \sum_u P_U(u) \frac{\left(d(P_{X|U}(\cdot|u), P_X)\right)^2}{\min P_X} + H(Y|X) \\ &= \frac{\epsilon^2}{\min P_X} + H(Y|X), \end{aligned}$$

where (a) follows by the reverse Pinsker inequality. ∎

In next section we study the special case where X is a deterministic function of Y, i.e., $H(X|Y) = 0$.

Special case: X is a deterministic function of Y

In this case we have

$$\begin{aligned} h_\epsilon^{w\ell}(P_{XY}) &= g_\epsilon^{w\ell}(P_{XY}) \qquad (49) \\ &= \max_{\substack{P_{U|Y}: X-Y-U \\ d(P_{X,U}(\cdot,u), P_X P_U(u)) \le \epsilon,\ \forall u \\ |\mathcal{U}| \le |\mathcal{Y}|}} I(Y;U) \\ &= \sup_{\substack{P_{U|Y}: d(P_{X,U}(\cdot,u), P_X P_U(u)) \le \epsilon,\ \forall u \\ |\mathcal{U}| \le |\mathcal{Y}|}} I(Y;U), \\ h_\epsilon^{\ell}(P_{XY}) &= g_\epsilon^{\ell}(P_{XY}) \qquad (50) \\ &= \max_{\substack{P_{U|Y}: X-Y-U \\ d(P_{X|U}(\cdot|u), P_X) \le \epsilon,\ \forall u \\ |\mathcal{U}| \le |\mathcal{Y}|}} I(Y;U) \\ &= \sup_{\substack{P_{U|Y}: d(P_{X|U}(\cdot|u), P_X) \le \epsilon,\ \forall u \\ |\mathcal{U}| \le |\mathcal{Y}|}} I(Y;U), \end{aligned}$$

since the Markov chain $X - Y - U$ holds. Consequently, by using Theorem 2 and (49) we have the next corollary.

Corollary 4. *For any $0 \le \epsilon \le \sqrt{2I(X;Y)}$ we have*

$$\max\{L_{h^{w\ell}}^1(\epsilon), L_{h^{w\ell}}^2(\epsilon), L_{g^{w\ell}}^1(\epsilon)\} \le g_\epsilon^{w\ell}(P_{XY}) \le U_{g^{w\ell}}(\epsilon).$$

We can see that the bounds in Corollary 4 are asymptotically optimal. The latter follows since in the high privacy regime, i.e., the leakage tends to zero, $U_{g^{w\ell}}(\epsilon)$ and $L_{h^{w\ell}}^1(\epsilon)$ both tend to $H(Y|X)$, which is the optimal solution to $g_0(P_{XY})$ when X is a deterministic function of Y, [16, Theorem 6]. Furthermore, by using Theorem 3 and (50) we obtain the next result.

Corollary 5. *For any $0 \le \epsilon < \frac{1}{2}\epsilon_2$ we have*

$$L_{g^\ell}^1(\epsilon) \le g_\epsilon^\ell(P_{XY}) \le \min\{U_{g^\ell}^1(\epsilon), U_{h^\ell}(\epsilon)\}.$$

Remark 9. *For deriving the upper bound $U_{h^\ell}(\epsilon)$ and lower bounds $L_{h^{w\ell}}^1(\epsilon)$ and $L_{h^{w\ell}}^2(\epsilon)$ we do not use the assumption that the leakage matrix $P_{X|Y}$ is of full row rank. Thus, these bounds hold for all $P_{X|Y}$ and all $\epsilon \ge 0$.*

Next result shows a property of the optimizers of $h_\epsilon^{w\ell}(P_{XY})$ and $h_\epsilon^\ell(P_{XY})$.

Proposition 7. *Let $\bar{U}_1$ and $\bar{U}_2$ be any optimizers of $h_\epsilon^{w\ell}(P_{XY})$ and $h_\epsilon^\ell(P_{XY})$, respectively. Then we have*

$$H(Y|X,\bar{U}_1) = H(Y|X,\bar{U}_2) = 0.$$

Proof: The proof follows similar arguments as for Lemma 7. In the proof of Lemma 7, instead of $\bar{U}$ use $\bar{U}_1$ and let U' be produced in a similar way. The only difference is that instead of showing $I(U;X) \le \epsilon$, we need to show that $d(P_{X,U}(\cdot,u), P_X P_U(u)) \le \epsilon$ for all u, where $U = (U', \bar{U}_1)$. The latter holds since U' is independent of $(\bar{U}_1, X)$ and $\bar{U}_1$ satisfies the strong privacy criterion 1. The same proof works for $\bar{U}_2$. ∎

In the next part, we study a numerical example to illustrate the new bounds.

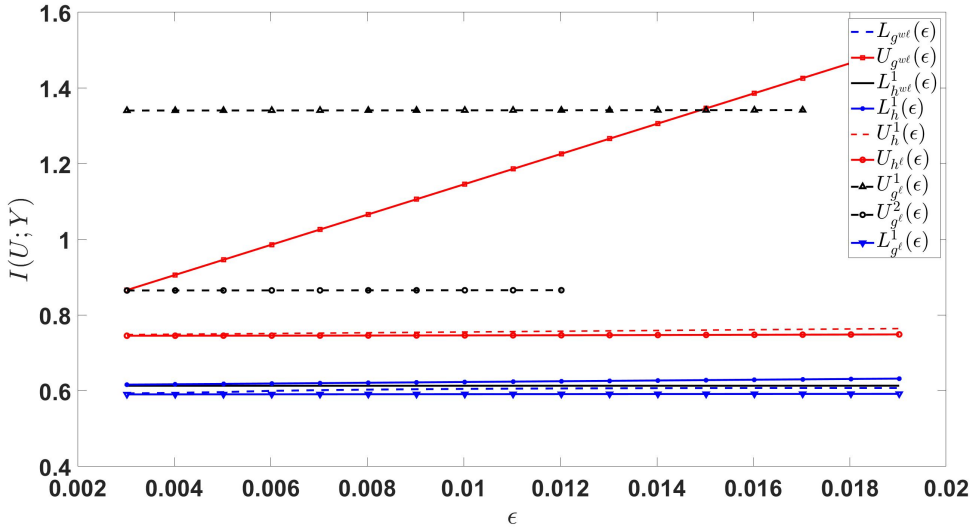

Fig. 3. Comparing the upper and lower bounds on $g_\epsilon(P_{XY})$, $h_\epsilon(P_{XY})$, $g^\ell_\epsilon(P_{XY})$, $h^\ell_\epsilon(P_{XY})$, $g^{w\ell}_\epsilon(P_{XY})$, and $h^{w\ell}_\epsilon(P_{XY})$. The upper bounds $U^1_{g^\ell}(\epsilon)$ and $U^2_{g^\ell}(\epsilon)$ are valid for $\epsilon < 0.0171$ and $\epsilon < 0.0121$, respectively. On the other hand, the upper bound $U_{h^\ell}(\epsilon)$ is valid for all $\epsilon \geq 0$.

Example

Let us consider RVs X and Y with joint distribution $P_{XY} = \begin{bmatrix} 0.693 & 0.027 & 0.108 & 0.072 \\ 0.006 & 0.085 & 0.004 & 0.005 \end{bmatrix}$. Using the definition of ϵ_2 in Proposition 6 we have $\epsilon_2 = 0.0341$. Fig. 3 illustrates the lower and upper bounds on $g_\epsilon(P_{XY})$, $h_\epsilon(P_{XY})$, $g^\ell_\epsilon(P_{XY})$, $h^\ell_\epsilon(P_{XY})$, $g^{w\ell}_\epsilon(P_{XY})$, and $h^{w\ell}_\epsilon(P_{XY})$ derived in Theorem 2, Theorem 6, and Theorem 7. As shown in Fig. 3, the upper bounds $U^1_{g^\ell}(\epsilon)$ and $U^2_{g^\ell}(\epsilon)$ are valid for $\epsilon < 0.0171$ and $\epsilon < 0.0121$, however the upper bound $U_{h^\ell}(\epsilon)$ is valid for all $\epsilon \geq 0$. In this example, considering the upper bounds in Theorem 7, we can see that for any ϵ the upper bound $U_{h^\ell}(\epsilon)$ is the smallest upper bound for $g^\ell_\epsilon(P_{XY})$. Furthermore, the upper and lower bounds on $g^\ell_\epsilon(P_{XY})$, $h^\ell_\epsilon(P_{XY})$, i.e., $L_{g^{w\ell}}(\epsilon)$, $U_{g^{w\ell}}(\epsilon)$ and $L^1_{h^{w\ell}}(\epsilon)$ obtained in Theorem 6 are illustrated. We can see that $L^1_{h^{w\ell}}(\epsilon) \geq L_{g^{w\ell}}(\epsilon)$. The lower bound $L^2_{h^{w\ell}}(\epsilon)$ is not shown in Fig. 3 since in this particular privacy range, the term $(1-\alpha)(\log(I(X;Y)+1)+4)$ outweighs the value of $H(Y|X)$. As a result, including it as a lower bound would not be meaningful or effective. Likewise, the lower bound $L^2_h(\epsilon)$ on $h_\epsilon(P_{XY})$ is also not shown in Fig. 3. More comparisons can be made by utilizing the relationships established in Section III, e.g., (19), (20), (22), (23), (25), and (26). For instance, we can see that $L_{g^{w\ell}}(\epsilon) \geq L_{g^\ell}(\epsilon)$. This result can be motivated by examining the relationship between $g^{w\ell}_\epsilon(P_{XY})$ and $g^\ell_\epsilon(P_{XY})$ as derived in equation (26). It is important to note that (26) does not imply $L_{g^{w\ell}}(\epsilon) \geq L_{g^\ell}(\epsilon)$. However, by considering the inequality $L_{g^{w\ell}}(\epsilon) \geq L_{g^\ell}(\epsilon)$, we can assess the behavior and effectiveness of the respective lower bounds utilizing (26). In other words, "good" lower and upper bounds on $g^\ell_\epsilon(P_{XY})$ and $g^{w\ell}_\epsilon(P_{XY})$ must satisfy (26). If they do not meet this requirement, it becomes necessary to consider both bounds on $g^\ell_\epsilon(P_{XY})$ and $g^{w\ell}_\epsilon(P_{XY})$ simultaneously. Along the same lines, we can also see that $U^1_h(\epsilon) \geq U_{h^\ell}(\epsilon)$ which is motivated by (22). This follows since in this privacy regime we have $\epsilon' = \frac{\epsilon^2}{\min P_X} \leq \epsilon$ ($\min P_X = 0.1$) which results in

$$h^\ell_\epsilon(P_{XY}) \leq h_{\epsilon'}(P_{XY}) \leq h_\epsilon(P_{XY}).$$

To compare $L^1_{h^{w\ell}}(\epsilon)$ and $L^1_h(\epsilon)$, using Theorem 2 and Theorem 6 we have

$$L^1_{h^{w\ell}}(\bar{\epsilon}) = L^1_h(\epsilon),$$

where $\bar{\epsilon} = \sqrt{2\epsilon}$. This result also fulfills (19) and can be seen in Fig. 3. Next, we provide an example that demonstrates the dominance of the lower bound $U^2_{g^\ell}(\epsilon)$ over other lower bounds, such as $U_{h^\ell}(\epsilon)$ and $U^1_{g^\ell}(\epsilon)$, for g^ℓ_ϵ.

