<http://www.diva-portal.org>

This is the published version of a paper presented at *2020 IEEE Information Theory Workshop (ITW)*.

Citation for the original published paper:

Zamani, A., Oechtering, T J., Skoglund, M. (2020)
Data Disclosure Mechanism Design with Non-zero Leakage
In:
<https://doi.org/10.1109/ITW46852.2021.9457592>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-299952>

Data Disclosure Mechanism Design with Non-zero Leakage

Amirreza Zamani, Tobias J. Oechtering, Mikael Skoglund

Division of Information Science and Engineering, KTH Royal Institute of Technology

Email: amizam@kth.se, oech@kth.se, mikael.skoglund@ee.kth.se

Abstract—We study an information-theoretic privacy problem, where an agent observes useful data Y and wants to reveal the information to a user. Since the useful data is correlated with sensitive data X , the agent employs a privacy mechanism to produce data U that can be disclosed. Thus, we study the privacy mechanism design that maximizes the revealed information about Y while satisfying an ℓ_1 -privacy criterion under the Markov chain $X - Y - U$. When a sufficiently small leakage is allowed, we show that the optimizer of the design problem has a specific structure which allows us to use a local approximation of mutual information. More specifically, we show that the optimizer vectors are perturbations of fixed distributions. By using this approximation the original optimization problem can be reduced to a linear programming problem and an approximate solution for privacy mechanism design can be obtained.

I. INTRODUCTION

The amount of data generated by software system, interconnected sensors that record and process signals from the physical world, robots and humans is growing rapidly. Direct disclosure of the raw data can cause privacy breaches through illegitimate inferences. Therefore, privacy mechanisms are needed to control the disclosure of the data.

Privacy mechanism design from an information theoretic perspective is receiving increased attention recently, c.f. [1]–[9]. Specifically, the concept of maximal leakage has been introduced in [1] and some bounds on the privacy utility trade-off have been derived. In [2], the concept of a privacy funnel is introduced where the privacy utility trade-off considering log-loss for privacy metric and distortion metric has been studied. The concept of differential privacy is introduced in [3], which aims to answer queries in a privacy preserving approach using minimizing the chance of identifying the membership in an statistical database. Fundamental limits of the privacy utility trade-off are studied in [4] measuring the leakage using estimation-theoretic guarantees. A source coding problem with secrecy is studied in [5]. In both [5] and [6], the privacy-utility trade-off is considered using expected distortion as a measure of utility and equivocation as measure of privacy.

In [7], the problem of maximizing utility, i.e., $I(U; Y)$, under the leakage constraint $I(U; X) \leq \epsilon$ and Markov chain $X - Y - U$ is studied and it is shown that under perfect privacy assumption, i.e., $\epsilon = 0$, the privacy mechanism design problem can be reduced to standard linear programming.

Our problem in the present paper is closely related to our previous work [8], where the problem of maximizing $I(U; Y)$ under a strong χ^2 -privacy criterion and Markov chain $X - Y - U$ is studied. It has been shown that for small ϵ , the original highly challenging optimization

Fig. 1. In this system model, disclosed data U is designed by a privacy mechanism which maximizes the information disclosed about Y and satisfies the strong ℓ_1 -privacy criterion.

problem can be approximated by the problem of finding the principal right-singular vector of a matrix using the concept of Euclidean information geometry. Furthermore, a geometrical interpretation of the privacy mechanism design has been provided. Also in [9], the concept of information geometry is adopted for the problem of privacy preserving data disclosure with specific utility task under perfect obfuscation constraint.

In this paper, we generalize the problem considered in [8] as shown in Fig. 1, where an agent tries to reveal some useful information to a user. Random variable (rv) Y denotes the useful data and is dependent on private data denoted by rv X , which may not be accessible by the agent. The useful data can not be disclosed directly since the agent wants to protect the private data. Thus, the agent employs a privacy mechanism to produce disclosed data denoted by rv U . U should disclose as much information about Y as possible and fulfill the privacy criterion simultaneously. In this work, we consider an element-wise ℓ_1 -privacy criterion, which we call “strong ℓ_1 -privacy criterion.” A strong χ^2 -privacy criterion is also considered in [8], where the privacy utility trade-off is studied using concepts of Euclidean information theory. A χ^2 -privacy criterion is also considered in [4], where upper bound and lower bounds on the privacy utility trade-off have been derived. The strong χ^2 -privacy criterion used in [8], is a point-wise criterion, while, in [4], an on average criterion is employed.

In [8], the leakage matrix, i.e., $P_{X|Y}$, is assumed to be invertible, however, in this work we generalize the scenario by letting $P_{X|Y}$ be a full row rank and fat matrix which belongs to $\mathcal{H}_{XY}$.

