Maximal α-Leakage for Quantum Privacy Mechanisms

Bo-Yu Yang [*1], Hsuan Yu [†1], and Hao-Chung Cheng [‡1–5]

[1] *Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan (R.O.C.)*
[2] *Graduate Institute of Communication Engineering, National Taiwan University*
[3] *Department of Mathematics, National Taiwan University*
[4] *Hon Hai (Foxconn) Quantum Computing Center, New Taipei City 236, Taiwan (R.O.C.)*
[5] *Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan (R.O.C.)*

Abstract

In this work, maximal α-leakage is introduced to quantify how much a quantum adversary can learn about any sensitive information of data upon observing its disturbed version via a quantum privacy mechanism. We first show that an adversary's maximal expected α-gain using optimal measurement is characterized by measured conditional Rényi entropy. This can be viewed as a parametric generalization of König *et al.*'s famous guessing probability formula [*IEEE Trans. Inf. Theory*, 55(9), 2009]. Then, we prove that the α-leakage and maximal α-leakage for a quantum privacy mechanism are determined by measured Arimoto information and measured Rényi capacity, respectively. Various properties of maximal α-leakage, such as data processing inequality and composition property are established as well. Moreover, we show that regularized α-leakage and regularized maximal α-leakage for identical and independent quantum privacy mechanisms coincide with α-tilted sandwiched Rényi information and sandwiched Rényi capacity, respectively.

1 Introduction

1.1 Background

In the era of Internet of Things, data are published and shared almost ubiquitously, which have significantly increased possible paths where private information can leak. For instance, political preferences may be unveiled through movie ratings [1]. Besides, publishing data while maintaining desirable privacy and utility guarantee has also come to issues in many fields such as business [2], healthcare [3], and so on. How can we prevent adversaries from inferring sensitive data via reverse engineering while preserving at least some levels of utility? A common approach is to perturb the data by properly designing a *privacy mechanism*: a stochastic map aiming to control the private information leakage to adversaries. Conventionally, there are two mainstream approaches: (i) *differential privacy* (DP) [4, 5], proposed by Dwork in 2006, guarantees the indistinguishability between neighboring data, addressing the paradox of learning less about an individual while gaining useful information about a population, which is further generalized to *pufferfish privacy* [6, 7] by Kifer *et al.* in 2014; (ii) *information-theoretic privacy*, which has been

[*] 106yang@gmail.com
[†] yu.sherry3@gmail.com
[‡] haochung.ch@gmail.com

investigated in variant settings (e.g. Shannon's mutual information [8] and information bottleneck function [9]), adopts the viewpoints from statistics and information-theoretic security to study how the privacy is quantified [10, 11, 12, 13]. Concurrently in 2017, Assodeh *et al.* proposed probability of correctly guessing [14, 15], whereas Issa *et al.* proposed the *maximal leakage* (MaxL) [16], both aiming to depict how much sensitive information adversary can infer from its released version. In 2018, Liao *et al.* further introduced α-*leakage* and *maximal* α-*leakage* [17] (both were extended to (α, β)-leakage [18] later) to generalize the aforementioned quantities, so as to capture various adversarial actions with a tunable parameter $\alpha \in [0, \infty]$ [19, 20]. More recently, Stratonovich's value of information [21] and maximal g-leakage [22] are proposed to subsume maximal leakage and maximal α-leakage, generalizing α-gain/loss depictions to arbitrary gain functions [23, 24, 25].

When quantum devices become mature in the future, it is natural to consider protecting sensitive data via a *quantum privacy mechanism*. To this end, we may construct quantum privacy mechanisms via encoding sensitive classical data into quantum states. On the other hand, a quantum adversary may observe the encoded quantum states by applying certain quantum measurement. Various methods and metrics have been proposed to help design quantum privacy mechanisms, including quantum differential privacy [26, 27, 28, 29, 30, 31] and quantum pufferfish privacy [32]; meanwhile, maximal quantum leakage [33], an extension of MaxL for quantum privacy mechanisms, has also been studied. Furthermore, Chen and Hanson *et al.* depict the guessing strategies of classical data encoded via quantum privacy mechanism to quantum states, which generalize the framework of guesswork proposed by Massey [34] to quantum state discrimination problem in [35, 36].

1.2 Problem Formulation

In this work, we consider the problem of information leakage formulated as follows (Fig. 1). Assume that random variable X on a finite set $\mathcal{X}$ with probability distribution p_X represents the non-sensitive data, which may be correlated with another random variable S, denoting some sensitive data. A user wants to share non-sensitive X with the service provider to gain utility while maintaining privacy, especially protecting information of sensitive S. The user may apply a quantum privacy mechanism, namely, a classical-quantum channel $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$, mapping each realization $x \in \mathcal{X}$ to a quantum state ρ_B^x as a perturbation.[1] In this manner, a fundamental question naturally arises:

How much information does a quantum system B leak about S?

To observe the released quantum data B, adversaries can guess the non-sensitive data X via a quantum measurement to obtain their inference $\hat{X}$ according to Born's rule [37]:

$$\Pr\{\hat{X} = X \mid X = x, B\} = \mathrm{Tr}[\rho_B^x \Pi_B^x], \tag{1}$$

where $X - B - \hat{X}$ forms a classical-quantum-classical Markov chain. Here, the collection $\{\Pi_B^x\}_{x\in\mathcal{X}}$ of positive semi-definite matrices satisfying unity of resolution: $\sum_x \Pi_B^x = \mathbb{1}_B$, i.e. *positive operator-valued measure* (POVM), represents a quantum measurement, which can be viewed as the quantum generalization of classical decisions.

[1] If $\{\rho_B^x\}_x$ share a common eigenbasis, then the map $x \mapsto \rho_B^x$ reduces to classical privacy mechanism, i.e. a stochastic transformation.

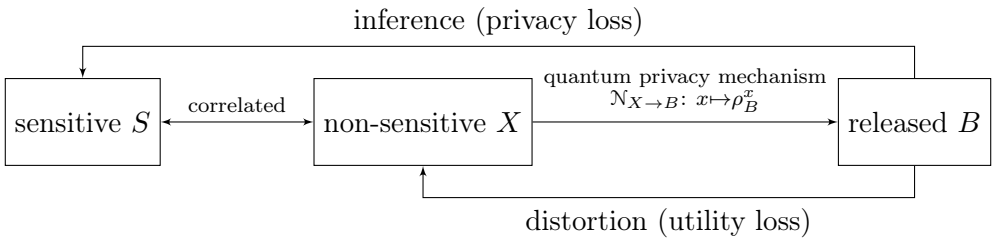

Figure 1: Information leakage framework with quantum privacy mechanisms. Here, systems S and X are classical, while system B is quantum. A quantum adversary may apply a quantum measurement on quantum system B to infer data X or S.

Liao *et al.* characterized various classical adversarial actions by proposing α-loss with $\alpha \in [1,\infty]$, interpolating between log-loss (for $\alpha = 1$) and 0-1 loss (also known as hard decision, for $\alpha = \infty$). Recently, Sypherd *et al.* [20] extended α-loss to the range $(0,\infty]$ to include exponential loss (for $\alpha = 1/2$) as well. Since α-leakage can be directly captured by the multiplicative gain increase upon an adversary observes the released data, motivated by Diaz *et al.* 's observation of maximal expected gain function [38, Eq. (8)], we introduce the *maximal expected α-gain* of a quantum ensemble $\{p_X(x), \rho_B^x\}_{x\in\mathcal{X}}$ (Def. 6) as

$$\mathsf{P}_\alpha(X|B)_\rho := \sup_{\text{POVM}\,\{\Pi_B^x\}_x} \mathbb{E}_{x\sim p_X}\left[\operatorname{Tr}\left[\rho_B^x (\Pi_B^x)^{\frac{\alpha-1}{\alpha}}\right]\right], \quad \forall \alpha \in [1,\infty] \tag{2}$$

to characterize how much an adversary correctly guesses the non-sensitive data X when using an optimal measurement strategy for observing the released quantum system B. If an adversary guesses X without observing B to obtain inference $\hat{X}$, we denote the corresponding (unconditional) maximal expected α-gain (Def. 6) as

$$\mathsf{P}_\alpha(X)_\rho := \sup_{p_{\hat{X}}:\hat{X}\perp X} \mathbb{E}_{x\sim p_X}\left[\Pr(\hat{X}=X|X=x)^{\frac{\alpha-1}{\alpha}}\right] = \sup_{p_{\hat{X}}} \mathbb{E}_{x\sim p_X}\left[p_{\hat{X}}(x)^{\frac{\alpha-1}{\alpha}}\right], \quad \forall \alpha \in [1,\infty], \tag{3}$$

where the condition $\hat{X} \perp X$ of the maximizer in the first expression springs from conditional independence of Markov chain $X - B - \hat{X}$; without the prior knowledge embedded in B, an adversary can only guess $\hat{X}$ independently.

In order to quantify how much information B leaks about the non-sensitive X, we introduce the *α-leakage* (Def. 7) from non-sensitive data X to quantum system B as the multiplicative increase of maximal expected gain upon observing B, written as

$$\mathcal{L}_\alpha(X\to B)_\rho := \frac{\alpha}{\alpha-1}\log\frac{\mathsf{P}_\alpha(X|B)_\rho}{\mathsf{P}_\alpha(X)_\rho}, \quad \forall \alpha \in (1,\infty], \tag{4}$$

which is a quantum generalization of the classical α-leakage [17]. Moreover, we define the *maximal α-leakage* (Def. 8) to quantify how much information leaks about any sensitive function S correlated to the non-sensitive X via the released quantum system B:

$$\mathcal{L}_\alpha^{\max}(X\to B)_\rho := \sup_{p_{S|X}:S-X-B} \mathcal{L}_\alpha(S\to B)_\rho, \quad \forall \alpha \in [1,\infty], \tag{5}$$

where the maximization is over all stochastic transformation $p_{S|X}$ such that $S - X - B$ forms a classical-classical-quantum Markov chain. The above quantity considers the maximal advantage of adversary for all sensitive data S through a quantum privacy mechanism.

1.3 Main Contributions

In this paper, we characterize the above-mentioned quantities via certain quantum information measures.

First, we show that the maximal expected α-gain is determined by the so-called *measured conditional Rényi entropy of order* α (Theorem 1):

$$\mathsf{P}_\alpha(X|B)_\rho = \mathrm{e}^{\frac{1-\alpha}{\alpha} H_\alpha^{\mathbb{M}}(X|B)_\rho}, \quad \forall\, \alpha \in [1,\infty]. \tag{6}$$

The detailed definition of the measured conditional Rényi entropy $H_\alpha^{\mathbb{M}}(X|B)_\rho$ with respect to a classical-quantum state ρ_{XB} can be found in Def. 3 later. This generalizes the classical result for characterizing the maximal expected α-gain by the conditional Rényi entropy over classical privacy mechanisms [17, Lemma 1]. Notably, for $\alpha = \infty$, our result recovers *guessing probability*, which was shown by König *et al.* [39] in 2009 as an operational meaning of min-entropy $H_\infty^{\mathbb{M}}(X|B)_\rho$. Hence, (6) is a parametric generalization of Ref. [39], and it provides an operational interpretation for measured conditional Rényi entropy $H_\alpha^{\mathbb{M}}(X|B)_\rho$ for all $\alpha \in [1,\infty]$.

Second, we prove that α-leakage is determined by the *measured Arimoto information of order* α (Theorem 2):

$$\mathcal{L}_\alpha(X \to B)_\rho = I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho. \tag{7}$$

The formal definition of measured Arimoto mutual information $I_\alpha^{\mathrm{A},\mathbb{M}}$ is provided in Def. 5 later.

Third, we prove that maximal α-leakage is determined by the *measured Rényi capacity* (Def. 7), which is equal to both *measured Arimoto capacity* (Def. 8) and *measured Rényi divergence radius* (Def. 6) for $\alpha \in (1,\infty]$, and by *measured Arimoto information* (Def. 5) for $\alpha = 1$:

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho = \begin{cases} I_1^{\mathrm{A},\mathbb{M}}(X:B)_\rho, & \alpha = 1; \text{ (8a)} \\ C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}) = C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}) = R_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}), & \alpha > 1. \text{ (8b)} \end{cases}$$

Note here that for $\alpha > 1$, the maximal α-leakage depends on the input distribution only through its support.

Moreover, we present some interesting properties for maximal α-leakage, such as data-processing inequality (DPI) for $\alpha \in [1,\infty]$ (Theorem 4, part 4) and composition property (Theorem 5). In the end, we discuss the asymptotic behavior of α-leakage and maximal α-leakage for n-fold product quantum privacy mechanisms $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$. We prove that, in the asymptotic limit $n \to \infty$, the maximal α-leakage is characterized by the *sandwiched Rényi divergence radius* (Theorem 7), i.e.

$$\lim_{n\to\infty} \frac{1}{n} \mathcal{L}_\alpha^{\max}(X^n \to B^n)_{\rho^n} = \lim_{n\to\infty} \frac{1}{n} C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}) = C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}) = R_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}), \quad \forall\, \alpha > 1. \tag{9}$$

Here, the underlying joint state $\rho_{X^nB^n}^n$ is induced by arbitrary fully-supported input distribution on non-sensitive data X^n and $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$. The left most equality of (9) directly comes from (8b). For the second equality, we show that the regularized measured Rényi capacity equals sandwiched Rényi capacity, which may be of independent interest (Proposition 8).

Our results hence provide operational meanings for measured Arimoto information, measured Rényi divergence radius, and sandwiched Rényi divergence radius.

This paper is organized as follows. Section 2 formally introduces measured quantities and their properties. In Section 3, we define α-leakage and maximal α-leakage for a quantum privacy mechanism, establish their equivalence to measured Arimoto information and measured Rényi

divergence radius, and derive various properties for maximal α-leakage. In Section 4, we discuss the implications of our results in privacy-utility tradeoffs and list some open problems.

For readability, we defer long proofs to Appendices B–I.

2 Preliminaries

2.1 Notation

Throughout this paper, we let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. Let $\mathcal{H}_X$ be an $|\mathcal{X}|$-dimensional Hilbert space (i.e. complex Euclidean space) with an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$. The outer product $|x\rangle\langle x|$ means an orthogonal projection onto the subspace spanned by vector $|x\rangle$. The set of probability distributions on $\mathcal{X}$ is denoted as $\mathcal{P}(\mathcal{X})$. For a probability distribution $p_X \in \mathcal{P}(\mathcal{X})$, we denote its support by $\mathtt{supp}(p_X)$. Let $\mathcal{S}(\mathcal{H}_B)$ be the set of density matrices (i.e. positive semi-definite matrices with unit trace) on Hilbert space $\mathcal{H}_B$, and $\mathbb{1}_B$ be the identity matrix on $\mathcal{H}_B$. The state of a quantum system B is modeled by some density matrix $\rho_B \in \mathcal{S}(\mathcal{H}_B)$. We denote $\mathtt{supp}(\rho_B)$ as the subspace spanned by the set of vectors in the eigenspace of ρ_B corresponding to positive eigenvalues. For two Hermitian matrices K and L on the same Hilbert space, we define the Hilbert–Schmidt inner product as $\langle K, L\rangle := \mathrm{Tr}\,[KL]$, where Tr is the standard trace.