Let $P_{XY} = \begin{bmatrix} 0.350 & 0.025 & 0.085 & 0.040 \\ 0.025 & 0.425 & 0.035 & 0.015 \end{bmatrix}$. In this case, $\epsilon_2 = 0.1994$. Fig. 4 illustrates the lower and upper bounds for $g^\ell_\epsilon(P_{XY})$ and $g^{w\ell}_\epsilon(P_{XY})$. We can see that for $\epsilon < 0.0705$, $U^2_{g^\ell}(\epsilon)$ is the smallest upper bound and for $\epsilon > 0.0705$, $U_{h^\ell}(\epsilon)$ is the smallest upper bound on $g^\ell_\epsilon(P_{XY})$. As a result both lower bounds $U_{h^\ell}(\epsilon)$ and $U^2_{g^\ell}(\epsilon)$ can be used for $g^\ell_\epsilon(P_{XY})$. Moreover, it can be seen that $U^1_{g^\ell}(\epsilon) \leq U_{g^{w\ell}}(\epsilon)$ and $L^1_{g^\ell}(\epsilon) \leq L_{g^{w\ell}}(\epsilon)$. Likewise, these inequalities can be examined using (20).

C. Privacy-utility trade-off with non-zero leakage and prioritized private data

In this part we find lower and upper bounds for $h^p_\epsilon(P_{X_1X_2Y})$ defined in (8). To find lower bounds we use similar techniques as used in Theorem 2 and Proposition 4, i.e., we use extended versions of FRL and SFRL for correlated (X_1, X_2) and U.

Theorem 8. *For any $0 \leq \epsilon$ and RVs (X_1, X_2, Y) distributed according to $P_{X_1X_2Y}$ supported on alphabets $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{Y}$ we have*

$$\max\{L^1_{h^p}(\epsilon), L^2_{h^p}(\epsilon), L^3_{h^p}(\epsilon)\} \leq h^p_\epsilon(P_{X_1X_2Y}) \leq U^1_{h^p}(\epsilon), \tag{51}$$

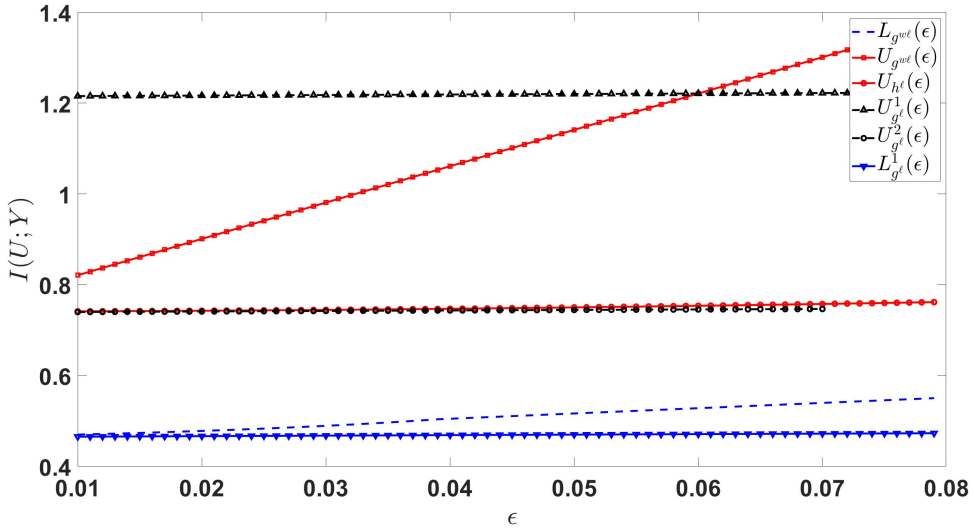

Fig. 4. Comparing the upper and lower bounds for $g^{w\ell}_{\epsilon}(P_{XY})$ and $g^{\ell}_{\epsilon}(P_{XY})$. The upper bounds $U^1_{g^{\ell}}(\epsilon)$ and $U^2_{g^{\ell}}(\epsilon)$ are valid for $\epsilon < 0.0997$ and $\epsilon < 0.0705$, respectively. However, the upper bound $U_{h^{\ell}}(\epsilon)$ is valid for all $\epsilon \geq 0$.

where

$$\begin{aligned}
L^1_{h^p}(\epsilon) &= \epsilon + H(Y|X_1,X_2) - H(X_1,X_2|Y),\\
L^2_{h^p}(\epsilon) &= \epsilon + H(Y|X_1,X_2) - \alpha H(X_2|Y)\\
&\quad - \left(\log(I(X_1,X_2;Y)+1)+4\right),\\
L^3_{h^p}(\epsilon) &= \epsilon + H(Y|X_1,X_2) - \alpha H(X_1,X_2|Y)\\
&\quad - (1-\alpha)\left(\log(I(X_1,X_2;Y)+1)+4\right),\\
U^1_{h^p}(\epsilon) &= \epsilon + H(Y|X_1,X_2),
\end{aligned}$$

with $\alpha = \frac{\epsilon}{H(X_2)}$.

Proof: The proof is provided in Appendix D. ∎

To compare the lower bounds $L^1_{h^p}(\epsilon)$, $L^2_{h^p}(\epsilon)$, and $L^3_{h^p}(\epsilon)$ we consider four cases as follows.

Case 1: Let X_1 be a deterministic function of Y, then we have $H(X_1,X_2|Y) = H(X_2|Y)$. Hence, in this case $L^1_{h^p}(\epsilon) \geq L^3_{h^p}(\epsilon) \geq L^2_{h^p}(\epsilon)$.

Case 2: Let X_2 be a deterministic function of Y and assume that $4 + H(Y) \leq H(X_1|Y)$. In this case, we have

$$\begin{aligned}
&L^2_{h^p}(\epsilon) - L^3_{h^p}(\epsilon)\\
&= \alpha\left(H(X_1|Y) - \log(I(X_1;Y) + H(X_2|X_1) + 1) - 4\right)\\
&\overset{(a)}{\geq} \alpha\left(H(X_1|Y) - I(X_1;Y) - H(X_2|X_1) - 4\right)\\
&\overset{(b)}{\geq} \alpha\left(H(X_1|Y) - I(X_1;Y) - H(Y|X_1) - 4\right)\\
&= \alpha\left(H(X_1|Y) - H(Y) - 4\right)\\
&\geq 0, \qquad (52)
\end{aligned}$$

where (a) follows since $\log(1+x) \leq x$ and in step (b) we use $H(X_2|X_1) \leq H(Y|X_1)$ since $H(X_2|Y) = 0$. Furthermore,

$$\begin{aligned}
&L^2_{h^p}(\epsilon) - L^1_{h^p}(\epsilon)\\
&= H(X_1|Y) - \log(I(X_1;Y) + H(X_2|X_1) + 1) - 4\\
&\geq H(X_1|Y) - I(X_1;Y) - H(X_2|X_1) - 4\\
&\geq H(X_1|Y) - I(X_1;Y) - H(Y|X_1) - 4\\
&= H(X_1|Y) - H(Y) - 4\\
&\geq 0. \qquad (53)
\end{aligned}$$

So, in this case $L^2_{h^p}(\epsilon) \geq \max(L^1_{h^p}(\epsilon), L^3_{h^p}(\epsilon))$.

Case 3: Let Y be independent of (X_1,X_2) and assume $H(X_1,X_2) \geq 4$. In this case we have $L^2_{h^p}(\epsilon) \geq L^1_{h^p}(\epsilon)$ and $L^3_{h^p}(\epsilon) \geq L^1_{h^p}(\epsilon)$.

Case 4: Let X_1 be a deterministic function of X_2 and $H(X_2|Y) \geq \log(I(X_2;Y)+1)+4$. A simple example can be letting $X_1 = f(X_2)$ and $H(X_2|Y) \geq \frac{H(X_2)}{2}+2$ which results in $H(X_2|Y) \geq \log(I(X_2;Y)+1)+4$ using $\log(x+1) \leq x$. Note that considering $X_1 = f(X_2)$ does not violate the priority constraint $I(U;X_1) \leq I(U;X_2)$, since in this case the inequality holds for any U. We have

$$\begin{aligned}
&L^3_{h^p}(\epsilon) - L^2_{h^p}(\epsilon)\\
&= \epsilon\frac{H(X_2|Y)}{H(X_2)} + \epsilon\frac{\log(I(X_2;Y)+1)+4}{H(X_2)} - \epsilon\frac{H(X_2|Y)}{H(X_2)}\\
&= \epsilon\frac{\log(I(X_2;Y)+1)+4}{H(X_2)}\\
&\geq 0.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&L^3_{h^p}(\epsilon) - L^1_{h^p}(\epsilon)\\
&= (1-\alpha)\left(H(X_2|Y) - \log(I(X_2;Y)+1) - 4\right)\\
&\geq 0.
\end{aligned}$$

Hence, in this case $L^3_{h^p}(\epsilon) \geq \max\{L^2_{h^p}(\epsilon), L^1_{h^p}(\epsilon)\}$. The upper bound $U^1_{h^p}(\epsilon)$ is attained whenever the pair (X_1,X_2) is a deterministic function of Y. In this case $U^1_{h^p}(\epsilon) = L^1_{h^p}(\epsilon)$.

