Information Theoretic Security for Encryption Based on Conditional Rényi Entropies

Mitsugu Iwamoto[1] and Junji Shikata[2]

[1] Center for Frontier Science and Engineering,
University of Electro-Communications, Japan
mitsugu@uec.ac.jp
[2] Graduate School of Environment and Information Sciences,
Yokohama National University, Japan
shikata@ynu.ac.jp

Abstract. In this paper, information theoretic cryptography is discussed based on conditional Rényi entropies. Our discussion focuses not only on cryptography but also on the definitions of conditional Rényi entropies and the related information theoretic inequalities. First, we revisit conditional Rényi entropies, and clarify what kind of properties are required and actually satisfied. Then, we propose security criteria based on Rényi entropies, which suggests us deep relations between (conditional) Rényi entropies and error probabilities by using several guessing strategies. Based on these results, unified proof of impossibility, namely, the lower bounds of key sizes is derived based on conditional Rényi entropies. Our model and lower bounds include the Shannon's perfect secrecy, and the min-entropy based encryption presented by Dodis, and Alimomeni and Safavi-Naini. Finally, new optimal symmetric key cryptography and almost optimal secret sharing schemes are proposed which achieve our lower bounds.

Keywords: Information Theoretic Cryptography, (Conditional) Rényi entropy, Error probability in guessing, Impossibility, Symmetric-key Encryption, Secret Sharing Schemes

1 Introduction

1.1 Motivation and Related Works

How to measure the quantities of information is an important issue not only in information theory, but also in cryptography because information measures in cryptography tell us not only the coding efficiency but also security level in terms of equivocation of secret information. Historically, Shannon entropy [1] is the measure of information theoretic cryptography. On the other hand, it is also important to evaluate the cardinality of a set in which a random variable takes values, i.e., Hartley entropy [2]. Furthermore, min-entropy [3] is also considered to be an important quantity in *guessing* the secret in the context of cryptography.

For instance, consider the case of symmetric-key encryption. As is well known by Shannon's seminal work [4], the perfect secrecy in symmetric-key encryption is formalized as $H(M) = H(M|C)$, where M and C are random variables which take values on sets of plaintexts and ciphertexts, respectively; and then, symmetric-key encryption with perfect secrecy implies the lower bound on secret-keys $H(K) \geq H(M)$ (Shannon's bound, Shannon's impossibility, [4]). Similarly, we also know that the number of key candidates can be no less than the cardinality of message set. Furthermore, Dodis [5] recently showed that the similar property also holds with respect to min-entropy. Namely, he showed the bound on secret-keys, $R_\infty(K) \geq R_\infty(M)$, for symmetric-key encryption with perfect secrecy. Also, Alimomeni and Safavi-Naini [6] introduced the *guessing secrecy*, formalized by $R_\infty(M) = R_\infty(M|C)$, and under which they derived the bound $R_\infty(K) \geq R_\infty(M)$, where $R_\infty(\cdot)$ and $R_\infty(\cdot|\cdot)$ are the min-entropy and the conditional min-entropy, respectively. Here, it is worth noting that the above results are proved utilizing *totally* different techniques. This fact is very interesting from the *theoretical* viewpoint, and it must be fruitful not only for cryptography but also for information theory if we can *unify* the above proofs and derive them as corollaries. In order to unify them, Rényi entropy [7] might be useful since it is considered to be a generalization of Shannon, min, and several other kinds of entropies as well as the cardinality.

However, unfortunately, we cannot expect Rényi entropies to satisfy rich properties like Shannon entropies, since Rényi entropies are obtained axiomatically from several relaxed postulates for Shannon entropy. Due to this fact, *subadditivity* does not hold for Rényi entropy although it is very fundamental property of Shannon entropy. Hence, it is not so easy to unify the above different kinds of proofs in terms of Rényi entropies. Even worse, the definition of *conditional* Rényi entropy is not uniquely determined. In order to understand the conditional Rényi entropies, the results by Teixeira et al. [8] are very useful. In [8], the relations among *three* different kinds of conditional Rényi entropies and *four* different kinds of conditional min-entropies are discussed. However, the authors missed to include another different definition of conditional Rényi entropies provided in [9, 17]. Moreover, they did not find the definitions of conditional Rényi entropies corresponding to several conditional min-entropies [10, 28] introduced in cryptographic contexts. Finding reasonable explanations for these min-entropies is also an important contribution since these relations actually bridges information theoretic conditional Rényi entropies and cryptographically important min-entropies.