A classical-quantum (c-q) state on $\mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_B)$ is $\rho_{XB} := \sum_{x\in\mathcal{X}} p_X(x)\,|x\rangle\langle x| \otimes \rho_B^x$, where $p_X \in \mathcal{P}(\mathcal{X})$ and each $\rho_B^x \in \mathcal{S}(\mathcal{H}_B)$. Namely, a c-q state ρ_{XB} represents a quantum ensemble $\{p_X(x), \rho_B^x\}_{x\in\mathcal{X}}$. Let $\{\Pi_B^x\}_{x\in\mathcal{X}}$ be a POVM, i.e. each Π_B^x is a positive semi-definite matrix on $\mathcal{H}_B$ and $\sum_{x\in\mathcal{X}} \Pi_B^x = \mathbb{1}_B$; equivalently, we also use $\Pi_{XB} := \sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \Pi_B^x$ to denote it for brevity. For a c-q state ρ_{XB}, the partial trace over system X is denoted by $\mathrm{Tr}_X[\rho_{XB}] = (\mathrm{Tr}_X \otimes \mathrm{id}_B)(\rho_{XB}) = \sum_{x\in\mathcal{X}} (\langle x| \otimes \mathbb{1}_B)\rho_{XB}(|x\rangle \otimes \mathbb{1}_B)$, where id_B is the identity map on system B.

For a power function $f : M \mapsto M^\alpha$ of a matrix M, we refer to f as acting on the support of M. Moreover, let $\mathcal{N}_{\mathcal{X}\to B} : \mathcal{X} \ni x \mapsto \rho_B^x \in \mathcal{S}(\mathcal{H}_B)$ be a c-q channel that maps each letter in the input set $x \in \mathcal{X}$ to a density matrix $\rho_B^x \in \mathcal{S}(\mathcal{H}_B)$. The exponential function is denoted by $(\cdot) \mapsto \mathrm{e}^{(\cdot)}$. The logarithmic function $(\cdot) \mapsto \log(\cdot)$ denotes natural logarithm.

2.2 Information-Theoretic Quantities

In this section, we briefly review the measured counterpart of entropic quantities, and refer readers to [40, 41] for more details.

Definition 1. *Let $\rho_{XB} = \sum_x p_X(x)\,|x\rangle\langle x| \otimes \rho_B^x \in \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be a c-q state.*

1. ***Order-α Rényi divergence*** [42]*: For $\alpha \in (0,1)\cup(1,\infty)$ and probability distributions $p_Y, q_Y \in \mathcal{P}(\mathcal{Y})$,*

$$D_\alpha(p_Y\|q_Y) := \frac{1}{\alpha-1}\log\sum_{y\in\mathcal{Y}} p_Y(y)^\alpha q_Y(y)^{1-\alpha}, \tag{10}$$

if $\alpha \in (0,1)$ or $\mathtt{supp}(p_Y) \subseteq \mathtt{supp}(q_Y)$; otherwise, it is defined as $+\infty$. In the limit $\alpha \to 1$, since Rényi divergence converges to Kullback–Leibler divergence, we denote it as $D_1(p_Y\|q_Y) \equiv D(p_Y\|q_Y) := \sum_{y\in\mathcal{Y}} p_Y(y)\log\frac{p_Y(y)}{q_Y(y)}$. As $\alpha\to\infty$, Rényi divergence converges to max-divergence

$$D_\infty(p_Y\|q_Y) := \lim_{\alpha\to\infty} D_\alpha(p_Y\|q_Y) = \sup_{y\in\mathcal{Y}} \log\frac{p_Y(y)}{q_Y(y)}. \tag{11}$$

The order-0 Rényi divergence is defined by taking limit $\alpha \to 0$.

2. ***Measured Rényi divergence*** [40]*: For $\alpha \in [0,\infty]$, density matrix ρ, and positive semi-definite matrix σ,*

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) := \sup_{(\mathcal{Y},\Pi)} D_\alpha\left(\{\mathrm{Tr}[\rho\Pi_y]\}_{y\in\mathcal{Y}}\|\{\mathrm{Tr}[\sigma\Pi_y]\}_{y\in\mathcal{Y}}\right). \tag{12}$$

The supremum is over all finite sets $\mathcal{Y}$ and POVMs $\{\Pi^y\}_{y\in\mathcal{Y}}$. For $\alpha \in (0,\infty)\backslash\{1\}$, sometimes we also express measured Rényi divergence as:

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) = \frac{1}{\alpha-1}\log Q_\alpha^{\mathbb{M}}(\rho\|\sigma), \tag{13}$$

where ***measured Rényi quasi divergence*** *is defined by*

$$Q_\alpha^{\mathbb{M}}(\rho\|\sigma) := \begin{cases} \sup\limits_{(\mathcal{Y},\Pi)} \sum\limits_{y\in\mathcal{Y}} (\mathrm{Tr}[\rho\Pi_y])^\alpha (\mathrm{Tr}[\sigma\Pi_y])^{1-\alpha}, & \alpha \in (1,\infty); \quad (14a) \\ \inf\limits_{(\mathcal{Y},\Pi)} \sum\limits_{y\in\mathcal{Y}} (\mathrm{Tr}[\rho\Pi_y])^\alpha (\mathrm{Tr}[\sigma\Pi_y])^{1-\alpha}, & \alpha \in (0,1). \quad (14b) \end{cases}$$

3. ***Measured conditional Rényi entropy*** [36, Eq. (64)]*: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$H_\alpha^{\mathbb{M}}(X|B)_\rho := -\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}(\rho_{XB}\|\mathbb{1}_X\otimes\sigma_B). \tag{15}$$

4. ***Measured Rényi information*** [43]*: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$I_\alpha^{\mathbb{M}}(X:B)_\rho := \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}(\rho_{XB}\|\rho_X\otimes\sigma_B). \tag{16}$$

5. ***Measured Arimoto information*** [44]*: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho := H_\alpha(X)_p - H_\alpha^{\mathbb{M}}(X|B)_\rho, \tag{17}$$

where

$$H_\alpha(X)_p := \frac{1}{1-\alpha}\log\sum_{x\in\mathcal{X}} p_X(x)^\alpha, \tag{18}$$

and for $\alpha = 1$ and ∞, $H_\alpha(X)_p$ is defined as taking limit $\alpha\to 1$ and $\alpha\to\infty$ respectively.

6. ***Measured Rényi divergence radius*** [43]*: For a c-q channel $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$ and $\alpha \in [0,\infty]$,*

$$R_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) := \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \sup_{x\in\mathcal{X}} D_\alpha^{\mathbb{M}}(\mathcal{N}(x)\|\sigma_B) \equiv \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \sup_{x\in\mathcal{X}} D_\alpha^{\mathbb{M}}(\rho_B^x\|\sigma_B), \tag{19}$$

which is the measured counterpart of the quantum Rényi divergence radius [45, Eq. (85)].

7. ***Measured Rényi capacity*** [43]*: For $\alpha \in [0,\infty]$ and a c-q channel $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$,*

$$C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) := \sup_{p_X\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathbb{M}}(X:B)_\rho. \tag{20}$$

8. ***Measured Arimoto capacity****: For $\alpha \in [0,\infty]$ and a c-q channel $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$,*

$$C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) := \sup_{p_X\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho. \tag{21}$$

9. ***Sandwiched Rényi divergence*** [46, 47]*: For density matrix ρ, positive semi-definite matrix σ, and $\alpha \in [0,1) \cup (1,\infty]$,*

$$D_\alpha^*(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] \tag{22}$$

if $\alpha \in [0,1)$ or $\mathtt{supp}(\rho) \subseteq \mathtt{supp}(\sigma)$; otherwise, it is defined as $+\infty$.

For $\alpha \in [1/2,1) \cup (1,\infty]$, D_α^* satisfies data-processing inequality [46, 47, 48] *.*

10. ***Umegaki's quantum relative entropy*** [49]*: For density matrix ρ and positive semi-definite matrix σ,*

$$D(\rho\|\sigma) := \mathrm{Tr}\left[\rho(\log\rho - \log\sigma)\right], \tag{23}$$

if $\mathtt{supp}(\rho) \subseteq \mathtt{supp}(\sigma)$; otherwise, it is defined as $+\infty$. Continuous extension of sandwiched Rényi divergence for $\alpha \to 1$ reduces to Umegaki's quantum relative entropy.

11. ***Sandwiched Rényi information*** [50, 51, 43]*: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$I_\alpha^*(X:B)_\rho := \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^*(\rho_{XB}\|\rho_X\otimes\sigma_B). \tag{24}$$

12. ***Sandwiched Arimoto information****: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$I_\alpha^{\mathrm{A},*}(X:B)_\rho := H_\alpha(X)_p + \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^*(\rho_{XB}\|\mathbb{1}_X\otimes\sigma_B). \tag{25}$$

13. ***Sandwiched Rényi divergence radius*** [45, Eq. (85)]*: For a c-q channel $\mathcal{N}_{\mathcal{X}\to B}: x \mapsto \rho_B^x$ and $\alpha \in [0,\infty]$,*

$$R_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}) := \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \sup_{x\in\mathcal{X}} D_\alpha^*(\mathcal{N}(x)\|\sigma_B) \equiv \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \sup_{x\in\mathcal{X}} D_\alpha^*(\rho_B^x\|\sigma_B). \tag{26}$$

14. ***Sandwiched Rényi capacity*** [43]*: For $\alpha \in [0,\infty]$ and a c-q state ρ_{XB},*

$$C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}) := \sup_{p_X\in\mathcal{P}(\mathcal{X})} I_\alpha^*(X:B)_\rho. \tag{27}$$

15. ***Sandwiched Arimoto capacity****: For $\alpha \in [0,\infty]$ and a c-q channel $\mathcal{N}_{\mathcal{X}\to B}: x \mapsto \rho_B^x$,*

$$C_\alpha^{\mathrm{A},*}(\mathcal{N}_{\mathcal{X}\to B}) := \sup_{p_X\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{A},*}(X:B)_\rho. \tag{28}$$

For $\alpha \in \{1/2,\infty\}$, measured Rényi divergence coincides with sandwiched Rényi divergence [39, 40, 52, 53]. In particular for $\alpha = +\infty$, it is named *maximum relative entropy* [54], i.e.

$$D_\infty^{\mathbb{M}}(\rho\|\sigma) = D_\infty^*(\rho\|\sigma) = \inf\{\lambda\in\mathbb{R} : \rho \le \mathrm{e}^\lambda\sigma\}. \tag{29}$$

For all $\alpha \in (1/2,\infty)$, we have the following strict relation [40, Thm. 6]:

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) < D_\alpha^*(\rho\|\sigma), \tag{30}$$

unless ρ commutes with σ.

Remark 1. We remark that in the classical case (i.e. $\{\rho_B^x\}_x$ mutually commute), the measured Arimoto information reduced to the classical Arimoto information introduced in Refs. [44, 55].

Definition 2 (α-tilted distribution [17, 55]). *Given a parameter $\alpha \in (0,\infty)$ and a probability distribution $p_X \in \mathcal{P}(\mathcal{X})$, the α-tilted distribution of p_X is defined as*

$$p_X^{(\alpha)}(x) := \frac{p_X(x)^\alpha}{\sum_{x\in\mathcal{X}} p_X(x)^\alpha}. \tag{31}$$

Lemma 1 (Variational formula of measured Rényi divergence [40, Lemma. 1, 3], [56, 57]). *For density matrices ρ, σ and $\alpha \in [0,\infty]$,*

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) = \begin{cases} \sup\limits_{\omega>0} \operatorname{Tr}\left[\rho\log\omega\right] - \log\operatorname{Tr}\left[\sigma\omega\right] = \sup\limits_{\omega>0} \operatorname{Tr}\left[\rho\log\omega\right] + 1 - \operatorname{Tr}\left[\sigma\omega\right], & \alpha = 1; \quad (32a) \\ \sup\limits_{\omega>0} \dfrac{1}{\alpha-1}\log\left(\left(\operatorname{Tr}\left[\rho\omega^{1-\frac{1}{\alpha}}\right]\right)^\alpha \left(\operatorname{Tr}\left[\sigma\omega\right]\right)^{1-\alpha}\right), & \alpha \neq 1. \quad (32b) \end{cases}$$

Lemma 2 (Data-processing inequality of measured Rényi divergence [56, Prop. 5.4]). *Let ρ, σ be density matrices and $\mathcal{N}$ be a fully quantum channel. Then, for all $\alpha \in [0,\infty]$,*

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) \geq D_\alpha^{\mathbb{M}}\left(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)\right). \tag{33}$$

Lemma 3 (Super-additivity of measured Rényi divergence [40, Eq. (65)]). *Let $\rho_A, \sigma_A \in \mathcal{S}(\mathcal{H}_A)$ and $\rho_B, \sigma_B \in \mathcal{S}(\mathcal{H}_B)$ be density matrices. For all $\alpha \in [0,\infty]$,*

$$D_\alpha^{\mathbb{M}}(\rho_A\otimes\rho_B\|\sigma_A\otimes\sigma_B) \geq D_\alpha^{\mathbb{M}}\left(\rho_A\|\sigma_A\right) + D_\alpha^{\mathbb{M}}\left(\rho_B\|\sigma_B\right). \tag{34}$$

Lemma 4 (Super-additivity of operational quantities [40, Sec. V]). *Let $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ and $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ be density matrices. Then, for all $\alpha \in [0,\infty]$,*

$$\inf_{\sigma_{AB}\in\mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)} D_\alpha^{\mathbb{M}}\left(\rho_A\otimes\rho_B\|\sigma_{AB}\right) \quad \geq \inf_{\sigma_A\in\mathcal{S}(\mathcal{H}_A)} D_\alpha^{\mathbb{M}}\left(\rho_A\|\sigma_A\right) + \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}\left(\rho_B\|\sigma_B\right). \tag{35}$$

Lemma 5 (Invariant maximal value over α-tilted distribution [58, Prop. 1]). *Let $\alpha \in (0,\infty)$. Given a distribution $p_X \in \mathcal{P}(\mathcal{X})$ and a continuous function $f : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$,*

$$\max_{p_X\in\mathcal{P}(\mathcal{X})} f(p_X) = \max_{p_X\in\mathcal{P}(\mathcal{X})} f\left(p_X^{(\alpha)}\right). \tag{36}$$

Lemma 6 (Equivalent expressions of measured Rényi capacity, Rényi divergence radius, and Arimoto capacity). *For $\alpha \geq 0$ and a c-q channel $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$, we have*

$$C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) = R_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) = C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}); \tag{37}$$

on the other hand, for $\alpha \geq 1/2$,

$$C_\alpha^{*}(\mathcal{N}_{\mathcal{X}\to B}) = R_\alpha^{*}(\mathcal{N}_{\mathcal{X}\to B}) = C_\alpha^{\mathrm{A},*}(\mathcal{N}_{\mathcal{X}\to B}). \tag{38}$$

Beigi and Tomamichel [43, Lemma 1] proved the equivalence of measured Rényi capacity and measured Rényi divergence radius for $\alpha \in (0,1)$ and stated that the proof for $\alpha \geq 1$ follows similarly. For completeness, we provide the proof for the $\alpha \geq 0$ case in Appendix B and further prove their equivalence to measured Arimoto capacity.