Next we compare the bounds obtained in Theorem 8 and Theorem 2. As we outlined in Section III, lower bounds on $h^p_{\epsilon}(P_{X_1X_2Y})$ can be used as lower bounds on $h^p_{\epsilon}(P_{X_1X_2Y})$ since we have $h^p_{\epsilon}(P_{X_1X_2Y}) \leq h_{\epsilon}(P_{X_1X_2Y})$. Let $X = (X_1,X_2)$, using Lemma 6, Theorem 2, and Theorem 8 we

have

$$L^1_{h^p}(\epsilon) = L^1_h(\epsilon), \tag{54}$$
$$U^1_{h^p}(\epsilon) = U^1_h(\epsilon). \tag{55}$$

In the following we consider two scenarios to compare $L^3_{h^p}(\epsilon)$ and $L^2_{h^p}(\epsilon)$ with $L^2_h(\epsilon)$.

Scenario 1: To compare $L^3_{h^p}(\epsilon)$ with $L^2_h(\epsilon)$, let us assume that $H(X_1, X_2|Y) \leq \log(I(X_1, X_2; Y) + 1) + 4$. A simple example can be considering X_1 and X_2 as binary RVs. In this case we have

$$\begin{aligned} &L^3_{h^p}(\epsilon) - L^2_h(\epsilon) \\ &= \beta \left(\log(I(X_1, X_2; Y) + 1) + 4 - H(X_1, X_2|Y)\right) \\ &\geq 0, \end{aligned}$$

where $\beta = \epsilon(\frac{1}{H(X_2)} - \frac{1}{H(X_1,X_2)})$.

Scenario 2: To compare $L^2_{h^p}(\epsilon)$ with $L^2_h(\epsilon)$, let us assume that X_2 is a deterministic function of Y and $H(X_1|Y) \geq \log(I(X_1, X_2; Y)+1)+4$. As we pointed out earlier a simple example is to let $4 + H(Y) \leq H(X_1|Y)$ which leads to $H(X_1|Y) \geq \log(I(X_1, X_2; Y) + 1) + 4$. In this case we have

$$\begin{aligned} &L^2_{h^p}(\epsilon) - L^2_h(\epsilon) \\ &= \beta \left(H(X_1|Y) - \log(I(X_1, X_2; Y) + 1) - 4\right) \\ &\geq 0, \end{aligned}$$

where $\beta = \frac{\epsilon}{H(X_1,X_2)}$. Moreover, by using (52), (53), and (54) we have

$$L^2_{h^p}(\epsilon) \geq \max\{L^2_h(\epsilon), L^3_{h^p}(\epsilon), L^1_{h^p}(\epsilon) = L^1_h(\epsilon)\}.$$

Consequently, in this scenario $L^2_{h^p}(\epsilon)$ can improve the lower bounds $L^1_h(\epsilon)$ and $L^2_h(\epsilon)$ that are derived for $h_\epsilon(P_{X_1X_2Y})$.

Using the comparisons between the lower bounds on $h_\epsilon(P_{X_1X_2Y})$ and $h^p_\epsilon(P_{X_1X_2Y})$, we can conclude that in some scenarios randomizing over X_2 can achieve better utilities compared to randomizing over (X_1, X_2).

Similar to Lemma 7 and Proposition 7 it can be shown that if $\tilde{U}$ is an optimizer of $h^p_\epsilon(P_{X_1X_2Y})$, then Y is a deterministic function of $\tilde{U}$ and (X_1, X_2).

Proposition 8. *Let $\tilde{U}$ be an optimizer of $h^p_\epsilon(P_{X_1X_2Y})$, then*

$$H(Y|X_1, X_2, \tilde{U}) = 0.$$

Proof: The proof follows similar arguments as in Lemma 7. Let $\tilde{U}$ be an optimizer of $h^p_\epsilon(P_{X_1,X_2,Y})$ and $H(Y|X_1, X_2, \tilde{U}) > 0$. Consequently, $I(X_1, X_2; \tilde{U}) \leq \epsilon$ and $I(X_1; \tilde{U}) \leq I(X_2; \tilde{U})$. Let U' be produced by FRL using $(X_1, X_2, \tilde{U})$ instead of X in Lemma 1 and same Y. Thus, $I(Y; U') > 0$ and by letting $U = (U', \tilde{U})$ and using similar arguments as in Lemma 7 we have $I(Y; U) > I(Y; \tilde{U})$. Furthermore,

$$\begin{aligned} I(X_1, X_2, U) &\overset{(a)}{=} I(X_1, X_2; \tilde{U}) \leq \epsilon, \\ I(X_1; U) &\overset{(b)}{=} I(X_1; \tilde{U}) \leq I(X_2; \tilde{U}) \overset{(c)}{=} I(X_2; U), \end{aligned}$$

where (a), (b) and (c) follow from the fact that U' is independent of $(X_1, X_2, \tilde{U})$. Thus, U achieves strictly larger utility than $\tilde{U}$ which contradicts the optimality of $\tilde{U}$. ■

V. CONCLUSION

Different information theoretic data disclosure problems have been studied in this work. The FRL and SRFL have been extended by relaxing the independence constraint and allowing certain amount of leakage using different privacy measures. It has been shown that by using extended versions of the FRL and SFRL lower bounds on privacy-utility trade-off functions can be derived. The results are useful since the proofs are constructive and therefore valuable for mechanism design and the bounds on optimality serve as a benchmark. Concepts from information geometry can be used to find lower bounds on privacy-utility trade-off functions considering the hidden private data scenario when per-letter privacy constraints (strong privacy criterions) are used.

ACKNOWLEDGMENT

The authors want to thank the Associate Editor Prof. Anand D. Sarwate and the anonymous reviewers for their careful and valuable assessment of the manuscript. The authors want to thank the anonymous reviewer who provided us a simpler necessary condition presented now in Theorem 1.

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.

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

[3] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *2016 Annual Conference on Information Science and Systems*, March 2016, pp. 234–239.

[4] I. Issa, S. Kamath, and A. B. Wagner, "Maximal leakage minimization for the shannon cipher system," in *2016 IEEE International Symposium on Information Theory*, 2016, pp. 520–524.

[5] H. Wang, L. Vo, F. P. Calmon, M. Médard, K. R. Duffy, and M. Varia, "Privacy with estimation guarantees," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8025–8042, Dec 2019.

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

[7] Z. Li, T. J. Oechtering, and D. Gündüz, "Privacy against a hypothesis testing adversary," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1567–1581, 2018.

[8] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, p. 15, 2016.

[9] B. Rassouli and D. Gündüz, "Optimal utility-privacy trade-off with total variation distance as a privacy measure," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 594–603, 2020.

[10] B. Rassouli and D. Gündüz, "On perfect privacy," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 177–191, 2021.

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

[12] A. Zamani, T. J. Oechtering, and M. Skoglund, "A design framework for strongly χ^2-private data disclosure," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2312–2325, 2021.

[13] A. Zamani, T. J. Oechtering, and M. Skoglund, "Data disclosure with non-zero leakage and non-invertible leakage matrix," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 165–179, 2022.

[14] Y. O. Basciftci, Y. Wang, and P. Ishwar, "On privacy-utility tradeoffs for constrained data release mechanisms," in *2016 Information Theory and Applications Workshop*, Jan 2016, pp. 1–6.

[15] Y. Wang, Y. O. Basciftci, and P. Ishwar, "Privacy-utility tradeoffs under constrained data release mechanisms," *arXiv preprint arXiv:1710.09295*, 2017.

[16] Y. Y. Shkel, R. S. Blum, and H. V. Poor, "Secrecy by design with applications to privacy and compression," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 824–843, 2021.

[17] F. P. Calmon, A. Makhdoumi, M. Medard, M. Varia, M. Christiansen, and K. R. Duffy, "Principal inertia components and applications," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5011–5038, Aug 2017.

[18] 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.

[19] A. Zamani, T. J. Oechtering, and M. Skoglund, "Bounds for privacy-utility trade-off with non-zero leakage," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 620–625.

[20] ——, "Bounds for privacy-utility trade-off with per-letter privacy constraints and non-zero leakage," in *2022 IEEE Information Theory Workshop (ITW)*, 2022, pp. 13–18.

[21] ——, "Multi-user privacy mechanism design with non-zero leakage," in *2023 IEEE Information Theory Workshop (ITW)*, 2023, pp. 401–405.

[22] A. Zamani, T. J. Oechtering, D. Gündüz, and M. Skoglund, "Cache-aided private variable-length coding with zero and non-zero leakage," *arXiv preprint arXiv:2306.13184*, 2023.

[23] M. A. Zarrabian, N. Ding, and P. Sadeghi, "On the lift, related privacy measures, and applications to privacy–utility trade-offs," *Entropy*, vol. 25, no. 4, p. 679, 2023.

[24] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8043–8066, Dec 2019.

[25] B. Rassouli, F. Rosas, and D. Gündüz, "Latent feature disclosure under perfect sample privacy," in *2018 IEEE International Workshop on Information Forensics and Security*, Dec 2018, pp. 1–7.

[26] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Estimation efficiency under privacy constraints," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1512–1534, March 2019.

[27] F. P. Calmon and N. Fawaz, "Privacy against statistical inference," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing*, Oct 2012, pp. 1401–1408.

[28] E. Nekouei, T. Tanaka, M. Skoglund, and K. H. Johansson, "Information-theoretic approaches to privacy in estimation and control," *Annual Reviews in Control*, 2019.

[29] H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 918–923, 1983.

[30] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *3rd Conf. Theory Cryptogr*. Berlin, Germany: Springer, 2006, pp. 265–284.

[31] C. Dwork, "Differential privacy," *Bugliesi M., Preneel B., Sassone V., Wegener I. (eds) Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science*, vol. 4052.

[32] P. Cuff and L. Yu, "Differential privacy as a mutual information constraint," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016, pp. 43–54.

[33] 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

[34] C. T. Li and A. El Gamal, "Strong functional representation lemma and applications to coding theorems," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 6967–6978, 2018.

[35] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[36] S. Borade and L. Zheng, "Euclidean information theory," in *2008 IEEE International Zurich Seminar on Communications*, 2008, pp. 14–17.

[37] S. L. Huang and L. Zheng, "Linear information coupling problems," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 1029–1033.

[38] Y. Polyanskiy and Y. Wu, "Lecture notes on information theory," *Lecture Notes for ECE563 (UIUC) and*, vol. 6, no. 2012-2016, p. 7, 2014.

[39] S. Asoodeh, F. Alajaji, and T. Linder, "Privacy-aware mmse estimation," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 1989–1993.

[40] T. Berger and R. W. Yeung, "Multiterminal source encoding with encoder breakdown," *IEEE Transactions on Information Theory*, vol. 35, no. 2, pp. 237–244, March 1989.