In [10] and [11], methods from Euclidean information theory are used to approximate KL divergence and mutual information in order to simplify the optimization problem. More specifically, in [10], this framework is employed for point-to-point channels and some broadcast channels. In this work, we show that the optimizer distributions have special geometry and therefore we can approximate the mutual information using information geometry concepts. We show that the main problem can be approximated by a standard

linear programming problem when small leakage is allowed. Furthermore, when the leakage tends to zero we obtain the same linear programming problem as in [7]. In more detail our contribution can be summarized as follows

(i) We present an information-theoretic disclosure problem using a strong ℓ_1 -privacy criterion in Section II.

(ii) We show that the distributions $P_{Y|U=u}$ can be decomposed into two parts, where the first part refers to the null space of $P_{X|Y}$ and the second part refers to the non-singular part of $P_{X|Y}$. Then, by using this decomposition we characterize the optimizer vectors of the main problem, which have a special geometry. Accordingly, by utilizing concepts from information geometry, we can approximate the main problem by a standard linear programming in Section III.

The paper is concluded in Section IV.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Let P_{XY} denote the joint distribution of discrete random variables X and Y defined on finite alphabets $\mathcal{X}$ and $\mathcal{Y}$ with $|\mathcal{X}| < |\mathcal{Y}|$. We represent P_{XY} by a matrix defined on $\mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|}$ and 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 assume that each element in vectors P_X and P_Y is non-zero. Furthermore, we represent the leakage matrix $P_{X|Y}$ by a matrix defined on $\mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|}$, which is assumed to be of full row rank. Thus, without loss of generality we represent $P_{X|Y}$ by two submatrices where the first submatrix is invertible, i.e., $P_{X|Y} = [P_{X|Y_1}, P_{X|Y_2}]$ such that $P_{X|Y_1}$ defined on $\mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|}$ is invertible. In the privacy problem with perfect privacy [7], it has been shown that information can be only revealed if $P_{X|Y}$ is not invertible. In [8], we designed a privacy mechanism for revealing information for an invertible $P_{X|Y}$, where a small leakage is allowed. The present work provides a framework where we unify these two results for a full row rank leakage matrix. RVs X and Y denote the private data and the useful data.

In this work, privacy is measured by the strong ℓ_1 -privacy criterion which we introduce next.

Definition 1. For private data $X \in \mathcal{X}$, the disclosed data $U \in \mathcal{U}$ satisfies the "strong ℓ_1 -privacy criterion" with leakage $\epsilon > 0$ if the joint pmf P_{XU} satisfies for all $u \in \mathcal{U}$ the following condition

$$\|P_{X|U=u} - P_X\|_1 = \sum_x |P_{X|U=u}(x) - P_X(x)| \leq \epsilon.$$

The norm is the ℓ_1 -norm.

Intuitively, for small ϵ , the privacy criterion means that the two distributions (vectors) $P_{X|U=u}$ and P_X are close to each other. This should hold for all $u \in \mathcal{U}$. The closeness of $P_{X|U=u}$ and P_X allows us to locally approximate the conditional entropy $H(Y|U)$.

Our goal is to design a privacy mechanism that produces disclosed data U that maximizes $I(U; Y)$ and satisfies the strong ℓ_1 -privacy criterion for sufficiently small ϵ . The relation between U and Y is described by the kernel $P_{U|Y}$ defined on $\mathbb{R}^{|\mathcal{U}| \times |\mathcal{Y}|}$. Therefore, the privacy mechanism

design problem can be stated as follows

$$\max_{P_{U|Y}} I(U; Y), \quad (1a)$$

$$\text{subject to: } X - Y - U, \quad (1b)$$

$$\|P_{X|U=u} - P_X\|_1 \leq \epsilon, \quad \forall u \in \mathcal{U}. \quad (1c)$$

We refer to problem (1) as the ϵ -private data disclosure problem.

Remark 1. For $\epsilon = 0$, (1) leads to the perfect privacy problem studied in [7].

Remark 2. The relation between the strong ℓ_1 -privacy criterion and the strong χ^2 -privacy criterion defined in [8, Definition 1] is as follows: The strong ℓ_1 -privacy criterion implies the strong χ^2 -privacy criterion with a different ϵ , since by using the fact that the ℓ_2 -norm is upper-bounded by the ℓ_1 -norm we have

$$\begin{aligned} \|P_{X|U=u} - P_X\|_2 &\leq \|P_{X|U=u} - P_X\|_1 \leq \epsilon \\ \left\| [\sqrt{P_X}^{-1}] (P_{X|U=u} - P_X) \right\|_2^2 &\leq \frac{1}{\min P_X} \|P_{X|U=u} - P_X\|_1^2 \\ \Rightarrow \chi^2(P_{X|U=u} \| P_X) &\leq (\epsilon')^2, \quad \forall u \in \mathcal{U}, \end{aligned}$$

