On Perfect Obfuscation: Local Information Geometry Analysis

Behrooz Razeghi*, Flavio. P. Calmon†, Deniz Gündüz‡, Slava Voloshynovskiy*
*University of Geneva
†Harvard University
‡Imperial College London

***Abstract*—We consider the problem of privacy-preserving data release for a specific utility task under perfect obfuscation constraint. We establish the necessary and sufficient condition to extract features of the original data that carry as much information about a utility attribute as possible, while not revealing any information about the sensitive attribute. This problem formulation generalizes both the information bottleneck and privacy funnel problems. We adopt a local information geometry analysis that provides useful insight into information coupling and trajectory construction of spherical perturbation of probability mass functions. This analysis allows us to construct the modal decomposition of the joint distributions, divergence transfer matrices, and mutual information. By decomposing the mutual information into orthogonal modes, we obtain the locally sufficient statistics for inferences about the utility attribute, while satisfying perfect obfuscation constraint. Furthermore, we develop the notion of perfect obfuscation based on χ^2-divergence and Kullback–Leibler divergence in the Euclidean information space.**

I. Introduction

Releasing an *optimal representation* of data for a given task while simultaneously assuring *privacy* of the individuals' identity and their associated data is one of the main challenges in the information-theory, signal processing, data mining and machine learning communities. An optimal representation is the most useful (sufficient), compressed (compact), and privacy-breaching (minimal) of data. Indeed, the optimal representation of data can be obtained subject to constraints on the target task and its computational and storage complexities.

We investigate the problem of privacy-preserving data release for a specific *utility task* and consider an obfuscation-utility trade-off model where both utility and obfuscation are measured under logarithmic loss. Consider two communication parties, a data owner and a utility service provider. The data owner observes a random variable X and acquires some utility, from the service provider, based on the information he discloses. Simultaneously, the data owner wishes to limit the amount of information revealed about a sensitive random variable S that it depends on X. Therefore, instead of revealing X directly to the service provider, the data owner releases a new representation, denoted by Z. The amount of information leaked to the service provider is measured by $\mathrm{I}(S;Z)$. In particular, the data owner is subjected to a constraint on the information complexity of representation that can be revealed to the service provider. This imposed information complexity is measured by $\mathrm{I}(X;Z)$. Moreover, in general, the utility acquired depends on a utility random variable U that is dependent on X and may be correlated to S. The amount of useful information revealed to the service provider is measured by $\mathrm{I}(U;Z)$. Therefore, considering Markov chain $(U,S) -\circ- X -\circ- Z$, our aim is to share a *sanitized representation* Z of *observed data* X, through a stochastic mapping $\mathbf{P}_{Z|X}$, while preserving information about *utility attribute* U and obfuscate information about *sensitive attribute* S. We called the stochastic mapping $\mathbf{P}_{Z|X}$ the *complexity-constraint obfuscation-utility-assuring mapping*.

Information theoretic (IT) privacy approaches [1]–[25], model and analyze privacy-utility trade-offs using the IT metrics to provide asymptomatic or non-asymptotic privacy-utility-guaranteed frameworks. Inspired from [2], in the most general form, the IT frameworks is based on the knowledge of specific 'private' variable (or data, attribute, information) and correlated non-private variable, and assumption of exact joint distribution or partial statistical knowledge of private and/or non-private data. In this setup, the goal is to design a privacy assuring mapping that transforms the pair of these variables into a new representation that achieves a specific application-based target utility, while simultaneously minimizing the information inferred about the private variable. In many applications, the data X is characterized over large (finite) alphabets while the attribute of interest, i.e., U, is characterized over small (finite) alphabets which results in $\mathrm{I}(U;X) \leq \mathrm{H}(U) \ll \mathrm{H}(X)$.

Focusing on the finite alphabets and considering local information geometry analysis, we develop the notion of perfect obfuscation based on χ^2-divergence and Kullback–Leibler (KL) divergence in the Euclidean information space. Under this analysis, we establish the necessary and sufficient condition to obtain representation Z of the original data X that maximizes the mutual information between utility attribute U and released representation Z, while simultaneously revealing no information about sensitive attribute S. We decompose statistical dependence between random variables U, S, X and Z by decomposing the corresponding mutual information $\mathrm{I}(X;Z)$, $\mathrm{I}(U;Z)$, and $\mathrm{I}(S;Z)$ into orthogonal modes. This model can be viewed as a generalization of two well-known bottleneck models, i.e., Information Bottleneck (IB) and Privacy Funnel (PF).

WIFS'2020, December, 6-9, 2020, New York, USA. XXX-X-XXXX-XXXX-X/XX/$XX.00

Throughout this paper, random variables are denoted by capital letters (e.g. X), deterministic values are denoted by small letters (e.g. x), alphabets (sets) are denoted by Calligraphic fonts (e.g. $\mathcal{X}$). Superscript $(.)^T$ stands for the transpose. For discrete random variable X, let consider a finite support set $\mathcal{X} \triangleq \{1, ..., |\mathcal{X}|\}$ with $2 \leq |\mathcal{X}| < +\infty$. We denote by $\mathcal{P}(\mathcal{X})$ the set of all possible probability distributions of a random variable X with range $\mathcal{X}$. We denote by $\mathbf{p}_X$ the probability mass function (pmf) vector with i-th entry equal to $p_X(i)$. $\mathrm{H}(\mathbf{p}_X) \coloneqq \mathbb{E}_{\mathbf{p}_X}[-\log \mathbf{p}_X]$ denotes Shannon entropy. The relative entropy is defined as $\mathrm{D}_{\mathrm{KL}}(\mathbf{p}_X \| \mathbf{q}_X) \coloneqq \mathbb{E}_{\mathbf{p}_X}\left[\log \frac{\mathbf{p}_X}{\mathbf{q}_X}\right]$.