On the other hand, the equivalence of sandwiched Rényi capacity and sandwiched Rényi divergence radius for $\alpha \geq 1/2$ is proved by Mosonyi and Ogawa [45, Prop. 4.2]. For completeness, we also prove their equivalence to sandwiched Arimoto capacity in Appendix B.

3 Information Leakage Measures and Main Results

In this section, we first introduce the notions of the *expected α-loss* and *expected α-gain*, characterizing how much information one loses or gains from observing a quantum ensemble. Subsequently, we propose the definitions of α-leakage and maximal α-leakage as measures of information leakage under a quantum privacy mechanism. These measures capture how much inference an adversary can draw when the data are released via a privacy mechanism. The purpose for a quantum adversary is to minimize expected α-loss, which is equivalent to maximize expected α-gain.[2]

We further prove that α-leakage is determined by measured Arimoto information and that maximal α-leakage is determined by measured Rényi capacity with a support constraint respectively. We also prove that regularized α-leakage and regularized maximal α-leakage for i.i.d. quantum privacy mechanisms are equivalent to sandwiched information and sandwiched capacity, respectively.

3.1 The Expected α-Loss and α-Gain

Focusing on the study of information leakage for quantum privacy mechanisms, we mainly discuss the range $\alpha \geq 1$ of α-loss. In fact, the range $\alpha \in (0,1)$ of α-loss was also defined in classical literature (see e.g. [20]). We define α-loss for $\alpha < 1$ in Appendix A.

Definition 3 (Expected α-loss)**.** *Consider a c-q state $\rho_{XB} = \sum_x p_X(x) \left|x\right\rangle\!\left\langle x\right| \otimes \rho_B^x$. For any $\alpha \in [1,\infty]$, we define an expected α-loss of POVM Π_{XB} under ρ_{XB} as:*

$\varepsilon_\alpha(\Pi_{XB})_\rho :=$

$$\begin{cases} \dfrac{\alpha}{\alpha-1}\left(1 - \operatorname{Tr}\left[\rho_{XB}\Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right]\right), & \text{if } \alpha > 1 \text{ or } \mathtt{supp}(\rho_{XB}) \subseteq \mathtt{supp}(\Pi_{XB}); \quad (39a) \\ -\operatorname{Tr}\left[\rho_{XB}\log\Pi_{XB}\right], & \text{if } \alpha = 1 \text{ and } \mathtt{supp}(\rho_{XB}) \subseteq \mathtt{supp}(\Pi_{XB}); \quad (39b) \\ +\infty, & \text{otherwise.} \quad (39c) \end{cases}$$

Remark 2. Note that expected α-loss defined in (39a) is always non-negative because $\Pi_{XB} \leq \Pi_{XB}^{\frac{\alpha-1}{\alpha}} \leq \mathbb{1}_{XB}$ for $\alpha \in [1,\infty]$.

Remark 3. For $\alpha = 1$, $\varepsilon_1(\Pi_{XB})_\rho = -\operatorname{Tr}\left[\rho_{XB}\log\Pi_{XB}\right]$ is known as *expected log-loss* [17]; as for $\alpha = \infty$,

$$\varepsilon_\infty(\Pi_{XB})_\rho = 1 - \sum_{x\in\mathcal{X}} p_X(x)\operatorname{Tr}\left[\rho_B^x\Pi_B^x\right] \tag{40}$$

corresponds to the error probability of discriminating the ensemble $\{p_X(x),\rho_B^x\}_{x\in\mathcal{X}}$ using the POVM $\{\Pi_B^x\}_{x\in\mathcal{X}}$ (see e.g. [59, Chapter 3]). Therefore, the expected α-loss ε_α serves as a tunable loss measure that interpolates between the various known loss (or risk) functions.

Remark 4. One may introduce an *α-loss operator* of a POVM Π_{XB} defined as

$$\ell_\alpha(\Pi_{XB}) := \begin{cases} \dfrac{\alpha}{\alpha-1}\left(\mathbb{1}_{XB} - \Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right), & \text{if } \alpha > 1 \text{ or } \Pi_{XB} > 0; \quad (41a) \\ -\log\Pi_{XB}, & \text{if } \alpha = 1 \text{ and } \Pi_{XB} > 0; \quad (41b) \\ +\infty, & \text{otherwise,} \quad (41c) \end{cases}$$

[2]The proposed *expected α-loss* and *expected α-gain* recovers the classical correspondence in the setting of classical privacy mechanisms [17, 38]. Instead of characterizing α-leakage by expected α-loss as in Ref. [17], we depict α-leakage via expected α-gain so as to obtain more insight.

such that $\varepsilon_\alpha(\Pi_{XB})_\rho = \langle \rho_{XB}, \ell_\alpha(\Pi_{XB})\rangle$. In the classical setting, such an α-loss operator is called α-loss function in [17, Def. 3].

Definition 4 (Minimal expected α-loss)**.** *For any c-q state $\rho_{XB} = \sum_x p_X(x)\,|x\rangle\langle x| \otimes \rho_B^x$ and $\alpha \in [1,\infty]$, we define the minimal expected α-loss for ρ_{XB} as:*

$$\varepsilon_\alpha(X|B)_\rho := \inf_{\mathrm{POVM}\,\Pi_{XB}} \varepsilon_\alpha(\Pi_{XB})_\rho = \begin{cases} \inf\limits_{\mathrm{POVM}\,\Pi_{XB}} \dfrac{\alpha}{\alpha-1}\left(1 - \mathrm{Tr}\left[\rho_{XB}\Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right]\right), & \text{for } \alpha > 1; \text{ (42a)} \\ \inf\limits_{\mathrm{POVM}\,\Pi_{XB}} -\mathrm{Tr}\left[\rho_{XB}\log\Pi_{XB}\right], & \text{for } \alpha = 1, \text{ (42b)} \end{cases}$$

where the minimization is over all POVMs *on $\mathcal{H}_B$, i.e. for all POVM Π_{XB} such that $\mathrm{Tr}_X[\Pi_{XB}] = \mathbb{1}_B$ and $\Pi_{XB} \geq 0$.*

On the other hand, we define the minimal expected α-loss for an inference $\hat{X}$ without observing B as

$$\varepsilon_\alpha(X)_\rho := \begin{cases} \inf\limits_{p_{\hat{X}}\in\mathcal{P}(\mathcal{X})} \dfrac{\alpha}{\alpha-1}\left(1 - \mathbb{E}_{x\sim p_X}\left[p_{\hat{X}}(x)^{\frac{\alpha-1}{\alpha}}\right]\right), & \text{for } \alpha > 1; \quad (43a) \\ \inf\limits_{p_{\hat{X}}\in\mathcal{P}(\mathcal{X})} -\mathbb{E}_{x\sim p_X}\left[\log p_{\hat{X}}(x)\right], & \text{for } \alpha = 1. \quad (43b) \end{cases}$$

Remark 5. A trivial guess $\Pi_{XB} = \sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \frac{\mathbb{1}_B}{|\mathcal{X}|}$ in $\varepsilon_\alpha(X|B)_\rho$ and $p_{\hat{X}}(x) = \frac{1}{|\mathcal{X}|}$ in $\varepsilon_\alpha(X)_\rho$ provides the following upper bound on the minimal expected α-loss:

$$\varepsilon_\alpha(X|B)_\rho \leq \frac{\alpha}{\alpha-1}\left(1 - |\mathcal{X}|^{\frac{1-\alpha}{\alpha}}\right) \quad (44)$$

$$\varepsilon_\alpha(X)_\rho \leq \frac{\alpha}{\alpha-1}\left(1 - |\mathcal{X}|^{\frac{1-\alpha}{\alpha}}\right). \quad (45)$$

Later in Theorem 1, we will provide explicit characterization of the minimal expected α-loss.

Definition 5 (Expected α-gain)**.** *Given a c-q state $\rho_{XB} = \sum_x p_X(x)\,|x\rangle\langle x| \otimes \rho_B^x$ and $\alpha \geq 1$, for POVM Π_{XB}, we define expected α-gain for ρ_{XB} as:*

$$g_\alpha(\Pi_{XB})_\rho := \mathrm{Tr}\left[\rho_{XB}\Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right]. \quad (46)$$

Definition 6 (Maximal expected α-gain)**.** *For any c-q state $\rho_{XB} = \sum_x p_X(x)\,|x\rangle\langle x| \otimes \rho_B^x$, POVM Π_{XB}, and $\alpha \in [1,\infty]$, we define the maximal expected α-gain for ρ_{XB} as:*

$$\mathsf{P}_\alpha(X|B)_\rho := \sup_{\mathrm{POVM}\,\Pi_{XB}} g_\alpha(\Pi_{XB})_\rho = \sup_{\mathrm{POVM}\,\{\Pi_B^x\}_{x\in\mathcal{X}}} \sum_{x\in\mathcal{X}} p_X(x)\,\mathrm{Tr}\left[\rho_B^x(\Pi_B^x)^{\frac{\alpha-1}{\alpha}}\right], \quad (47)$$

where the maximization is over all POVMs *on $\mathcal{H}_B$, i.e. for all Π_{XB} such that $\mathrm{Tr}_X[\Pi_{XB}] = \mathbb{1}_B$ and $\Pi_{XB} \geq 0$.*

On the other hand, we define the maximal expected α-gain for an inference $\hat{X}$ without observing B as

$$\mathsf{P}_\alpha(X)_\rho := \sup_{p_{\hat{X}}:\hat{X}\perp X} \mathbb{E}_{x\sim p_X}\left[\Pr(\hat{X} = X|X = x)^{\frac{\alpha-1}{\alpha}}\right] = \sup_{p_{\hat{X}}} \mathbb{E}_{x\sim p_X}\left[p_{\hat{X}}(x)^{\frac{\alpha-1}{\alpha}}\right], \quad (48)$$

where the condition $\hat{X} \perp X$ of the supremum in the first expression springs from conditional independence of Markov chain $X - B - \hat{X}$. Without the prior knowledge embedded in B, an adversary can only guess $\hat{X}$ independently.

Remark 6. For $\alpha = \infty$, the quantity $\mathsf{P}_\infty(X|B)_\rho$ is equal to the maximal success (guessing) probability of discriminating the ensemble $\{p_X(x), \rho_B^x\}_{x\in\mathcal{X}}$. Moreover, it is straightforward to calculate the maximal success (guessing) probability without observing the system B, i.e. $\mathsf{P}_\infty(X)_\rho = \max_{x\in\mathcal{X}} p_X(x)$, placing all weights on the symbol with the maximal prior probability.

Remark 7. Note that for POVM $\Pi_{XB} > 0$, minimal expected α-loss and maximal expected α-gain are related by

$$\mathsf{P}_\alpha(X|B)_\rho = 1 - \frac{\alpha-1}{\alpha}\varepsilon_\alpha(X|B)_\rho. \tag{49}$$

Moreover, maximal expected α-gain lies in the interval $[0,1]$ because $\Pi_{XB} \leq \Pi_{XB}^{\frac{\alpha-1}{\alpha}} \leq \mathbb{1}_{XB}$ for $\alpha \in [1,\infty]$.

Our first result is the following characterization of the maximal expected α-gain for a c-q state ρ_{XB} via the measured conditional Rényi entropy. This thereby provides an operational interpretation for measured conditional Rényi entropy.

Theorem 1 (Characterization of the maximal expected α-gain)**.** *For any classical-quantum state ρ_{XB} and $\alpha \in [1,\infty]$, the maximal expected α-gain in Def. 6 is given by*

$$\mathsf{P}_\alpha(X|B)_\rho = \mathrm{e}^{\frac{1-\alpha}{\alpha}H_\alpha^{\mathbb{M}}(X|B)_\rho}, \tag{50}$$

where $H_\alpha^{\mathbb{M}}(X|B)_\rho$ was introduced in Def. 1. 3.

A detail proof for Theorem 1 is provided in Appendix C.

Remark 8. For $\alpha = \infty$ and recalling the max relative entropy in (29), the error exponent in (50) is reduced to the so-called min-entropy introduced by König, Renner, and Schaffner [39, Thm. 1], i.e.,

$$\begin{aligned} -\log \mathsf{P}_\infty(X|B)_\rho &= H_\infty^*(X|B)_\rho && (51)\\ &= -\inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_\infty^*(\rho_{XB} \,\|\, \mathbb{1}_X \otimes \sigma_B) && (52)\\ &= -\inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \inf\{\lambda \in \mathbb{R} : \rho_{XB} \leq \mathrm{e}^\lambda(\mathbb{1}_X \otimes \sigma_B)\}. && (53) \end{aligned}$$

Via (49), the minimal expected α-loss is expressed as

$$\varepsilon_\alpha(X|B)_\rho = \begin{cases} \dfrac{\alpha}{\alpha-1}\left(1 - \mathrm{e}^{\frac{1-\alpha}{\alpha}H_\alpha^{\mathbb{M}}(X|B)_\rho}\right), & \text{for } \alpha > 1, \quad (54a)\\ H_1^{\mathbb{M}}(X|B)_\rho, & \text{for } \alpha = 1. \quad (54b) \end{cases}$$

We remark that (54a) and (54b) recover the classical results given in [17, Lemma 1] and [38, Proposition 1].

3.2 Main Results: α-Leakage and Maximal α-Leakage

Let $\mathcal{N}_{\mathcal{X}\to B} : x \mapsto \rho_B^x$ denote a classical-quantum privacy mechanism. To quantify the multiplicative increase in the maximal expected gain of data X when observing its released version B via quantum privacy mechanisms, we introduce the following information-leakage measures: α-leakage and maximal α-leakage.

Definition 7 (α-leakage)**.** *Given $\alpha \in [1,\infty]$ and a c-q state ρ_{XB}, the α-leakage from X to B is defined for $\alpha > 1$ as*

$$\mathcal{L}_\alpha(X \to B)_\rho := \frac{\alpha}{\alpha-1}\log\frac{\mathsf{P}_\alpha(X|B)_\rho}{\mathsf{P}_\alpha(X)_\rho}. \tag{55}$$

When $\alpha = 1$, the α-leakage is

$$\mathcal{L}_1(X \to B)_\rho := \lim_{\alpha\to 1}\mathcal{L}_\alpha(X \to B)_\rho = \varepsilon_1(X)_\rho - \varepsilon_1(X|B)_\rho \tag{56}$$

by continuous extension.

Remark 9. Def. 7 is an extension of classical α-leakage introduced by Liao *et al.* [17, Def. 5] to the scenarios of quantum privacy mechanisms. Note that the denominator of the logarithmic term in (55) is maximal expected gain of a decision without additional information apart from X, while the numerator of the logarithmic term in (55) is maximal expected gain of a decision having access to quantum system B [38]. In particular, for the case of $\alpha = \infty$, the proposed α-leakage recovers the correctly-guessing information leakage that Asoodeh *et al.* [15] proposed:

$$\mathcal{L}_{\infty}(X \to B)_{\rho} := \log \frac{\mathsf{P}_{\infty}(X|B)_{\rho}}{\mathsf{P}_{\infty}(X)_{\rho}}. \tag{57}$$

Theorem 2 (Characterization of α-leakage)**.** *For $\alpha \in [1, \infty]$, α-leakage defined in Def. 7 can be expressed as*

$$\mathcal{L}_{\alpha}(X \to B)_{\rho} = I_{\alpha}^{\mathrm{A},\mathbb{M}}(X : B)_{\rho}, \tag{58}$$

where $I_{\alpha}^{\mathrm{A},\mathbb{M}}(X : B)_{\rho}$ was introduced in Def. 1. 5.

