α-leakage Interpretation of Rényi Capacity

Ni Ding
University of Auckland
New Zealand
ni.ding@auckland.ac.nz

Farhad Farokhi
University of Melbourne
Australia
farhad.farokhi@unimelb.edu.au

Tao Guo, Yinfei Xu and Xiang Zhang
Southeast University
China
{taoguo, yinfeixu, xiangzhang369}@seu.edu.cn

***Abstract*—For $\tilde{f}(t) = \exp\big(\frac{\alpha-1}{\alpha}t\big)$, this paper shows that the Sibson mutual information is an α-leakage averaged over the adversary's $\tilde{f}$-mean relative information gain (on the secret) at elementary event of channel output Y as well as the joint occurrence of elementary channel input X and output Y. This interpretation is used to derive a sufficient condition that achieves a δ-approximation of ϵ-upper bounded α-leakage. A Y-elementary α-leakage is proposed, extending the existing pointwise maximal leakage to the overall Rényi order range $\alpha \in [0,\infty)$. Maximizing this Y-elementary leakage over all attributes U of channel input X gives the Rényi divergence. Further, the Rényi capacity is interpreted as the maximal $\tilde{f}$-mean information leakage over both the adversary's malicious inference decision and the channel input X (represents the adversary's prior belief). This suggests an alternating max-max implementation of the existing generalized Blahut-Arimoto method.**

I. Introduction

The quantification of privacy leakage was originally proposed in a Bayesian inference setting [1] as follows. Compute the information on the secret corresponding respectively to the prior belief (before the adversary observes channel output Y) and the posterior belief (after the adversary observes channel output Y). Take the privacy leakage as the increase in the adversary's information on the secret from prior to posterior belief. Such measure ranges from the mutual information [1], [2], maximal leakage in [3] and α-leakage [4], a Rényi measure for order $\alpha \in [1,\infty)$ tunable between mutual information (average leakage) and maximal leakage. In [5], a Rényi cross entropy is proposed, extending the definition of α-leakage to the entire Rényi order range $\alpha \in [0,\infty)$ which well interprets Arimoto mutual information as a privacy risk assessment.

Independently, recent work [6] proposes another α-leakage following the intuitions of Rényi measures in [7]. Instead of separate calculations for prior and posterior, the elementary (instance-wise) relative information gain jointly incurred is obtained, and the $\tilde{f}$-mean[1] of all elementary measures quantifies the information leakage. Here, $\tilde{f}(t) = \exp\big(\frac{\alpha-1}{\alpha}t\big)$. It is shown in [6, Theorem 1] that the maximum of this $\tilde{f}$-mean measure over all adversary's estimation decisions (on the secret) equals to Rényi divergence. This, via the expression [5, Eq. (20)], further interprets Sibson mutual information as a $\tilde{f}$-mean of Y-elementary leakage, the relative information gain incurred at each instance y of channel output Y.

[1]For invertible and continuous f, the f-mean of X is $\bar{X} = f^{-1}(\mathbb{E}[f(X)])$, which is also called quasi-arithmetic mean.

Following [6], we reveal in this paper that the Sibson mutual information is also a $\tilde{f}$-mean of XY-elementary leakage at the joint occurrence (x,y) of channel input X and output Y. Using exponential Chebyshev tail bound, we derive a sufficient condition in terms of Sibson mutual information that guarantees a δ-approximation of both Y- and XY-elementary leakage is upper bounded by ϵ. We reveal a $\tilde{f}$-mean interpretation of α-leakage measure proposed in [5, Definition], [3], [4], and derive the corresponding Y-elementary leakage and show that the maximum of it over all attributes U of channel input X is the α-leakage measured by Rényi divergence. This result extends the existing pointwise maximal leakage (PML) [8] to the overall Rényi order range $\alpha \in [0,\infty)$,[2] where PML corresponds to the case $\alpha = \infty$. We further reveal that the Rényi capacity C_α is the maximal $\tilde{f}$-mean information leakage over both the adversary's estimation decision on secret and the channel input X, which represents the adversary's prior belief. This parallels the existing definition of Rényi capacity as the max-min of conditional Rényi divergence and suggests an alternating max-max implementation of the existing generalized Blahut-Arimoto method [9] for computing C_α.

A. Notation

Random variables (r.v.) are denoted by capital Roman letters while their elementary events, instances, or realizations are denoted by lowercase Roman letters. Sets are denoted by calligraphic Roman letters. For instance, x in an instance or realization of r.v. X and $\mathcal{X}$ refers to its alphabet. We assume finite countable alphabet in this paper. Let $P_X(x)$ be the probability of outcome $X = x$ and $P_X = (P_X(x)\colon x \in \mathcal{X})$ be the probability mass function. The support of P_X is defined as $\text{supp}(P_X) = \{x\colon P_X(x) > 0\}$. The expected value of $f(X)$ with respect to (w.r.t.) probability mass function P_X is $\mathbb{E}_{P_X}[f(X)] = \sum_{x\in\mathcal{X}} P_X(x)f(x)$. The probability distribution of Y conditioned on the event $X = x$ is denoted by $P_{Y|X=x} = (P_{Y|X}(y|x)\colon y \in \mathcal{Y})$. Furthermore, $P_{Y|X} = (P_{Y|X}(y|x)\colon x \in \mathcal{X}, y \in \mathcal{Y})$. For r.v. X, we denote X_α the α-scaled X with probability $P_{X_\alpha}(x) = P_X^\alpha(x)/\sum_x P_X^\alpha(x)$ for all $x \in \mathcal{X}$.