II. Perfect Information Obfuscation Model

Given the observed data X the defender (data owner) wishes to release a representation Z for a utility task U while keeping another attribute S as sensitive. Let us assume that $\mathbf{P}_{U,S,X}$ is fixed and known by both defender and adversary, and $(U,S) -\!\circ\!- X -\!\circ\!- Z$. We consider the non-interactive, one-shot regime, where the data owner discloses the representation Z once, and no additional information is released. The general objective is to obtain stochastic map $\mathbf{P}_{Z|X} : \mathcal{X} \to \mathcal{Z}$ such that $\mathbf{P}_{U|Z} \approx \mathbf{P}_{U|X}, \forall Z \in \mathcal{Z}, \forall U \in \mathcal{U}, \forall X \in \mathcal{X}$, while $\mathbf{p}_{S|Z=z} \approx \mathbf{p}_S, \forall z \in \mathcal{Z}, \forall S \in \mathcal{S}$. This means that the posterior distribution of the utility attribute U are similar given the released representation Z and original data X, while the posterior distribution of the sensitive attribute S are independent of the released representation Z. One can raises the question whether it is *feasible* that the defender releases a representation Z such that $\mathrm{I}(S;Z) = \mathbb{E}_{\mathbf{p}_Z}\left[\mathrm{D}_{\mathrm{KL}}\left(\mathbf{p}_{S|Z=z} \| \mathbf{p}_S\right)\right] = 0$, i.e., $S \perp\!\!\!\perp Z$, while $\mathrm{I}(U;Z) > 0$, i.e., $U \not\perp\!\!\!\perp Z$. This is a fundamental problem in information-theoretic privacy which is known as data disclosure under *perfect privacy* regime. We will refer to this notion as *perfect information obfuscation*.

To pave our way, let us shortly review the previous models which are specific cases of our model. Consider the Markov chain $U -\!\circ\!- X -\!\circ\!- Z$. This gives us the celebrated Information Bottleneck (IB) problem [26], where $\mathrm{I}(U;Z)$ is referred to as the useful released information (relevance of U) and $\mathrm{I}(X;Z)$ is referred to as the information complexity (description length). The goal of IB model is to find a representation Z of X such that Z is maximally informative about U while being minimally informative about X. We now consider the Markov chain $S -\!\circ\!- X -\!\circ\!- Z$. This gives us the well known Privacy Funnel (PF) problem [8], where $\mathrm{I}(S;Z)$ is referred to as the disclosed sensitive information, and $\mathrm{I}(X;Z)$ is referred to as the useful information. The goal of PF model is to obtain a representation Z of X that minimizes information between sensitive data S and disclosed representation Z while maximizes the amount of information between non-private (useful) data X and disclosed representation Z.

Considering the PF model, the optimal obfuscation-utility coefficient for a given distribution $\mathbf{P}_{S,X}$ is defined as [10]:

$$\nu^*(\mathbf{P}_{S,X}) \coloneqq \inf_{\substack{\mathbf{P}_{Z|X}: \\ S -\!\circ\!- X -\!\circ\!- Z}} \frac{\mathrm{I}(S;Z)}{\mathrm{I}(X;Z)}. \tag{1}$$

They showed that $\nu^*(\mathbf{P}_{S,X})$ is related to the smallest principal component of $\mathbf{P}_{S,X}$, and obtained the necessary and sufficient conditions under which $\nu^*(\mathbf{P}_{S,X}) = 0$. In [17], they studied a similar problem, however, they formulated the objective as of utility maximization under privacy leakage constraint. Hence, the optimal obfuscation-utility coefficient for a given distribution $\mathbf{P}_{S,X}$ is defined as:

$$g_\gamma(\mathbf{P}_{S,X}) = \sup_{\substack{\mathbf{P}_{Z|X}: \\ S -\!\circ\!- X -\!\circ\!- Z \\ \mathrm{I}(S;Z) \leq \gamma}} \mathrm{I}(X;Z), \tag{2}$$

where perfect information obfuscation is said to be feasible if $g_0(\mathbf{P}_{S,X}) > 0$.

We consider the Markov model $(U,S) -\!\circ\!- X -\!\circ\!- Z$ which subsumes both IB and PF objectives. In this case, the functional (2) can be generalized as:

$$g_\gamma(\mathbf{P}_{U,S,X}) = \sup_{\substack{\mathbf{P}_{Z|X}: \\ (U,S) -\!\circ\!- X -\!\circ\!- Z \\ \mathrm{I}(X;Z) \leq R \\ \mathrm{I}(S;Z) \leq \gamma}} \mathrm{I}(U;Z). \tag{3}$$

In particular, we study the necessary and sufficient conditions under which $g_0(\mathbf{P}_{U,S,X}) > 0$ under local information geometry analysis. To this goal, let us define the non-trivial perfect information obfuscation as follows.

Definition 1 (Non-trivial Perfect Information Obfuscation)**.** For a pair of random variables (U,S,X), we say that non-trivial perfect information obfuscation is *feasible* if there exists a random variable Z, that satisfies the following conditions:

1) $(U,S) -\!\circ\!- X -\!\circ\!- Z$ forms a Markov chain.
2) S and Z are independent, i.e., $S \perp\!\!\!\perp Z$.
3) U and Z are not independent, i.e., $U \not\perp\!\!\!\perp Z$.

Note that this definition subsumes the notion of perfect privacy addressed in [17] as well as the notion of weakly independent introduced in [27].

We assume that we observe data X and the distribution $\mathbf{p}_X$ is fixed. Hence, our purpose in non-trivial perfect information obfuscation problem is to construct a trajectory of perturbed pmfs such that a change along that direction changes $\mathbf{p}_U$, while keeps $\mathbf{p}_S$ unchanged. We establish the necessary and sufficient condition for the *existence* of $\mathrm{I}(U;Z) > 0$, under a perfect obfuscation regime.

Lemma 1. Without loss of optimality we can restrict the size of $\mathcal{Z}$ in (3) to $|\mathcal{Z}| \leq |\mathcal{X}| + 2$.