A detailed proof of Theorem 2 is provided in Appendix D.

In addition to capturing how much an adversary learns about X from B, in practice, we are often more interested in quantifying how much information leaks via B for any function S of X. With this goal, now we introduce the definition of maximal α-leakage.

Definition 8 (Maximal α-leakage)**.** *Given a joint c-q state $\rho_{XB} \in \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_B)$, the maximal α-leakage from X to B is defined as*

$$\mathcal{L}_{\alpha}^{\max}(X \to B)_{\rho} := \sup_{p_{S|X}: S - X - B} \mathcal{L}_{\alpha}(S \to B)_{\rho} \tag{59}$$

for $\alpha \in [1, \infty]$, where S denotes any function of X and takes values from an arbitrary finite set.

Theorem 3 (Characterization of the maximal α-leakage)**.** *For $\alpha \in [1, \infty]$, the maximal α-leakage defined in Def. 8 can be expressed as*

$$\mathcal{L}_{\alpha}^{\max}(X \to B)_{\rho} = \begin{cases} I_1^{\mathrm{A},\mathbb{M}}(X : B)_{\rho} = \mathcal{L}_1(X \to B)_{\rho}, & \alpha = 1; \quad \text{(60a)} \\ C_{\alpha}^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X) \to B}) = C_{\alpha}^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X) \to B}) = R_{\alpha}^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X) \to B}), & \alpha > 1. \quad \text{(60b)} \end{cases}$$

A detailed proof of Theorem 3 is in Appendix E.

When $\alpha = 1$, maximal α-leakage reduces to measured Arimoto mutual information $I_1^{\mathrm{A},\mathbb{M}}(X : B)_{\rho}$; when $\alpha = \infty$, maximal α-leakage reduces to maximal leakage proposed in [33].

Remark 10. Theorem 3 provides an operational meaning of measured Rényi capacity. Note that when $\alpha = 1$, maximal α-leakage depends on input probability distribution p_X; when $\alpha > 1$, maximal α-leakage depends on input probability distribution only through its support $\mathtt{supp}(p_X)$.

As stated in Lemma 6, maximal α-leakage is also equal to measured Arimoto capacity and measured Rényi divergence radius.

3.3 Properties of Maximal α-Leakage

To further analyze the performance of quantum privacy mechanisms, now we explore some properties of maximal α-leakage.

Theorem 4. *Denote by* $\bar{\rho}_{XB} = \sum_{x\in\mathcal{X}} \bar{p}_X(x)|x\rangle\langle x| \otimes \rho_B^x$ *for any input distribution* $\bar{p}_X \in \mathcal{P}(\mathcal{X})$ *as an optimization variable. For* $\alpha \in [1,\infty]$ *and given a c-q state* $\rho_{XB} = \sum_{x\in\mathcal{X}} p_X(x)|x\rangle\langle x| \otimes \rho_B^x$*, maximal* α*-leakage*

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho = \begin{cases} I_1^{\mathrm{A},\mathbb{M}}(X:B)_\rho, & \alpha = 1; \quad (61a) \\ \sup\limits_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \inf\limits_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}(\bar{\rho}_{XB} \| \bar{p}_X \otimes \sigma_B), & \alpha > 1. \quad (61b) \end{cases}$$

has the following properties:

1. *is a concave program of optimization variable* $\bar{p}_X$ *for* $\alpha > 1$*, and a concave function of input distribution* p_X *for* $\alpha = 1$*;*
2. *is quasi-convex in* ρ_B^x *given optimization variable* $\bar{p}_X$ *for* $\alpha > 1$ *or given input distribution* p_X *for* $\alpha = 1$*;*
3. *is non-decreasing in* α*;*
4. *satisfies data-processing inequality;*
5. $$0 \le \mathcal{L}_\alpha^{\max}(X \to B)_\rho \le \begin{cases} \log(|\mathtt{supp}(p_X)|), & \text{for } \alpha > 1; \quad (62a) \\ H_1(X)_p, & \text{for } \alpha = 1; \quad (62b) \end{cases}$$

A detailed proof of Theorem 4 is provided in Appendix F.

Sometimes, an adversary can access more than one released version of non-sensitive data X. The following theorem shows that even an adversary receives multiple independently released data B, they cannot obtain information about sensitive data S more than marginal sums of maximal α-leakage. This behavior is also known as composition property [60, Sec. 3] [16, Lemma. 6] [17, Thm. 5].

Theorem 5 (Composition property)**.** *Given a probability distribution* $p_X \in \mathcal{P}(\mathcal{X})$ *and quantum privacy mechanisms* $\mathcal{N}_{\mathcal{X}\to B_1} : x \mapsto \rho_{B_1}^x$ *and* $\mathcal{N}_{\mathcal{X}\to B_2} : x \mapsto \rho_{B_2}^x$*, for any* $\alpha \in [1,\infty]$*, the maximal* α*-leakage from* X *to* B_1B_2 *is bounded above by*

$$\mathcal{L}_\alpha^{\max}(X \to B_1, B_2)_\rho \le \mathcal{L}_\alpha^{\max}(X \to B_1)_\rho + \mathcal{L}_\alpha^{\max}(X \to B_2)_\rho, \quad (63)$$

where $\rho_{XB_1B_2} := \sum_{x\in\mathcal{X}} p_X(x)\,|x\rangle\langle x| \otimes \rho_{B_1}^x \otimes \rho_{B_2}^x$.

A detailed proof of Theorem 5 is provided in Appendix G.

3.4 Asymptotic behaviors of α-leakage and maximal α-leakage

In previous sections, we have considered α-leakage and maximal α-leakage for a quantum privacy mechanism in the *one-shot setting*; namely, the underlying data S and X and the quantum privacy mechanism $\mathcal{N}_{X\to B}$ are used only once. Here, we discuss the asymptotic behaviors of α-leakage and maximal α-leakage when non-sensitive data X^n are released via i.i.d. quantum privacy mechanisms:

$$\mathcal{N}_{\mathcal{X}\to B}^{\otimes n} : x_1x_2\cdots x_n \mapsto \rho_{B_1}^{x_1} \otimes \rho_{B_2}^{x_2} \otimes \cdots \otimes \rho_{B_n}^{x_n} =: \rho_{B^n}^{x^n}. \quad (64)$$

More precisely, we will study both $\mathcal{L}_\alpha$ and $\mathcal{L}_\alpha^{\max}$ under $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$ and normalize the quantities by n to study the average information leakage when $n \to \infty$, which is termed *regularization* in quantum information theory.

While the i.i.d. assumption is not often adopted in information-theoretic security because one cannot restrict quantum adversaries to attack only in an i.i.d. manner; nevertheless, in the scenario of information leakage, the quantum privacy mechanism is employed by the system designer, and an i.i.d. quantum privacy mechanism $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$ is arguably easier to perform than a general n-shot privacy mechanism $\mathcal{N}_{\mathcal{X}^n\to B^n}$: that may output a multipartite entangled state on system B^n.

When non-sensitive data X^n are i.i.d. as well, the following theorem shows that the asymptotic behavior of α-leakage is characterized by sandwiched Rényi information but under the α-tilted distribution $p_X^{(\alpha)}$. Note that the one-shot α-leakage $\mathcal{L}_\alpha(X\to B)_\rho$ is characterized by measured Arimoto information $I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho$ (Theorem 2), which is in general not identical to measured Rényi information $I_\alpha^{\mathbb{M}}(X:B)_\rho$ nor sandwiched Rényi information $I_\alpha^*(X:B)_\rho$. Hence, the regularization of $\mathcal{L}_\alpha$ converges to a different quantity as expected (Theorem 6). The interested readers are referred to Lemmas 8 and 9 in Appendix H for more details.

Theorem 6 (Characterization of regularized α-leakage)**.** *Let $\rho_{XB}=\sum_{x\in\mathcal{X}}p_X(x)|x\rangle\langle x|\otimes\rho_B^x$. For $\alpha\geq 1$, then regularized α-leakage from i.i.d. X^n to B^n under $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$ is given by*

$$\lim_{n\to\infty}\frac{1}{n}\mathcal{L}_\alpha(X^n\to B^n)_{\rho^{\otimes n}}=I_\alpha^*\left(X:B\right)_{\rho^{(\alpha)}}, \tag{65}$$

where the α-tilted distribution $p_X^{(\alpha)}$ is introduced in Def. 2 and we denote

$$\rho_{XB}^{(\alpha)}\equiv\sum_{x\in\mathcal{X}}p_X^{(\alpha)}(x)\,|x\rangle\langle x|\otimes\rho_B^x. \tag{66}$$

Theorem 6 follows from the fact that regularized measured Rényi divergence is given by the sandwiched Rényi divergence [41]. The detailed proof is deferred to Appendix H.

Note that the n-shot α-leakage $\mathcal{L}_\alpha(X^n\to B^n)_{\rho^{\otimes n}}$ is evaluated under i.i.d. non-sensitive data X^n as input. Surprisingly, below we show that when calculating the regularized maximal α-leakage $\mathcal{L}_\alpha^{\max}(X^n\to B^n)_{\rho^n}$ for $\alpha>1$, the privacy mechanism $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$ works for general correlated data X^n. Hence, the i.i.d. assumption on X^n is no longer needed.

Theorem 7 below establishes the regularized maximal α-leakage for $\alpha>1$ and provides an operational meaning for sandwiched Rényi capacity $C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B})$.

Theorem 7 (Characterization of regularized maximal α-leakage)**.** *For any $\alpha>1$, the regularized maximal α-leakage from X^n with any arbitrary distribution $p_{X^n}\in\mathcal{P}(\mathcal{X}^n)$ that has full support to B^n under $\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}$ is given by*

$$\lim_{n\to\infty}\frac{1}{n}\mathcal{L}_\alpha^{\max}(X^n\to B^n)_{\rho^n}=C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B})=R_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}), \tag{67}$$

where $\rho_{X^nB^n}^n:=\sum_{x^n\in\mathcal{X}^n}p_{X^n}(x^n)|x^n\rangle\langle x^n|\otimes\rho_{B^n}^{x^n}$.

Note that Theorem 3 already characterizes that

$$\frac{1}{n}\mathcal{L}_\alpha^{\max}(X^n\to B^n)_{\rho^n}=\frac{1}{n}C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}^{\otimes n})=\frac{1}{n}C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}),\quad\forall\, n\in\mathbb{N}. \tag{68}$$

Below we show that the regularized measured Rényi capacity is given by the sandwiched Rényi capacity $C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B})$. This concludes Theorem 7.

Proposition 8 (Equivalence of regularized Arimoto capacity, regularized Rényi capacity, and sandwiched capacity)**.** *Let $\mathcal{N}_{X\to B} : x \mapsto \rho_B^x$ denote a classical-quantum channel. For any $\alpha \geq \frac{1}{2}$, the following identities hold:*

$$\lim_{n\to\infty} \frac{1}{n} C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}) = \lim_{n\to\infty} \frac{1}{n} C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}^{\otimes n}) = C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}) = R_\alpha^*(\mathcal{N}_{\mathcal{X}\to B}). \tag{69}$$

The proof is deferred to Appendix I.

4 Discussion

In this work, we propose the definition of α-leakage and maximal α-leakage for a quantum privacy mechanism based on maximal expected α-gain. We prove that (i) one-shot α-leakage is determined by measured Arimoto information (Theorem 2); (ii) one-shot maximal α-leakage is determined by measured Rényi capacity, measured Arimoto capacity, and measured Rényi divergence radius (Theorem 3); (iii) regularized α-leakage is characterized by α-tilted sandwiched Rényi information (Theorem 6); (iv) regularized maximal α-leakage is characterized by one-shot sandwiched Rényi capacity (Theorem 7). Also, we derive various properties (Theorem 4, 5) for maximal α-leakage, such as DPI and sub-additivity under the same quantum privacy mechanism. The established characterizations apply to the so-called *privacy-utility trade-off* (PUT) scenario of classical data protected by quantum privacy mechanisms, which are depicted as below.

The goal of a privacy mechanism is to preserve information leakage subject to some desired level of utility. By adopting maximal α-leakage $\mathcal{L}_\alpha^{\max}(X \to B)_\rho$ (Def. 8) as the privacy metric and any quantum-classical distortion $d(X,B)_\rho$ (see e.g. [61]) as the utility metric bounded by the maximal permitted distortion δ, we can model this privacy-utility tradeoff (PUT) problem as the optimization problem below: for any probability distribution p_X on $\mathcal{X}$,

$$\begin{cases} \min\limits_{x\mapsto\rho_B^x} & \mathcal{L}_\alpha^{\max}(X\to B)_\rho & \text{(70a)} \\ \text{subject to} & d(X,B)_\rho \leq \delta, & \text{(70b)} \end{cases}$$

the optimal PUT is denoted as

$$\mathrm{PUT}(\delta)_\rho := \inf_{d(X,B)_\rho \leq \delta} \mathcal{L}_\alpha^{\max}(X\to B)_\rho. \tag{71}$$

We have shown that for $\alpha > 1$, the objective function of maximal α-leakage $\mathcal{L}_\alpha^{\max}(X \to B)$ is concave in optimization variable $\bar{p}_X$ and quasi-convex in ρ_B^x given $\bar{p}_X$ in Theorem 4. Therefore, if the distortion function $d(X,B)_\rho$ is convex in ρ_B^x, this problem belongs to quasi-convex programs [62]. However, given the released quantum state B, does our framework differ from measuring it and then tackling it directly with classical techniques? To address this issue, we can consider the Markov chain $X - B - Y$, where Y is a classical state obtained by measuring B. Observing that $\mathcal{H}_Y \subseteq \mathcal{H}_B$, we immediately obtain

$$\mathrm{PUT}(\delta)_{\rho_{XB}} \leq \mathrm{PUT}(\delta)_{\rho_{XY}} \tag{72}$$

due to contraction of the constraint set. Therefore, we suppose that the classical method cannot outperform our quantum privacy mechanism in PUT problem with the privacy metric of maximal α-leakage.

Future research directions for this work are abundant, especially for those tasks requiring privacy analysis. Here we list some open problems:

1. This work studies the information leakage via a quantum privacy mechanism with a c-c-q Markov chain. The fully-quantum case (q-q-q) or other frameworks (e.g. c-q-q, q-c-q Markov chain) have not been explore yet.

2. Liao *et al.* [17] proposed the definition of f-leakage and maximal f-leakage, and formulated them as PUT problems. On the other hand, Hiai and Mosonyi defined measured f-divergence [41], one can also consider extending classical f-leakage and maximal f-leakage to PUT problem under a quantum privacy mechanism by measured f-divergence.