[2]In this paper, for Rényi order range such as $[0,\infty)$, we keep the ∞ side open. But, it should be clear that all results apply to the order $\alpha = \infty$.

II. α-Leakage: Maximum Information Gain

Estimation theory can be used to define information leakage [1], [3], [4], [8]. A privacy-preserving channel can be modeled via conditional probability $P_{Y|X}$. Here, the channel output Y can be accessed by all users, including malicious or adversarial entities, while X is the raw input, which may contain sensitive information. An adversary can construct an estimate of X, denoted by $\hat{X}$, based on the output Y using conditional probability $P_{\hat{X}|Y}$. This induces a Markov chain $X-Y-\hat{X}$. In this context, the adversary's prior belief about the secrete r.v. X is captured by P_X while the adversary's posterior belief is given by $P_{\hat{X}|Y}$ and can be used to measure information gain. The adversary will seek the optimal estimation $P^*_{\hat{X}|Y}$ that maximizes the information gain, which incurs the maximal privacy leakage on the channel input X.

A. Leakage Measure by Divergence

For $f(t)=\exp((\alpha-1)t)$, Rényi divergence is defined in [7] as an f-mean relative information gain as

$$\begin{aligned}&D_\alpha(P_{X|Y=y}\|P_X)\\&=\frac{1}{\alpha-1}\log\sum_x P_{X|Y}(x|y)\exp\Big((\alpha-1)\iota_{P_{X|Y=y}\|P_X}(x)\Big)\end{aligned}$$

where $\iota_{P_{X|Y=y}\|P_X}(x)=\log\frac{P_{X|Y}(x|y)}{P_X(x)}$ measures the information increase in $P_{X|Y}$ from P_X for each elementary event x.

Let P_X be the reference probability when we measure the information gain. A $\tilde{f}$-mean relative information gain measure for $\tilde{f}(t)=\exp\big(\frac{\alpha-1}{\alpha}t\big)$ is proposed in [6] as follows. Let

$$\iota_{P_{\hat{X}|Y=y}\|P_X}(x)=\log\frac{P_{\hat{X}|Y}(x,y)}{P_X(x)} \tag{1}$$

be the elementary information gain incurred by the adversary's estimation decision $P_{\hat{X}|Y}$ relative to P_X. For each x, the appearance frequency of $\iota_{P_{\hat{X}|Y=y}\|P_X}(x)$ is determined by the posterior probability $P_{X|Y}(x|y)$. The $\tilde{f}$-mean of $\iota_{P_{\hat{X}|Y=y}\|P_X}$ is

$$\begin{aligned}&\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\\&\quad=\frac{\alpha}{\alpha-1}\log\sum_x P_{X|Y}(x|y)\exp\bigg(\frac{\alpha-1}{\alpha}\iota_{P_{\hat{X}|Y=y}\|P_X}(x)\bigg)\\&\quad=\frac{\alpha}{\alpha-1}\log\sum_x P_{X|Y}(x|y)\Big(\frac{P_{\hat{X}|Y}(x|y)}{P_X(x)}\Big)^{\frac{\alpha-1}{\alpha}}\end{aligned} \tag{2}$$

for all $\alpha\in(0,1)\cup(1,\infty)$. At extended orders,

$$\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)=\begin{cases}\log\min\limits_x\frac{P_{\hat{X}|Y}(x|y)}{P_X(x)} & \alpha=0,\\ \sum\limits_x P_{X|Y}(x|y)\log\frac{P_{\hat{X}|Y}(x|y)}{P_X(x)} & \alpha=1,\\ \log\sum\limits_x P_{X|Y}(x|y)\frac{P_{\hat{X}|Y}(x|y)}{P_X(x)} & \alpha=\infty.\end{cases}$$

We call $\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)$ the $\tilde{f}$-mean conditional divergence between $P_{\hat{X}|Y=y}$ and P_X given $P_{X|Y=y}$. It is proved in [6, Theorem 1] that the maximization of this $\tilde{f}$-mean information gain gives the f-mean relative information: for all $\alpha\in[0,\infty)$,

$$\begin{aligned}&D_\alpha(P_{X|Y=y}\|P_X)\\&\quad=\max_{P_{\hat{X}|Y=y}}\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big),\quad\forall y\end{aligned} \tag{3}$$

with the maximizer

$$P^*_{\hat{X}|Y}(x|y)=\frac{P^\alpha_{X|Y}(x|y)/P_X^{\alpha-1}(x)}{\sum_x P^\alpha_{X|Y}(x|y)/P_X^{\alpha-1}(x)},\ \forall x,y. \tag{4}$$

There is a proof of (3) in [6] by convex optimization. We will show another proof by a decomposition of the unconditional $\tilde{f}$-mean relative information measure in Section IV-A. The interpretation of (3) is that the Rényi divergence $D_\alpha(P_{X|Y=y}\|P_X)$ quantifies the adversary's maximal information gain on X after observing the channel output outcome $Y=y$ and therefore indicates the maximal leakage from X at each instance y.