where $\epsilon' = \frac{\epsilon}{\sqrt{\min P_X}}$ and $\min P_X = \min_{x \in \mathcal{X}} P_X(x)$. Thus, as shown in [8] by using an inequality between KL-divergence and the strong χ^2 -privacy criterion, the mutual information $I(U, X)$ is bounded by $(\epsilon')^2$. Furthermore, by using the concept of Euclidean information theory, $(\epsilon')^2$ can be strengthened and replaced by $\frac{1}{2}(\epsilon')^2 + o((\epsilon')^2)$ for small ϵ' [8, Proposition 2]. Thus, (1c) implies $I(X; U) \leq \frac{1}{2}(\epsilon')^2 + o((\epsilon')^2)$ for small ϵ' .

Proposition 1. It suffices to consider U such that $|\mathcal{U}| \leq |\mathcal{Y}|$. Furthermore, a maximum can be used in (1a) since the corresponding supremum is achieved.

Proof: The proof is based on Fenchel-Eggleston-Carathéodory's Theorem [12]. For more details see [13]. ■

III. PRIVACY MECHANISM DESIGN

In this section we show that the problem defined in (1) can be approximated by a quadratic problem. Furthermore, it is shown that the quadratic problem can be converted to a standard linear program. By using (1c), we can rewrite the conditional distribution $P_{X|U=u}$ as a perturbation of P_X . Thus, for any $u \in \mathcal{U}$, similarly as in [8], [10], [11], we can write $P_{X|U=u} = P_X + \epsilon \cdot J_u$, where $J_u \in \mathbb{R}^{|\mathcal{X}|}$ is the perturbation vector that has the following three properties:

$$\sum_{x \in \mathcal{X}} J_u(x) = 0, \quad \forall u, \quad (2)$$

$$\sum_{u \in \mathcal{U}} P_U(u) J_u(x) = 0, \quad \forall x, \quad (3)$$

$$\sum_{x \in \mathcal{X}} |J_u(x)| \leq 1, \quad \forall u. \quad (4)$$

The first two properties ensure that $P_{X|U=u}$ is a valid probability distribution and the third property follows from (1c). In the following lemma we derive two important properties of $\text{Null}(P_{X|Y})$.

Lemma 1. Let β be a vector in $\mathbb{R}^{|\mathcal{X}|}$. Then, $\beta \in \text{Null}(P_{X|Y})$ if and only if $\beta \in \text{Null}(M)$, where $M \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|}$ is constructed as follows: Let V be the matrix of right eigenvectors of $P_{X|Y}$, i.e., $P_{X|Y} = U\Sigma V^T$ and $V = [v_1, v_2, \dots, v_{|\mathcal{Y}|}]$, then M is defined as

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

Furthermore, if $\beta \in \text{Null}(P_{X|Y})$, then $1^T \beta = 0$.

Proof: Since the rank of $P_{X|Y}$ is $|\mathcal{X}|$, every vector β in $\text{Null}(P_{X|Y})$ can be written as linear combination of $\{v_{|\mathcal{X}|+1}, \dots, v_{|\mathcal{Y}|}\}$ and since each vector in $\{v_{|\mathcal{X}|+1}, \dots, v_{|\mathcal{Y}|}\}$ is orthogonal to the rows of M we conclude that $\beta \in \text{Null}(M)$. If $\beta \in \text{Null}(M)$, β is orthogonal to the vectors $\{v_1, \dots, v_{|\mathcal{X}|}\}$ and thus, $\beta \in \text{linspan}\{v_{|\mathcal{X}|+1}, \dots, v_{|\mathcal{Y}|}\} = \text{Null}(P_{X|Y})$. Furthermore, for every $\beta \in \text{Null}(P_{X|Y})$, we can write $1^T \beta = 1^T P_{X|Y} \beta = 0$. ■

The next lemma shows that if the Markov chain $X - Y - U$ holds and J_u satisfies the three properties (2), (3), and (4), then $P_{Y|U=u}$ lies in a convex polytope.

Lemma 2. For sufficiently small $\epsilon > 0$, for every $u \in \mathcal{U}$, the vector $P_{Y|U=u}$ belongs to the convex polytope $\mathbb{S}_u$ defined as

$$\mathbb{S}_u \triangleq \left\{ y \in \mathbb{R}^{|\mathcal{Y}|} \mid My = MP_Y + \epsilon M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}, y \geq 0 \right\},$$

where $\begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix} \in \mathbb{R}^{|\mathcal{Y}|}$ and J_u satisfies (2), (3), and (4).