Finally, note that constructing a unified framework of information theoretic cryptography based on conditional Rényi entropies is not only theoretically interesting but also practically important, because measuring the security by (conditional) Rényi entropies offers us a new security criteria. We should note that the defining security by conditional Rényi entropies instead of conditional Shannon entropies, the security criteria is weaken compared to perfect secrecy, and hence, the wide class of cryptosystems can be discussed in terms of Rényi entropies. In particular, discussing min-entropy criteria is very important since the attacker will guess the key with the highest probability (called *guessing secrecy*). From this viewpoint, the security should be measured by min-entropy instead of Shannon entropy. Although this fact was pointed out by Alimomeni and Safavi-Naini [6], the construction of encryption satisfying the guessing secrecy criteria is not provided in the literature. Hence, it is an very interesting open problem to design information theoretic cryptography under Rényi entropies security criteria as well as the constructions meeting tightly the lower bounds of key size measured by Rényi entropies or min-entropies.

1.2 Our Contributions and Organization of This Paper

Conditional Rényi entropies, revisited (Sections 2 and 3) In [8], Teixeira et al. analyzed the relations among exiting conditional Rényi entropies. However, their analyses are not sufficient in three aspects. First, they do not care about the implications of their results deeply. Recall that Rényi entropies are originally discovered [7] axiomatically, and a lot of nice properties are known for Shannon entropy which is a special case of Rényi entropy. Then, it is necessary to discuss conditional Rényi entropies from axiomatic and/or technological viewpoints. Second, the analysis in [8] missed to include two important conditional Rényi entropies due to Arimoto [17] and Hayashi [9] denoted by $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$, respectively, which are introduced in information theoretic and/or cryptographic contexts . Third, cryptographically important conditional min-entropies are not sufficiently analyzed in [8] since they cannot be not obtained from the conditional Rényi entropies discussed in [8].

Based on the above motivations, we will discuss what kind of properties should be investigated in this paper from the axiomatic, information theoretic, and cryptographic viewpoints. Our analysis also includes $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$. In Sect. 2.3, we start our discussion from the postulates required for Shannon and Rényi entropies, and discuss what kind of properties should be required and/or are interested. Then, we consider the relation between conditional Rényi entropies and conditional min-entropies. We clarify that the conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ related to the conditional min entropies useful in cryptographic context. As a result, we conclude that non-negativity, monotonicity, conditioning reduces entropy (CRE), data processing inequality (DPI) are hopefully required, but the chain rule might not be satisfied. Actually, we will show in Sect. 2.4 that the chain rule does not hold generally in the case of (conditional) Rényi entropies.

Sections 3.1–3.3 are devoted to show that the above inequalities actually hold. Furthermore, we show an extension of Fano's inequality [12] for conditional Rényi entropies in Section 3.4, which will be useful in the forthcoming discussion as well as the inequalities discussed in Sections 3.1–3.3.

Proposal of security criteria based on conditional Rényi entropies (Section 4) In this paper, we propose security criteria based on conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$. Our motivation and significance for proposing it lies in the following two points.

The first point lies in realistic significance which is deeply related to guessing probability by adversaries. Owing to theoretical results about the conditional Rényi entropies in Sections 2 and 3, we will show that conditional Rényi entropies, $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$, play an important role to derive a lower bound on failure of guessing by adversaries, and it turns out that our security criteria is a sufficient condition to make it reasonably large enough. Our way of thinking of this is deeply related to the approach to show the converse of channel coding theorem by Shannon [1] and the recent one to show the converse of channel coding theorem in finite blocklength regime [30, 31] in information theory.

The second point lies in mathematical importance for generalizing Shannon's impossibility (or Shannon's bounds) $H(K) \geq H(M)$ in symmetric-key encryption with perfect secrecy. For details about this contribution, see below.

Generalizing Shannon's impossibility in encryption and secret sharing (Sections 5-7) One of our main purpose in this paper is to generalize Shannon's impossibility (or Shannon's bound) $H(K) \geq H(M)$ in perfectly secure symmetric-key encryption so that all known bounds (i.e., the Shannon's, Dodis's, and Alimomeni and Safavi-Naini's bounds) are captured in our generic bound. By utilizing information-theoretic results about conditional Rényi entropies obtained in Sections 2 and 3, we extend Shannon's impossibility result for encryption by a generic and unified proof technique, and it turns out that our new bound includes all the bounds mentioned above (i.e., the bounds by Shannon, Dodis, and Alimomeni and Safavi-Naini) as special cases. In addition, we apply our discussion in encryption to the case of secret sharing protocols, and we show similar results even for secret sharing protocols. Furthermore, we slightly extend our bound in terms of conditional Rényi entropies to the one under a class of conditional entropy functions which is naturally characterized from axiomatic consideration in Section 2.3.