B. Maximal Leakage by Sibson Mutual Information

Given channel $P_{Y|X}$ and input distribution P_X, the Sibson mutual information [10]

$$I^{\mathrm{S}}_\alpha(P_X,P_{Y|X})=\frac{\alpha}{\alpha-1}\log\sum_{y\in\mathcal{Y}}\Big(\sum_{x\in\mathcal{X}}P_X(x)P^\alpha_{Y|X}(y|x)\Big)^{\frac{1}{\alpha}}.$$

is originally defined as the information radius of f-mean Rényi divergence.[3] We show below that the Sibson mutual information can be viewed as a maximal leakage measure in terms of $\tilde{f}$-mean relative information gain:[4] for all $\alpha\in[0,\infty)$,

$$\begin{aligned}&I^{\mathrm{S}}_\alpha(P_X;P_{Y|X})\\&=\frac{\alpha}{\alpha-1}\log\mathbb{E}_{P_Y}\big[\exp\big(\frac{\alpha-1}{\alpha}D_\alpha(P_{X|Y=y}\|P_X)\big)\big]\\&=\frac{\alpha}{\alpha-1}\log\mathbb{E}_{P_Y}\big[\exp\big(\frac{\alpha-1}{\alpha}\\&\qquad\max_{P_{\hat{X}|Y=y}}\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\big)\big]\\&=\begin{cases}-\min\limits_{P_{\hat{X}|Y}}\log\max\limits_y\exp\big(-\\\qquad\tilde{D}_0\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\big) & \alpha=0\\ \frac{\alpha}{\alpha-1}\min\limits_{P_{\hat{X}|Y}}\log\mathbb{E}_{P_Y}\big[\exp\big(\frac{\alpha-1}{\alpha}\\\qquad\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\big)\big] & \alpha\in(0,1)\\ \max\limits_{P_{\hat{X}|Y}}\mathbb{E}_{P_Y}\Big[\tilde{D}_1\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\Big] & \alpha=1\\ \frac{\alpha}{\alpha-1}\max\limits_{P_{\hat{X}|Y}}\log\mathbb{E}_{P_Y}\big[\exp\big(\frac{\alpha-1}{\alpha}\\\qquad\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\big)\big] & \alpha\in(1,\infty)\\ \max\limits_{P_{\hat{X}|Y}}\log\mathbb{E}_{P_Y}\big[\exp\big(\\\qquad\tilde{D}_\infty\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\big)\big] & \alpha=\infty\end{cases}\end{aligned} \tag{5}$$

(6)

[3] Information radius, as defined in [10, Section 2], is a probability distribution that minimizes f-mean Rényi divergence from a given set of probabilities.

[4] The equation array (5) to (8) was first presented in [6] without (6), which is a necessary step to prove (7) by separable optimizations.

$$= \max_{P_{\hat{X}|Y}} \frac{\alpha}{\alpha-1} \log \mathbb{E}_{P_Y}\Big[\exp\Big(\frac{\alpha-1}{\alpha} \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)\Big)\Big] \tag{7}$$

$$= \max_{P_{\hat{X}|Y}} \frac{\alpha}{\alpha-1} \log \sum_{x,y} P_{Y|X}(y|x)P_X(x)\Big(\frac{P_{\hat{X}|Y}(x|y)}{P_X(x)}\Big)^{\frac{\alpha-1}{\alpha}} \tag{8}$$

Eq. (5) was first presented in [5, Eqs. (19) and (20)] showing that the Sibson mutual information is the $\tilde{f}$-mean of Rényi divergence. There is a conversion from minimization to maximization in (6). It is because the measure $\exp\big(\frac{\alpha-1}{\alpha}D_\alpha(P_{X|Y=y}\|P_X)\big)$ changes from information loss to information gain when α increases above 1. A similar situation can be found in the proof of [5, Theorem 1]. Eq. (8) actually formulates a $\tilde{f}$-mean of XY-elementary leakage, which will be utilized to derive a tail bound for attaining the δ-approximation of ϵ-leakage in Section III.

C. Elementary (Pointwise) Information Leakage

Independently, a Y-elementary leakage measure, called *pointwise maximal leakage (PML)*,[5] is proposed in [8], quantifying the adversary's maximal information gain incurred at each elementary event of channel output Y. The idea is based on the definition of information gain in [3], [4], [8] related to Arimoto conditional entropy and Arimoto mutual information [11].

While the PML only defines the Y-elementary leakage at $\alpha = \infty$. We propose in this section a generalized measure in the overall Rényi order range $\alpha \in [0,\infty)$. This is done by borrowing the definition of Rényi cross entropy in [5], a measure that interprets Arimoto conditional entropy in privacy leakage, revising the α-leakage in [4] such that it is well defined by the intuition of Rényi measures including order $\alpha = 0$. The resulting measure is a new Y-elementary leakage, relates to Arimoto mutual information, but proved to be upper bounded by the Rényi divergence $D_\alpha(P_{X|Y=y}\|P_X)$.