Proof: By using the Markov chain $X - Y - U$ we have

$$\epsilon J_u = P_{X|U=u} - P_X = P_{X|Y} [P_{Y|U=u} - P_Y]. \quad (5)$$

Let $\alpha = P_{Y|U=u} - P_Y$. By partitioning α in two parts $[\alpha_1 \ \alpha_2]^T$ with sizes $|\mathcal{X}|$ and $|\mathcal{Y}| - |\mathcal{X}|$, respectively, from (5) we obtain

$$P_{X|Y_1} \alpha_1 + P_{X|Y_2} \alpha_2 = \epsilon J_u,$$

which implies using invertibility of $P_{X|Y_1}$ that

$$\alpha_1 = \epsilon P_{X|Y_1}^{-1} J_u - P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2,$$

thus,

$$P_{Y|U=u} = P_Y + \epsilon \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix} + \begin{bmatrix} -P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2 \\ \alpha_2 \end{bmatrix}.$$

Note that the vector $\begin{bmatrix} -P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2 \\ \alpha_2 \end{bmatrix}$ belongs to $\text{Null}(P_{X|Y})$ since we have

$$\begin{aligned} P_{X|Y} \begin{bmatrix} -P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2 \\ \alpha_2 \end{bmatrix} &= [P_{X|Y_1} \ P_{X|Y_2}] \begin{bmatrix} -P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2 \\ \alpha_2 \end{bmatrix} \\ &= -P_{X|Y_1} P_{X|Y_1}^{-1} P_{X|Y_2} \alpha_2 + P_{X|Y_2} \alpha_2 \\ &= \mathbf{0} \in \mathbb{R}^{|\mathcal{X}|}, \end{aligned}$$

and thus by Lemma 1 it belongs to $\text{Null}(M)$. So, we have

$$MP_{Y|U=u} = MP_Y + \epsilon M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}.$$

Consequently, if the Markov chain $X - Y - U$ holds and the perturbation vector J_u satisfies (2), (3), and (4), then we have $P_{Y|U=u} \in \mathbb{S}_u$. ■

Lemma 3. Any vector α in $\mathbb{S}_u$ is a standard probability vector. Also, for any pair (U, Y) , for which $P_{Y|U=u} \in \mathbb{S}_u$, $\forall u \in \mathcal{U}$ with J_u satisfying (2), (3), and (4), we can have $X - Y - U$ and $P_{X|U=u} - P_X = \epsilon \cdot J_u$.

Proof: The proof is provided in Appendix A. ■

Theorem 1. We have the following equivalency

$$\min_{P_{U|Y}: X-Y-U} H(Y|U) = \min_{P_U, P_{Y|U=u} \in \mathbb{S}_u, \forall u \in \mathcal{U}, \sum_u P_U(u) P_{Y|U=u} = P_Y, J_u \text{ satisfies (2), (3), and (4)}} H(Y|U). \quad (6)$$

Proof: The proof follows directly from Lemmas 2 and 3. ■

In next proposition we discuss how $H(Y|U)$ is minimized over $P_{Y|U=u} \in \mathbb{S}_u$ for all $u \in \mathcal{U}$.

Proposition 2. Let $P_{Y|U=u}^*$, $\forall u \in \mathcal{U}$ be the minimizer of $H(Y|U)$ over the set $\{P_{Y|U=u} \in \mathbb{S}_u, \forall u \in \mathcal{U} \mid \sum_u P_U(u) P_{Y|U=u} = P_Y\}$, then $P_{Y|U=u}^* \in \mathbb{S}_u$ for all $u \in \mathcal{U}$ must belong to extreme points of $\mathbb{S}_u$.

Proof: The proof builds on the concavity property of the entropy. For more details see [13]. ■

In order to solve the minimization problem found in (6), we propose the following procedure: In the first step, we find the extreme points of $\mathbb{S}_u$, which is an easy task. Since the extreme points of the sets $\mathbb{S}_u$ have a particular geometry, in the second step, we locally approximate the conditional entropy $H(Y|U)$ so that we end up with a quadratic problem with quadratic constraints over $P_U(\cdot)$ and J_u that can be easily solved.