3. For $\alpha = 1$, we find that both 1-leakage and maximal 1-leakage are determined by the *measured Arimoto mutual information*

$$I_1^{\mathrm{A},\mathbb{M}}(X;B)_\rho = H_1(X)_p + \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_1^{\mathbb{M}}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B) \tag{73}$$

   Note that one can also define the *measured mutual information* as

$$I_1^{\mathbb{M}}(X;B)_\rho = \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_1^{\mathbb{M}}\left(\rho_{XB} \| \rho_X \otimes \sigma_B\right). \tag{74}$$

   In the classical setting, the two quantities coincide. It would be interesting to see if they are identical in the quantum scenario.

4. In Ref. [36], it was pointed out that the measured Rényi conditional entropy $H_\alpha^{\mathbb{M}}(X|B)_\rho$ is efficiently computable via the variational expressions [40]. It would be interesting to find a concrete algorithm for computing this minimax optimization.

Acknowledgment

H.-C. Cheng is supported by the Young Scholar Fellowship (Einstein Program) of the National Science and Technology Council, Taiwan (R.O.C.) under Grants No. NSTC 112-2636-E-002-009, No. NSTC 112-2119-M-007-006, No. NSTC 112-2119-M-001-006, No. NSTC 112-2124-M-002-003, by the Yushan Young Scholar Program of the Ministry of Education, Taiwan (R.O.C.) under Grants No. NTU-112V1904-4 and by the research project "Pioneering Research in Forefront Quantum Computing, Learning and Engineering" of National Taiwan University under Grant No. NTU-CC-112L893405 and NTU-CC-113L891605. H.-C. Cheng acknowledges the support from the "Center for Advanced Computing and Imaging in Biomedicine (NTU-113L900702)" through The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

Appendices

A The Expected α-Loss when $\alpha < 1$

Definition 9 (Expected α-loss)**.** *Consider a c-q state $\rho_{XB} := \sum_x p_X(x) |x\rangle\langle x| \otimes \rho_B^x$. For any $\alpha \in (0, \infty]$, we define an expected α-loss of POVM Π_{XB} under ρ_{XB} as:*

$$\varepsilon_\alpha(\Pi_{XB})_\rho :=$$

$$\begin{cases} \dfrac{\alpha}{\alpha-1}\left(1 - \operatorname{Tr}\left[\rho_{XB}\Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right]\right), & \text{if } \alpha > 1 \text{ or } \mathtt{supp}(\rho_{XB}) \subseteq \mathtt{supp}(\Pi_{XB}); \quad (75a) \\ -\operatorname{Tr}\left[\rho_{XB}\log\Pi_{XB}\right], & \text{if } \alpha = 1 \text{ and } \mathtt{supp}(\rho_{XB}) \subseteq \mathtt{supp}(\Pi_{XB}); \quad (75b) \\ +\infty, & \text{otherwise.} \quad (75c) \end{cases}$$

Remark 11. Note that expected α-loss defined in (75a) is always non-negative. For instance, when $\alpha \in (0,1)$, since $\Pi_{XB}^{\beta} \geq \mathbb{1}_{XB}$ for any $\beta < 0$, one may see that both $\frac{\alpha}{\alpha-1}$ and $1 - \mathrm{Tr}[\rho_{XB}\Pi_{XB}^{\frac{\alpha-1}{\alpha}}]$ are negative, producing a positive value of expected α-loss.

Remark 12. For $\alpha = 1/2$, $\varepsilon_{1/2}(\Pi_{XB}) = \mathrm{Tr}\left[\rho_{XB}\Pi_{XB}^{-1}\right] - 1$ is called *expected exponential-loss* [20].

Remark 13. To extend the definition of α-loss to $(0,1)$, one may introduce an *α-loss operator* of a POVM Π_{XB} defined as

$$\ell_\alpha(\Pi_{XB}) := \begin{cases} \dfrac{\alpha}{\alpha-1}\left(\mathbb{1}_{XB} - \Pi_{XB}^{\frac{\alpha-1}{\alpha}}\right), & \text{if } \alpha > 1 \text{ or } \Pi_{XB} > 0; \quad (76a) \\ -\log \Pi_{XB}, & \text{if } \alpha = 1 \text{ and } \Pi_{XB} > 0; \quad (76b) \\ +\infty, & \text{otherwise}, \quad (76c) \end{cases}$$

such that $\varepsilon_\alpha(\Pi_{XB})_\rho = \langle \rho_{XB}, \ell_\alpha(\Pi_{XB})\rangle$.

B Proof of Lemma 6

Proof. For $\alpha \geq 0$, the equivalence of measured Rényi capacity and measured Rényi divergence radius is given by

$$C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}) = \sup_{p_X \in \mathcal{P}(\mathcal{X})} I_\alpha^{\mathbb{M}}(X:B)_\rho \quad (77)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}(\rho_{XB} \| \rho_X \otimes \sigma_B) \quad (78)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log \mathbb{E}_{x\sim p_X}\left[Q_\alpha^{\mathbb{M}}(\rho_B^x \| \sigma_B)\right] \quad (79)$$

$$\overset{(a)}{=} \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \sup_{p_X \in \mathcal{P}(\mathcal{X})} \frac{1}{\alpha-1} \log \mathbb{E}_{x\sim p_X}\left[Q_\alpha^{\mathbb{M}}(\rho_B^x \| \sigma_B)\right] \quad (80)$$

$$= \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \sup_{x \in \mathcal{X}} D_\alpha^{\mathbb{M}}(\rho_B^x \| \sigma_B) =: R_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathcal{X}\to B}), \quad (81)$$

where (a) results from the minimax theorem [45, Lemma 2.7], following the $\alpha \in [\frac{1}{2}, 1)$ case proved in [63, Lemma 1].

Meanwhile, measured Arimoto capacity ($(t) = \mathbb{M}$) and sandwiched Arimoto capacity ($(t) = *$) can be expressed as

$$C_\alpha^{\mathrm{A},(t)}(\mathcal{N}_{\mathcal{X}\to B}) = \sup_{p_X \in \mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{A},(t)}(X:B)_\rho \quad (82)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} D_\alpha^{(t)}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B) + H_\alpha(X)_p \quad (83)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X(x)^\alpha Q_\alpha^{(t)}(\rho_B^x \| \sigma_B) - \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X(x)^\alpha \quad (84)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} \frac{p_X(x)^\alpha}{\sum_{x\in\mathcal{X}} p_X(x)^\alpha} Q_\alpha^{(t)}(\rho_B^x \| \sigma_B) \quad (85)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X^{(\alpha)}(x) Q_\alpha^{(t)}(\rho_B^x \| \sigma_B) \quad (86)$$

$$\overset{(b)}{=} \sup_{p_X \in \mathcal{P}(\mathcal{X})} \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X(x) Q_\alpha^{(t)}(\rho_B^x \| \sigma_B) \quad (87)$$

$$= \sup_{p_X \in \mathcal{P}(\mathcal{X})} I_\alpha^{(t)}(X:B)_\rho =: C_\alpha^{(t)}(\mathcal{N}_{\mathcal{X}\to B}), \quad (88)$$

where (b) comes from Lemma 5 since the objective function in (87) is continuous (see e.g. [50, Lemma 1]) in p_X.

Thus, we show the equivalence not only between measured Rényi capacity, measured Rényi divergence radius and measured Arimoto capacity for $\alpha \geq 0$, but also the equivalence between sandwiched Rényi capacity, sandwiched Rényi divergence radius and sandwiched Arimoto capacity for $\alpha \geq 1/2$. □

C Proof of Theorem 1

Proof. We prove Theorem 1 via Lagrange duality and the variational expression in Lemma 1. Note that we only consider Definitions 5 and 6 for $\alpha \geq 1$ (because later we only focus on the α-leakage in Definition 7 for such a range), we remark that Theorem 1 holds for $\alpha \in [1/2, \infty]$. In the following, we first prove the $\alpha \in [1/2, \infty]\backslash\{1\}$ case. Recalling Definition 6 of the maximal expected α-gain, we first modify the optimization region:

$$\frac{\alpha}{\alpha-1}\log \mathsf{P}_\alpha(X|B)_\rho \overset{\text{(a)}}{=} \sup_{\substack{\omega_{XB}>0\\ \mathrm{Tr}_X[\omega_{XB}]=\mathbb{1}_B}} \frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right] \tag{89}$$

$$\overset{\text{(b)}}{=} \sup_{\substack{\omega_{XB}>0\\ \mathrm{Tr}_X[\omega_{XB}]\leq\mathbb{1}_B}} \frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right], \tag{90}$$

where (a) is because the objective function is continuous and hence we can take the interior of the finite-dimensional POVMs ω_{XB}; (b) follows from the fact that $(\cdot)^{\frac{\alpha-1}{\alpha}}$ is operator monotone increasing for $\alpha > 1$ and operator monotone decreasing for $\alpha \in [1/2, 1)$ [64, §4].

We introduce Lagrange operator $W_B \geq 0$ to rewrite (90) into its dual problem:

$$\inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right] - \left(\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]] - \mathrm{Tr}[W_B]\right). \tag{91}$$

Note that strong duality theorem [65, Prop. 5.3.1] holds since there exists at least an interior point (e.g. $\omega_{XB} = \frac{\mathbb{1}_X}{2|\mathcal{X}|}\otimes\mathbb{1}_B$) satisfying the constraint $\mathrm{Tr}_X[\omega_{XB}] \leq \mathbb{1}_B$.

Next, we rewrite the penalty terms in (91) as follows

$$\inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right] + \big((1-\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]) - (1-\mathrm{Tr}[W_B])\big) \tag{92}$$

$$\overset{\text{(c)}}{=} \inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right] + \big((-\log \mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]) - (1-\mathrm{Tr}[W_B])\big) \tag{93}$$

$$\overset{\text{(d)}}{=} \inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log \mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right] + \big((-\log \mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]) - (-\log \mathrm{Tr}[W_B])\big). \tag{94}$$

Here we first show that the penalty $1-\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]$ can be changed to $-\log \mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]$ in identity (c). Observe that (93) is invariant under the substitution $\omega_{XB} \mapsto \gamma\omega_{XB}$ for $\gamma > 0$. Further, note that the optimizer $W_B^\star$ never attains value at zero in (93); otherwise, the optimization value of expression (93) escalates to $+\infty$. Because $\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]] > 0$ for $W_B > 0$ and $\omega_{XB} > 0$, we can impose a normalization constraint $\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]] = 1$ on the following

equation:

$$\inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right]+\big((-\log\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]])-(1-\mathrm{Tr}[W_B])\big) \tag{95}$$

$$=\inf_{W_B\geq 0}\sup_{\substack{\omega_{XB}>0\\ \mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]]=1}}\frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right]+\big((-\log\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]])-(1-\mathrm{Tr}[W_B])\big) \tag{96}$$

$$\leq\inf_{W_B\geq 0}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right]+\big((1-\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]])-(1-\mathrm{Tr}[W_B])\big), \tag{97}$$

where the inequality comes from relaxation of the constraint set and $-\log x=1-x$ when $x=1$. On the other hand, since we have $\log(x+1)\leq x$ for all $x>-1$, $-\log\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]\geq 1-\mathrm{Tr}[W_B\,\mathrm{Tr}_X[\omega_{XB}]$, we can also lower-bound (93) by (92). Thus, we have proved the identity (c). Now we can apply the same technique to further prove identity (d) by observing that (94) is invariant under the substitution $W_B\mapsto\zeta W_B$ for any $\zeta>0$. We remark that a similar technique was used in the proof of [40, Lemma 1].

Finally, let $\sigma_B=\frac{W_B}{\mathrm{Tr}[W_B]}$ in (94), we obtain

$$\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right]-\log\mathrm{Tr}[\sigma_B\,\mathrm{Tr}_X[\omega_{XB}]] \tag{98}$$

$$=\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}\sup_{\omega_{XB}>0}\frac{\alpha}{\alpha-1}\log\mathrm{Tr}\left[\rho_{XB}\omega_{XB}^{\frac{\alpha-1}{\alpha}}\right]-\log\mathrm{Tr}[(\mathbb{1}_X\otimes\sigma_B)\omega_{XB}] \tag{99}$$

$$\overset{\text{(e)}}{=}\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}D_\alpha^{\mathbb{M}}(\rho_{XB}\|\mathbb{1}_X\otimes\sigma_B) \tag{100}$$

$$=-H_\alpha^{\mathbb{M}}(X|B)_\rho, \tag{101}$$

where (e) is from the variational formula of measured Rényi divergence (Lemma 1).

□

D Proof of Theorem 2

Proof. For $\alpha\in(1,\infty]$, by definition, the left-hand side of (58) can be written as

$$\mathcal{L}_\alpha(X\to B)_\rho=\frac{\alpha}{\alpha-1}\log\mathsf{P}_\alpha(X|B)_\rho-\frac{\alpha}{\alpha-1}\log\sup_{p_{\hat{X}}}\sum_{x\in\mathcal{X}}p_X(x)p_{\hat{X}}(x)^{\frac{\alpha-1}{\alpha}} \tag{102}$$

$$\overset{\text{(a)}}{=}\left(-H_\alpha^{\mathbb{M}}(X|B)_\rho\right)-\left(-H_\alpha(X)_p\right)=I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho, \tag{103}$$

where (a) follows from Theorem 1 indicates that the first term of (102) can be expressed as the measured conditional Rényi entropy, whereas the latter term in (102) can be simplified by the KKT condition [66, Chapter 5.5.3].

On the other hand, the $\alpha=1$ case can be proved as follows.

$$\mathcal{L}_1(X\to B)_\rho=-\varepsilon_1(X|B)_\rho+\varepsilon_1(X)_\rho \tag{104}$$

$$\overset{\text{(b)}}{=}\left(-H_1^{\mathbb{M}}(X|B)_\rho\right)-\left(-H_1(X)_p\right)=I_1^{\mathrm{A},\mathbb{M}}(X:B)_\rho, \tag{105}$$

where (b) follows from the similar Lagrange multiplier method as the $\alpha>1$ proof in Appendix C:

$$\varepsilon_1(X|B)_\rho=\inf_{\text{POVM }\Pi_{XB}}-\mathrm{Tr}[\rho_{XB}\log\Pi_{XB}] \tag{106}$$

$$=-\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}\sup_{\Pi_{XB}>0}\left(\mathrm{Tr}\left[\rho_{XB}\log\Pi_{XB}\right]-\log\mathrm{Tr}[(\mathbb{1}_X\otimes\sigma_B)\Pi_{XB}]\right) \tag{107}$$

$$=H_1^{\mathbb{M}}(X|B)_\rho. \tag{108}$$

□

E Proof of Theorem 3

Proof. For the $\alpha = 1$ case, by Definition 8 and Theorem 2, we have

$$\mathcal{L}_1^{\max}(X \to B)_\rho = \sup_{p_{S|X}: S-X-B} I_1^{\mathrm{A},\mathbb{M}}(S:B)_\rho \leq I_1^{\mathrm{A},\mathbb{M}}(X:B)_\rho, \tag{109}$$

where the inequality is from Lemma 2. To prove the achievability of (109), we set $\rho_{SB} = \rho_{XB}$, and thus $\mathcal{L}_1^{\max}(X \to B)_\rho = I_1^{\mathrm{A},\mathbb{M}}(X:B)_\rho$.

Now we consider the scenario of $\alpha > 1$. This proof consists of two parts: the upper bound is optimality, whereas the lower bound is achievability.