For U being the sensitive attribute of channel input X, assume the adversary makes a soft decision $P_{\hat{U}|Y}$ to get the estimate $\hat{U}$. The information leakage is quantified on the Markov chain $U - X - Y - \hat{U}$. It can be quantified by the Arimoto mutual information $I_\alpha^{\mathrm{A}}(P_U, P_{Y|U})$: for all $\alpha \in [0,\infty)$,

$$\begin{aligned}\mathcal{L}_\alpha(U\to Y) &= \frac{\alpha}{1-\alpha}\log\mathbb{E}_{P_Y}\Big[\exp\Big(\frac{1-\alpha}{\alpha}\Big(\min_{P_{\hat{U}}} H_\alpha(P_U,P_{\hat{U}}) \\ &\quad - \min_{P_{\hat{U}|Y=y}} H_\alpha(P_{U|Y=y},P_{\hat{U}|Y=y})\Big)\Big)\Big] \\ &= I_\alpha^{\mathrm{A}}(P_U;P_{Y|U}).\end{aligned} \tag{9}$$

Here,

$$H_\alpha(P_U,P_{\hat{U}}) = \frac{\alpha}{1-\alpha}\log\sum_u P_U(u)P_{\hat{U}}^{\frac{\alpha-1}{\alpha}}(u)$$

[5]The word 'pointwise' is used in [8], while we use 'elementary' as it refers to the elementary/atomic event in the sample space.

is the Rényi cross entropy proposed in [5, Eq. (7)] as a $\tilde{f}$-mean uncertainty measure incurred by $P_{\hat{U}}$ w.r.t. the probability P_U. The minimization of this Rényi cross entropy over $P_{\hat{U}}$ gives Rényi entropy: $\min_{P_{\hat{U}}} H_\alpha(P_U,P_{\hat{U}}) = H_\alpha(P_U)$ [5, Theorem 1].

Eq (9) in fact rewrites [5, Definition 1] viewing $\mathcal{L}_\alpha(U\to Y)$ as a $\tilde{f}$-mean leakage measure. Based on this interpretation, we propose a Y-elementary leakage incurred at each channel output instance y as

$$\mathcal{L}_\alpha(U\to y) = \min_{P_{\hat{U}}} H_\alpha(P_U,P_{\hat{U}}) - \min_{P_{\hat{U}|Y=y}} H_\alpha(P_{U|Y=y},P_{\hat{U}|Y=y}) \tag{10}$$

$$= \begin{cases} -\log\frac{\min\limits_{P_{\hat{U}|Y=y}}\max\limits_{u\in\mathrm{supp}(P_{U|Y=y})} P_{\hat{U}|Y}^{-1}(u|y)}{\min\limits_{P_{\hat{U}}}\max\limits_{u\in\mathrm{supp}(P_U)} P_{\hat{U}}^{-1}(u)} & \alpha=0 \\ \frac{\alpha}{\alpha-1}\log\frac{\min\limits_{P_{\hat{U}|Y=y}}\mathbb{E}_{P_{U|Y=y}}\left[P_{\hat{U}|Y}^{\frac{\alpha-1}{\alpha}}(u|y)\right]}{\min\limits_{P_{\hat{U}}}\mathbb{E}_{P_U}\left[P_{\hat{U}}^{\frac{\alpha-1}{\alpha}}(u)\right]} & \alpha\in(0,1) \\ \log\frac{\max\limits_{P_{\hat{U}|Y=y}}\prod\limits_u P_{\hat{U}|Y}(u|y)^{P_{U|Y}(u|y)}}{\max\limits_{P_{\hat{U}}}\prod\limits_u P_{\hat{U}}(u)^{P_U(u)}} & \alpha=1 \\ \frac{\alpha}{\alpha-1}\log\frac{\max\limits_{P_{\hat{U}|Y=y}}\mathbb{E}_{P_{U|Y=y}}\left[P_{\hat{U}|Y}^{\frac{\alpha-1}{\alpha}}(u|y)\right]}{\max\limits_{P_{\hat{U}}}\mathbb{E}_{P_U}\left[P_{\hat{U}}^{\frac{\alpha-1}{\alpha}}(u)\right]} & \alpha\in(1,\infty) \\ \log\frac{\max\limits_{P_{\hat{U}|Y=y}}\mathbb{E}_{P_{U|Y=y}}\left[P_{\hat{U}|Y}(u|y)\right]}{\max\limits_{P_{\hat{U}}}\mathbb{E}_{P_U}\left[P_{\hat{U}}(u)\right]} & \alpha=\infty \end{cases} \tag{11}$$

$$= H_\alpha(P_U) - H_\alpha(P_{U|Y=y}) = \frac{1}{\alpha-1}\log\frac{\sum_u P_{U|Y}^\alpha(u|y)}{\sum_u P_U^\alpha(u)}, \qquad \forall\alpha\in[0,\infty). \tag{12}$$