2 Conditional Rényi Entropies, Revisited

2.1 Preliminaries: Rényi Entropies and α-divergence

Definition 1 (Rényi entropy, [7]) *Let X be a random variable taking values on a finite set $\mathcal{X}$. For a real number $\alpha \geq 0$, the Rényi entropy of order α is defined by*[3]

$$R_\alpha(X) := \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} P_X(x)^\alpha.$$

It is well known that many information measures such as Hartley entropy, Shannon entropy, collision entropy, and min-entropies are special cases of Rényi entropy. Namely, they are respectively obtained by $R_0(X) = \log|\mathcal{X}|$, $R_1(X) := \lim_{\alpha \to 1} R_\alpha(X) = H(X)$, $R_2(X) = -\log \Pr\{X = X'\}$, and $R_\infty(X) := \lim_{\alpha \to \infty} R_\alpha(X) = \min_{x \in \mathcal{X}}\{-\log P_X(x)\}$, where X and X' are independently and identically distributed (i.i.d.) random variables, and $H(X) := -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$ is Shannon entropy.

In the forthcoming discussion, the *α-divergence* (also called Rényi divergence of order α or the normalized Chernoff α-divergence) is important.

[3] Throughout of the paper, the base of logarithm is e. Note that the base of logarithm is not essential since the same arguments hold for arbitrary base of logarithm.

Definition 2 **(α-divergence)** *Let X and Y be random variables taking values on a finite set $\mathcal{X}$. For a real number $\alpha \geq 0$, the α-divergence is defined by*

$$D_\alpha(X\|Y) = D_\alpha(P_X(\cdot)\|P_Y(\cdot)) = \frac{1}{\alpha - 1} \log \sum_{x \in \mathcal{X}} \frac{P_X(x)^\alpha}{P_Y(x)^{\alpha-1}}. \tag{1}$$

In particular, binary α-divergence is analogously defined as $d_\alpha(p\|q) := D_\alpha([p, 1-p]\|[q, 1-q]) = (\alpha-1)^{-1} \log \left\{p^\alpha q^{1-\alpha} + (1-p)^\alpha (1-q)^{1-\alpha}\right\}$.

The α-divergence is considered as an generalization of Kullback-Leibler divergence defined by $D(X\|Y) := \sum_{x \in \mathcal{X}} P_X(x) \log(P_X(x)/P_Y(x))$ since it holds that $\lim_{\alpha \to 1} D_\alpha(X\|Y) = D(X\|Y)$. Note that the α-divergence is nonnegative for all $\alpha \geq 0$. We also note that α-divergence is equal to 0 if and only if $P_X(\cdot) = P_Y(\cdot)$, similarly to Kullback-Leibler divergence.

2.2 Definitions of Conditional Rényi Entropies

Similarly to Shannon entropy, it is natural to consider the *conditional* Rényi entropies. However, several definitions of conditional Rényi entropies have been proposed, e.g.,[17], [18], [19, 20], [21], and [9]. In particular, relations and properties are discussed in [8] among three kinds of conditional Rényi entropies such as

$$R_\alpha^{\mathsf{C}}(X|Y) := \sum_{y \in \mathcal{Y}} P_Y(y) R_\alpha(X|Y = y) \tag{2}$$

$$R_\alpha^{\mathsf{JA}}(X|Y) := R_\alpha(XY) - R_\alpha(Y) \tag{3}$$

$$R_\alpha^{\mathsf{RW}}(X|Y) := 1/(1-\alpha) \max_{y \in \mathcal{Y}} \log \sum_{x \in \mathcal{X}} P_{X|Y}(x|y)^\alpha \tag{4}$$

defined in [18], [19, 20], and [21], respectively. The definitions $R_\alpha^{\mathsf{C}}(X|Y)$ and $R_\alpha^{\mathsf{JA}}(X|Y)$ can be interpreted as extensions of conditional Shannon entropy since they are analogues of $H(X|Y) := \sum_{y \in \mathcal{Y}} P_Y(y) H(X|Y = y)$ and $H(X|Y) := H(XY) - H(Y)$, respectively. The third definition $R_\alpha^{\mathsf{RW}}(X|Y)$ is obtained by letting $\varepsilon = 0$ of the conditional smooth Rényi entropy [21].

Moreover, there are two conditional Rényi entropies are known other than the above. They are defined as

$$R_\alpha^{\mathsf{A}}(X|Y) := \frac{\alpha}{1-\alpha} \log \sum_{y \in \mathcal{Y}} P_Y(y) \left\{ \sum_{x \in \mathcal{X}} P_{X|Y}(x|y)^\alpha \right\}^{1/\alpha} \tag{5}$$

$$R_\alpha^{\mathsf{H}}(X|Y) := \frac{1}{1-\alpha} \log \sum_{y \in \mathcal{Y}} P_Y(y) \sum_{x \in \mathcal{X}} P_{X|Y}(x|y)^\alpha \tag{6}$$

which are introduced in [17] and [9], respectively. Both of these conditional Rényi entropies are outside the scope of [8].