A. Finding $\mathbb{S}_u^*$ (Extreme points of $\mathbb{S}_u$)

In this part we find the extreme points of $\mathbb{S}_u$ for each $u \in \mathcal{U}$. As argued in [7], the extreme points of $\mathbb{S}_u$ are the basic feasible solutions of $\mathbb{S}_u$. Basic feasible solutions of $\mathbb{S}_u$ can be found in the following manner. Let Ω be the set of indices which corresponds to $|\mathcal{X}|$ linearly independent columns of M , i.e., $|\Omega| = |\mathcal{X}|$ and $\Omega \subset \{1, \dots, |\mathcal{Y}|\}$. Let $M_\Omega \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|}$ be the submatrix of M with columns indexed by the set Ω . It can be seen that M_Ω is an invertible matrix since $\text{rank}(M) = |\mathcal{X}|$. Then, if all elements of the vector $M_\Omega^{-1}(MP_Y + \epsilon M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix})$ are non-negative, the vector $V_\Omega^* \in \mathbb{R}^{|\mathcal{Y}|}$, which is defined in the following, is a basic feasible solution of $\mathbb{S}_u$. Assume that $\Omega = \{\omega_1, \dots, \omega_{|\mathcal{X}|}\}$, where $\omega_i \in \{1, \dots, |\mathcal{Y}|\}$ and all elements are arranged in an increasing order. The ω_i -th element of V_Ω^* is defined as i -th element of $M_\Omega^{-1}(MP_Y + \epsilon M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix})$, i.e., for $1 \leq i \leq |\mathcal{X}|$ we have

$$V_\Omega^*(\omega_i) = (M_\Omega^{-1} MP_Y + \epsilon M_\Omega^{-1} M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix})(i). \quad (7)$$

Other elements of V_Ω^* are set to be zero. In the next proposition we show two properties of each vector inside V_Ω^* .

Proposition 3. Let $\Omega \subset \{1, \dots, |\mathcal{Y}|\}$, $|\Omega| = |\mathcal{X}|$. For every Ω we have $1^T (M_\Omega^{-1} MP_Y) = 1$. Furthermore, $1^T (M_\Omega^{-1} M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}) = 0$.

Proof: The proof is provided in Appendix B. ■

Next we define a set $\mathcal{H}_{XY}$ which includes all leakage matrices $P_{X|Y}$ having a property that allows us to approximate the conditional entropy $H(Y|U)$.

Definition 2. Let $\mathcal{P}(\mathcal{X})$ be the standard probability simplex defined as $\mathcal{P}(\mathcal{X}) = \{x \in \mathbb{R}^{|\mathcal{X}|} | 1^T x = 1, x(i) \geq 0, 1 \leq i \leq |\mathcal{X}|\}$ and let $\mathcal{P}'(\mathcal{X})$ be the subset of $\mathcal{P}(\mathcal{X})$ defined as $\mathcal{P}'(\mathcal{X}) = \{x \in \mathbb{R}^{|\mathcal{X}|} | 1^T x = 1, x(i) > 0, 1 \leq i \leq |\mathcal{X}|\}$. For every $\Omega \subset \{1, \dots, |\mathcal{Y}|\}$, $|\Omega| = |\mathcal{X}|$, $\mathcal{H}_{XY}$ is the set of all leakage matrices defined as follows

$$\mathcal{H}_{XY} = \{P_{X|Y} \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|} | \text{if } t_\Omega \in \mathcal{P}(\mathcal{X}) \Rightarrow t_\Omega \in \mathcal{P}'(\mathcal{X})\},$$

where $t_\Omega = M_\Omega^{-1} M P_Y$. In other words, for any $P_{X|Y} \in \mathcal{H}_{XY}$, if t_Ω is a probability vector, then it consists only of nonzero elements for every Ω .

Example 1. Let $P_{X|Y} = \begin{bmatrix} 0.3 & 0.8 & 0.5 \\ 0.7 & 0.2 & 0.5 \end{bmatrix}$. By using SVD of $P_{X|Y}$, M can be found as $\begin{bmatrix} -0.5556 & -0.6016 & 0.574 \\ 0.6742 & -0.73 & 0.1125 \end{bmatrix}$. Possible sets Ω are $\Omega_1 = \{1, 2\}$, $\Omega_2 = \{1, 3\}$, and $\Omega_3 = \{2, 3\}$. By calculating $M_{\Omega_i}^{-1} M P_Y$ we have $M_{\Omega_1}^{-1} M P_Y = \begin{bmatrix} 0.7667 \\ 0.2333 \end{bmatrix}$, $M_{\Omega_2}^{-1} M P_Y = \begin{bmatrix} 0.4167 \\ 0.5833 \end{bmatrix}$, $M_{\Omega_3}^{-1} M P_Y = \begin{bmatrix} -0.2778 \\ 1.2778 \end{bmatrix}$. Since the elements of the first two probability vectors are positive, $P_{X|Y} \in \mathcal{H}_{XY}$.

Remark 3. For $P_{X|Y} \in \mathcal{H}_{XY}$, $M_\Omega M P_Y > 0$ implies $M_\Omega^{-1} M P_Y + \epsilon M_\Omega^{-1} M \begin{bmatrix} P_{X|Y_1} J_u \\ 0 \end{bmatrix} > 0$, furthermore, if the vector $M_\Omega M P_Y$ contains a negative element, the vector $M_\Omega^{-1} M P_Y + \epsilon M_\Omega^{-1} M \begin{bmatrix} P_{X|Y_1} J_u \\ 0 \end{bmatrix}$ contains a negative element as well (i.e., not a feasible distribution), since we assumed ϵ is sufficiently small.