Eq. (10) quantifies the Y-elementary information leakage as the difference of the best prior and posterior uncertainty reduction at the adversary for each elementary channel output y. In (11), the cases when $\alpha = 0$ and $\alpha = 1$ can be simplified to $\log\frac{\max\limits_{P_{\hat{U}|Y=y}}\min\limits_{u\in\mathrm{supp}(P_{U|Y=y})} P_{\hat{U}|Y}(u|y)}{\max\limits_{P_{\hat{U}}}\min\limits_{u\in\mathrm{supp}(P_U)} P_{\hat{U}}(u)}$ and $\max\limits_{P_{\hat{U}|Y=y}} \mathbb{E}_{P_{U|Y=y}}[\log P_{\hat{U}|Y=y}(\cdot|y)] - \max\limits_{P_{\hat{U}}}\mathbb{E}_{P_U}[\log P_{\hat{U}}(\cdot)]$, respectively. The segmented eq. (11) tries to express each case in terms of multiplicative/fractional posterior vs. prior gain/loss, which coincides with the idea in [3], [4], [8]. Eq. (11) also shows a conversion from minimization to maximization for 1-crossing α, similar to (6). The reason is the change from uncertainty to certainty measure when α crosses 1: for $\alpha > 1$, $P_{\hat{U}}^{\frac{\alpha-1}{\alpha}}(u)$ measures certainty (incurred by $P_{\hat{U}}$); for $\alpha < 1$, $(P_{\hat{U}}^{-1}(u))^{\frac{1-\alpha}{\alpha}}$ measures uncertainty.

Proposition 1: The maximal leakage over all attributes U of X for given P_X and $P_{Y|X}$ over Markov chain $U - X - Y - \hat{U}$ is

$$\sup_{P_{U|X}} \mathcal{L}_\alpha(U\to y) = D_\alpha(P_{X|Y=y}\|P_X), \quad \forall\alpha\in[0,\infty) \tag{13}$$

Proof: Let $P_{U_\alpha}(u) = \frac{P_U^\alpha(u)}{\sum_{u\in} P_U^\alpha(u)}$ for all u and $P_{U_\alpha|Y}(u|y) = \frac{P_{Y|U}(y|u)P_{U_\alpha}(u)}{P_Y(y)}$ for all u, y.[6] Then, $\mathcal{L}_\alpha(U \to y) = D_\alpha(P_{U_\alpha|Y=y}\|P_{U_\alpha})$ and

$$\begin{aligned}\sup_{P_{U|X}} \mathcal{L}_\alpha(U \to y) &= \sup_{P_{U_\alpha|X}} D_\alpha(P_{U_\alpha|Y=y}\|P_{U_\alpha}) \\ &= \sup_{P_{U|X}} D_\alpha(P_{U|Y=y}\|P_U) \qquad (14)\\ &= D_\alpha(P_{X|Y=y}\|P_X), \qquad (15)\end{aligned}$$

where (14) is simply a change of the notation of the decision variable and (15) is because of the data processing inequality of Rényi divergence [12, Theorem 1]. ∎

The so-called pointwise maximal leakage (PML) [8, Definition 1] first defines [8, Eq. (3)]

$$\mathcal{L}_\infty(U \to y) = \log \frac{\max\limits_{P_{\hat{U}|Y=y}} \mathbb{E}_{P_{U|Y=y}}\left[P_{\hat{U}|Y}(u|y)\right]}{\max\limits_{P_{\hat{U}}} \mathbb{E}_{P_U}\left[P_{\hat{U}}(u)\right]} \qquad (16)$$

$$= \log \frac{\max\limits_{P_{\hat{U}|Y=y}} \Pr\Big(\hat{U} = U|Y = y\Big)}{\max\limits_{P_{\hat{U}}} \Pr\Big(\hat{U} = U\Big)}, \qquad (17)$$

the Y-elementary leakage on U at order $\alpha = \infty$; then takes the supremum $\sup_{P_{U|X}} \mathcal{L}_\infty(U \to y)$ and call it PML. It is then proved in [8, Eq. (3)] that PML equals the ∞-order Rényi divergence $\sup_{P_{U|X}} \mathcal{L}_\infty(U \to y) = D_\infty(P_{X|Y=y}\|P_X)$, which is Proposition 1 in the case of $\alpha = \infty$.

Remark 1: $\mathcal{L}_\alpha(U \to y)$ could be negative, which does not disqualify it from being a Y-elementary leakage measure, as the elementary relative information gain, e.g, $\iota_{P_{X|Y=y}\|P_X}$ in Rényi divergence, is not necessarily nonnegative. However, $\sup_{P_{U|X}} \mathcal{L}_\alpha(U \to y) \geq 0$ always due to the nonnegativity of Rényi divergence.

Remark 2: Proposition 1 is consistent with the existing result $I_\alpha^{\mathrm{A}}(P_U; P_{Y|U}) = I_\alpha^{\mathrm{S}}(P_{U_\alpha}; P_{Y|U}) \leq I_\alpha^{\mathrm{S}}(P_X; P_{Y|X})$ [13].

III. δ-Approximation of α-Leakage: Tail Bound

With the $\tilde{f}$-mean expression, we are ready to propose the δ-approximation of α-leakage, similar to the idea of (ϵ, δ)-differential privacy [14].