[41] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 163–179, 1975.

[42] R. Ahlswede and J. Körner, *On common information and related characteristics of correlated information sources*. Springer, 2006.

[43] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[44] I. Sason and S. Verdú, "f -divergence inequalities," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 5973–6006, 2016.

[45] T.-Y. Liu, I. Wang *et al.*, "Robust privatization with non-specific tasks and the optimal privacy-utility tradeoff," *arXiv preprint arXiv:2010.10081*, 2020.

[46] T.-Y. Liu and I.-H. Wang, "Privacy-utility tradeoff with nonspecific tasks: Robust privatization and minimum leakage," in *2020 IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–5.

APPENDIX A

Proofs for Section II and Section III:

Proof of Proposition 1: For each $u \in \mathcal{U}$ we have

$$\begin{aligned}
\mathcal{L}^1(X;U=u) &= d(P_{X|U=u}(\cdot|u), P_X) \\
&= d(P_{X|Y}P_{Y|U=u}(\cdot|u), P_{X|Y}P_Y) \\
&= \sum_x |\sum_y P_{X|Y}(x,y)(P_{Y|U=u}(y) - P_Y(y))| \\
&\overset{(a)}{\leq} \sum_x \sum_y P_{X|Y}(x,y)|P_{Y|U=u}(y) - P_Y(y)| \\
&= \sum_y \sum_x P_{X|Y}(x,y)|P_{Y|U=u}(y) - P_Y(y)| \\
&= \sum_y |P_{Y|U=u}(y) - P_Y(y)| \\
&= d(P_{Y|U=u}(\cdot|u), P_Y) = \mathcal{L}^1(Y;U=u),
\end{aligned}$$

where (a) follows from the triangle inequality. Furthermore, we can multiply all the above expressions by the term $P_U(u)$ and we obtain

$$\mathcal{L}^2(X;U=u) \leq \mathcal{L}^2(Y;U=u).$$

Proof of (18): We have

$$\begin{aligned}
\epsilon \geq I(U;X) &= \sum_u P_U(u) D(P_{X|U}(\cdot|u), P_X) \\
&\overset{(a)}{\geq} \sum_u \frac{P_U(u)}{2} \left(d(P_{X|U}(\cdot|u), P_X)\right)^2 \\
&\geq \frac{\left(P_U(u) d(P_{X|U}(\cdot|u), P_X)\right)^2}{2}
\end{aligned}$$

where (a) follows by the Pinsker's inequality [44].

Proof of (21): We have

$$\begin{aligned}
I(X;U) &\overset{(a)}{\leq} \sum_u P_U(u) \frac{\left(d(P_{X|U}(\cdot|u), P_X)\right)^2}{\min P_X} \\
&\leq \sum_u P_U(u) \left(\frac{\epsilon^2}{\min P_X}\right) \\
&= \frac{\epsilon^2}{\min P_X},
\end{aligned}$$

where (a) follows by the reverse Pinsker's inequality [44].

APPENDIX B

Proofs for Privacy-utility trade-off with non-zero leakage*:*

Proof of Lemma 4: Let $\tilde{U}$ be the RV found by FRL and let $W = \begin{cases} X, & \text{w.p. } \alpha \\ c, & \text{w.p. } 1-\alpha \end{cases}$, where c is a constant which does not belong to the support of X and Y and $\alpha = \frac{\epsilon}{H(X)}$. We show that $U = (\tilde{U}, W)$ satisfies the conditions. We have

$$\begin{aligned} I(X;U) &= I(X;\tilde{U},W) \\ &= I(\tilde{U};X) + I(X;W|\tilde{U}) \\ &\overset{(a)}{=} H(X) - H(X|\tilde{U},W) \\ &= H(X) - \alpha H(X|\tilde{U},X) - (1-\alpha)H(X|\tilde{U},c) \\ &= H(X) - (1-\alpha)H(X) = \alpha H(X) = \epsilon, \end{aligned}$$

where in (a) we used the fact that X and $\tilde{U}$ are independent. Furthermore,

$$\begin{aligned} H(Y|X,U) &= H(Y|X,\tilde{U},W) \\ &= \alpha H(Y|X,\tilde{U}) + (1-\alpha)H(Y|X,\tilde{U},c) \\ &= H(Y|X,\tilde{U}) = 0. \end{aligned}$$

In the last line we used the fact that $\tilde{U}$ is produced by FRL.

Proof of Lemma 5: Let $\tilde{U}$ be the RV found by SFRL and W be the same RV which is used to prove Lemma 4. It is sufficient to show that $I(X;U|Y) \leq \alpha H(X|Y) + (1-\alpha)\left[\log(I(X;Y)+1)+4\right]$ since all other properties are already proved in Lemma 3. We have

$$\begin{aligned} I(X;\tilde{U},W|Y) &= I(X;\tilde{U}|Y) + I(X,W|\tilde{U},Y) \\ &\overset{(a)}{=} I(X;\tilde{U}|Y) + \alpha H(X|\tilde{U},Y) \\ &= I(X;\tilde{U}|Y) + \alpha(H(X|Y) - I(X;\tilde{U}|Y)) \\ &= \alpha H(X|Y) + (1-\alpha)I(X;\tilde{U}|Y) \\ &\overset{(b)}{\leq} \alpha H(X|Y) + (1-\alpha)[\log(I(X;Y)+1)+4], \end{aligned}$$

where in step (a) we used the fact that

$$\begin{aligned} I(X,W|\tilde{U},Y) &= H(X|\tilde{U},Y) - H(X|W,\tilde{U},Y) \\ &= H(X|\tilde{U},Y) - (1-\alpha)H(X|\tilde{U},Y) \\ &= \alpha H(X|\tilde{U},Y), \end{aligned}$$

and (b) follows since $\tilde{U}$ is produced by SFRL.

Lemma 8. *For any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$, where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite, there exists RV U such that it satisfies* (11)*,* (12)*, and*

$$H(U) \leq \sum_{x\in\mathcal{X}} H(Y|X=x) + \epsilon + h(\alpha)$$

with $\alpha = \frac{\epsilon}{H(X)}$ and $h(\cdot)$ denotes the binary entropy function.

Proof: Let $U = (\tilde{U}, W)$ where W is the same RV used in Lemma 4 and $\tilde{U}$ is produced by FRL which has the same construction as used in proof of [16, Lemma 1]. Thus, by using [16, Lemma 2] we have

$$H(\tilde{U}) \leq \sum_{x\in\mathcal{X}} H(Y|X=x),$$

therefore,

$$\begin{aligned} H(U) = H(\tilde{U},W) &\leq H(\tilde{U}) + H(W), \\ &\leq \sum_{x\in\mathcal{X}} H(Y|X=x) + H(W), \end{aligned}$$

where,

$$\begin{aligned} H(W) &= -(1-\alpha)\log(1-\alpha) - \sum_{x\in\mathcal{X}} \alpha P_X(x)\log(\alpha P_X(x)), \\ &= h(\alpha) + \alpha H(X), \end{aligned}$$

which completes the proof. ∎

Proof of Theorem 1: For proving the first part let $h_\epsilon(P_{XY}) > \epsilon$. Using (27) we have

$$\begin{aligned} \epsilon < h_\epsilon(P_{XY}) &\leq H(Y|X) + \sup_{U: I(X;U)\leq\epsilon} I(X;U) \\ &= H(Y|X) + \epsilon \Rightarrow 0 < H(Y|X). \end{aligned}$$

For the second part let $H(Y|X) > 0$. Using [33, Lemma 1] we have

$$h_\epsilon(P_{XY}) \geq g_\epsilon(P_{XY}) \geq \epsilon\frac{H(Y)}{I(X;Y)}. \tag{56}$$

Using $H(Y|X) > 0$ we have $\frac{H(Y)}{I(X;Y)} > 1$ which results in $h_\epsilon(P_{XY}) > \epsilon$.

Proof of Theorem 2: $L_h^3(\epsilon)$ can be derived by using [33, Remark 2], since we have $h_\epsilon(P_{XY}) \geq g_\epsilon(P_{XY}) \geq L_h^3(\epsilon)$. For deriving $L_h^1(\epsilon)$, let U be produced by EFRL. Thus, using the construction of U as in Lemma 4 we have $I(X,U) = \epsilon$ and $H(Y|X,U) = 0$. Then, using (27) we obtain

$$\begin{aligned} h_\epsilon(P_{XY}) &\geq I(U;Y) \\ &= I(X;U) + H(Y|X) - H(Y|U,X) - I(X;U|Y) \\ &= \epsilon + H(Y|X) - H(X|Y) + H(X|Y,U) \\ &\geq \epsilon + H(Y|X) - H(X|Y) = L_h^1(\epsilon). \end{aligned}$$

For deriving $L_h^2(\epsilon)$, let U be produced by ESFRL. Thus, by using the construction of U as in Lemma 5 we have $I(X,U) = \epsilon$, $H(Y|X,U) = 0$ and $I(X;U|Y) \leq \alpha H(X|Y) + (1-\alpha)\left(\log(I(X;Y)+1)+4\right)$. Then, by using (27) we obtain

$$\begin{aligned} h_\epsilon(P_{XY}) &\geq I(U;Y) \\ &= I(X;U) + H(Y|X) - H(Y|U,X) - I(X;U|Y) \\ &= \epsilon + H(Y|X) - I(X;U|Y) \\ &\geq \epsilon + H(Y|X) - \alpha H(X|Y) \\ &\quad + (1-\alpha)\left(\log(I(X;Y)+1)+4\right) = L_h^2(\epsilon). \end{aligned}$$

Let X be a deterministic function of Y. In this case, set $\epsilon = 0$ in $L_h^1(\epsilon)$ so that we obtain $h_0(P_{XY}) \geq H(Y|X)$. Furthermore, by using (27) we have $h_0(P_{XY}) \leq H(Y|X)$. Moreover, since X is a deterministic function of Y, the Markov chain $X - Y - U$ holds and we have $h_0(P_{XY}) = g_0(P_{XY}) = H(Y|X)$. Therefore, $L_h^3(\epsilon)$ can be rewritten as

$$\begin{aligned} L_h^3(\epsilon) &= \epsilon\frac{H(Y)}{H(X)} + H(Y|X)\left(\frac{H(X)-\epsilon}{H(X)}\right), \\ &= \epsilon\frac{H(Y)}{H(X)} + (H(Y)-H(X))\left(\frac{H(X)-\epsilon}{H(X)}\right), \\ &= H(Y) - H(X) + \epsilon. \end{aligned}$$