Proof. The proof is based on Fenchel–Eggleston strengthening of Carathéodory's Theorem [28]. □

III. Local Information Geometry Analysis

To get insight into the trajectory construction, we adopt the local information geometry analysis [29]–[33] that provides geometrically appealing interpretation. Consider any reference pmf $\mathbf{p}_X \in \mathcal{P}^\circ(\mathcal{X})$ in the relative interior of the probability simplex in $\mathbb{R}^{|\mathcal{X}|}$, where $\mathcal{P}^\circ(\mathcal{X}) \triangleq \{\mathbf{p}_X \in \mathcal{P}(\mathcal{X}) : p_X(x) > 0, \forall x \in \mathcal{X}\}$ denotes the relative interior of $\mathcal{P}(\mathcal{X})$. Consider a perturbed pmf $\mathbf{r}_X^{(\epsilon)} = \mathbf{p}_X + \epsilon\, \mathbf{h}_X \in \mathcal{P}(\mathcal{X})$ from $\mathbf{p}_X$, for some

small[1] value ϵ, where $\mathbf{h}_X$ is an additive perturbation vector of dimension $|\mathcal{X}|$, satisfying $\sum_x h_X(x) = 0$. The second order Taylor expansion of KL divergence can be written as:

$$\mathrm{D_{KL}}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) = -\sum_x p_X(x) \log \frac{r_X^{(\epsilon)}(x)}{p_X(x)} \tag{4a}$$

$$= -\sum_x p_X(x) \log\left(1 + \epsilon \, \frac{h_X(x)}{p_X(x)}\right) \tag{4b}$$

$$= \frac{1}{2}\,\epsilon^2 \sum_x \frac{1}{p_X(x)} h_X^2(x) + o\left(\epsilon^2\right) \tag{4c}$$

$$= \mathrm{D}_{\chi^2}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) + o\left(\epsilon^2\right). \tag{4d}$$

where $o\left(\epsilon^2\right)$ denotes the Bachmann-Landau asymptotic little-o notation[2], and $\mathrm{D}_{\mathcal{X}^2}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right)$ denotes χ^2-divergence between $\mathbf{p}_X$ and $\mathbf{r}_X^{(\epsilon)}$, defined as follows:

$$\mathrm{D}_{\chi^2}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) \triangleq \sum_{x \in \mathcal{X}} \frac{\left(p_X(x) - r_X^{(\epsilon)}(x)\right)^2}{p_X(x)}. \tag{5}$$

Considering (4c), one can view $\sum_x h_X^2(x)/p_X(x)$ as a weighted norm square of the perturbation vector $\mathbf{h}_X$, i.e., KL divergence is locally a weighted Euclidean metric[3]. Note that, in general, $\mathrm{D_{KL}}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) \neq \mathrm{D_{KL}}\left(\mathbf{r}_X^{(\epsilon)} \| \, \mathbf{p}_X\right)$, however, these divergences are equal up to the first order approximations, i.e., they are locally symmetric. Since by replacing the weights $p_X(x)$ in this norm by any other distribution in the neighborhood, the first order approximation remains the same. Therefore, we have $\mathrm{D_{KL}}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) = \mathrm{D_{KL}}\left(\mathbf{r}_X^{(\epsilon)} \| \, \mathbf{p}_X\right) + o\left(\epsilon^2\right)$. This means that they resemble the standard Euclidean metric within a local neighborhood of pmfs around a reference pmf (i.e., from the center of the local neighborhood) in $\mathcal{P}^\circ(\mathcal{X})$.

We now go one step further and instead of additive perturbation $\mathbf{r}_X^{(\epsilon)} = \mathbf{p}_X + \epsilon \ \mathbf{h}_X$, define the spherical perturbations for our analysis. Consider any reference pmf $\mathbf{p}_X \in \mathcal{P}^\circ(\mathcal{X})$, and any other pmf $\mathbf{r}_X \in \mathcal{P}(\mathcal{X})$. We can define the *spherical perturbation* vector of $\mathbf{r}_X$ from $\mathbf{p}_X$ as $\mathbf{k}_X \triangleq (\mathbf{r}_X - \mathbf{p}_X)\,\mathsf{diag}\left(\sqrt{\mathbf{p}_X}\right)^{-1}$, where $\sqrt{\mathbf{p}_X}$ denotes the entry-wise square root of $\mathbf{p}_X$, and $\mathsf{diag}\left(\sqrt{\mathbf{p}_X}\right)$ denotes a diagonal matrix with principal entries equal to $\sqrt{\mathbf{p}_X}$. Now, we can construct a trajectory of spherically perturbation pmfs as follows:

$$\mathbf{r}_X^{(\epsilon)} = \mathbf{p}_X + \epsilon \, \mathbf{k}_X \, \mathsf{diag}(\sqrt{\mathbf{p}_X}) \tag{6a}$$

$$= (1-\epsilon)\,\mathbf{p}_X + \epsilon\,\mathbf{r}_X, \tag{6b}$$

where $\epsilon \in (0,1)$ controls closeness of $\mathbf{r}_X^{(\epsilon)}$ and $\mathbf{p}_X$. The second equation expresses $\mathbf{r}_X^{(\epsilon)}$ as a convex combination of $\mathbf{p}_X$ and $\mathbf{r}_X$. Note that $\mathbf{k}_X$ in (6a) is a normalized perturbation vector and provides the direction of our trajectory. Furthermore, considering the constraint $\sum_x h_X(x) = 0$ we can verify that $\mathbf{k}_X$ in (6a) satisfies the orthogonality constraint $(\mathrm{C1}) : \mathbf{k}_X^T \sqrt{\mathbf{p}_X} = 0$. Finally, we can rewrite the quadratic approximation of KL divergence as a scaled Euclidean norm of $\mathbf{k}_X$. We have:

$$\mathrm{D_{KL}}\left(\mathbf{p}_X \| \, \mathbf{r}_X^{(\epsilon)}\right) = \frac{1}{2}\,\epsilon^2 \, \|\mathbf{k}_X\|_2^2 + o\left(\epsilon^2\right). \tag{7}$$

Note that using this local approximation we can construct inner products as well as orthogonal perturbations and projections in the Euclidean space.