$R_\alpha^{\mathsf{A}}(X|Y)$ is used in [17] to show that the strong converse of channel coding theorem. $R_\alpha^{\mathsf{H}}(X|Y)$ is defined in [9] to derive an upper bound of leaked information in universal privacy amplification.

Not only the conditional Rényi entropies discussed in [8] but also $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ is non-negative and is bounded by $\log |\mathcal{X}|$. Note that $R_\alpha^{\mathsf{A}}(X|Y) = 0$ and $R_\alpha^{\mathsf{H}}(X|Y) = 0$ hold if and only if every x is obtained from a certain $y \in \operatorname{supp} P_Y$ deterministically. On the other hand, $R_\alpha^{\mathsf{A}}(X|Y) = R_\alpha^{\mathsf{H}}(X|Y) = \log |\mathcal{X}|$ holds, if X and Y are statistically independent and X is uniformly distributed on $\mathcal{X}$. The proofs are not so hard and we omit them (Proofs for $R_\alpha^{\mathsf{A}}(X|Y)$, see [17]). Note that the following fundamental relations hold with respect to $R_\alpha^{\mathsf{H}}(X|Y)$ and $R_\alpha^{\mathsf{A}}(X|Y)$.

Theorem 1 *For a fixed real number $\alpha \geq 0$, the probability distributions P_Y, and the conditional probability distribution $P_{X|Y}$, it holds that*

$$R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha^{\mathsf{A}}(X|Y). \tag{7}$$

Proof. See Appendix A.1. □

2.3 Fundamental Requirements for Conditional Rényi entropies

Here, we discuss fundamental properties required to conditional Rényi entropies from axiomatic, information theoretic, and cryptographic viewpoints. In this section, Rényi entropies are not restricted to each definitions, and hence, it is denoted by $R_\alpha(X|Y)$.

Axiomatic Consideration Recall that Rényi entropy is axiomatically obtained, namely, Rényi entropy is the unique quantity (up to a constant factor) that satisfies weakened postulates for Shannon entropy [7]. According to [7], the postulates that characterize the Shannon entropy are, (a) $H(X)$ is a symmetric function with respect to each probability in a probability distribution of X; (b) $H(X)$ is a continuous function of P_X; (c) $H(X) = 1$ if X is a uniform binary random variable, and; (d) the *chain rule*, i.e., $H(XY) = H(Y) + H(X|Y)$ holds[4], where $H(X|Y) := \sum_y H(X|Y = y) = -\sum_{x,y} P_{XY}(x, y) \log P_{X|Y}(x|y)$. Then, Rényi entropy is obtained by (a)–(c) and, instead of (d), $H(XY) = H(X) + H(Y)$ if X and Y are statistically independent.

Based on this derivation, it might be acceptable to require conditional Rényi entropies to satisfy (a)–(c) with conditioned random variables. Namely,

- $R_\alpha(X|Y)$ is symmetric with respect to $\{P_{X|Y}(x|y)\}_{x\in\mathcal{X}}$ for each $y \in \mathcal{Y}$, and $\{P_Y(y)\}_{y\in\mathcal{Y}}$.
- $R_\alpha(X|Y)$ is a continuous function with respect to $P_{XY}(\cdot, \cdot)$.
- $R_\alpha(X|Y) = 1$ if a binary random variable X is uniformly distributed for given Y, i.e., $P_{X|Y}(1|y) = P_{X|Y}(0|y) = 1/2$ for all $y \in \operatorname{supp} Y$, where $\operatorname{supp} Y := \{y \in \mathcal{Y} \mid P_Y(y) > 0\}$.

All conditional Rényi entropies in this paper satisfy the above properties although we omit their proof.

Since the postulate (d) is replaced with $H(XY) = H(X) + H(Y)$, it is natural that Rényi entropies do not satisfy the chain rule. Actually, it is pointed out in [8, Theorem 5] that $R^{\mathsf{C}}_\alpha(X|Y)$ and $R^{\mathsf{RW}}_\alpha(X|Y)$ do not satisfy the chain rule for arbitrary $\alpha \neq 1$[5]. We will see in Section 2.4 that the chain rules also do not hold for $R^{\mathsf{A}}_\alpha(X|Y)$ and $R^{\mathsf{H}}_\alpha(X|Y)$.