$L_h^2(\epsilon)$ can be rewritten as follows

$$L_h^2(\epsilon) = H(Y|X) + \epsilon - (1 - \frac{\epsilon}{H(X)})(\log(H(X)+1)+4).$$

Thus, if $H(X|Y)=0$, then $L_h^1(\epsilon) = L_h^3(\epsilon) \geq L_h^2(\epsilon)$. Now we show that $L_h^1(\epsilon) = L_h^3(\epsilon)$ is tight. By using (27) we have

$$\begin{aligned} I(U;Y) &\overset{(a)}{=} I(X;U) + H(Y|X) - H(Y|U,X), \\ &\leq \epsilon + H(Y|X) = L_h^1(\epsilon) = L_h^3(\epsilon). \end{aligned}$$

where (a) follows since X is deterministic function of Y which leads to $I(X;U|Y)=0$. Thus, if $H(X|Y)=0$, the lower bound in (30) is tight. Now suppose that the lower bound $L_h^1(\epsilon)$ is tight and X is not a deterministic function of Y. Let $\tilde{U}$ be produced by FRL using the construction of [16, Lemma 1]. As argued in the proof of [16, Th. 6], there exists $x \in \mathcal{X}$ and $y_1, y_2 \in \mathcal{Y}$ such that $P_{X|\tilde{U},Y}(x|\tilde{u},y_1) > 0$ and $P_{X|\tilde{U},Y}(x|\tilde{u},y_2) > 0$ which results in $H(X|Y,\tilde{U}) > 0$. Let $U = (\tilde{U}, W)$ where W is defined in Lemma 4. For such U we have

$$\begin{aligned} &H(X|Y,U) = (1-\alpha)H(X|Y,\tilde{U}) > 0, \\ &\Rightarrow I(U;Y) \overset{(a)}{=} \epsilon + H(Y|X) - H(X|Y) + H(X|Y,U) \\ &\qquad > \epsilon + H(Y|X) - H(X|Y). \end{aligned}$$

where in (a) we used the fact that such U satisfies $I(X;U)=\epsilon$ and $H(Y|X,U)=0$. The last line is a contradiction with tightness of $L_h^1(\epsilon)$, since we can achieve larger values, thus, X needs to be a deterministic function of Y.

Lemma 9. *For any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$, then if U satisfies $I(X;U) \leq \epsilon$, $H(Y|X,U)=0$ and $|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+1]\,[|\mathcal{X}|+1]$, we have*

$$\begin{aligned} \sup_U H(U) \geq &\alpha H(Y|X) + (1-\alpha)(\max_{x\in\mathcal{X}} H(Y|X=x)) \\ &+ h(\alpha) + \epsilon \geq H(Y|X) + h(\alpha) + \epsilon, \end{aligned}$$

where $\alpha = \frac{\epsilon}{H(X)}$ and $h(\cdot)$ corresponds to the binary entropy.

Proof: Let $U = (\tilde{U}, W)$ where $W = \begin{cases} X, & \text{w.p. } \alpha \\ c, & \text{w.p. } 1-\alpha \end{cases}$, and c is a constant which does not belong to the support of X, Y and $\tilde{U}$, furthermore, $\tilde{U}$ is produced by FRL. Using (27) and [16, Lemma 3] we have

$$\begin{aligned} H(\tilde{U}|Y) &= H(\tilde{U}) - H(Y|X) + I(X;\tilde{U}|Y) \\ &\overset{(a)}{\geq} \max_{x\in\mathcal{X}} H(Y|X=x) - H(Y|X) \\ &\quad + H(X|Y) - H(X|Y,\tilde{U}), \end{aligned} \tag{57}$$

where (a) follows from [16, Lemma 3]. Furthermore, in the first line we used $I(X;\tilde{U})=0$ and $H(Y|\tilde{U},X)=0$. Using (27) we obtain

$$\begin{aligned} H(U) &\overset{(a)}{=} H(U|Y) + H(Y|X) - H(X|Y) + \epsilon + H(X|Y,U), \\ &\overset{(b)}{=} H(W|Y) + \alpha H(\tilde{U}|Y,X) + (1-\alpha)H(\tilde{U}|Y) \\ &\quad + H(Y|X) - H(X|Y) + \epsilon + (1-\alpha)H(X|Y,\tilde{U}), \\ &\overset{(c)}{=} (\alpha-1)H(X|Y) + h(\alpha) + \alpha H(\tilde{U}|Y,X) + \epsilon \\ &\quad + (1-\alpha)H(\tilde{U}|Y) + H(Y|X) + (1-\alpha)H(X|Y,\tilde{U}), \\ &\overset{(d)}{\geq} (\alpha-1)H(X|Y) + h(\alpha) + \alpha H(\tilde{U}|Y,X) \\ &\quad + (1-\alpha)(\max_{x\in\mathcal{X}} H(Y|X=x) - H(Y|X) + H(X|Y) \\ &\quad - H(X|Y,\tilde{U})) + H(Y|X) + \epsilon + (1-\alpha)H(X|Y,\tilde{U}) \\ &= \alpha H(Y|X) + (1-\alpha)(\max_{x\in\mathcal{X}} H(Y|X=x)) + h(\alpha) + \epsilon. \end{aligned}$$

In step (a) we used $I(U;X)=\epsilon$ and $H(Y|X,U)=0$ and in step (b) we used $H(U|Y) = H(W|Y) + H(\tilde{U}|Y,W) = H(W|Y) + \alpha H(\tilde{U}|Y,X) + (1-\alpha)H(\tilde{U}|Y)$ and $H(X|Y,U) = H(X|Y,\tilde{U},W) = (1-\alpha)H(X|Y,\tilde{U})$. In step (c) we used the fact that $P_{W|Y} = \begin{cases} \alpha P_{X|Y}(x|\cdot) & \text{if } w=x, \\ 1-\alpha & \text{if } w=c, \end{cases}$ since $P_{W|Y}(w=x|\cdot) = \frac{P_{W,Y}(w=x,\cdot)}{P_Y(\cdot)} = \frac{P_{Y|W}(\cdot|w=x)P_W(w=x)}{P_Y(\cdot)} = \frac{P_{Y|X}(\cdot|x)\alpha P_X(x)}{P_Y(\cdot)} = \alpha P_{X|Y}(x|\cdot)$, furthermore, $P_{W|Y}(w=c|\cdot) = 1-\alpha$. Hence, after some calculation we obtain $H(W|Y) = h(\alpha) + \alpha H(X|Y)$. Finally, step (d) follows from (57). ∎

Remark 10. *The constraint $|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+1]\,[|\mathcal{X}|+1]$ in Lemma 9 guarantees that $\sup_U H(U) < \infty$.*

Proof of Theorem 3:

- i $\Rightarrow$ ii: Using Lemma 6 we have $H(Y|X) + \epsilon = g_\epsilon(P_{XY}) \leq h_\epsilon(P_{XY}) \leq H(Y|X) + \epsilon$. Thus, $g_\epsilon(P_{XY}) = h_\epsilon(P_{XY})$.
- ii $\Rightarrow$ iii: Let $\bar{U}$ be an optimizer of $g_\epsilon(P_{XY})$. Thus, the Markov chain $X - Y - \bar{U}$ holds and we have $I(X;\bar{U}|Y) = 0$. Furthermore, since $g_\epsilon(P_{XY}) = h_\epsilon(P_{XY})$ this $\bar{U}$ achieves $h_\epsilon(P_{XY})$. Thus, by using Lemma 7 we have $H(Y|\bar{U},X)=0$ and according to (27)

$$\begin{aligned} I(\bar{U};Y) &= I(X;\bar{U}) + H(Y|X) - H(Y|\bar{U},X) - I(X;\bar{U}|Y) \\ &= I(X;\bar{U}) + H(Y|X). \end{aligned} \tag{58}$$

We claim that $\bar{U}$ must satisfy $I(X;Y|\bar{U}) > 0$ and $I(X;\bar{U}) = \epsilon$. For the first claim assume that $I(X;Y|\bar{U})=0$, hence the Markov chain $X - \bar{U} - Y$ holds. Using $X - \bar{U} - Y$ and $H(Y|\bar{U},X)=0$ we have $H(Y|\bar{U}) = 0$, hence Y and $\bar{U}$ become independent. Using (58)

$$\begin{aligned} H(Y) = I(Y;\bar{U}) &= I(X;\bar{U}) + H(Y|X), \\ \Rightarrow I(X;\bar{U}) &= I(X;Y). \end{aligned}$$

The last line is a contradiction since by assumption we have $I(X;\bar{U}) \leq \epsilon < I(X;Y)$. Thus, $I(X;Y|\bar{U}) > 0$. For proving the second claim assume that $I(X;\bar{U}) = \epsilon_1 < \epsilon$. Let $U = (\bar{U}, W)$ where $W =$

$\begin{cases} Y, & \text{w.p. } \alpha \\ c, & \text{w.p. } 1-\alpha \end{cases}$, and c is a constant that $c \notin \mathcal{X} \cup \mathcal{Y} \cup \bar{\mathcal{U}}$ and $\alpha = \frac{\epsilon-\epsilon_1}{I(X;Y|\bar{U})}$. We show that $\frac{\epsilon-\epsilon_1}{I(X;Y|\bar{U})} < 1$. By the assumption we have

$$\frac{\epsilon-\epsilon_1}{I(X;Y|\bar{U})} < \frac{I(X;Y)-I(X;\bar{U})}{I(X;Y|\bar{U})} \overset{(a)}{\leq} 1,$$

where step (a) follows since $I(X;Y) - I(X;\bar{U}) - I(X;Y|\bar{U}) = I(X;Y) - I(X;Y,\bar{U}) \leq 0$. It can be seen that such U satisfies $H(Y|X,U)=0$ and $I(X;U|Y)=0$ since

$$\begin{aligned} H(Y|X,U) &= \alpha H(Y|X,\bar{U},Y) \\ &\quad + (1-\alpha)H(Y|X,\bar{U}) = 0, \\ I(X;U|Y) &= H(X|Y) - H(X|Y,\bar{U},W) \\ &= H(X|Y) - \alpha H(X|Y,\bar{U}) \\ &\quad - (1-\alpha)H(X|Y,\bar{U}) \\ &= H(X|Y) - H(X|Y) \\ &= 0, \end{aligned}$$

where in deriving the last line we used the Markov chain $X - Y - \bar{U}$. Furthermore,

$$\begin{aligned} I(X;U) &= I(X;\bar{U},W) = I(X;\bar{U}) + I(X;W|\bar{U}) \\ &= I(X;\bar{U}) + \alpha H(X|\bar{U}) - \alpha H(X|\bar{U},Y) \\ &= I(X;\bar{U}) + \alpha I(X;Y|\bar{U}) \\ &= \epsilon_1 + \epsilon - \epsilon_1 \\ &= \epsilon, \end{aligned}$$

and

$$\begin{aligned} I(Y;U) &= I(X;U) + H(Y|X) - H(Y|U,X) - I(X;U|Y) \\ &= \epsilon + H(Y|X). \end{aligned}$$

Thus, if $I(X;\bar{U}) = \epsilon_1 < \epsilon$ we can substitute $\bar{U}$ by U for which $I(U;Y) > I(\bar{U};Y)$. This is a contraction and we conclude that $I(X;\bar{U}) = \epsilon$ which proves the second claim. Hence, (58) can be rewritten as

$$I(\bar{U};Y) = \epsilon + H(Y|X).$$

As a result $h_\epsilon(P_{XY}) = \epsilon + H(Y|X)$ and the proof is completed.