Upper bound (optimality): Subsequently, we fix the distribution p_X and the collection of states $\{\rho_B^x\}_{x\in\mathcal{X}}$. We first rewrite the maximal α-leakage as the following:

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho = \sup_{p_{S|X}: S-X-B} I_\alpha^{\mathrm{A},\mathbb{M}}(S:B)_\rho \tag{110}$$

$$= \sup_{\substack{p_{\bar{S}\bar{X}}: p_{\bar{X}} = p_X \\ \bar{S}-\bar{X}-B}} I_\alpha^{\mathrm{A},\mathbb{M}}(\bar{S}:B)_\rho, \tag{111}$$

where the supremum in (110) is over the set of Markov classical-classical-quantum states ρ_{SXB} of the form:

$$\left\{ \rho_{SXB} = \sum_s \sum_{x\in\mathcal{X}} p_{S|X=x}(s|x)|s\rangle\langle s| \otimes p_X(x)|x\rangle\langle x| \otimes \rho_B^x : \forall \{p_{S|X=x}\}_{x\in\mathcal{X}} \subseteq \mathcal{P}(\mathcal{S}) \right\}, \tag{112}$$

whereas the supremum in (111) is over the set of states $\rho_{\bar{S}\bar{X}B}$ of the form:

$$\left\{ \rho_{\bar{S}\bar{X}B} = \sum_s \sum_{x\in\bar{\mathcal{X}}} p_{\bar{S}}(s)\,|s\rangle\langle s| \otimes p_{\bar{X}|\bar{S}}(x|s)\,|x\rangle\langle x| \otimes \rho_B^x : \forall\ p_{\bar{S}} \in \mathcal{P}(\bar{\mathcal{S}}), p_{\bar{X}|\bar{S}} \text{ s.t. } p_{\bar{X}} = p_X \right\}. \tag{113}$$

By inspection, the sets (112) and (113) coincide.

Next, we relax the constraint $p_{\bar{X}} = p_X$ for the stochastic transformation $p_{\bar{X}|\bar{S}}$ in (113) to all distributions on the support of p_X, i.e.

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho \leq \sup_{p_{\bar{X}|\bar{S}}(\cdot|s)\in\mathcal{P}(\mathtt{supp}(p_X))} \sup_{p_{\bar{S}}: \bar{S}-\bar{X}-B} I_\alpha^{\mathrm{A},\mathbb{M}}(\bar{S}:B)_\rho \tag{114}$$

$$\overset{\text{(a)}}{=} \sup_{p_{\bar{X}|\bar{S}}(\cdot|s)\in\mathcal{P}(\mathtt{supp}(p_X))} \sup_{p_{\bar{S}}: \bar{S}-\bar{X}-B} I_\alpha^{\mathbb{M}}(\bar{S}:B)_\rho \tag{115}$$

$$\overset{\text{(b)}}{=} \sup_{p_{\bar{X}}\in\mathcal{P}(\mathtt{supp}(p_X))} I_\alpha^{\mathbb{M}}(\bar{X}:B)_\rho = C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}) \tag{116}$$

$$\overset{\text{(c)}}{=} \sup_{p_{\bar{X}}\in\mathcal{P}(\mathtt{supp}(p_X))} I_\alpha^{\mathrm{A},\mathbb{M}}(\bar{X}:B)_\rho = C_\alpha^{\mathrm{A},\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}) \tag{117}$$

$$\overset{\text{(d)}}{=} R_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}), \tag{118}$$

where (a) holds from that the supremum of measured Arimoto information equals the supremum of measured Rényi information (Lemma 6); (b) results from the data processing inequality (Lemma 2) of measured Rényi information for the Markov chain $\bar{S} - \bar{X} - B$, where the equality

is attained when $\bar{S} = \bar{X}$.[3] The last two lines, (c) and (d), are again direct consequences of Lemma 6.

Lower bound (achievability): We bound the maximal α-leakage from below by considering a random variable S, where $S - X - B$ forms a classical-classical-quantum Markov chain and $H_1(X|S) = 0$. Let us apply the similar method as in [17]: let $\mathcal{S}$ and $\mathcal{X}$ be a bijection so that $\mathcal{S}$ is composed of $\mathcal{S}_x$ (i.e. $\mathcal{S} = \bigcup_{x\in\mathcal{X}} \mathcal{S}_x$), with $S = s \in \mathcal{S}_x$ if and only if $X = x$. This way, we can construct a bijective function $f : x \mapsto \mathbb{R}$ by $S \sim P_S$ for a random variable X distributed over support $\mathcal{X}$ as

$$f(x) = \sum_{s\in\mathcal{S}_x} P_S(s)^\alpha. \tag{119}$$

Then, a probability distribution $p_{\bar{X}}(x)$ over support $\mathcal{X}$ can be formed by (119) as

$$p_{\bar{X}}(x) = \frac{f(x)}{\sum_{x\in\mathcal{X}} f(x)} = \frac{\sum_{s\in\mathcal{S}_x} P_S(s)^\alpha}{\sum_{x\in\mathcal{X}} \sum_{s\in\mathcal{S}_x} P_S(s)^\alpha} \qquad \forall x \in \mathcal{X}. \tag{120}$$

And ρ_B^s is constructed as

$$\rho_B^s = \begin{cases} \rho_B^x, & \text{for } s \in \mathcal{S}_x \quad (121a) \\ 0, & \text{otherwise.} \quad (121b) \end{cases}$$

Setting $\sigma_{SB} \equiv \sum_{s\in\mathcal{S}} \frac{1}{|\mathcal{S}|} |s\rangle\langle s| \otimes \sigma_B$, we have

$$\begin{aligned}
I_\alpha^{\mathbb{M}}(\bar{X} : B)_\rho &= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^{\mathbb{M}}(\rho_{\bar{X}B}\|\rho_{\bar{X}} \otimes \sigma_B) && (122)\\
&= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log \mathbb{E}_{x\sim p_{\bar{X}}} \left[Q_\alpha^{\mathbb{M}}(\rho_B^x\|\sigma_B)\right] && (123)\\
&= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} \frac{\sum_{s\in\mathcal{S}_x} P_S(s)^\alpha}{\sum_{x\in\mathcal{X}} \sum_{s\in\mathcal{S}_x} P_S(s)^\alpha} Q_\alpha^{\mathbb{M}}(\rho_B^x\|\sigma_B) && (124)\\
&= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log \sum_{s\in\mathcal{S}} P_S(s)^\alpha Q_\alpha^{\mathbb{M}}(\rho_B^s\|\sigma_B) + H_\alpha(S)_p && (125)\\
&= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log \left[\sup_{\{\Pi_B^s\}_s} \sum_{s\in\mathcal{S}} \left(P_S(s) \operatorname{Tr}\left[\rho_B^s (\Pi_B^s)^{\frac{\alpha-1}{\alpha}}\right]\right)^\alpha \left(\frac{1}{|\mathcal{S}|} \operatorname{Tr}\left[\sigma_B \Pi_B^s\right]\right)^{1-\alpha} \right] \\
&\quad - \log|\mathcal{S}| + H_\alpha(S)_p && (126)\\
&= \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha-1} \log Q_\alpha^{\mathbb{M}}(\rho_{SB}\|\sigma_{SB}) - \log|\mathcal{S}| + H_\alpha(S)_p && (127)\\
&= I_\alpha^{\mathrm{A},\mathbb{M}}(S : B)_\rho. && (128)
\end{aligned}$$

[3]Indeed the data-processing operation of Lemma 2 is expressed in terms of a channel. For the Markov chain $\bar{S} - \bar{X} - B$ of the form in (113), we can regard the stochastic transformation $p_{\bar{S}|\bar{X}}$ as a channel $\mathcal{N}_{\mathrm{supp}(p_X)\to\bar{S}} = \sum_s \sum_{x\in\mathrm{supp}(p_X)} p_{\bar{S}|X}(s|x)|s\rangle\langle x| \cdot |x\rangle\langle s|$. Hence, Lemma 2 implies the following inequality $I_\alpha^{\mathbb{M}}(\bar{S} : B)_\rho = \inf_{\sigma_B} D_\alpha^{\mathbb{M}}(\rho_{\bar{S}B}\|\rho_{\bar{S}} \otimes \sigma_B) = \inf_{\sigma_B} D_\alpha^{\mathbb{M}}\left((\mathcal{N}_{\mathrm{supp}(p_X)\to\bar{S}} \otimes \mathrm{id}_B)(\rho_{\bar{X}B})\|(\mathcal{N}_{\mathrm{supp}(p_X)\to\bar{S}} \otimes \mathrm{id}_B)(\rho_{\bar{X}} \otimes \sigma_B)\right) \le \inf_{\sigma_B} D_\alpha^{\mathbb{M}}(\rho_{\bar{X}B}\|\rho_{\bar{X}} \otimes \sigma_B) = I_\alpha^{\mathbb{M}}(\bar{X} : B)_\rho$.

Therefore,

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho \tag{129}$$

$$= \sup_{p_{S|X}: S-X-B} I_\alpha^{\mathrm{A},\mathbb{M}}(S:B)_\rho \tag{130}$$

$$\geq \sup_{\substack{p_{S|X}: S-X-B \\ H_1(X|S)=0}} I_\alpha^{\mathrm{A},\mathbb{M}}(S:B)_\rho \tag{131}$$

$$= \sup_{p_{\bar{X}} \in \mathcal{P}(\mathtt{supp}(p_X))} I_\alpha^{\mathbb{M}}(\bar{X}:B)_\rho \equiv C_\alpha^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}). \tag{132}$$

The last equality holds since for any $p_{\bar{X}} \in \mathcal{P}(\mathtt{supp}(p_X))$, we can construct corresponding $P_S(s)$ for $s \in \mathcal{S}$ that satisfies (119) and (120) by bijection. Thus the supremum over $S - X - B$ and $H_1(X|S) = 0$ in (131) is equivalent to the supremum over $p_{\bar{X}}$ in (132). By combining (118) and (132), we proved Theorem 3. We remark that a similar technique was used in the proof of [17, Thm. 2]. □

F Proof of Theorem 4

Proof.

The proof of part 1: For $\alpha > 1$, the maximal α-leakage can be expressed as

$$\mathcal{L}_\alpha^{\max}(X \to B)_\rho = \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \frac{1}{\alpha - 1} \log \left(\sum_{x \in \mathcal{X}} \bar{p}_X(x) Q_\alpha^{\mathbb{M}}(\rho_B^x \| \sigma_B) \right), \tag{133}$$

where the objective function in (133) is concave in the optimization variable $\bar{p}_X$.

The proof for $\alpha = 1$ follows from the concavity of measured Arimoto information (Def. 5)

$$I_\alpha^{\mathrm{A},\mathbb{M}}(X:B)_\rho = H_\alpha(X)_p + \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_1^{\mathbb{M}}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B). \tag{134}$$

When $\alpha = 1$, it is straightforward that Shannon entropy $H_1(X)$ is concave function of p_X. The concavity of $\inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_1^{\mathbb{M}}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B)$ is from that the variational expression (Lemma 1) is linear in p_X:

$$D_1^{\mathbb{M}}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B) = \sup_{\omega_{XB} > 0} \mathrm{Tr}\left[\rho_{XB} \log \omega_{XB}\right] + 1 - \mathrm{Tr}\left[(\mathbb{1}_X \otimes \sigma_B)\omega_{XB}\right]. \tag{135}$$

Besides, the infimum of arbitrary concave (linear in (135)) functions is still concave. Therefore, the maximal 1-leakage is a concave function of p_X.

The proof of part 2: For two c-q states ρ_{XB} and τ_{XB} which follow the same distribution p_X and $\alpha \in [1, \infty]$, we assume $D_\alpha^{\mathbb{M}}(\rho_{XB} \parallel \rho_X \otimes \sigma_B) \geq D_\alpha^{\mathbb{M}}(\tau_{XB} \parallel \rho_X \otimes \sigma_B)$ without loss of generality. Here we prove the $\alpha = 1$ case. For a c-q state $\beta\rho_{XB} + (1-\beta)\tau_{XB}$ (with $0 \leq \beta \leq 1$) by Definition 1 and Lemma 1, we observe that

$$\inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} D_1^{\mathbb{M}}(\beta\rho_{XB} + (1-\beta)\tau_{XB} \| \rho_X \otimes \sigma_B) \tag{136}$$

$$= \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \sup_{\Pi_{XB} > 0} \mathrm{Tr}\left[(\beta\rho_{XB} + (1-\beta)\tau_{XB}) \log \Pi_{XB}\right] - \log \mathrm{Tr}\left[(\rho_X \otimes \sigma_B)\Pi_{XB}\right] \tag{137}$$

$$\leq \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \sup_{\Pi_{XB} > 0} \mathrm{Tr}\left[(\beta\rho_{XB} + (1-\beta)\rho_{XB}) \log \Pi_{XB}\right] - \log \mathrm{Tr}\left[(\rho_X \otimes \sigma_B)\Pi_{XB}\right] \tag{138}$$

$$= D_1^{\mathbb{M}}(\rho_{XB} \parallel \rho_X \otimes \sigma_B), \tag{139}$$

where the inequality follows from our assumption. Note that the quasi-convexity is preserved under taking infimum [66]. Thus, combine this with Theorem 3, maximal 1-leakage is quasi-convex in ρ_B^x given p_X. The proof of maximal α-leakage for $\alpha > 1$ follows similarly by using Lemma 1 and Theorem 3.

The proof of part 3: Since the first-order derivative of α in the measured Rényi information is non-negative, we immediately see that maximal α-leakage is non-decreasing in α.

The proof of part 4: We can view Markov chain $S - X - B$ as two channels from $S \to X$ and $X \to B$, then by Lemma 2 and Theorem 3, maximal-α leakage satisfies data processing inequality.

The proof of part 5: The lower bound directly follows from the non-negativity of measured Rényi divergence.

To derive the upper bound for $\alpha > 1$, we bound maximal α-leakage by maximal ∞-leakage using monotonicity with respect to tunable parameter α first (Theorem 4, part 3), and adopt the divergence radius expression of maximal ∞-leakage to obtain an optimization problem of maximum relative entropy (defined in Eq. (29)):

$$\begin{aligned}
&\mathcal{L}_\alpha^{\max}(X \to B)_\rho \le \mathcal{L}_\infty^{\max}(X \to B)_\rho && (140)\\
&= R_\infty^{\mathbb{M}}(\mathcal{N}_{\mathtt{supp}(p_X)\to B}) && (141)\\
&= \sup_{\tilde{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \log\inf\{\gamma : \tilde{\rho}_{XB} \le \gamma \tilde{\rho}_X \otimes \sigma_B\} && (142)\\
&= \inf_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} \log\inf\{\gamma : \rho_B^x \le \gamma \sigma_B, \quad \forall\, x \in \mathtt{supp}(p_X)\} && (143)\\
&= \log\inf\{\mathrm{Tr}\,[M] : \rho_B^x \le M, \quad \forall\, x \in \mathtt{supp}(p_X)\} && (144)\\
&\le \log(|\mathtt{supp}(p_X)|), && (145)
\end{aligned}$$

where $\tilde{\rho}_{XB} = \sum_{x\in\mathcal{X}} \tilde{p}_X|x\rangle\langle x| \otimes \rho_B^x$ for every dummy input distribution $\tilde{p}_X$, and in the last inequality we simply take a feasible solution: $M = \sum_{x\in\mathtt{supp}(p_X)} \rho_B^x$.