Instead, we consider several fundamental properties related to chain rule. Note that, *monotonicity*, i.e., $H(XY) \geq H(X)$ is derived from the chain rule since *non-negativity* holds for conditional Shannon entropies. Hence, the non-negativity for conditional Rényi entropies and monotonicity for Rényi entropies are important. In fact, it is known that the monotonicity holds for Rényi entropies. Hence, we are interested in the monotonicity for the conditional Rényi entropies. Namely, it is desirable to satisfy that $R_\alpha(X|Z) \leq R_\alpha(XY|Z)$ for random variables X, Y, and Z. This inequality for conditional Shannon entropies are introduced in [13, (13.9) in Lemma 13.6] as a useful one. Hence, we will investigate the following properties:

- (Non-negativity) $R_\alpha(X|Y) \geq 0$ for all random variables X and Y.
- (Conditioned monotonicity) $R_\alpha(X|Z) \leq R_\alpha(XY|Z)$ for random variables X, Y, and Z, where the equality holds if $Y = f(X, Z)$ for some (deterministic) mapping f.

It is easy to show that the conditional Rényi entropies in this paper satisfy the non-negativity, the proofs of them are omitted.

It is also known that Rényi entropies *do not satisfy* the *subadditivity* since only the additivity for independent random variables is required instead of the postulate (d) for Rényi entropies. Subadditivity for Shannon entropy is written as $H(XY) \leq H(X) + H(Y)$, which is equivalent to $H(X|Y) \leq H(X)$. This inequality is called as "*Conditioning reduces entropy*" [11], CRE for short. Note that CRE states that the entropy of random variable X decreases if some information Y related to X is revealed. On the other hand, monotonicity implies that the entropy of X increases if some information is added.

[4] This form of the chain rule is inductively obtained by using the postulate (d) in [7, p. 547].

[5] In the case of $\alpha = 1$, conditional Rényi entropies coincide with conditional Shannon entropy, and hence, chain rule is of course satisfied. In addition, it is obvious that $R^{\mathsf{JA}}_\alpha(X|Y)$ also satisfies the chain rule since it is defined to satisfy the chain rule.

Furthermore, we can consider an inequality $I(X;Z|Y) \geq 0$, which is a direct consequence of CRE, i.e., $H(X|YZ) \leq H(X|Y)$. This property is often used in proving information theoretic inequality, e.g., see Proof I in Section 5.2. Also, combining this inequality with the chain rule, we can prove that Shannon entropy is a *polymatroid* function [22].

In the case of Shannon entropy, $I(X;Z|Y) = 0$ holds when X, Y, and Z form a *Markov chain* in this order [11], in symbols $X \leftrightarrow Y \leftrightarrow Z$. Moreover, we note that stronger inequality than $H(X|YZ) \leq H(X|Y)$ is known for Shannon entropy if $X \leftrightarrow Y \leftrightarrow Z$. In this case, it holds that $H(X|Z) \leq H(X|Y)$, which is equivalent to $I(X;Z) \geq I(X;Y)$, called *Data Processing inequality* (DPI).

Summarizing, we will investigate in the following properties:

- (CRE) $R_\alpha(X|Y) \leq R_\alpha(X)$ for all random variables X and Y, where the equality holds if X and Y are independent.
- (DPI) If random variables X, Y, and Z form a Markov chain, it holds that $R_\alpha(X|Y) \geq R_\alpha(X|Z)$, where the equality holds if there exists a surjective mapping $f : \mathcal{Y} \to \mathcal{Z}$.

Independency Similarly to conditional Shannon entropy, conditional Rényi entropy is hopefully a measure of dependency between random variables X and Y. First, we consider the case where random variables X and Y are independent. In this case, $R_\alpha(X|Y) = R_\alpha(X)$ is necessary to be satisfied. Note that $R_\alpha(X|Y) = R_\alpha(X)$ holds if Y is deterministic since every random variable is independent from the deterministic event. Next, let us consider more general case, i.e., random variables X and Y are mutually correlated. In such a case, we often use the mutual information $I(X;Y) := H(X) - H(X|Y) = H(Y) - H(Y|X)$. It is very fundamental property that $I(X;Y) \geq 0$, i.e., $H(X|Y) \geq H(X)$, and here, we find CRE again.

Relation to other entropies Rényi entropy is an extension of many information measures such as Shannon entropy, min-entropy, and Hartley entropy, collision entropy, etc. In particular, from a cryptographic viewpoint, Shannon and min-entropies are prominently important. Hence, it is better if $R_\alpha(X|Y)$ satisfies the following properties:

(i) $\lim_{\alpha \to 1} R_\alpha(X|Y) = H(X|Y)$.
(ii) Conditional Rényi entropy of order α converges to conditional min-entropies if $\alpha \to \infty$.