- iii $\Rightarrow$ i: Let $\bar{U}$ be the optimizer of $h_\epsilon(P_{XY})$ and $h_\epsilon(P_{XY}) = H(Y|X) + \epsilon$. Using Lemma 7 we have $H(Y|\bar{U},X) = 0$. By using (27) we must have $I(X;\bar{U}|Y) = 0$ and $I(X;\bar{U}) = \epsilon$. We conclude that for this $\bar{U}$, the Markov chain $X - Y - \bar{U}$ holds and as a result $\bar{U}$ achieves $g_\epsilon(P_{XY})$ and we have $g_\epsilon(P_{XY}) = H(Y|X) + \epsilon$.

Proof of (32)***:*** To prove (32) we use the Wyner notion of common information. Using (17) we know that $C(X;Y) \geq I(X;Y)$, hence, it remains to show that $I(X;Y)$ is achievable. Let $H(X|Y)=0$ and $U=X$ which satisfies $I(X;Y|U)=0$. We have

$$I(U;X,Y) = I(X;X,Y) = H(X) \overset{(a)}{=} I(X,Y).$$

where (a) follows by $H(X|Y)=0$.

Proof of Lemma 7: Let $\bar{U}$ be an optimizer of $h_\epsilon(P_{XY})$ and assume that $H(Y|X,\bar{U}) > 0$. Consequently, we have $I(X;\bar{U}) \leq \epsilon$. Let U' be founded by FRL with $(X,\bar{U})$ instead of X in Lemma 1 and same Y, that is $I(U';X,\bar{U})=0$ and $H(Y|X,\bar{U},U')=0$. Using [16, Th. 5] we have

$$I(Y;U') > 0,$$

since we assumed $H(Y|X,\bar{U}) > 0$. Let $U = (\bar{U},U')$ and we first show that U satisfies $I(X;U) \leq \epsilon$. We have

$$\begin{aligned} I(X;U) &= I(X;\bar{U},U') = I(X;\bar{U}) + I(X;U'|\bar{U}), \\ &= I(X;\bar{U}) + H(U'|\bar{U}) - H(U'|\bar{U},X), \\ &= I(X;\bar{U}) + H(U') - H(U') \leq \epsilon, \end{aligned}$$

where in last line we used the fact that U' is independent of the pair $(X,\bar{U})$. Finally, we show that $I(Y;U) > I(Y,\bar{U})$ which is a contradiction with optimality of $\bar{U}$. We have

$$\begin{aligned} I(Y;U) &= I(Y;\bar{U},U') = I(Y;U') + I(Y;\bar{U}|U'), \\ &= I(Y;U') + I(Y,U';\bar{U}) - I(U';\bar{U}) \\ &= I(Y;U') + I(Y,\bar{U}) + I(U';\bar{U}|Y) - I(U';\bar{U}) \\ &\overset{(a)}{\geq} I(Y;U') + I(Y,\bar{U}) \\ &\overset{(b)}{>} I(Y,\bar{U}), \end{aligned}$$

where in (a) follows since $I(U';\bar{U}|Y) \geq 0$ and $I(U';\bar{U}) = 0$. Step (b) follows since $I(Y;U') > 0$. Thus, the obtained contradiction completes the proof.

Proof of Theorem 4: L_1^0 and L_2^0 can be obtained by letting $\epsilon = 0$ in Theorem 2. U_1^0 which has been derived in [16, Th. 7] can be obtained by (27). U_1^0 can be derived as follows. Since X and U are independent, (27) can be rewritten as

$$I(Y;U) = H(Y|X) - H(Y|U,X) - I(X;U|Y),$$

thus, using Lemma 3

$$\begin{aligned} h_0(P_{XY}) &\leq H(Y|X) - \inf_{H(Y|U,X)=0,\ I(X;U)=0} I(X;U|Y) \\ &= H(Y|X) - \psi(X \to Y) \\ &\leq H(Y|X) + \\ &\sum_{y\in\mathcal{Y}} \int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt \\ &+ I(X;Y). \end{aligned}$$

For $|\mathcal{Y}| = 2$ using Lemma 3 we have $\psi(X \to Y) = -\sum_{y\in\mathcal{Y}} \int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt - I(X;Y)$ and let $\bar{U}$ be the RV that attains this bound. Thus,

$$\begin{aligned} &I(\bar{U};Y) = H(Y|X) + \\ &\sum_{y\in\mathcal{Y}} \int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt + \\ &I(X;Y). \end{aligned}$$

Therefore, $\bar{U}$ attains U_2^0 and $h_0(P_{XY}) = U_0^2$.

Proof of (33)***:*** Since X is a deterministic function of Y, for any $y \in \mathcal{Y}$ we have

$$P_{Y|X}(y|x) = \begin{cases} \frac{P_Y(y)}{P_X(x)}, & x = f(y) \\ 0, & \text{else} \end{cases},$$

thus,

$$\sum_{y\in\mathcal{Y}} \int_0^1 \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt =$$
$$\sum_{y\in\mathcal{Y}} \int_0^{\frac{P_Y(y)}{P_X(x=f(y))}} \mathbb{P}_X\{P_{Y|X}(y|X) \geq t\} \log(\mathbb{P}_X\{P_{Y|X}(y|X) \geq t\}) dt =$$
$$\sum_{y\in\mathcal{Y}} \frac{P_Y(y)}{\mathbb{P}_X\{x = f(y)\}} \mathbb{P}_X\{x = f(y)\} \log(\mathbb{P}_X\{x = f(y)\}) =$$
$$\sum_{y\in\mathcal{Y}} P_Y(y) \log(\mathbb{P}_X\{x = f(y)\}) =$$
$$\sum_{y\in\mathcal{Y}} P_X(x) \log(P_X(x)) =$$
$$-H(X) = -I(X;Y),$$

where in last line we used $\sum_{y\in\mathcal{Y}} P_Y(y) \log(\mathbb{P}_X\{x = f(y)\}) = \sum_{x\in\mathcal{X}} \sum_{y:x=f(y)} P_Y(y) \log(\mathbb{P}_X\{x = f(y)\}) = \sum_{x\in\mathcal{X}} P_X(x) \log(P_X(x))$.

APPENDIX C

Proofs for Privacy-utility trade-off with non-zero leakage and per-letter privacy constraints*:*

Lemma 10. *For any $0 \leq \epsilon < \sqrt{2I(X;Y)}$ and any pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite, there exists a RV U supported on $\mathcal{U}$ such that X and U satisfy the strong privacy criterion 1, i.e., we have*

$$d(P_{X,U}(\cdot,u), P_X P_U(u)) \leq \epsilon, \ \forall u, \tag{59}$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X) = 0, \tag{60}$$

and

$$|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+1]\,[|\mathcal{X}|+1]\,. \tag{61}$$

Proof: Let U be found by the EFRL, where we let the leakage be $\frac{\epsilon^2}{2}$. Thus, we have

$$\begin{aligned} \frac{\epsilon^2}{2} &= I(U;X) \\ &= \sum_u P_U(u) D(P_{X|U}(\cdot|u), P_X) \\ &\overset{(a)}{\geq} \sum_u \frac{P_U(u)}{2} \left(d(P_{X|U}(\cdot|u), P_X)\right)^2 \\ &\overset{(b)}{\geq} \sum_u \frac{P_U(u)^2}{2} \left(d(P_{X|U}(\cdot|u), P_X)\right)^2 \\ &\geq \frac{P_U(u)^2}{2} \left(d(P_{X|U}(\cdot|u), P_X)\right)^2 \\ &= \frac{\left(d(P_{X,U}(\cdot,u), P_X P_U(u))\right)^2}{2}, \end{aligned}$$

where $D(\cdot,\cdot)$ corresponds to KL-divergence. Furthermore, (a) follows by the Pinsker's inequality [44] and (b) follows since $0 \leq P_U(u) \leq 1$. Using the last line we obtain

$$d(P_{X,U}(\cdot,u), P_X P_U(u)) \leq \epsilon, \ \forall u.$$

The other constraints can be obtained by using Lemma 4. ■

Remark 11. *RV U, which is specified by the FRL (Lemma 1), satisfies all constraints in Lemma 10. However, as we show later, it achieves less utility compared to the RV U which is used in the proof of Lemma 10. Furthermore, we can add constraints such as $0 < I(U;X)$ and $0 < \epsilon < \sqrt{2I(X;Y)}$ to Lemma 10 while the RV U found by the FRL does not satisfy them.*

Lemma 11. *For any $0 \leq \epsilon < \sqrt{2I(X;Y)}$ and pair of RVs (X,Y) distributed according to P_{XY} supported on alphabets $\mathcal{X}$ and $\mathcal{Y}$ where $|\mathcal{X}|$ is finite and $|\mathcal{Y}|$ is finite or countably infinite with $I(X,Y) < \infty$, there exists a RV U defined on $\mathcal{U}$ such that X and U satisfy the strong privacy criterion 1, i.e., we have*

$$d(P_{X,U}(\cdot,u), P_X P_U(u)) \leq \epsilon, \ \forall u,$$

Y is a deterministic function of (U,X), i.e., we have

$$H(Y|U,X) = 0,$$

$I(X;U|Y)$ can be upper bounded as follows

$$I(X;U|Y) \leq \alpha H(X|Y) + (1-\alpha)[\log(I(X;Y)+1)+4]\,, \tag{62}$$

and $|\mathcal{U}| \leq [|\mathcal{X}|(|\mathcal{Y}|-1)+2]\,[|\mathcal{X}|+1]\,$, where $\alpha = \frac{\epsilon^2}{2H(X)}$.