From Proposition 3 and Remark 3, we conclude that each basic feasible solution of $\mathbb{S}_u$, i.e., $V_{\Omega_u}^*$, can be written as summation of a standard probability vector (built by $M_\Omega^{-1} M P_Y$) and a perturbation vector (built by $\epsilon M_\Omega^{-1} M \begin{bmatrix} P_{X|Y_1} J_u \\ 0 \end{bmatrix}$). This is the key property to locally approximate $H(Y|U)$.

Remark 4. The number of basic feasible solutions of $\mathbb{S}_u$ is at most $\binom{|\mathcal{Y}|}{|\mathcal{X}|}$, thus we have at most $\binom{|\mathcal{Y}|}{|\mathcal{X}|}^{|\mathcal{Y}|}$ optimization problems with variables $P_U(\cdot)$ and J_u .

B. Quadratic Optimization Problem

In this part, we approximate $H(Y|U)$ for an obtained basic feasible solution which leads to the new quadratic problem. In the following we use the Bachmann-Landau notation where $o(\epsilon)$ describes the asymptotic behaviour of a function $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ which satisfies that $\frac{f(\epsilon)}{\epsilon} \rightarrow 0$ as $\epsilon \rightarrow 0$.

Lemma 4. Assume $P_{X|Y} \in \mathcal{H}_{XY}$ and $V_{\Omega_u}^*$ is an extreme point of the set $\mathbb{S}_u$, then for $P_{Y|U=u} = V_{\Omega_u}^*$ we have

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

with $b_u = l_u (M_{\Omega_u}^{-1} M P_Y)$, $a_u = l_u (M_{\Omega_u}^{-1} M (1:|\mathcal{X}|) P_{X|Y_1}^{-1}) \in \mathbb{R}^{1 \times |\mathcal{X}|}$, $l_u = [\log(M_{\Omega_u}^{-1} M P_Y(i))]_{i=1:|\mathcal{X}|} \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|}$, and $M_{\Omega_u}^{-1} M P_Y(i)$ stands for i -th ($1 \leq i \leq |\mathcal{X}|$) element of the vector $M_{\Omega_u}^{-1} M P_Y$. Furthermore, $M(1:|\mathcal{X}|)$ stands for submatrix of M with first $|\mathcal{X}|$ columns.

Proof: The result corresponds to first order term of Taylor expansion of $\log(1+x)$. For more details see [13]. ■

In the following we use the notation $f(x) \cong g(x)$ which denotes $f(x) = g(x) + o(x)$, i.e., $g(x)$ is the first order term of the Taylor expansion of $f(x)$.

Theorem 2. Let $P_{X|Y} \in \mathcal{H}_{XY}$ and $V_{\Omega_u}^* \in \mathbb{S}_u^*$, $u \in \{1, \dots, |\mathcal{Y}|\}$. For sufficiently small ϵ , the minimization problem in (6) can be approximated as follows

$$\min_{P_U(\cdot), \{J_u, u \in \mathcal{U}\}} - \left(\sum_{u=1}^{|\mathcal{Y}|} P_u b_u + \epsilon P_u a_u J_u \right) \quad (8)$$

subject to:

$$\begin{aligned} \sum_{u=1}^{|\mathcal{Y}|} P_u V_{\Omega_u}^* &= P_Y, \quad \sum_{u=1}^{|\mathcal{Y}|} P_u J_u = 0, \quad P_U(\cdot) \geq 0, \\ \sum_{i=1}^{|\mathcal{X}|} |J_u(i)| &\leq 1, \quad \sum_{i=1}^{|\mathcal{X}|} J_u(i) = 0, \quad \forall u \in \mathcal{U}. \end{aligned}$$

Proof: The proof is based on Proposition 2 and Lemma 4. For $P_{Y|U=u} = V_{\Omega_u}^*$, $u \in \{1, \dots, |\mathcal{Y}|\}$, $H(Y|U)$ can be approximated as $H(Y|U) = \sum_u P_u H(P_{Y|U=u}) \cong \sum_{u=1}^{|\mathcal{Y}|} P_u b_u + \epsilon P_u a_u J_u$, where b_u and a_u are defined as in Lemma 4. ■

Remark 5. The weights P_u , which satisfy the constraints $\sum_{u=1}^{|\mathcal{Y}|} P_u V_{\Omega_u}^* = P_Y$ and $P_U(\cdot) \geq 0$, form a standard probability distribution, since the sum of elements in each vector $V_{\Omega_u}^* \in \mathbb{S}_u^*$ equals to one.