Similarly to conditional Rényi entropies, we can find several definitions of conditional min-entropies. Among them, the average conditional min-entropy

$$R_\infty^{\mathsf{avg}}(X|Y) := -\log \mathbb{E}_Y \left[\max_x P_{X|Y}(x|Y) \right] \tag{8}$$

proposed in [10] is important from a cryptographic viewpoint, e.g., [10, 23–27] . Also, we can find the *worst case* conditional min-entropy (e.g., in the analysis of physically unclonable functions (PUFs), see [28]).

$$R_\infty^{\mathsf{wst}}(X|Y) := -\log \max_{\substack{x \in \mathcal{X} \\ y \in \operatorname{supp} P_Y}} P_{X|Y}(x|y). \tag{9}$$

Here we note that the conditional Rényi entropies $R_\alpha^{\mathsf{C}}(X|Y)$, $R_\alpha^{\mathsf{JA}}(X|Y)$, and $R_\alpha^{\mathsf{RW}}(X|Y)$ do not satisfy either (i) or (ii) shown above. Namely, it is pointed out in [8] that,

- $\lim_{\alpha \to \infty} R_\alpha^{\mathsf{RW}}(X|Y) = R_\infty^{\mathsf{wst}}(X|Y)$ but $\lim_{\alpha \to 1} R_\alpha^{\mathsf{RW}}(X|Y) \neq H(X|Y)$,
- $\lim_{\alpha \to 1} R_\alpha^{\mathsf{N}}(X|Y) = H(X|Y)$ but $\lim_{\alpha \to \infty} R_\alpha^{\mathsf{N}}(X|Y) \neq R_\infty^{\mathsf{avg}}(X|Y), R_\infty^{\mathsf{wst}}(X|Y)$ for $\mathsf{N} \in \{\mathsf{C}, \mathsf{JA}\}$.

In the above sense, $R_\alpha^{\mathsf{N}}(X|Y)$, $\mathsf{N} \in \{\mathsf{C}, \mathsf{JA}, \mathsf{RW}\}$ do not satisfy our requirements for conditional Rényi entropies. In addition, note that (8) is not sufficiently analyzed in [8] since the conditional Rényi entropies corresponding to $R_\infty^{\mathsf{avg}}(X|Y)$ is not provided in the literature while it plays important roles in many cryptographic applications, One of the reasons why we focus on $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ is that the conditional Rényi entropy $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ missing in [8] actually bridge the conditional Shannon entropy and the conditional min-entropy appeared in cryptography as shown below:

Theorem 2 *For random variables X and Y, following relations are satisfied:*

(i) $\lim_{\alpha\to 1} R_\alpha^{\mathsf{A}}(X|Y) = \lim_{\alpha\to 1} R_\alpha^{\mathsf{H}}(X|Y) = H(X|Y)$.
(ii) $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{A}}(X|Y) = R_\infty^{\mathsf{avg}}(X|Y)$*, and* $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(X|Y) = R_\infty^{\mathsf{wst}}(X|Y)$.

Proof. The proof of $\lim_{\alpha\to 1} R_\alpha^{\mathsf{A}}(X|Y) = H(X|Y)$ is provided in [17]. For the rest of the proofs, see Appendix A.2. □

Therefore, in this paper, we will mainly focus on the properties of conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$.

2.4 Chain Rule and Weak Chain Rule for Conditional Rényi entropies

Based on the discussion of the previous section, it might be hard to expect that the conditional Rényi entropies satisfy the *chain rule*. According to [8], we can readily know that the chain rule will not hold *with equality* if the conditionally Rényi entropies satisfies CRE since, by defining the conditional Rényi entropy as $R_\alpha^{\mathsf{JA}}(X|Y) := R_\alpha(XY) - R_\alpha(Y)$ [19,20], it does not satisfy CRE. Hence, we aim to relax the requirement so that the chain rule holds *with inequality.*

Definition 3 *For* $\mathsf{N} \in \{\mathsf{C}, \mathsf{RW}, \mathsf{A}, \mathsf{H}\}$*, we say that the conditional Rényi entropy* $R_\alpha^{\mathsf{N}}(X|Y)$ *satisfies weak chain rule if, for arbitrarily fixed* $\alpha \geq 0$*, either* $R_\alpha(XY) \geq R_\alpha^{\mathsf{N}}(X|Y) + R_\alpha(Y)$ *or* $R_\alpha(XY) \leq R_\alpha^{\mathsf{N}}(X|Y) + R_\alpha(Y)$ *holds for arbitrarily random variables X and Y. These conditions are equivalent to* $R_\alpha^{\mathsf{JA}}(X|Y) \geq R_\alpha^{\mathsf{N}}(X|Y)$ *and* $R_\alpha^{\mathsf{JA}}(X|Y) \leq R_\alpha^{\mathsf{N}}(X|Y)$*, respectively.*