Proof: Let U be found by the ESFRL, where we let the leakage be $\frac{\epsilon^2}{2}$. The first constraint in this statement can be obtained by using the same proof as Lemma 10. Furthermore, (62) can be derived using Lemma 5. ■

Remark 12. *RV U produced by the SFRL (Lemma 2) does not satisfy* (62) *in general. However, in case of satisfying* (62)*, by using similar arguments for comparing the attained utility by FRL, SFRL and the extended versions, it achieves less or equal utility compared to the RV U which is used in the proof of Lemma 11. Similarly, we later show that the RV found by*

proof of Lemma 11 strictly improves the utility for non-zero leakage.

Proof of Proposition 4: For deriving $L^1_{h^{w\ell}}(\epsilon)$ let U be produced as in the proof of Lemma 10. Thus, $I(X;U)=\frac{\epsilon^2}{2}$ and U satisfies (59) and (60). We have

$$\begin{aligned} h^{w\ell}_\epsilon(P_{XY}) &\geq I(U;Y) \\ &= I(X;U)+H(Y|X)-I(X;U|Y)-H(Y|X,U) \\ &= \frac{\epsilon^2}{2}+H(Y|X)-H(X|Y)+H(X|Y,U) \\ &\geq \frac{\epsilon^2}{2}+H(Y|X)-H(X|Y). \end{aligned}$$

Next for deriving $L^2_{h^{w\ell}}(\epsilon)$ let U be produced by Lemma 11. Hence, $I(X;U)=\frac{\epsilon^2}{2}$ and U satisfies (59), (60), and (62). We obtain

$$\begin{aligned} h^{w\ell}_\epsilon(P_{XY}) \geq I(U;Y) &= \frac{\epsilon^2}{2}+H(Y|X)-I(X;U|Y) \\ &\geq \frac{\epsilon^2}{2}+H(Y|X)-\alpha H(X|Y) \\ &\quad -(1-\alpha)\left(\log(I(X;Y)+1)+4\right). \end{aligned}$$

Proof of Theorem 6: Lower bounds on $g^{w\ell}_\epsilon(P_{XY})$ and $h^{w\ell}_\epsilon(P_{XY})$ are derived in (46) and Proposition 4, respectively. Furthermore, inequality $g^{w\ell}_\epsilon(P_{XY})\leq h^{w\ell}_\epsilon(P_{XY})$ holds since $h^{w\ell}_\epsilon(P_{XY})$ has less constraints. To prove the upper bound on $g^{w\ell}_\epsilon(P_{XY})$, i.e., $U_{g^{w\ell}(\epsilon)}$, let U satisfy $X-Y-U$ and $P_U(u)d(P_{X|U}(\cdot|u),P_X)\leq\epsilon$, then we have

$$\begin{aligned} I(U;Y) &= I(X;U)+H(Y|X)-I(X;U|Y)-H(Y|X,U) \\ &\overset{(a)}{=} I(X;U)+H(Y|X)-H(Y|X,U) \\ &\leq I(X;U)+H(Y|X) \\ &= \sum_u P_U(u)D(P_{X|U}(\cdot|u),P_X)+H(Y|X) \\ &\overset{(b)}{\leq} \sum_u P_U(u)\frac{\left(d(P_{X|U}(\cdot|u),P_X)\right)^2}{\min P_X}+H(Y|X) \\ &\overset{(c)}{\leq} \sum_u P_U(u)\frac{d(P_{X|U}(\cdot|u),P_X)}{\min P_X}|\mathcal{X}|+H(Y|X) \\ &\overset{(d)}{\leq} \frac{\epsilon|\mathcal{Y}||\mathcal{X}|}{\min P_X}+H(Y|X), \end{aligned}$$

where (a) follows by the Markov chain $X-Y-U$, (b) follows by the reverse Pinsker inequality [44, (23)] and (c) holds since $d(P_{X|U}(\cdot|u),P_X)=\sum_{i=1}^{|\mathcal{X}|}|P_{X|U}(x_i|u)-P_X(x_i)|\leq|\mathcal{X}|$. Latter holds since for each u and i, $|P_{X|U}(x_i|u)-P_X|\leq 1$. Moreover, (d) holds since by Proposition 5 without loss of optimality we can assume $|\mathcal{U}|\leq|\mathcal{Y}|$. In other words (d) holds since by Proposition 5 we have

$$\begin{aligned} g^{w\ell}_\epsilon(P_{XY}) &= \sup_{\substack{P_{U|Y}:X-Y-U \\ d(P_{X,U}(\cdot,u),P_XP_U(u))\leq\epsilon,\ \forall u}} I(Y;U) \\ &= \max_{\substack{P_{U|Y}:X-Y-U \\ d(P_{X,U}(\cdot,u),P_XP_U(u))\leq\epsilon,\ \forall u \\ |\mathcal{U}|\leq|\mathcal{Y}|}} I(Y;U). \end{aligned} \tag{63}$$

Proof of Proposition 6: By using [13, Proposition 2], it suffices to assume $|\mathcal{U}|\leq|\mathcal{Y}|$. Using [13, Proposition 3], let us consider $|\mathcal{Y}|$ extreme points that achieves the minimum in [13, Theorem 2] as V_{Ω_j} for $j\in\{1,..,|\mathcal{Y}|\}$. Let $|\mathcal{X}|$ non-zero elements of V_{Ω_j} be $a_{ij}+\epsilon b_{ij}$ for $i\in\{1,..,|\mathcal{X}|\}$ and $j\in\{1,..,|\mathcal{Y}|\}$, where a_{ij} and b_{ij} can be found in [13, (6)]. As a summary for $i\in\{1,..,|\mathcal{X}|\}$ and $j\in\{1,..,|\mathcal{Y}|\}$ we have $\sum_i a_{ij}=1$, $\sum_i b_{ij}=0$, $0\leq a_{ij}\leq 1$, and $0\leq a_{ij}+\epsilon b_{ij}\leq 1$. We obtain

$$\begin{aligned} \max I(U;Y) &= H(Y)\sum_j P_j\sum_i(a_{ij}+\epsilon b_{ij})\log(a_{ij}+\epsilon b_{ij}), \\ &= H(Y)+\sum_j P_j\sum_i(a_{ij}+\epsilon b_{ij})(\log(a_{ij})+\log(1+\epsilon\frac{b_{ij}}{a_{ij}})). \end{aligned}$$

In [13, Theorem 2], we have used the Taylor expansion to derive the approximation of the equivalent problem. From the Taylor's expansion formula we have

$$\begin{aligned} f(x) = f(a)+\frac{f'(a)}{1!}(x-a)+ \\ \frac{f''(a)}{2!}(x-a)^2+...+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_{n+1}(x), \end{aligned}$$

where

$$\begin{aligned} R_{n+1}(x) &= \int_a^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)dt \\ &= \frac{f^{(n+1)}(\zeta)}{(n+1)!}(x-a)^{n+1}, \end{aligned} \tag{64}$$

for some $\zeta\in[a,x]$. In [13] we approximated the terms $\log(1+\frac{b_{ij}}{a_{ij}}\epsilon)$ by $\frac{b_{ij}}{a_{ij}}\epsilon+o(\epsilon)$. Using (64), there exists an $\zeta_{ij}\in[0,\epsilon]$ such that the error of approximating the term $\log(1+\epsilon\frac{a_{ij}}{b_{ij}})$ is as follows

$$R^{ij}_2(\epsilon)=-\frac{1}{2}\left(\frac{\frac{b_{ij}}{a_{ij}}}{1+\frac{b_{ij}}{a_{ij}}\zeta_{ij}}\right)^2\epsilon^2=-\frac{1}{2}\left(\frac{b_{ij}}{a_{ij}+b_{ij}\zeta_{ij}}\right)^2\epsilon^2.$$

Thus, the error of approximation is as follows

$$\begin{aligned} \text{Approximation error} &= \sum_{ij}P_j(a_{ij}+\epsilon b_{ij})R^{ij}_2(\epsilon)+\sum_{ij}P_j\frac{b^2_{ij}}{a_{ij}}\epsilon^2 \\ &= -\sum_{ij}P_j(a_{ij}+\epsilon b_{ij})\frac{1}{2}\left(\frac{b_{ij}}{a_{ij}+b_{ij}\zeta_{ij}}\right)^2\epsilon^2+\sum_{ij}P_j\frac{b^2_{ij}}{a_{ij}}\epsilon^2 \end{aligned} \tag{65}$$

An upper bound on approximation error can be obtained as follows

$$\begin{aligned} &|\text{Approximation error}| \leq \\ &|\sum_{ij}P_j(a_{ij}+\epsilon b_{ij})\frac{1}{2}\left(\frac{b_{ij}}{a_{ij}+b_{ij}\zeta_{ij}}\right)^2\epsilon^2|+|\sum_{ij}P_j\frac{b^2_{ij}}{a_{ij}}\epsilon^2|. \end{aligned} \tag{66}$$