On the other hand, we can bound $\mathcal{L}_1^{\max}(X \to B)_\rho$ by

$$\begin{aligned}
\mathcal{L}_1^{\max}(X \to B)_\rho &= H_1(X)_p - H_1^{\mathbb{M}}(X|B)_\rho && (146)\\
&= H_1(X)_p + \inf_{\sigma_B} D_1^{\mathbb{M}}(\rho_{XB}\|\mathbb{1}_X \otimes \sigma_B) && (147)\\
&\le H_1(X)_p + D_1^{\mathbb{M}}(\rho_{XB}\|\mathbb{1}_X \otimes \rho_B) && (148)\\
&\overset{(*)}{\le} H_1(X)_p + D_1^{\mathbb{M}}(\rho_{XB}\|\rho_{XB}) = H_1(X)_p, && (149)
\end{aligned}$$

where $(*)$ follows from the reduction criterion of separability [67, Sec. III].

□

G Proof of Theorem 5

Proof. Throughout the proof, we denote by

$$\bar{\rho}_{XB} = \sum_{x\in\mathcal{X}} \bar{p}_X|x\rangle\langle x| \otimes \rho_B^x \tag{150}$$

for any input distribution $\bar{p}_X \in \mathcal{P}(\mathcal{X})$ as a dummy variable in optimization.

For $\alpha \in [1, \infty)$,

$$\mathcal{L}_\alpha^{\max}(X \to B_1, B_2)_{\bar{\rho}} = \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} I_\alpha^{\mathbb{M}}(X : B_1 B_2)_{\bar{\rho}} \tag{151}$$

$$= \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \inf_{\sigma_{B_1}, \sigma_{B_2}} D_\alpha^{\mathbb{M}}(\bar{\rho}_{XB_1B_2} \| \bar{\rho}_X \otimes \sigma_{B_1} \otimes \sigma_{B_2}) \tag{152}$$

$$\overset{(a)}{=} \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \left[\inf_{\sigma_{B_1}} D_\alpha^{\mathbb{M}}(\bar{\rho}_{XB_1} \| \bar{\rho}_X \otimes \sigma_{B_1}) + \inf_{\sigma_{B_2}} D_\alpha^{\mathbb{M}}(\rho_{B_2} \| \sigma_{B_2}) \right] \tag{153}$$

$$\overset{(b)}{\leq} \sup_{\bar{p}_{X_1} \in \mathcal{P}(\mathtt{supp}(p_X))} \inf_{\sigma_{B_1}} D_\alpha^{\mathbb{M}}(\bar{\rho}_{X_1B_1} \| \bar{\rho}_{X_1} \otimes \sigma_{B_1}) + \sup_{\bar{p}_{X_2} \in \mathcal{P}(\mathtt{supp}(p_X))} \inf_{\sigma_{B_2}} D_\alpha^{\mathbb{M}}(\bar{\rho}_{X_2B_2} \| \bar{\rho}_{X_2} \otimes \sigma_{B_2}) \tag{154}$$

$$= \mathcal{L}_\alpha^{\max}(X \to B_1)_\rho + \mathcal{L}_\alpha^{\max}(X \to B_2)_\rho, \tag{155}$$

where: (a) additivity holds from Lemma 3 since the conditional independency of B_1 and B_2 given X in $B_1 - X - B_2$ implies that the optimal measurement $\Pi = \Pi_1 \otimes \Pi_2$ of joint measured Rényi divergence $D_\alpha^{\mathbb{M}}(\rho_{XB_1B_2} \| \rho_X \otimes \sigma_{B_1} \otimes \sigma_{B_2})$ is separable; (b) is from that optimal probability distributions p_{X_1}, p_{X_2} may differ for the two corresponding subproblems.

For $\alpha = \infty$, since maximal α-leakage and measured Rényi information is monotone non-decreasing in α (Theorem 4, part 3), we can exchange the order of taking limit $\alpha \to \infty$ and taking supremum with respect to $\bar{p}_X$ such that

$$\lim_{\alpha \to \infty} \mathcal{L}_\alpha^{\max}(X \to B_1, B_2)_\rho = \sup_{\alpha \in (1,\infty)} \mathcal{L}_\alpha^{\max}(X \to B_1, B_2)_\rho \tag{156}$$

$$= \sup_{\alpha \in (1,\infty)} \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} I_\alpha^{\mathbb{M}}(X : B_1 B_2)_{\bar{\rho}} \tag{157}$$

$$= \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \sup_{\alpha \in (1,\infty)} I_\alpha^{\mathbb{M}}(X : B_1 B_2)_{\bar{\rho}} \tag{158}$$

$$= \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} \lim_{\alpha \to \infty} I_\alpha^{\mathbb{M}}(X : B_1 B_2)_{\bar{\rho}} \tag{159}$$

$$= \sup_{\bar{p}_X \in \mathcal{P}(\mathtt{supp}(p_X))} I_\infty^{\mathbb{M}}(X : B_1 B_2)_{\bar{\rho}}. \tag{160}$$

Now we can apply the similar derivation as in (151) on (160) to obtain composition property for $\alpha = \infty$.

□

H Proof of Theorem 6

We start with the equivalence of regularized measured Rényi divergence and sandwiched Rényi divergence. Next, we prove the equivalence of regularized measured Arimoto information and sandwiched Rényi information (under the α-tilted distribution) in Lemma 9. Finally, with these building blocks at hand, we can prove Theorem 6 and Theorem 7.

Lemma 7 (Regularized measured Rényi divergence [41, Eq. (4.34)])**.** *Given density matrices ρ and σ. For $\alpha \geq \frac{1}{2}$, we have*

$$\lim_{n \to \infty} \frac{1}{n} D_\alpha^{\mathbb{M}} \left(\rho^{\otimes n} \| \sigma^{\otimes n} \right) = D_\alpha^*(\rho \| \sigma). \tag{161}$$

For $\alpha = 1$ case, see Ref. [68, Thm. 2.1], for $\alpha > 1$, see Ref. [52, Thm. 3.7], and for $[\frac{1}{2}, 1)$, see Ref. [69, Prop. 8].

Beigi and Tomamichel [43] proved the equivalence of regularized measured Rényi information and sandwiched Rényi information, which we list as Lemma 8. However, it is the measured Arimoto information that plays a role in characterizing α-leakage $\mathcal{L}_\alpha(X \to B)_\rho$ in this paper. Lemma 9 below shows that its regularization is given by the sandwiched Rényi information, albeit evaluated with the α-tilted distribution $p_X^{(\alpha)}$ instead of the original input distribution p_X.

Lemma 8 (Regularized measured Rényi information [43, Lemma 6])**.** *For a c-q state ρ_{XB} and $\alpha \geq \frac{1}{2}$, we have*

$$\lim_{n\to\infty} \frac{1}{n} I_\alpha^{\mathbb{M}} \left(X^n : B^n\right)_{\rho^{\otimes n}} = I_\alpha^*(X:B)_\rho. \tag{162}$$

Lemma 9 (Regularized measured Arimoto information)**.** *For a c-q state $\rho_{XB} = \sum_{x\in\mathcal{X}} p_X(x)|x\rangle\langle x| \otimes \rho_B^x$ and $\alpha \geq \frac{1}{2}$, we have*

$$\lim_{n\to\infty} \frac{1}{n} I_\alpha^{\mathrm{A},\mathbb{M}} \left(X^n : B^n\right)_{\rho^{\otimes n}} = I_\alpha^*(X:B)_{\rho^{(\alpha)}}, \tag{163}$$

where $\rho_{XB}^{(\alpha)} \equiv \sum_x p_X^{(\alpha)}(x)\,|x\rangle\langle x| \otimes \rho_B^x$ and $p_X^{(\alpha)}$ is the α-tilted distribution (Def. 2).

Proof. For $\alpha \geq 1/2$, we can expand the definition of regularized measured Arimoto information as

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} I_\alpha^{\mathrm{A},\mathbb{M}} \left(X^n : B^n\right)_{\rho^{\otimes n}} && (164)\\
&= \lim_{n\to\infty} \frac{1}{n} (H_\alpha(X^n)_p - H_\alpha^{\mathbb{M}}(X|B)_{\rho^{\otimes n}}) && (165)\\
&\overset{\text{(a)}}{=} H_\alpha(X)_p + \lim_{n\to\infty} \frac{1}{n} \inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}_{B^n})} D_\alpha^{\mathbb{M}}(\rho_{XB}^{\otimes n} \| \mathbb{1}_X^{\otimes n} \otimes \sigma_{B^n}) && (166)\\
&\overset{\text{(b)}}{=} H_\alpha(X)_p + \lim_{n\to\infty} \frac{1}{n} \inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}_B^{\otimes n})} D_\alpha^{\mathbb{M}}(\rho_{XB}^{\otimes n} \| \mathbb{1}_X^{\otimes n} \otimes \sigma_{B^n}) && (167)\\
&\overset{\text{(c)}}{=} H_\alpha(X)_p + \inf_{\sigma_B} D_\alpha^*(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B) && (168)\\
&\overset{\text{(d)}}{=} \frac{-1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X(x)^\alpha + \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} p_X(x)^\alpha Q_\alpha^*(\rho_B^x \| \sigma_B) && (169)\\
&\overset{\text{(e)}}{=} \inf_{\sigma_B} \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} \frac{p_X(x)^\alpha}{\sum_{x\in\mathcal{X}} p_X(x)^\alpha} Q_\alpha^*(\rho_B^x \| \sigma_B) && (170)\\
&= \inf_{\sigma_B} \frac{1}{\alpha-1} \log \mathbb{E}_{x\sim p_X^{(\alpha)}} \left[Q_\alpha^*(\rho_B^x \| \sigma_B)\right] = I_\alpha^*(X:B)_{\rho^{(\alpha)}} && (171)
\end{aligned}$$

where the first term of (a) follows from the additivity of H_α and $p_{X^n} = p_X^{\otimes n}$. The second term of (b) first follows from the observation that the minimizer of σ_{B^n} is invariant under permutations of subsystems, which is applied in [43, Lemma 6]; then, by quantum de Finetti theorem in [70], the state σ_B^n can be substituted by $\sigma_B^{\otimes n}$ with sub-linear cost, which vanishes when $n \to \infty$. (c) is from Lemma 7. In (d), for density matrices ρ and σ, the sandwiched quasi Rényi divergence is defined as

$$Q_\alpha^*(\rho\|\sigma) := \begin{cases} \mathrm{Tr}\left[(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}})^\alpha\right], & \text{for } \alpha \in (0,1) \text{ or for } \alpha \in (1,\infty),\ \mathtt{supp}(\rho) \subseteq \mathtt{supp}(\sigma); \quad (172a)\\ \infty, & \text{otherwise.} \quad (172b) \end{cases}$$

(e) is from direct-sum property of sandwiched Rényi divergence, and one can refer to [71, Proposition 7.29]. □

Thus, we obtain Theorem 6 by combining Lemma 5, Lemma 9, and Theorem 2.

I Proof of Proposition 8

The first equality of (69) directly follows from Theorem 3. We establish the second equality of (69) as follows. Throughout the proof, we denote by

$$\bar{\rho}^n_{X^nB^n} = \sum_{x^n\in\mathcal{X}^n} \bar{p}_{X^n}|x^n\rangle\langle x^n| \otimes \rho^{x^n}_{B^n} \tag{173}$$

for any input distribution $\bar{p}_{X^n} \in \mathcal{P}(\mathcal{X}^n)$ as a dummy variable in optimization.

The upper bound of regularized Rényi capacity is straightforward. We have, for any $n \in \mathbb{N}$,

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}C^{\mathbb{M}}_\alpha\left(\mathcal{N}^{\otimes n}_{\mathcal{X}\to B}\right) &= \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_{X^n}}\inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}^{\otimes n})} D^{\mathbb{M}}_\alpha(\bar{\rho}_{X^nB^n}\|\bar{\rho}_{X^n}\otimes\sigma_{B^n}) && (174)\\
&\le \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_{X^n}}\inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}^{\otimes n})} D^{*}_\alpha(\bar{\rho}_{X^nB^n}\|\bar{\rho}_{X^n}\otimes\sigma_{B^n}) && (175)\\
&= \lim_{n\to\infty}\frac{1}{n}C^{*}_\alpha\left(\mathcal{N}^{\otimes n}_{\mathcal{X}\to B}\right) && (176)\\
&\overset{(*)}{=} C^{*}_\alpha\left(\mathcal{N}_{\mathcal{X}\to B}\right), && (177)
\end{aligned}$$

where the inequality comes from (30), and $(*)$ comes from the additivity of sandwiched Rényi capacity [72, Lemma 6] for i.i.d. c-q channels $\mathcal{N}^{\otimes n}_{\mathcal{X}\to B}$.