Proposition 1 ([8]) *Let X and Y be random variables taking values in finite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then, it holds that* $R_\alpha^{\mathsf{JA}}(X|Y) \geq R_\alpha^{\mathsf{RW}}(X|Y)$ *if* $\alpha > 1$*,* $R_\alpha^{\mathsf{JA}}(X|Y) \leq R_\alpha^{\mathsf{RW}}(X|Y)$*, otherwise. On the other hand, the values of* $R_\alpha^{\mathsf{JA}}(X|Y)$ *and of* $R_\alpha^{\mathsf{C}}(X|Y)$ *are incomparable.*

Proposition 1 implies that only $R_\alpha^{\mathsf{RW}}(X|Y)$ satisfies the weak chain rule. However, similarly to $R_\alpha^{\mathsf{C}}(X|Y)$, we can show that $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ do not satisfy the chain rule even in a weak sense.

Proposition 2 *For* $\mathsf{N} \in \{\mathsf{A}, \mathsf{H}\}$*, the values of* $R_\alpha^{\mathsf{JA}}(X|Y)$ *and* $R_\alpha^{\mathsf{N}}(X|Y)$ *are incomparable. Namely, for a fixed α, there exist probability distributions P_{XY} and $P_{X'Y'}$ satisfying* $R_\alpha(XY) > R_\alpha^{\mathsf{N}}(X|Y) + R_\alpha(Y)$ *and* $R_\alpha(X'Y') < R_\alpha^{\mathsf{N}}(X'|Y') + R_\alpha(Y')$*.*

This proposition can be verified by the following example in a binary alphabet case:

Example 1 *Consider the following two cases:*

Case I. $P_{XY}(0,0) = 1/2$, $P_{XY}(0,1) = 1/8$, $P_{XY}(1,0) = 1/4$, *and* $P_{XY}(1,1) = 1/8$.
Case II. $P_{XY}(0,0) = 3/8$, $P_{XY}(0,1) = 1/4$, $P_{XY}(1,0) = 5/16$, *and* $P_{XY}(1,1) = 1/16$.

The graph of $\varphi^{\mathsf{N}}(\alpha) := R_\alpha(XY) - R_\alpha^{\mathsf{N}}(X|Y) - R_\alpha(Y)$ *for* $\mathsf{N} \in \{\mathsf{A}, \mathsf{H}\}$ *are depicted in Fig. 1–(a),(b) in Appendix, which means that* $R_\alpha(XY) > R_\alpha^{\mathsf{N}}(X|Y) + R_\alpha(Y)$ *holds only when* $0 \leq \alpha < 1$ *with Cases A, but* $R_\alpha(XY) < R_\alpha^{\mathsf{N}}(X|Y) + R_\alpha(Y)$ *holds only when* $0 \leq \alpha < 1$ *with Case B. Recall that, in the case of* $\alpha = 1$*, Rényi entropies coincide with Shannon entropies. Hence, in this case, the chain rule, i.e.,* $\varphi^{\mathsf{N}}(1) = 0$*, holds.*

3 Information Theoretic Inequalities for Rényi Entropies

As is pointed out in Theorem 2, the conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ are related to cryptographically meaningful min-entropies. Furthermore, in this section, we show that several important inequalities are satisfied by these conditional Rényi entropies, which is another reason why we are focusing on them.

3.1 Conditioning Reduces Entropy

First, we discuss "*conditioning reduces entropy*" (CRE, [11]), which is formulated as, in the case of Shannon entropies, $H(X) \geq H(X|Y)$ for arbitrary random variables X and Y. It is well known that CRE is very useful and fundamental property in proving information theoretic inequalities. However, it is known that several definitions of Rényi entropies do not satisfy CRE. Actually, it is pointed out in [8] that $R_\alpha^{\mathsf{C}}(X|Y)$, $R_\alpha^{\mathsf{JA}}(X|Y)$, and $R_\alpha^{\mathsf{RW}}(X|Y)$ given by (2)–(4), respectively, *do not satisfy* CRE in general[6].

Fortunately, however, we will point out in this section that $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$, which are outside the scope of [8], satisfy CRE in general, and in particular, we can observe that $R_\infty^{\mathsf{avg}}(X|Y)$ and $R_\infty^{\mathsf{wst}}(X|Y)$ satisfy CRE, though it is possible to show it directly. This fact is very useful to show the forthcoming results in symmetric key cryptography in Section 5, and gives one of the reasons why the conditional min-entropies $R_\infty^{\mathsf{avg}}(X|Y)$ and $R_\infty^{\mathsf{wst}}(X|Y)$ work well.