Remark 1. In [27], the authors defined the notion of *weakly independence* for a pair of random variables $(S, X) \in \mathcal{S} \times \mathcal{X}$ $(|\mathcal{S}|, |\mathcal{X}| < \infty)$ as existence of a random variable Z such that: (i) $S - \circ - X - \circ - Z$ forms a Markov chain, (ii) S and Z are independent, and (iii) X and Z are not independent. They showed that such a random variable Z exists if and only if the columns of $\mathbf{P}_{S|X}$ are linearly dependent. Inspired by this notion of weakly independent, the authors in [17], [24] carefully studied and analyzed perfect obfuscation problem where the goal is to release the useful information X while keeping S as private. Here we extend both, and establish the notion of weakly dependence based on KL-divergence and χ^2-divergence.

Using the local information approximation, we can write the conditional distributions $\mathbf{p}_{X|Z=z}$ as perturbation of $\mathbf{p}_X$, i.e., we have:

$$\mathbf{p}_{X|Z=z} = \mathbf{p}_X + \epsilon \, \mathbf{k}_{X|z} \, \mathsf{diag}(\sqrt{\mathbf{p}_X}). \tag{8}$$

We just need to ensure that $\mathbf{p}_{X|Z=z}$, for different values z, be a valid probability distribution and satisfy the marginal constraints. Hence, we additionally required $(\mathrm{C2}) : \sum_z p_Z(z)\, k_{X|z}(x) \sqrt{p_X(x)} = 0, \forall x \in \mathcal{X}$ which guarantees that marginal pmf of X is preserved, i.e., $\sum_z p_Z(z) p_{X|Z=z} = p_X$. Therefore, our purpose in non-trivial perfect obfuscation problem under local information geometry analysis is to design the latent distribution $\mathbf{p}_Z$ and the conditional distributions $\mathbf{p}_{X|Z=z}$, for different values of z, such that: (i) the constraints $(\mathrm{C1})$ and $(\mathrm{C2})$ are satisfied, (ii) $S \perp\!\!\!\perp Z$, and (iii) $U \not\perp\!\!\!\perp Z$.

Proposition 1. For perfect obfuscation data released model $(U, S) - \circ - X - \circ - Z$ under local information geometry analysis, the non-trivial perfect obfuscation is feasible if and only if for all $z \in \mathcal{Z}$ we simultaneously have:

$$\mathbf{W}_S \ \mathbf{k}_{X|z} \, \mathsf{diag}(\sqrt{\mathbf{p}_X}) = \mathbf{0}, \tag{9a}$$

$$\mathbf{W}_U \ \mathbf{k}_{X|z} \, \mathsf{diag}(\sqrt{\mathbf{p}_X}) \neq \mathbf{0}, \tag{9b}$$

where $\mathbf{W}_S \coloneqq \mathbf{P}_{S|X} : \mathcal{X} \to \mathcal{S}$ and $\mathbf{W}_U \coloneqq \mathbf{P}_{U|X} : \mathcal{X} \to \mathcal{U}$ are fixed probability transition kernels, with dimension $|\mathcal{S}| \times |\mathcal{X}|$ and $|\mathcal{U}| \times |\mathcal{X}|$, respectively.

Proof. To ensure perfect obfuscation, we need $\mathbf{p}_{S|Z=z} = \mathbf{p}_S, \forall S \in \mathcal{S}, z \in \mathcal{Z}$. Considering the Markov chain $S - \circ - X - \circ - Z$, we have:

$$\mathbf{p}_{S|Z=z} \ = \ \mathbf{W}_S \, \mathbf{p}_{X|Z=z} = \mathbf{W}_S \ \mathbf{p}_X + \epsilon \, \mathbf{W}_S \ \mathbf{h}_{X|z}$$

$$= \ \mathbf{p}_S + \epsilon \ \mathbf{W}_S \ \mathbf{k}_{X|z} \, \mathsf{diag}(\sqrt{\mathbf{p}_X}), \ \forall z \in \mathcal{Z}. \tag{10}$$

[1] We assume that $\epsilon \neq 0$ is small enough such that $\mathbf{r}_X^{(\epsilon)}$ is a valid pmf. Note that for larger values of ϵ it may not be entry-wise non-negative.

[2] $\lim_{\epsilon \to 0} o\left(\epsilon^2\right)/\epsilon^2 = 0$

[3] Note that all the well-defined f-divergences are locally equivalent to χ^2-divergence measure to within a constant scale factor. Moreover, note that they locally behave like a Fisher information metric on the statistical manifold.

Therefore, $S \perp\!\!\!\perp Z$, if and only if $\mathbf{W}_S\ \mathbf{k}_{X|z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big) = \mathbf{0}, \forall z \in \mathcal{Z}$. Analogously, considering the Markov chain $U - \circ - X - \circ - Z$, we have:

$$\begin{aligned}\mathbf{p}_{U|Z=z} &= \mathbf{W}_U\,\mathbf{p}_{X|Z=z} = \mathbf{W}_U\ \mathbf{p}_X + \epsilon\,\mathbf{W}_U\ \mathbf{h}_{X|z} \\ &= \mathbf{p}_U + \epsilon\,\mathbf{W}_U\ \mathbf{k}_{X|z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big),\ \forall z \in \mathcal{Z}. \quad (11)\end{aligned}$$

Hence, if we can find the perturbation direction such that $\mathbf{W}_S\ \mathbf{k}_{X|z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big) = \mathbf{0}$ and $\mathbf{W}_U\ \mathbf{k}_{X|z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big) \neq \mathbf{0}$, for some $z \in \mathcal{Z}$, the non-trivial solution, i.e., $\mathrm{I}(U;Z) > 0$, is possible. Conversely, we have a non-trivial solution only if there exists a random variable Z and a valid perturbation vector $\mathbf{k}_{X|z}$ such that a change along that direction changes $\mathbf{p}_U$, while keeping $\mathbf{p}_S$ unchanged. This implies (9). □