Remark 6. If we set $\epsilon = 0$, (8) becomes the same linear programming problem as presented in [7].

Proposition 4. All constraints in (8) are feasible.

Proof: Let $J_u = 0$ for all $u \in \mathcal{U}$. In this case all sets $\mathbb{S}_u$ become the same sets named $\mathbb{S}$. Since the set $\mathbb{S}$ is an at most $|\mathcal{Y}| - 1$ dimensional polytope and $P_Y \in \mathbb{S}$, P_Y can be written as convex combination of some extreme points in $\mathbb{S}$. Thus, the constraint $\sum_{u=1}^{|\mathcal{Y}|} P_u V_{\Omega_u}^* = P_Y$ is feasible. Furthermore, by choosing $J_u = 0$ for all $u \in \mathcal{U}$, all other constraints in (8) are feasible too. ■

In next section we show that (8) can be converted into a linear programming problem.

C. Equivalent linear programming problem

Let $\eta_u = P_u M_{\Omega_u}^{-1} M P_Y + \epsilon M_{\Omega_u}^{-1} M (1:|\mathcal{X}|) P_{X|Y_1}^{-1} (P_u J_u)$ for all $u \in \mathcal{U}$, where $\eta_u \in \mathbb{R}^{|\mathcal{X}|}$. η_u corresponds to multiple of non-zero elements of the extreme point $V_{\Omega_u}^*$. P_u and J_u can be found as follows

$$\begin{aligned} P_u &= \mathbf{1}^T \cdot \eta_u, \\ J_u &= \frac{P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u} [\eta_u - (\mathbf{1}^T \eta_u) M_{\Omega_u}^{-1} M P_Y]}{\epsilon (\mathbf{1}^T \cdot \eta_u)}, \end{aligned}$$

Thus, (8) can be rewritten as a linear programming problem using η_u as follows. For the cost function it can be shown

$$-\left(\sum_{u=1}^{|\mathcal{Y}|} P_u b_u + \epsilon P_u a_u J_u\right) = -\sum_u b_u (\mathbf{1}^T \eta_u) - \epsilon \sum_u a_u [P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u} [\eta_u - (\mathbf{1}^T \eta_u) M_{\Omega_u}^{-1} M P_Y]],$$

which is a linear function of elements of η_u for all $u \in \mathcal{U}$. Non-zero elements of the vector $P_u V_{\Omega_u}^*$ are equal to the elements of η_u , i.e., we have $P_u V_{\Omega_u}^*(\omega_i) = \eta_u(i)$. Thus, the constraint $\sum_{u=1}^{|\mathcal{Y}|} P_u V_{\Omega_u}^* = P_Y$ can be rewritten as a linear function of elements of η_u . The constraints $\sum_{u=1}^{|\mathcal{Y}|} P_u J_u = 0$ and $P_u \geq 0 \forall u$ as well as $\sum_{i=1}^{|\mathcal{X}|} J_u(i) = 0$ are reformulated as $\sum_u P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u} [\eta_u - (\mathbf{1}^T \cdot \eta_u) M_{\Omega_u}^{-1} M P_Y] = 0$, $\sum_i \eta_u(i) \geq 0$, $\forall u$ and $\mathbf{1}^T P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u} [\eta_u - (\mathbf{1}^T \cdot \eta_u) M_{\Omega_u}^{-1} M P_Y] = 0$, respectively. Furthermore, the last constraint, i.e., $\sum_{i=1}^{|\mathcal{X}|} |J_u(i)| \leq 1$, can be rewritten as

$$\sum_i |(P_{X|Y_1} M(1:|\mathcal{X}|)^{-1} M_{\Omega_u} [\eta_u - (\mathbf{1}^T \eta_u) M_{\Omega_u}^{-1} M P_Y])(i)| \leq \epsilon (\mathbf{1}^T \eta_u), \forall u.$$

Thus, all constraints can be rewritten as linear functions of elements of η_u for all u .

IV. CONCLUSION

It has been shown that the information-theoretic disclosure control problem can be decomposed if the leakage matrix is of full row rank and fat which belongs to $\mathcal{H}_{XY}$. Furthermore, information geometry can be used to approximate $H(Y|U)$, allowing us to simplify the optimization problem for sufficiently small ℓ_1 -privacy leakage. In particular, the new optimization problem can be transferred to a linear programming problem with details as found in [13].