We first derive an XY-elementary leakage from $D_\alpha(P_{X|Y=y}\|P_X)$. The maximal relative information gain incurred by the optimal decision $P^*_{\hat{X}|Y}$ for each elementary event (x, y) with the joint probability $P_{Y|X} \otimes P_X(x, y) = P_{Y|X}(y|x)P_X(x)$ is

$$\iota_{P^*_{\hat{X}|Y}\|P_X}(x, y) = \log \frac{P^*_{\hat{X}|Y}(x|y)}{P_X(x)} \qquad (18)$$

$$= \log \frac{P^\alpha_{X|Y}(x|y)/P^\alpha_X(x)}{\sum_x P^\alpha_{X|Y}(x|y)/P_X^{\alpha-1}(x)} \qquad (19)$$

[6]Note that $P_{U_\alpha|Y}(u|y)$ is a conditional probability as for Markov chain $U - X - Y$ with fixed P_X and $P_{Y|X}$, we have $P_X(x) = \sum_u P_{X|U}(x|u)P_{U_\alpha}(u)$ and so $\sum_u P_{Y|U}(y|u)P_{U_\alpha}(u) = \sum_x P_{Y|X}(y|x)\sum_u P_{X|U}(x|u)P_{U_\alpha}(u) = \sum_x P_{Y|X}(y|x)P_X(x) = P_Y(y)$.

By Eq. (5), the Sibson mutual information defines the moment generating function of Y- and XY-elementary leakage, respectively, as

$$\exp\left(\frac{\alpha-1}{\alpha} I_\alpha^{\mathrm{S}}(P_X, P_{Y|X})\right)$$

$$= \mathbb{E}_{P_Y}\left[\exp\left(\frac{\alpha-1}{\alpha} D_\alpha(P_{X|Y=y}\|P_X)\right)\right] \qquad (20)$$

$$= \mathbb{E}_{P_{Y|X}\otimes P_X}\left[\exp\left(\frac{\alpha-1}{\alpha} \iota_{P^*_{\hat{X}|Y}\|P_X}(X, Y)\right)\right]. \qquad (21)$$

A tail bound on Y- and XY-elementary leakage can be directly obtained by the exponential Chebyshev's inequality: for all $\alpha \in (1, \infty)$,

$$\Pr\big(D_\alpha(P_{X|Y=y}\|P_X) > \epsilon\big) \leq \exp\left(\frac{\alpha-1}{\alpha}(I_\alpha^{\mathrm{S}}(P_X, P_{Y|X}) - \epsilon)\right), \qquad (22)$$

$$\Pr\Big(\iota_{P^*_{\hat{X}|Y}\|P_X}(X, Y) > \epsilon\Big) \leq \exp\left(\frac{\alpha-1}{\alpha}(I_\alpha^{\mathrm{S}}(P_X, P_{Y|X}) - \epsilon)\right). \qquad (23)$$

Following the idea of [8, Definition 5], the *δ-approximation of privacy leakage* refers to the situation when the probability of elementary leakage being greater than a positive ϵ is no greater than $(1 - \delta)$. By (22) and (23), enforcing

$$\exp\left(\frac{\alpha-1}{\alpha}(I_\alpha^{\mathrm{S}}(P_X, P_{Y|X}) - \epsilon)\right) \leq 1 - \delta, \qquad (24)$$

for any $\alpha \in (1, \infty)$ sufficiently achieves the δ-approximation of ϵ-upper bounded α-leakage such that $\Pr\big(D_\alpha(P_{X|Y=y}\|P_X) > \epsilon\big) \leq 1 - \delta$ and $\Pr\Big(\iota_{P^*_{\hat{X}|Y}\|P_X}(X, Y) > \epsilon\Big) \leq 1 - \delta$.

IV. Capacity of α-Leakage

Similar to the channel capacity in communication systems, we refer to the capacity of information leakage as the maximal privacy leaked via channel $P_{Y|X}$ over all channel input P_X. With the results in Section II-B, we show in this section that the Rényi capacity is an α-leakage capacity. This result will be formally derived with a $\tilde{f}$-mean relative information interpretation of the generalized Blahut-Arimoto algorithm in [9] for computing the Rényi capacity.

A. Unconditional $\tilde{f}$-mean Relative Information

For $\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big)$ being a conditional $\tilde{f}$-mean divergence, we define the corresponding unconditional divergence by equating the frequency probability to reference probability $P_{X|Y=y} = P_X$, which also removes the dependence on Y in the decision variable:

$$\begin{aligned}\tilde{D}_\alpha\big(P_{\hat{X}}\|P_X\big) &= \tilde{D}_\alpha\big(P_{\hat{X}}\|P_X|P_X\big) \\ &= \frac{\alpha}{\alpha-1}\log \sum_{x\in\mathcal{X}} P_X(x)\Big(\frac{P_{\hat{X}}(x)}{P_X(x)}\Big)^{\frac{\alpha-1}{\alpha}} \\ &= \frac{\alpha}{\alpha-1}\log \sum_{x\in\mathcal{X}} P_X^{\frac{1}{\alpha}}(x) P_{\hat{X}}^{\frac{\alpha-1}{\alpha}}(x). \qquad (25)\end{aligned}$$

This $\tilde{f}$-mean divergence is closely related to the Rényi divergence: for all $\alpha \in [0,\infty)$,

$$\tilde{D}_\alpha\big(P_{\hat{X}}\|P_X\big) = -D_{\frac{1}{\alpha}}(P_X\|P_{\hat{X}}). \tag{26}$$

Therefore, $\max \tilde{D}_\alpha\big(P_{\hat{X}}\|P_X\big) = -\min D_{\frac{1}{\alpha}}(P_X\|P_{\hat{X}}) = 0$.