Definition 2 (Divergence Transfer Matrix)**.** Given the random variables $X \in \mathcal{X}$ and $Z \in \mathcal{Z}$ with joint pmf $\mathbf{P}_{X,Z} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z})$, with conditional pmfs $\mathbf{P}_{X|Z} \in \mathcal{P}(\mathcal{X} \mid \mathcal{Z})$ and marginal pmfs satisfying $\mathbf{p}_X \in \mathcal{P}^\circ(\mathcal{X})$ and $\mathbf{p}_Z \in \mathcal{P}^\circ(\mathcal{X})$, the divergence transfer matrix associated with $\mathbf{P}_{X,Z}$ is defined as follows:

$$\begin{aligned}\mathbf{B}_{X,Z} = \mathbf{B}(\mathbf{P}_{X,Z}) &\triangleq \mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big)^{-1}\ \mathbf{P}_{X,Z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_Z}\big)^{-1} \\ &= \mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big)^{-1}\ \mathbf{P}_{X|Z}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_Z}\big). \quad (12)\end{aligned}$$

Note that based on the above definition $\mathbf{B}_{X,Z}^T = \mathsf{diag}\big(\sqrt{\mathbf{p}_Z}\big)^{-1}\ \mathbf{P}_{Z|X}\,\mathsf{diag}\big(\sqrt{\mathbf{p}_X}\big)$. We now express the Singular Value decomposition (SVD) of $\mathbf{B}_{X,Z}$ as:

$$\mathbf{B}_{X,Z} = \sum_{i=1}^{K} \sigma_i^{XZ}\,\boldsymbol{\psi}_i^Z\,(\boldsymbol{\psi}_i^Z)^T, \quad (13)$$

where $K \coloneqq \min\{|\mathcal{X}|, |\mathcal{Z}|\}$, σ_i^{XZ} denotes the i-th singular value, and where $\boldsymbol{\psi}_i^Z$ and $\boldsymbol{\psi}_i^X$ are the corresponding left (output) and right (input) singular vectors. By convention, suppose that $\sigma_1^{XZ} \geq \sigma_2^{XZ} \geq \cdots \geq \sigma_K^{XZ}$. Likewise consider SVD of $\mathbf{B}_{U,X}$, $\mathbf{B}_{S,X}$, $\mathbf{B}_{U,Z}$ and $\mathbf{B}_{S,Z}$.

Proposition 2 (Local Approximation of Information Measures)**.** Under the local approximation conditions, the information complexity $\mathrm{I}(X;Z)$, utility information $\mathrm{I}(U;Z)$, and information leakage $\mathrm{I}(S;Z)$ can recast as:

$$\mathrm{I}(X;Z) = \frac{1}{2}\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\,\mathbf{k}_{X|z}\,\|_2^2 + o\left(\epsilon^2\right) \quad (14a)$$

$$= \frac{1}{2}\left(\|\mathbf{B}_{X,Z}\|_{\mathrm{F}}^2 - 1\right) + o\left(\epsilon^2\right), \quad (14b)$$

$$\mathrm{I}(U;Z) = \frac{1}{2}\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{B}_{U,X}\,\mathbf{k}_{X|z}\|_2^2 + o\left(\epsilon^2\right) \quad (14c)$$

$$= \frac{1}{2}\left(\|\mathbf{B}_{U,Z}\|_{\mathrm{F}}^2 - 1\right) + o\left(\epsilon^2\right), \quad (14d)$$

$$\mathrm{I}(S;Z) = \frac{1}{2}\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{B}_{S,X}\,\mathbf{k}_{X|z}\|_2^2 + o\left(\epsilon^2\right) \quad (14e)$$

$$= \frac{1}{2}\left(\|\mathbf{B}_{S,Z}\|_{\mathrm{F}}^2 - 1\right) + o\left(\epsilon^2\right), \quad (14f)$$

where $\mathbf{B}_{U,X}$ and $\mathbf{B}_{U,Z}$ are defined analogous to (12).

Proof. See Appendix A. □

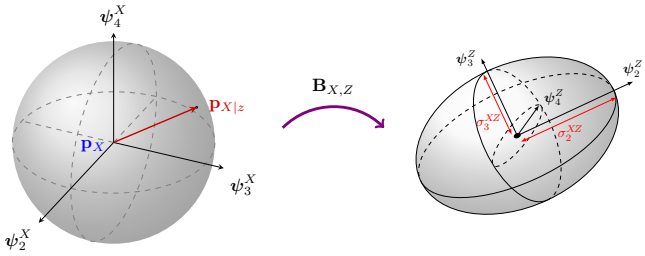

Fig. 1: The information geometry associated with the divergence transfer matrix $\mathbf{B}_{X,Z}$. Visualization for second, third and forth singular vectors, ignoring the first invalid direction. $\mathbf{B}_{X,Z}$ maps a local divergence sphere in $\mathcal{P}(\mathcal{X})$ to a local divergence ellipsoid in $\mathcal{P}(\mathcal{Z})$.

The local approximation (14) gives a nice geometric interpretation. Consider a local divergence sphere in $\mathcal{P}(\mathcal{X})$ constructed as (8). The divergence transfer matrix $\mathbf{B}_{X,Z}$ maps a local divergence sphere in $\mathcal{P}(\mathcal{X})$ to a local divergence ellipsoid in $\mathcal{P}(\mathcal{Z})$ (Fig. 1). Noting that the Markov chain $U - \circ - X - \circ - Z$ implies:

$$p_U(u) = \sum_{x\in\mathcal{X}} p_{U|X}(u\,|\,x)\,p_X(x), \quad (15a)$$

$$p_{U|Z}(u\,|\,z) = \sum_{x\in\mathcal{X}} p_{U|X}(u\,|\,x)\,p_{X|Z}(x\,|\,z). \quad (15b)$$

Analogously, consider the likewise relation for $p_S(s)$ and $p_{S|Z}(s\,|\,z)$. Therefore, the information perturbation vector $\mathbf{k}_{X|z}$ maps to the associated information perturbation vectors $\mathbf{k}_{U|z} \triangleq \mathbf{B}_{U,X}\mathbf{k}_{X|z}$ and $\mathbf{k}_{S|z} \triangleq \mathbf{B}_{S,X}\mathbf{k}_{X|z}$. In other words, the local geometry of $p_{X|Z}$ in the simplex $\mathcal{P}(\mathcal{X})$ induces a corresponding local geometry for $p_{U|Z}$ and $p_{S|Z}$.