By using the definition of ϵ_2 in Proposition 5 we have $\epsilon<\epsilon_2$ implies $\epsilon<\frac{\min_{ij}a_{ij}}{\max_{ij}|b_{ij}|}$, since $\min_{ij}a_{ij}=\min_{y,\Omega\in\Omega^1}M^{-1}_\Omega MP_Y(y)$ and $\max_{ij}|b_{ij}|<\max_{\Omega\in\Omega^1}|\sigma_{\max}(H_\Omega)|$. By using the upper

bound $\epsilon < \frac{\min_{ij} a_{ij}}{|\max_{ij} b_{ij}|}$ we can bound the second term in (66) by 1, since we have

$$\begin{aligned}|\sum_{ij} P_j \frac{b_{ij}^2}{a_{ij}}\epsilon^2| &< |\sum_{ij} P_j \frac{b_{ij}^2}{a_{ij}} \left(\frac{\min_{ij} a_{ij}}{\max_{ij}|b_{ij}|}\right)^2| \\ &< |\sum_{ij} P_j \min_{ij} a_{ij}| = |\mathcal{X}| \min_{ij} a_{ij} \\ &\overset{(a)}{<} 1,\end{aligned}$$

where (a) follows from $\sum_i a_{ij} = 1,\ \forall j \in \{1, .., |\mathcal{Y}|\}$.
If we use $\frac{1}{2}\epsilon_2$ as an upper bound on ϵ, we have $\epsilon < \frac{1}{2}\frac{\min_{ij} a_{ij}}{\max_{ij}|b_{ij}|}$. We show that by using this upper bound the first term in (66) can be upper bounded by $\frac{1}{2}$. We have

$$\begin{aligned}&\frac{1}{2}|\sum_{ij} P_j(a_{ij} + \epsilon b_{ij}) \left(\frac{b_{ij}}{a_{ij} + b_{ij}\zeta_{ij}}\right)^2 \epsilon^2| \overset{(a)}{<} \\ &\frac{1}{2}|\sum_{ij} P_j(a_{ij} + \epsilon b_{ij}) \left(\frac{|b_{ij}|}{a_{ij} - \epsilon|b_{ij}|}\epsilon\right)^2| \overset{(b)}{<} \\ &\frac{1}{2}|\sum_{ij} P_j(a_{ij} + \epsilon b_{ij})| < \frac{1}{2},\end{aligned}$$

where (a) follows from $0 \le \zeta_{ij} \le \epsilon,\ \forall i,\ \forall j$, and (b) follows from $\frac{|b_{ij}|}{a_{ij} - \epsilon|b_{ij}|}\epsilon < 1$ for all i and j. The latter can be shown as follows

$$\frac{|b_{ij}|}{a_{ij} - \epsilon|b_{ij}|}\epsilon < \frac{|b_{ij}|}{a_{ij} - \frac{1}{2}\frac{\min_{ij} a_{ij}}{\max_{ij}|b_{ij}|}|b_{ij}|}\epsilon < \frac{b_{ij}}{\frac{1}{2}\min_{ij} a_{ij}}\epsilon < 1.$$

For $\epsilon < \frac{1}{2}\epsilon_2$ the term $a_{ij} - \epsilon|b_{ij}|$ is positive and there is no need of absolute value for this term. Thus, $\epsilon < \frac{1}{2}\epsilon_2$ implies the following upper bound

$$|\text{Approximation error}| < \frac{3}{4}.$$

Furthermore, by following similar steps if we use the upper bound $\epsilon < \frac{1}{2}\frac{\epsilon_2}{\sqrt{|\mathcal{X}|}}$ instead of $\epsilon < \frac{1}{2}\epsilon_2$, the upper bound on error can be strengthened by

$$|\text{Approximation error}| < \frac{1}{2(2\sqrt{|\mathcal{X}|} - 1)^2} + \frac{1}{4|\mathcal{X}|}.$$

APPENDIX D

Proofs for Privacy-utility trade-off with non-zero leakage and prioritized private data*:*

Proof of Theorem 8: The upper bound can be obtained using the key equation in (27), since the total leakage $I(U; X_1, X_2)$ is bounded by ϵ. The first lower bound $L^1_{h^{12}}(\epsilon)$ can be obtained as follows. Let $\bar{U}$ be found by FRL with $X = (X_1, X_2)$. Moreover, let $U = (\bar{U}, W)$ with $W = \begin{cases} X_2, & \text{w.p. } \alpha \\ c, & \text{w.p. } 1-\alpha \end{cases}$, where c is a constant which does not belong to $\mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{Y}$ and $\alpha = \frac{\epsilon}{H(X_2)}$. We have

$$\begin{aligned}I(U; X_1, X_2) &= I(\bar{U}, W; X_1, X_2) \overset{(a)}{=} I(W; X_1, X_2) \\ &= H(X_1,X_2) - \alpha H(X_1|X_2) - (1-\alpha)H(X_1,X_2) \\ &= \alpha H(X_2) = \epsilon,\end{aligned}$$

where (a) follows since $\bar{U}$ is independent of (X_1, X_2, W). For the other leakage constraint we have

$$\begin{aligned}I(U; X_2) &= I(W; X_2) \\ &= \alpha H(X_2) \\ &\ge \alpha I(X_1; X_2) \\ &= H(X_1) - \alpha H(X_1|X_2) - (1-\alpha)H(X_1) \\ &= I(U; X_1),\end{aligned}$$

and by using (27), we obtain

$$\begin{aligned}h^p_\epsilon(P_{X_1X_2Y}) &\ge I(U; Y) \\ &= \epsilon + H(Y|X_1, X_2) - I(X_1, X_2; U|Y) \\ &\ge \epsilon + H(Y|X_1, X_2) - H(X_1, X_2|Y).\end{aligned}$$

The bounds $L^2_{h^p}(\epsilon)$ and $L^3_{h^p}(\epsilon)$ can be obtained as follows. Let $\bar{U}$ be found by SFRL with $X = (X_1, X_2)$. Moreover, let $U = (\bar{U}, W)$ with $W = \begin{cases} X_2, & \text{w.p. } \alpha \\ c, & \text{w.p. } 1-\alpha \end{cases}$, where c is a constant which does not belong to $\mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{Y}$ and $\alpha = \frac{\epsilon}{H(X_2)}$. Similarly we have $I(U; X_1, X_2) = \epsilon$ and $I(U; X_2) \ge I(U; X_1)$. Next, we expand $I(U; X_1, X_2|Y)$.

$$\begin{aligned}&I(U; X_1, X_2|Y) = I(\bar{U}; X_1, X_2|Y) + I(W; X_1, X_2|Y, \bar{U}) \\ &= I(\bar{U}; X_1, X_2|Y) + H(X_1, X_2|Y, \bar{U}) - H(X_1, X_2|Y, \bar{U}, W) \\ &= I(\bar{U}; X_1, X_2|Y) + \alpha H(X_1, X_2|Y, \bar{U}) - \alpha H(X_1|Y, \bar{U}, X_2) \\ &= I(\bar{U}; X_1, X_2|Y) - \alpha H(X_1|Y, \bar{U}, X_2) \\ &+ \alpha \left(H(X_1, X_2|Y) - I(\bar{U}; X_1, X_2|Y)\right) \\ &= (1-\alpha) I(\bar{U}; X_1, X_2|Y) + \alpha H(X_1, X_2|Y) - \alpha H(X_1|Y, \bar{U}, X_2).\end{aligned} \tag{67}$$

In the following we bound (67) in two ways. We have

$$\begin{aligned}(67) &= (1-\alpha) I(\bar{U}; X_1, X_2|Y) + \alpha H(X_2|Y) + \alpha I(X_1; \bar{U}|Y, X_2) \\ &= I(\bar{U}; X_1, X_2|Y) + \alpha H(X_2|Y) - \alpha I(\bar{U}; X_2|Y) \\ &\overset{(a)}{\le} \log(I(X_1, X_2; Y) + 1) + 4 + \alpha H(X_2|Y).\end{aligned} \tag{68}$$

Furthermore,

$$\begin{aligned}(67) &\le (1-\alpha) I(\bar{U}; X_1, X_2|Y) + \alpha H(X_1, X_2|Y) \\ &\overset{(b)}{\le} (1-\alpha)\left(\log(I(X_1, X_2; Y) + 1) + 4\right) + \alpha H(X_1, X_2|Y).\end{aligned} \tag{69}$$

Inequalities (a) and (b) follow since $\bar{U}$ is produced by SFRL, so that $I(\bar{U}; X_1, X_2|Y) \le \log(I(X_1, X_2; Y) + 1) + 4$. Using (68), (69) and key equation in (27) we have

$$\begin{aligned}&h^p_\epsilon(P_{X_1X_2Y}) \ge I(U; Y) \overset{(c)}{\ge} \\ &\epsilon + H(Y|X_1, X_2) - (\log(I(X_1, X_2; Y) + 1) + 4 + \alpha H(X_2|Y)) \\ &= L^2_{h^p}(\epsilon),\end{aligned}$$

and

$$\begin{aligned}&h^p_\epsilon(P_{X_1X_2Y}) \ge I(U; Y) \\ &\overset{(d)}{\ge} \epsilon + H(Y|X_1, X_2) \\ &- \left((1-\alpha)(\log(I(X_1, X_2; Y) + 1) + 4) + \alpha H(X_1, X_2|Y)\right) \\ &= L^3_{h^p}(\epsilon).\end{aligned}$$

In steps (c) and (d) we used $H(Y|X_1, X_2, U) = 0$. The latter follows by definition of W and the fact that $\bar{U}$ is produced by SFRL.

Amirreza Zamani (Member IEEE) received the B. Tech. degree in electrical engineering from the University of Tehran, Iran, in 2016, the M.Sc. degree from Sharif university of Technology, Iran, in 2018. He is presently a Ph.D. student at KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include statistical inference, information theory, information-theoretic privacy and security.

Tobias J. Oechtering (S01-M08-SM12) received his Dipl-Ing degree in Electrical Engineering and Information Technology in 2002 from RWTH Aachen University, Germany, his Dr-Ing degree in Electrical Engineering in 2007 from the Technische Universität Berlin, Germany. In 2008 he joined KTH Royal Institute of Technology, Stockholm, Sweden and has been a Professor since 2018. In 2009, he received the "Förderpreis 2009 from the Vodafone Foundation.

Dr. Oechtering is currently Senior Editor of IEEE Transactions on Information Forensic and Security since May 2020 and served previously as Associate Editor for the same journal since June 2016, and IEEE Communications Letters during 2012-2015. He has served on numerous technical program committees for IEEE sponsored conferences, and he was general co-chair for IEEE ITW 2019. His research interests include privacy and physical layer security, communication and information theory, statistical learning and signal processing, as well as communications for networked control.

Mikael Skoglund (S'93-M'97-SM'04-F'19) received the Ph.D. degree in 1997 from Chalmers University of Technology, Sweden. In 1997, he joined the Royal Institute of Technology (KTH), Stockholm, Sweden, where he was appointed to the Chair in Communication Theory in 2003. At KTH, he heads the Division of Information Science and Engineering, as well as the Department of Intelligent Systems.

Dr. Skoglund has worked on problems in source-channel coding, coding and transmission for wireless communications, Shannon theory, information-theoretic security, information theory for statistics and learning, information and control, and signal processing. He has authored and co-authored around 200 journal and more than 400 conference papers.

Dr. Skoglund is a Fellow of the IEEE. During 2003–08 he was an associate editor for the IEEE Transactions on Communications. In the interval 2008–12 he was on the editorial board for the IEEE Transactions on Information Theory and starting in the Fall of 2021 he he joined it once again. He has served on numerous technical program committees for IEEE sponsored conferences, he was general co-chair for IEEE ITW 2019 and TPC co-chair for IEEE ISIT 2022. He is an elected member of the IEEE Information Theory Society Board of Governors.