B. Decompositions by $\tilde{D}_\alpha\big(P_{\hat{X}}\|P_X\big)$

We show below two decompositions by the proposed unconditional $\tilde{f}$-mean.

Proposition 2: For all $\alpha \in [0,\infty)$,

$$H_\alpha(P_X, P_{\hat{X}}) = H_\alpha(P_X) - \tilde{D}_\alpha(P_{\hat{X}}\|P_{X_\alpha}) \tag{27}$$

$$\tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big) = D_\alpha(P_{X|Y=y}\|P_X) + \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P^*_{\hat{X}|Y=y}\big) \tag{28}$$

Proposition 2 is obtained by simple algebraic works (proof omitted). By (27), $\min_{P_{\hat{X}}} H_\alpha(P_X, P_{\hat{X}}) = H_\alpha(P_X) - \max_{P_{\hat{X}}} \tilde{D}_\alpha(P_{\hat{X}}\|P_{X_\alpha}) = H_\alpha(P_X)$, which is another proof of [5, Theorem 1]. For $\alpha = 1$, $\tilde{D}_1(P_{\hat{X}}\|P_X) = -D_1(P_X\|P_{\hat{X}})$, where $D_1(\cdot\|\cdot)$ is the Kullback-Leibler divergence and we have (27) being $H_1(P_X, P_{\hat{X}}) = H_1(P_X) + D_1(P_X\|P_{\hat{X}})$, the well-known Shannon order decomposition of the cross entropy. By (28), $\max_{P_{\hat{X}|Y=y}} \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{X|Y=y}\big) = D_\alpha(P_{X|Y=y}\|P_X) + \max_{P_{\hat{X}|Y=y}} \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P^*_{\hat{X}|Y=y}\big) = D_\alpha(P_{X|Y=y}\|P_X)$, which provides another proof of (3).

C. Generalized Blahurt-Arimoto Method [9]

A generalized[7] Blahut-Arimoto method have been developed to compute Rényi capacity [9]. We first demonstrate that maximizing $\tilde{f}$-mean information gain $\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X)$ over channel input P_X and estimation decision $P_{\hat{X}|Y}$ results in Rényi capacity. To do so, we may use the the decomposition in Eq. (30), derived in a similar fashion to (28) and the Sibson identity [10], [13], [17]–[19]. We can then use the generalized Blahut-Arimoto method [9] for maximization of $\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X)$ iteratively.

Let us restate [9, Theorem 1-(2)] in terms of the $\tilde{f}$-mean relative information gain measure.

Theorem 1: For all $\alpha \in [0,\infty)$,

$$\max_{P_X} \tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X) = \log \sum_{x\in\mathcal{X}} \Big(\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) P_{\hat{X}|Y}^{\frac{\alpha-1}{\alpha}}(x|y)\Big)^{\frac{\alpha}{\alpha-1}} \tag{29}$$

with the maximizer

$$P^*_X(x) = \frac{\big(\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) P_{\hat{X}|Y}^{\frac{\alpha-1}{\alpha}}(x|y)\big)^{\frac{\alpha}{\alpha-1}}}{\sum_{x\in\mathcal{X}}\big(\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) P_{\hat{X}|Y}^{\frac{\alpha-1}{\alpha}}(x|y)\big)^{\frac{\alpha}{\alpha-1}}}, \forall x.$$

Proof: The maximand in (29) is

$$\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X\big) = \frac{\alpha}{\alpha-1}\log \sum_{x\in\mathcal{X}} P_X^{\frac{1}{\alpha}}(x) \sum_{y\in\mathcal{Y}} P_{Y|X}(y|x) P_{\hat{X}|Y}^{\frac{\alpha-1}{\alpha}}(x|y).$$

With $P^*_X(x)$, we have the decomposition

$$\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X\big) = \tilde{D}_\alpha\big(P^*_X\|P_X\big) + \tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P^*_X|P_{Y|X}\otimes P^*_X). \tag{30}$$

As $\max_{P_X} \tilde{D}_\alpha\big(P^*_X\|P_X\big) = 0$, $\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X\big)$ reaches maximum at $P_X = P^*_X$

$$\tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P^*_X|P_{Y|X}\otimes P^*_X) = \log \sum_x \Big(\sum_y P_{Y|X}(y|x) P_{\hat{X}|Y}^{\frac{\alpha-1}{\alpha}}(x|y)\Big)^{\frac{\alpha}{\alpha-1}}. \qquad \blacksquare$$

Following Theorem 1, Rényi capacity[8] [11], [21] is

$$C_\alpha(P_{Y|X}) = \max_{P_X} I^{\mathrm{S}}_\alpha(P_X, P_{Y|X}) = \max_{P_X}\max_{P_{\hat{X}|Y}} \tilde{D}_\alpha(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X). \tag{31}$$