Under the local information approximation (14) and satisfying the constraints (C1) and (C2) of perturbation construction, and neglecting $o\left(\epsilon^2\right)$ terms, the optimization problem under perfect obfuscation constraint can recast as:

$$\max_{\mathbf{p}_Z, \mathbf{p}_{X|Z}} \quad \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{B}_{U,X}\,\mathbf{k}_{X|z}\|_2^2 \quad (16a)$$

$$\text{s.t.} \quad \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{k}_{X|z}\|_2^2 \leq R', \quad (16b)$$

$$\sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{B}_{S,X}\,\mathbf{k}_{X|z}\|_2^2 = 0, \quad (16c)$$

or equivalently, as:

$$\max_{\mathbf{p}_Z, \mathbf{p}_{X|Z}} \|\mathbf{B}_{U,Z}\|_{\mathrm{F}}^2 \quad \text{s.t.} \quad \|\mathbf{B}_{X,Z}\|_{\mathrm{F}}^2 \leq R'',\ \|\mathbf{B}_{S,Z}\|_{\mathrm{F}}^2 = 1, \quad (17)$$

where $R' = 2R/\epsilon^2$ and $R'' = 2R + 1$. Note that the pmf $\mathbf{p}_Z$ does not affect the optimization and can be removed from (16). Hence SVD solves the optimization problem (16) by finding $\mathbf{p}_{X|Z}$. Finally, note that by construction $\mathrm{I}(X;Z) \leq \frac{1}{2}\epsilon^2$. Hence, as long as $R \leq \frac{1}{2}\epsilon^2$, we can relax the associated constraint in local information geometry analysis.

To get insight of the optimization problem (16), let us ignore the constraints (16b) and (16c). Note that based on the constraint (C1) the valid normalized perturbation $\mathbf{k}_{X|z}$ must be orthogonal to $\sqrt{\mathbf{p}_X}$, hence $\sqrt{\mathbf{p}_X}$ (the right singular vector of $\mathbf{B}_{U,X}$ corresponding to the largest singular value) is an invalid direction to perturb pmf. Letting σ_2^{UX} be the second largest singular value of $\mathbf{B}_{U,X}$, we have $\|\mathbf{B}_{U,X}\,\mathbf{k}_{X|z}\|^2 \leq (\sigma_2^{\mathrm{UX}})^2\,\|\mathbf{k}_{X|z}\|^2$. Therefore, under this assumption, the optimal solution to (16a) is to choose the perturbation $\mathbf{k}_{X|z}$ to be along the right singular vector of $\mathbf{B}_{U,X}$ corresponding to the second largest singular value. Note that all the unit norm right singular vectors of $\mathbf{B}_{U,X}$ which are orthogonal to $\sqrt{\mathbf{p}_X}$ are a valid perturbation. This means that any linear combination of these singular vectors also valid candidates for $\{\mathbf{k}_{X|z}, z \in \mathcal{Z}\}$.

Let $\mathsf{Range}\,(\mathbf{B}_{U,X}) = \{\mathbf{B}_{U,X}\,\mathbf{k}_{X|z} \mid \mathbf{k}_{X|z} \in \mathbb{R}^{|\mathcal{X}|}\} \subseteq \mathbb{R}^{|\mathcal{U}|}$ denotes the range-space of $\mathbf{B}_{U,X}$, and $\mathsf{Null}\,(\mathbf{B}_{S,X}) = \{\mathbf{k}_{X|z} \mid \mathbf{B}_{S,X}\mathbf{k}_{X|z} = \mathbf{0}\} \subseteq \mathbb{R}^{|\mathcal{X}|}$ denotes the null-space of $\mathbf{B}_{S,X}$. We now have the following proposition.

Proposition 3. For perfect obfuscation data released model $(U,S) \multimap\!\circ\!\!-\!\!-X \multimap\!\circ\!\!-\!\!-Z$ under local information geometry analysis, the non-trivial perfect information obfuscation is feasible if and only if:

$$\dim\left(\,\mathsf{Range}\,(\mathbf{B}_{U,X}) \cap \mathsf{Null}\,(\mathbf{B}_{S,X})\,\right) > 0. \tag{18}$$

Proof. The proof follows by using Proposition 1 and noting that Divergence Transfer Matrix $\mathbf{B}_{U,X}$ (likewise $\mathbf{B}_{S,X}$) is an equivalent representation for $\mathbf{P}_{U,X}$ (likewise $\mathbf{P}_{S,X}$) and, in turn, $\mathbf{P}_{U|X}$ (likewise $\mathbf{P}_{S|X}$). □

We now relate the solutions of (16) to locally sufficient statistics for inferences about utility attribute U based on Z. Let us consider an arbitrary embedding (feature) $f : \mathcal{X} \to \mathbb{R}$ and let $g : \mathcal{Z} \to \mathbb{R}$ be the embedding (feature) induced by f through conditional expectation with respect to $\mathbf{p}_{X|Z=z}$. We have:

$$g(z) = \mathbb{E}\left[f(X) \mid Z = z\right], \quad z \in \mathcal{Z}. \tag{19}$$