APPENDIX A

For the first claim it is sufficient to show that for any $\gamma \in \mathbb{S}_u$, we have $1^T \gamma = 1$. Since $\gamma \in \mathbb{S}_u$, we have $M(\gamma - P_Y - \epsilon \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}) = 0$ and from Lemma 1 we obtain $1^T(\gamma - P_Y - \epsilon \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}) = 0$ which yields $1^T \gamma = 1$. In the last conclusion we used the fact that $1^T \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix} = 0$, which is true since by using (2) we have $1_{|\mathcal{Y}|}^T \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix} = 1_{|\mathcal{X}|}^T P_{X|Y_1}^{-1} J_u \stackrel{(a)}{=} 1^T J_u = 0$, where (a) comes from $1_{|\mathcal{X}|}^T P_{X|Y_1}^{-1} = 1_{|\mathcal{X}|}^T \Leftrightarrow 1_{|\mathcal{X}|}^T P_{X|Y_1} = 1_{|\mathcal{X}|}^T$ and noting that columns of $P_{X|Y_1}$ are vector distributions. If $P_{Y|U=u} \in \mathbb{S}_u$ for all $u \in \mathcal{U}$ then we can have $X - Y - U$ and

$$\begin{aligned} M(P_{Y|U=u} - P_Y - \epsilon \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}) &= 0, \\ \Rightarrow P_{X|Y}(P_{Y|U=u} - P_Y - \epsilon \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}) &= 0, \\ \Rightarrow P_{X|U=u} - P_X &= \epsilon \cdot P_{X|Y} \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix} = \epsilon \cdot J_u, \end{aligned}$$

where in last line we use Markovity of $X - Y - U$. Furthermore, J_u satisfies (2), (3), and (4) for all $u \in \mathcal{U}$, thus, the privacy criterion defined in (1c) holds.

APPENDIX B

Consider the set $\mathbb{S} = \{y \in \mathbb{R}^{|\mathcal{Y}|} | My = M P_Y\}$. Any element in $\mathbb{S}$ has sum elements equal to one since we have $M(y - P_Y) = 0 \Rightarrow y - P_Y \in \text{Null}(P_{X|Y})$ and from Lemma 1 we obtain $1^T(y - P_Y) = 0 \Rightarrow 1^T y = 1$. The basic solutions of $\mathbb{S}$ are W_{Ω}^* defined as follows. Let $\Omega = \{\omega_1, \dots, \omega_{|\mathcal{X}|}\}$, where $\omega_i \in \{1, \dots, |\mathcal{Y}|\}$. Then

$$W_{\Omega}^*(\omega_i) = M_{\Omega}^{-1} M P_Y(i),$$

other elements of W_{Ω}^* are zero. Thus, the sum over all elements of $M_{\Omega}^{-1} M P_Y$ equal to one, since each element in $\mathbb{S}$ has sum element equal to one. For the second statement consider the set $\mathbb{S}' = \left\{y \in \mathbb{R}^{|\mathcal{Y}|} | My = M P_Y + \epsilon M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}\right\}$. As argued before basic solutions of $\mathbb{S}'$ are V_{Ω}^* , where

$$V_{\Omega}^*(\omega_i) = (M_{\Omega}^{-1} M P_Y + \epsilon M_{\Omega}^{-1} M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix})(i).$$

Here elements of V_{Ω}^* can be negative or non-negative. From Lemma 2, each element in $\mathbb{S}'$ has sum elements equal to one. Thus, by using the first statement of this proposition sum over all elements of the vector $M_{\Omega}^{-1} M \begin{bmatrix} P_{X|Y_1}^{-1} J_u \\ 0 \end{bmatrix}$ equal to zero.

REFERENCES

- [1] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *2016 Annual Conference on Information Science and Systems (CISS)*, March 2016, pp. 234–239.
- [2] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard, "From the information bottleneck to the privacy funnel," in *2014 IEEE Information Theory Workshop (ITW 2014)*. IEEE, 2014, pp. 501–505.
- [3] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of cryptography conference*. Springer, 2006, pp. 265–284.
- [4] 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.
- [5] 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.
- [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] B. Rassouli and D. Gündüz, "On perfect privacy," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 2551–2555.
- [8] A. Zamani, T. J. Oechtering, and M. Skoglund, "A design framework for epsilon-private data disclosure," *arXiv preprint arXiv:2009.01704*, 3 Sep 2020.
- [9] B. Razeghi, F. P. Calmon, D. Gndz, and S. Voloshynovskiy, "On perfect obfuscation: Local information geometry analysis," in *2020 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2020, pp. 1–6.
- [10] S. Borade and L. Zheng, "Euclidean information theory," in *2008 IEEE International Zurich Seminar on Communications*. IEEE, 2008, pp. 14–17.
- [11] S.-L. Huang and L. Zheng, "Linear information coupling problems," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 1029–1033.
- [12] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.
- [13] A. Zamani, T. J. Oechtering, and M. Skoglund, "Data disclosure mechanism design with non-zero leakage," 2020, <https://people.kth.se/~oech/ITWlong.pdf>.