This max-max optimization parallels the well-known max-min approach $C_\alpha(P_{Y|X}) = \max_{P_X}\min_{Q_Y} D_\alpha(P_{Y|X}\|Q_Y|P_X)$ and immediately suggests an iterative algorithm of alternating maximizations, which is proposed in [9, pp.667] and can be written as

$$P_{\hat{X}|Y} := \arg\max_{P_{\hat{X}|Y}} \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X\big), \tag{32}$$

$$P_X := \arg\max_{P_X} \tilde{D}_\alpha\big(P_{\hat{X}|Y=y}\|P_X|P_{Y|X}\otimes P_X\big). \tag{33}$$

The sequence $\{\tilde{D}_\alpha(P^{(t)}_{\hat{X}|Y=y}\|P^{(t)}_X|P_{Y|X}\otimes P^{(t)}_X)\colon t=0,1,\dots\}$ converges monotonically to $C_\alpha(P_{Y|X})$ [9, Th. 3].

V. Conclusion

In this paper, we used $\tilde{f}$-mean α-leakage interpretation of the Sibson mutual information to formulate Y- and XY-elementary information leakage measures. We developed a sufficient condition to achieve δ-approximation of ϵ-bounded Y- and XY-elementary leakage. We also proposed an alternative Y-elementary information leakage to generalize existing pointwise maximal leakage. We proved that Rényi capacity captures the maximal α-leakage over all adversary's inference decisions and channel inputs. This enabled us to use the generalized Blahurt-Arimoto algorithm for computing the $\tilde{f}$-mean information gain measure.

The following directions are viable for future work. The relationship between f-mean and $\tilde{f}$-mean measures, particularly for $\alpha \in [0,1)$, can be further explored. The dependence of optimal estimation decision $P^*_{\hat{X}|Y}(x|y)$ to α and its interpretation in terms of information leakage can also be discussed. Parallel to (3), another well known variational expression is $D_\alpha(P_{X|Y=y}\|P_X) = \frac{1}{1-\alpha}\min_{P_{\hat{X}|Y=y}} \big\{\alpha D_1(P_{\hat{X}|Y=y}\|P_{X|Y=y}) + (1 -$

[7]It is shown in [9] that for $\alpha = 1$ the generalized Blahut-Arimoto method reduces to Blahut-Arimoto method [15], [16] for computing Shannon capacity.

[8]There are different versions of Rényi capacity and (31) is one of them. See [20, Sec.4.3] and [13].

$\alpha) D_1(P_{\hat{X}|Y=y} \| P_X)\big\}$ [12, Theorem 30]. It is interesting to see whether similar interpretations hold for information leakage as well.

REFERENCES

[1] F. du Pin Calmon and N. Fawaz, "Privacy against statistical inference," in *Proceedings of 50th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2012, pp. 1401–1408.

[2] 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, Jun. 2013.

[3] 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, Mar. 2020.

[4] J. Liao, O. Kosut, L. Sankar, and F. du Pin 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.

[5] N. Ding, M. A. Zarrabian, and P. Sadeghi, "A cross entropy interpretation of rényi entropy for α-leakage," in *Proceedings of IEEE International Symposium on Information Theory*, Athens, 2024. [Online]. Available: https://arxiv.org/abs/2401.15202

[6] ——, "α-leakage by rényi divergence and Sibson mutual information," *arXiv preprint arXiv:2405.00423*, 2024.

[7] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, vol. 4. University of California Press, 1961, pp. 547–562.

[8] S. Saeidian, G. Cervia, T. J. Oechtering, and M. Skoglund, "Pointwise maximal leakage," *IEEE Transactions on Information Theory*, vol. 69, no. 12, pp. 8054–8080, Dec. 2023.

[9] S. Arimoto, "Computation of random coding exponent functions," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 665–671, Nov. 1976.

[10] R. Sibson, "Information radius," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 14, no. 2, pp. 149–160, 1969.

[11] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," *Topics in information theory*, 1977.

[12] T. van Erven and P. Harremoes, "Rényi divergence and Kullback-Leibler divergence," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3797–3820, Jul. 2014.

[13] S. Verdú, "Error exponents and α-mutual information," *Entropy*, vol. 23, no. 2, p. 199, Feb. 2021.

[14] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy." *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211–407, 2014.

[15] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 14–20, Jan. 1972.

[16] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 460–473, Jul. 1972.

[17] B. Nakiboglu, "The Rényi capacity and center," *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 841–860, Feb. 2019.

[18] Y. Polyanskiy and S. Verdu, "Arimoto channel coding converse and Rényi divergence," in *Proceedings of 48th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2010, pp. 1327–1333.

[19] M. Hayashi and V. Y. F. Tan, "Equivocations, exponents, and second-order coding rates under various rényi information measures," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 975–1005, Feb. 2017.

[20] G. Aishwarya and M. Madiman, "Conditional Rényi entropy and the relationships between Rényi capacities," *Entropy*, vol. 22, no. 5, p. 526, May 2020.

[21] I. Csiszár, "A class of measures of informativity of observation channels," *Periodica Mathematica Hungarica*, vol. 2, no. 1–4, pp. 191–213, Mar. 1972.