We can recast (19) as:

$$\begin{aligned} g(z) &= \frac{1}{p_Z(z)} \sum_{x\in\mathcal{X}} p_{X,Z}(x,z) f(x) \\ &= \frac{1}{\sqrt{p_Z(z)}} \sum_{x\in\mathcal{X}} B_{X,Z}(x,z)\sqrt{p_X(x)} f(x), \end{aligned} \tag{20}$$

where $B_{X,Z}(x,z) = \frac{p_{X,Z}(x,z)}{\sqrt{p_X(x)}\sqrt{p_Z(z)}}, \forall x \in \mathcal{X}, z \in \mathcal{Z}$ is the (x,z)-th entry of $\mathbf{B}_{X,Z}$. We now define $\xi^X(x) \coloneqq \sqrt{p_X(x)} f(x)$ and $\xi^Z(z) \coloneqq \sqrt{p_Z(z)} g(z)$, $\forall x \in \mathcal{X}, z \in \mathcal{Z}$. Then we can express (20) as:

$$\xi^Z(z) = \sum_{x\in\mathcal{X}} B_{X,Z}(x,z)\,\xi^X(x). \tag{21}$$

The vectors $\boldsymbol{\xi}^X$ and $\boldsymbol{\xi}^Z$ whose x-th and z-th entries are $\xi^X(x), \forall x \in \mathcal{X}$ and $\xi^Z(z), \forall z \in \mathcal{Z}$, respectively, can be referred as feature vectors associated with the feature functions f and g.

According to (13) and proof of proposition 2, we have $B_{X,Z}(x,z) = \sum_{i=1}^{K} \sigma_i^{XZ}\psi_i^X(x)\psi_i^Z(z) = \sqrt{p_X(x)}\sqrt{p_Z(z)} + \sum_{i=2}^{K} \sigma_i^{XZ}\psi_i^X(x)\psi_i^Z(z)$. We now define features $f_i^* : \mathcal{X} \to \mathbb{R}$ and $g_i^* : \mathcal{X} \to \mathbb{R}$, for $i = 2, 3, ..., K$, as follows:

$$f_i^*(x) \coloneqq \frac{\psi_i^X(x)}{\sqrt{p_X(x)}}, \quad g_i^*(z) \coloneqq \frac{\psi_i^Z(z)}{\sqrt{p_Z(z)}}. \tag{22}$$

Hence we have:

$$B_{X,Z}(x,z) = \sqrt{p_X(x)}\sqrt{p_Z(z)}\left(1 + \sum_{i=2}^{K} \sigma_i^{XZ} f_i^*(x) g_i^*(z)\right). \tag{23}$$

Noting that $P_{X,Z}(x,z) = B_{X,Z}(x,z)\sqrt{p_X(x)}\sqrt{p_Z(z)} = p_X(x)p_Z(z)\big(1 + \sum_{i=2}^{K} \sigma_i^{XZ} f_i^*(x) g_i^*(z)\big)$, we have modal decomposition of joint distributions, conditional distributions, and mutual information in terms of feature functions $(f_i^*, g_i^*), i = 2, 3, ..., K$. Hence, the valid perturbation directions in optimization problem (16) give us the corresponding valid feature functions, as well as, associated locally normalized sufficient statistics for inferences about U based on Z, under perfect obfuscation constraint.

IV. Conclusion

Adopting a local information geometry analysis and considering mutual information as both obfuscation and utility measure, we studied a data released mechanism for a given utility task, and under perfect obfuscation constraint. The addressed model subsumes both the Information Bottleneck model and the Privacy Funnel model. We studied the notion of perfect obfuscation based on χ^2-divergence and Kullback–Leibler divergence in the Euclidean information space. Furthermore, we characterized the necessary and sufficient conditions under which a non-trivial solution is feasible.

V. Acknowledgment

The authors of [34] independently and simultaneity addressed a similar analysis to ours. At the same time, both works contain some differences that will be explained in consecutive works of both groups.

Appendix A
Proof of Proposition 2

Proof.

$$\mathrm{I}\,(X;Z) = \sum_z p_Z(z)\,\mathrm{D_{KL}}\left(\mathbf{p}_{X|Z=z} \,\|\, \mathbf{p}_X\right) \tag{24a}$$

$$= \frac{1}{2}\,\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{k}_{X|z}\|_2^2 + o\left(\epsilon^2\right) \tag{24b}$$

$$= \frac{1}{2}\epsilon^2 \sum_{z,x} p_Z(z)\left(\frac{p_{X|Z}(x \mid z) - p_X(x)}{\epsilon\,\sqrt{p_X(x)}}\right)^2 + o\left(\epsilon^2\right) \tag{24c}$$

$$= \frac{1}{2}\sum_{z,x}\left(\frac{p_{X,Z}(x,z) - p_X(x)p_Z(z)}{\sqrt{p_X(x)}\sqrt{p_Z(z)}}\right)^2 + o\left(\epsilon^2\right) \tag{24d}$$

$$= \frac{1}{2}\|\mathbf{B}_{X,Z} - \sqrt{\mathbf{p}_X}\sqrt{\mathbf{p}_Z}^T\|_{\mathrm{F}}^2 + o\left(\epsilon^2\right) \tag{24e}$$

$$= \frac{1}{2}\left(\|\mathbf{B}_{X,Z}\|_{\mathrm{F}}^2 - 1\right) + o\left(\epsilon^2\right), \tag{24f}$$

$$\mathrm{I}\left(U;Z\right) = \sum_{z} p_Z(z)\, \mathrm{D}_{\mathrm{KL}}\left(\mathbf{p}_{U|Z=z} \,\|\, \mathbf{p}_U\right) \tag{25a}$$

$$= \frac{1}{2}\,\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\cdot \left\|\mathsf{diag}(\sqrt{\mathbf{p}_U})^{-1}\mathbf{W}_U\mathsf{diag}(\sqrt{\mathbf{p}_X})\,\mathbf{k}_{X|z}\right\|_2^2 + o\big(\epsilon^2\big) \tag{25b}$$

$$= \frac{1}{2}\,\epsilon^2 \sum_{z\in\mathcal{Z}} p_Z(z)\,\|\mathbf{B}_{U,X}\,\mathbf{k}_{X|z}\|_2^2 + o\left(\epsilon^2\right) \tag{25c}$$