For the lower bound, let $|\mathrm{spec}(\sigma)|$ denote the number of mutually different eigenvalues of σ, and let $\mathcal{P}_M(\cdot) := \sum_i P_i \cdot P_i$ be the *pinching map* with respect to Hermitian M which has eigen-projections $\{P_i\}_i$. Then, we calculate

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}C^{\mathbb{M}}_\alpha\left(\mathcal{N}^{\otimes n}_{\mathcal{X}\to B}\right) &= \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_{X^n}\in\mathcal{P}(\mathcal{X}^n)}\inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}^{\otimes n})} D^{\mathbb{M}}_\alpha(\bar{\rho}_{X^nB^n}\|\bar{\rho}_{X^n}\otimes\sigma_{B^n}) && (178)\\
&\overset{(a)}{\ge} \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})}\inf_{\sigma_{B^n}\in\mathcal{S}(\mathcal{H}^{\otimes n})} D^{\mathbb{M}}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma_{B^n}) && (179)\\
&\overset{(b)}{=} \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} D^{\mathbb{M}}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}) && (180)\\
&\overset{(c)}{\ge} \lim_{n\to\infty}\frac{1}{n}\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} D^{*}_\alpha\left(\mathcal{P}_{\bar{\rho}^{\otimes n}_X\otimes\sigma^{\star}_{B^n}}(\bar{\rho}^{\otimes n}_{XB})\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}\right) && (181)\\
&\overset{(d)}{\ge} \lim_{n\to\infty}\frac{1}{n}\left[\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} D^{*}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}) - 2\log|\mathrm{spec}(\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n})|\right] && (182)\\
&\overset{(e)}{=} \lim_{n\to\infty}\frac{1}{n}\left[\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} D^{*}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}) - \log\mathrm{poly}(n)\right] && (183)\\
&\overset{(f)}{\ge} \lim_{n\to\infty}\frac{1}{n}\left[\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} I^{*}_\alpha(X^n;B^n)_{\bar{\rho}^{\otimes n}} - \log\mathrm{poly}(n)\right] && (184)\\
&\overset{(g)}{=} \lim_{n\to\infty}\sup_{\bar{p}_X\in\mathcal{P}(\mathcal{X})} I^{*}_\alpha(X:B)_{\bar{\rho}} - \frac{1}{n}\log\mathrm{poly}(n) && (185)\\
&= C^{*}_\alpha\left(\mathcal{N}_{\mathcal{X}\to B}\right), && (186)
\end{aligned}$$

where (a) is from restricting the optimization $\bar{p}_{X^n} \in \mathcal{P}(\mathcal{X}^n)$ to i.i.d. inputs $\bar{p}^{\otimes n}_X$; in (b) we denote $\sigma^{\star}_{B^n} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ as the permutation-invariant minimizer to the map $\sigma_{B^n} \mapsto D^{\mathbb{M}}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma_{B^n})$ (which depends on $\bar{p}_X$; see e.g. [43, Eq. (58)]); (c) from data-processing inequality of applying the pinching map $\mathcal{P}_{\bar{\rho}^{\otimes n}_X\otimes\sigma^{\star}_{B^n}}$ to force the resulting states on system B^n to commute (see e.g. [43, Lemma 6]), i.e.

$$D^{\mathbb{M}}_\alpha(\bar{\rho}^{\otimes n}_{XB}\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}) \ge D^{\mathbb{M}}_\alpha\left(\mathcal{P}_{\bar{\rho}^{\otimes n}_X\otimes\sigma^{\star}_{B^n}}(\bar{\rho}^{\otimes n}_{XB})\|\bar{\rho}^{\otimes n}_{X}\otimes\omega_{B^n}\right) = D^{*}_\alpha\left(\mathcal{P}_{\bar{\rho}^{\otimes n}_X\otimes\sigma^{\star}_{B^n}}(\bar{\rho}^{\otimes n}_{XB})\|\bar{\rho}^{\otimes n}_{X}\otimes\sigma^{\star}_{B^n}\right). \tag{187}$$

(d) is because measured Rényi divergence can be lower-bounded by sandwiched Rényi divergence subtracted by a logarithmic number of spectrum [69, Lemma 3]:

$$D_\alpha^{\mathbb{M}}(\rho\|\sigma) \geq D_\alpha^*(\mathcal{P}_\sigma(\rho)\|\sigma) \geq D_\alpha^*(\rho\|\sigma) - \begin{cases} \log|\mathrm{spec}(\sigma)|, & \text{if } \alpha \in [0,2]; \quad (188a) \\ 2\log|\mathrm{spec}(\sigma)|, & \text{if } \alpha > 2, \quad (188b) \end{cases}$$

which results from pinching inequality [73, Lemma 3.10]:

$$\rho \leq |\mathrm{spec}(\sigma)|\mathcal{P}_\sigma(\rho); \tag{189}$$

(e) is from the number of distinct eigenvalues of the permutation-invariant minimizer $|\mathrm{spec}(\bar{\rho}_X^{\otimes n} \otimes \sigma_{B^n}^\star)|$ grows polynomially with n, since from Schur duality, denoting d as the dimension of finite-dimension Hilbert space $\mathcal{H}_X \otimes \mathcal{H}_B$, and let $\Lambda_n = \{\lambda = (\lambda_1, ..., \lambda_d) | \lambda_1 \geq ... \geq \lambda_d \geq 0 \text{ and } \sum_{i=1}^n \lambda_i = n\}$ be partitions of n into $\leq d$ parts, $\bar{\rho}_X^{\otimes n} \otimes \sigma_{B^n}^\star$ can be represented in Schur basis P_λ and Q_λ^d (see e.g. [74, Eq. (5.16), (5.18)][75, Eq. (48)]):

$$\sum_{\lambda \in \Lambda_n} |\lambda\rangle\langle\lambda| \otimes \sigma_{Q_\lambda^d} \otimes \mathbb{1}_{P_\lambda}, \tag{190}$$

where one can upper-bound these quantities as follows [74, Sec. 6.2]:

$$|\Lambda_n| \leq (n+1)^d \in \mathrm{poly}(n); \tag{191}$$

$$\dim Q_\lambda^d \leq (n+d)^{\frac{d(d-1)}{2}} \in \mathrm{poly}(n). \tag{192}$$

Note that this holds for any $\bar{p}_X \in \mathcal{P}(\mathcal{X})$. We remark that a similar analysis of this step was also used in [75, Lemma 2.4].

(f) follows from the minimization in the definition of measured Rényi information; (g) from additivity of sandwiched Rényi information under product states [45, Lemma 4.8]. We remark that a similar technique was applied in the proof of [45, Thm. 5.14].

Since both the upper bound and the lower bound are given by $C_\alpha^*(\mathcal{N}_{\mathcal{X}\to B})$, this concludes the proof.

References

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

[2] M. Hu, "Cambridge analytica's black box," *Big Data & Society*, vol. 7, no. 2, p. 2053951720938091, 2020.

[3] K. Abouelmehdi, A. Beni-Hessane, and H. Khaloufi, "Big healthcare data: preserving security and privacy," *Journal of big data*, vol. 5, no. 1, pp. 1–18, 2018.

[4] C. Dwork, "Differential privacy," in *International colloquium on automata, languages, and programming*. Springer, 2006, pp. 1–12.

[5] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[6] D. Kifer and A. Machanavajjhala, "Pufferfish: A framework for mathematical privacy definitions," *ACM Transactions on Database Systems (TODS)*, vol. 39, no. 1, pp. 1–36, 2014.

[7] T. Nuradha and Z. Goldfeld, "Pufferfish privacy: An information-theoretic study," *IEEE Transactions on Information Theory*, 2023.

[8] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

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

[10] 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.

[11] 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.

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

[13] H. Wang and F. P. Calmon, "An estimation-theoretic view of privacy," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 886–893.

[14] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Privacy-aware guessing efficiency," in *2017 ieee international symposium on information theory (isit)*. IEEE, 2017, pp. 754–758.

[15] ——, "Estimation efficiency under privacy constraints," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1512–1534, 2018.

[16] 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, 2019.

[17] 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, 2019.

[18] A. Gilani, G. R. Kurri, O. Kosut, and L. Sankar, "(α, β)-leakage: A unified privacy leakage measure," *arXiv preprint arXiv:2304.07456*, 2023.

[19] J. Liao, L. Sankar, O. Kosut, and F. P. Calmon, "Maximal α-leakage and its properties," in *2020 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 2020, pp. 1–6.

[20] T. Sypherd, M. Diaz, J. K. Cava, G. Dasarathy, P. Kairouz, and L. Sankar, "A tunable loss function for robust classification: Calibration, landscape, and generalization," *IEEE Transactions on Information Theory*, vol. 68, no. 9, pp. 6021–6051, 2022.

[21] A. Kamatsuka, T. Yoshida, and T. Matsushima, "A generalization of the stratonovich's value of information and application to privacy-utility trade-off," in *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2022, pp. 1999–2004.

[22] G. R. Kurri, L. Sankar, and O. Kosut, "An operational approach to information leakage via generalized gain functions," *IEEE Transactions on Information Theory*, 2023.

[23] 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, 2023.

[24] A. Zamani, T. J. Oechtering, and M. Skoglund, "Private variable-length coding with non-zero leakage," *arXiv preprint arXiv:2310.19122*, 2023.

[25] ——, "New privacy mechanism design with direct access to the private data," *arXiv preprint arXiv:2309.09033*, 2023.

[26] L. Zhou and M. Ying, "Differential privacy in quantum computation," in *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*. IEEE, 2017, pp. 249–262.

[27] S. Aaronson and G. N. Rothblum, "Gentle measurement of quantum states and differential privacy," in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, 2019, pp. 322–333.

[28] W. M. Watkins, S. Y.-C. Chen, and S. Yoo, "Quantum machine learning with differential privacy," *Scientific Reports*, vol. 13, no. 1, p. 2453, 2023.

[29] A. Angrisani, M. Doosti, and E. Kashefi, "Differential privacy amplification in quantum and quantum-inspired algorithms," *arXiv preprint arXiv:2203.03604*, 2022.

[30] Y. Du, M.-H. Hsieh, T. Liu, S. You, and D. Tao, "Quantum differentially private sparse regression learning," *IEEE Transactions on Information Theory*, vol. 68, no. 8, pp. 5217–5233, 2022.

[31] C. Hirche, C. Rouzé, and D. S. França, "Quantum differential privacy: An information theory perspective," *IEEE Transactions on Information Theory*, 2023.

[32] T. Nuradha, Z. Goldfeld, and M. M. Wilde, "Quantum pufferfish privacy: A flexible privacy framework for quantum systems," *arXiv preprint arXiv:2306.13054*, 2023.

[33] F. Farokhi, "Maximal quantum information leakage," *arXiv preprint arXiv:2307.12529*, 2023.

[34] J. L. Massey, "Guessing and entropy," in *Proceedings of 1994 IEEE International Symposium on Information Theory*. IEEE, 1994, p. 204.

[35] W. Chen, Y. Cao, H. Wang, and Y. Feng, "Minimum guesswork discrimination between quantum states," *arXiv preprint arXiv:1410.5180*, 2014.

[36] E. P. Hanson, V. Katariya, N. Datta, and M. M. Wilde, "Guesswork with quantum side information," *IEEE Transactions on Information Theory*, vol. 68, no. 1, pp. 322–338, 2021.

[37] A. M. GLEASON, "Measures on the closed subspaces of a hilbert space," *Journal of Mathematics and Mechanics*, vol. 6, no. 6, pp. 885–893, 1957. [Online]. Available: http://www.jstor.org/stable/24900629

[38] 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.

[39] R. König, R. Renner, and C. Schaffner, "The operational meaning of min-and max-entropy," *IEEE Transactions on Information theory*, vol. 55, no. 9, pp. 4337–4347, 2009.

[40] M. Berta, O. Fawzi, and M. Tomamichel, "On variational expressions for quantum relative entropies," *Letters in Mathematical Physics*, vol. 107, pp. 2239–2265, 2017.

[41] F. Hiai and M. Mosonyi, "Different quantum f-divergences and the reversibility of quantum operations," *Reviews in Mathematical Physics*, vol. 29, no. 07, p. 1750023, 2017.

[42] A. Rényi, "On measures of entropy and information," *Proc. 4th Berkeley Symp. on Math. Statist. Probability*, vol. 1, pp. 547–561, 1962.

[43] S. Beigi and M. Tomamichel, "Lower bounds on error exponents via a new quantum decoder," *arXiv preprint arXiv:2310.09014*, 2023.

[44] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," in *Topics in Information Theory, Proc. Coll. Math. Soc. János Bolyai*, 1975, pp. 41–52.

[45] M. Mosonyi and T. Ogawa, "Strong converse exponent for classical-quantum channel coding," *Communications in Mathematical Physics*, vol. 355, pp. 373–426, 2017.

[46] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum rényi entropies: A new generalization and some properties," *Journal of Mathematical Physics*, vol. 54, no. 12, Dec. 2013. [Online]. Available: http://dx.doi.org/10.1063/1.4838856

[47] M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy," *Communications in Mathematical Physics*, vol. 331, no. 2, p. 593–622, Jul. 2014. [Online]. Available: http://dx.doi.org/10.1007/s00220-014-2122-x

[48] S. Beigi, "Sandwiched rényi divergence satisfies data processing inequality," *Journal of Mathematical Physics*, vol. 54, no. 12, 2013.

[49] H. Umegaki, "Conditional expectation in an operator algebra," *Tohoku Mathematical Journal, Second Series*, vol. 6, no. 2-3, pp. 177–181, 1954.

[50] H.-C. Cheng, L. Gao, and M.-H. Hsieh, "Properties of noncommutative rényi and Augustin information," *Communications in Mathematical Physics*, feb 2022.

[51] K. Li and Y. Yao, "Operational interpretation of the sandwiched r\'enyi divergences of order 1/2 to 1 as strong converse exponents," *arXiv preprint arXiv:2209.00554*, 2022.

[52] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum rényi relative entropies," *Communications in Mathematical Physics*, vol. 334, no. 3, p. 1617–1648, Dec. 2014. [Online]. Available: http://dx.doi.org/10.1007/s00220-014-2248-x

[53] C. A. Fuchs, "Distinguishability and accessible information in quantum theory," *arXiv preprint quant-ph/9601020*, 1996.

[54] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2816–2826, Jun 2009.

[55] S. Verdú, "α-mutual information," in *2015 Information Theory and Applications Workshop (ITA)*. IEEE, feb 2015.

[56] F. Hiai, *Quantum f-Divergences in von Neumann Algebras*, 1st ed., ser. Mathematical Physics Studies. Singapore, Singapore: Springer, Jan. 2021.

[57] H.-C. Cheng and L. Gao, "On strong converse theorems for quantum hypothesis testing and channel coding," *arXiv preprint quant-ph/2403.13584*, 2024. [Online]. Available: https://arxiv.org/abs/2403.13584

[58] A. Kamatsuka, Y. Ishikawa, K. Kazama, and T. Yoshida, "New algorithms for computing sibson capacity and arimoto capacity," *arXiv preprint arXiv:2401.14241*, 2024.

[59] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[60] P. Kairouz, S. Oh, and P. Viswanath, "The composition theorem for differential privacy," in *International conference on machine learning*. PMLR, 2015, pp. 1376–1385.

[61] N. Datta, M.-H. Hsieh, M. M. Wilde, and A. Winter, "Quantum-to-classical rate distortion coding," *Journal of Mathematical Physics*, vol. 54, no. 4, 2013.

[62] A. Agrawal and S. Boyd, "Disciplined quasiconvex programming," *Optimization Letters*, vol. 14, pp. 1643–1657, 2020.

[63] J. Schindler and A. Winter, "Continuity bounds on observational entropy and measured relative entropies," *arXiv preprint arXiv:2302.00400*, 2023.

[64] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*. Springer International Publishing, 2014. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-04150-6

[65] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995, vol. 4.

[66] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[67] M. l Horodecki and P. l Horodecki, "Reduction criterion of separability and limits for a class of protocols of entanglement distillation," *arXiv preprint arXiv:9708015v3*, 1999.

[68] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Communications in mathematical physics*, vol. 143, pp. 99–114, 1991.

[69] M. Hayashi and M. Tomamichel, "Correlation detection and an operational interpretation of the rényi mutual information," *Journal of Mathematical Physics*, vol. 57, no. 10, Oct. 2016. [Online]. Available: http://dx.doi.org/10.1063/1.4964755

[70] C. M. Caves, C. A. Fuchs, and R. Schack, "Unknown quantum states: The quantum de finetti representation," *Journal of Mathematical Physics*, vol. 43, no. 9, p. 4537–4559, Sep. 2002. [Online]. Available: http://dx.doi.org/10.1063/1.1494475

[71] S. Khatri and M. M. Wilde, "Principles of quantum communication theory: A modern approach," 2021.

[72] M. K. Gupta and M. M. Wilde, "Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity," *Communications in Mathematical Physics*, vol. 334, pp. 867–887, 2015.

[73] M. Hayashi, *Quantum information theory.* Springer, 2016.

[74] A. W. Harrow, "Applications of coherent classical communication and the schur transform to quantum information theory," *arXiv preprint quant-ph/0512255*, 2005.

[75] M. Berta, F. G. Brandao, and C. Hirche, "On composite quantum hypothesis testing," *Communications in Mathematical Physics*, vol. 385, no. 1, pp. 55–77, 2021.