$$= \frac{1}{2}\epsilon^2 \sum_{z,u} p_Z(z)\left(\frac{p_{U|Z}(u\mid z) - p_U(u)}{\epsilon\,\sqrt{p_U(u)}}\right)^2 + o\left(\epsilon^2\right) \tag{25d}$$

$$= \frac{1}{2}\sum_{z,x}\left(\frac{p_{X,Z}(x,z) - p_X(x)p_Z(z)}{\sqrt{p_X(x)}\sqrt{p_Z(z)}}\right)^2 + o\left(\epsilon^2\right) \tag{25e}$$

$$= \frac{1}{2}\|\mathbf{B}_{U,Z} - \sqrt{\mathbf{p}_U}\sqrt{\mathbf{p}_Z}^T\|_{\mathrm{F}}^2 + o\left(\epsilon^2\right) \tag{25f}$$

$$= \frac{1}{2}\left(\|\mathbf{B}_{U,Z}\|_{\mathrm{F}}^2 - 1\right) + o\left(\epsilon^2\right). \tag{25g}$$

The equalities (24f) and (25g) follow by noticing that the largest singular value of divergence transfer matrices $\mathbf{B}_{X,Z}$ and $\mathbf{B}_{U,Z}$ are 1, i.e., their spectral norm is equal to one. Note that $\mathbf{B}_{X,Z}$ and $\mathbf{B}_{U,Z}$ originate from the column stochastic transition matrices of conditional probabilities $\mathbf{P}_{X|Z}$ and $\mathbf{P}_{U|Z}$, respectively. Therefore, the corresponding right (input) singular vectors are as follows:

$$\boldsymbol{\psi}_1^X = \mathbf{B}_{X,Z}\,\sqrt{\mathbf{p}_Z} = \sigma_1^{\mathrm{XZ}}\sqrt{\mathbf{p}_X} = \sqrt{\mathbf{p}_X}, \tag{26a}$$

$$\boldsymbol{\psi}_1^U = \mathbf{B}_{U,Z}\,\sqrt{\mathbf{p}_Z} = \sigma_1^{\mathrm{UZ}}\sqrt{\mathbf{p}_U} = \sqrt{\mathbf{p}_U}. \tag{26b}$$

The local approximation of information leakage $\mathrm{I}(S;Z)$ derivation follows similar lines as (25). □

REFERENCES

[1] I. S. Reed, "Information theory and privacy in data banks," in *Proceedings of the June 4-8, 1973, national computer conference and exposition*. ACM, 1973, pp. 581–587.

[2] 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.

[3] A. Evfimievski, J. Gehrke, and R. Srikant, "Limiting privacy breaches in privacy preserving data mining," in *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 2003, pp. 211–222.

[4] D. Rebollo-Monedero, J. Forne, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 11, pp. 1623–1636, 2009.

[5] F. du Pin Calmon and N. Fawaz, "Privacy against statistical inference," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2012, pp. 1401–1408.

[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] F. P. Calmon, M. Varia, M. Médard, M. M. Christiansen, K. R. Duffy, and S. Tessaro, "Bounds on inference," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2013, pp. 567–574.

[8] A. Makhdoumi and N. Fawaz, "Privacy-utility tradeoff under statistical uncertainty," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2013, pp. 1627–1634.

[9] S. Asoodeh, F. Alajaji, and T. Linder, "Notes on information-theoretic privacy," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2014, pp. 1272–1278.

[10] F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1796–1800.

[11] S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "Managing your private and public data: Bringing down inference attacks against your privacy," *IEEE Jour. of Selected Topics in Sig. Proc.*, vol. 9, no. 7, pp. 1240–1255, 2015.

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

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

[14] K. Kalantari, L. Sankar, and O. Kosut, "On information-theoretic privacy with general distortion cost functions," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 2865–2869.

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

[16] 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, 2018.

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

[18] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "Privacy under hard distortion constraints," in *IEEE Inf. Theory Workshop (ITW)*, 2018.

[19] H. Hsu, S. Asoodeh, F. du Pin Calmon, and N. Fawaz, "Information-theoretic privacy watchdogs," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019.

[20] J. Liao, O. Kosut, L. Sankar, and F. du Pin Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Tran. on Information Theory*, vol. 65, pp. 8043–8066, 2019.

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

[22] T. Xiao and A. Khisti, "Maximal information leakage based privacy preserving data disclosure mechanisms," in *2019 16th Canadian Workshop on Information Theory (CWIT)*. IEEE, 2019, pp. 1–6.

[23] M. Diaz, H. Wang, F. P. Calmon, and L. Sankar, "On the robustness of information-theoretic privacy measures and mechanisms," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 1949–1978, 2019.

[24] B. Rassouli, F. E. Rosas, and D. Gündüz, "Data disclosure under perfect sample privacy," *IEEE Transactions on Information Forensics and Security*, 2019.

[25] 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, 2019.

[26] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *IEEE Allerton*, 2000.

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

[28] H. G. Eggleston, *Convexity*. CUP Archive, 1958, no. 47.

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

[30] A. Makur, "A study of local approximations in information theory," Master's thesis, Massachusetts Institute of Technology, 2015.

[31] S.-L. Huang, A. Makur, G. W. Wornell, and L. Zheng, "On universal features for high-dimensional learning and inference," *arXiv preprint arXiv:1911.09105*, 2019.

[32] A. Makur, "Information contraction and decomposition," Ph.D. dissertation, Massachusetts Institute of Technology, 2019.

[33] A. Makur, G. W. Wornell, and L. Zheng, "On estimation of modal decompositions," in *IEEE International Symposium on Information Theory (ISIT)*, 2020.

[34] A. Zamani, T. J. Oechtering, and M. Skoglund, "A design framework for epsilon-private data disclosure," *arXiv preprint arXiv:2009.01704*, 2020.