In the following, we will focus on CRE with respect to $R_\alpha^{\mathsf{H}}(X|Y)$ since CRE for $R_\alpha^{\mathsf{A}}(X|Y)$ is proved in [17, 29]. In order to understand CRE for Rényi entropy $R_\alpha^{\mathsf{H}}(X|Y)$, we introduce a conditional α-divergence defined by the same idea with $R_\alpha^{\mathsf{H}}(X|Y)$ in the following form.

Definition 4 ([30]) *Let X_1, X_2, and Y be random variables taking values on $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{Y}$, respectively. Assume that the probability distributions of these random variables are given by $P_{X_1|Y}(\cdot|y) = W(\cdot|y)Q(y)$, $P_{X_2|Y}(\cdot|y) = V(\cdot|y)Q(y)$ for all $y \in \mathcal{Y}$ with a probability distribution $Q(\cdot)$ and conditional probability distributions $W(\cdot|\cdot)$ and $V(\cdot|\cdot)$.*

Then, for a real number $\alpha \geq 0$, define the conditional α-divergence $D_\alpha(X_1\|X_2|Y)$ to be

$$D_\alpha(X_1\|X_2|Y) := D_\alpha(W\|V|Q) = \frac{1}{\alpha-1}\log\sum_{x,y}\frac{W(x|y)^\alpha}{V(x|y)^{\alpha-1}}Q(y). \tag{10}$$

Similarly to the conditional Rényi entropies, $D_\alpha(X_1\|X_2|Y)$ satisfies the fundamental properties of conditional α-divergence. For a real number $\alpha \geq 0$, the conditional α-divergence satisfies the following properties:

Proposition 3 *Let X_1, X_2, and Y be random variables following the probability distributions $P_{X_1|Y}(\cdot|y) = W(\cdot|y)Q(y)$, $P_{X_2|Y}(\cdot|y) = V(\cdot|y)Q(y)$ for all $y \in \mathcal{Y}$.*

Then, for a real number $\alpha \geq 0$, α-divergence $D_\alpha(X_1\|X_2|Y)$ satisfies the following properties:

(i) $D_\alpha(X_1\|X_2|Y) \geq 0$, *where the equality holds if and only if* $W(\cdot|y) = V(\cdot|y)$ *for all* $y \in \operatorname{supp} Q$.
(ii) $\lim_{\alpha\to1} D_\alpha(X_1\|X_2|Y) = D(X_1\|X_2|Y) := \sum_{x,y} Q(y)W(x|y)\log(W(x|y)/V(x|y))$.

Proof. The property (ii) is pointed out in [30] without proof. We provide the formal proofs for (i) and (ii) in Appendix A.3 for readers' convenience. □

Our derivation of CRE is immediately obtained from the following relation:

Theorem 3 *Let X, Y, and Z be random variables taking values on finite sets $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, respectively. For all $\alpha \geq 0$, it holds that*

$$R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y) = D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha}) \tag{11}$$

where $P_{X_\alpha}(x) := P_X(x)^\alpha/\sum_{\tilde{x}} P_X(\tilde{x})^\alpha$ for $x \in \mathcal{X}$.

While this theorem follows from the identity obtained by [30, eq. (21)], by letting $Q_{AB}(a,b) = P_A(a)P_U(b)$ where U follows the uniform distribution, the direct proof is given as follows:

Proof. Observe that

$$\begin{aligned}\left\{\sum_x P_X(x)^\alpha\right\}^{-1}\sum_{x,y}P_Y(y)P_{X|Y}(x|y)^\alpha &= \left\{\sum_x P_X(x)^\alpha\right\}^{-1}\sum_{x,y}P_{XY}(x,y)^\alpha P_Y(y)^{1-\alpha}\\ &= \sum_{x,y}\frac{P_X(x)^\alpha}{\sum_x P_X(x)^\alpha}P_{Y|X}(y|x)^\alpha P_Y(y)^{1-\alpha}. \end{aligned} \tag{12}$$

[6] We can show that CRE is satisfied by $R_\alpha^{\mathsf{RW}}(X|Y)$ in the case of $\alpha > 1$. See Prop. 15 of Section 7.

Taking the logarithms of both sides of (12) and multiplying $-1/(1-\alpha)$, we obtain (11). □

This relation (11) is an analogue of the well-known definition of the mutual information: $I(X;Y) := H(X) - H(X|Y)$ since the mutual information can be written as

$$I(X;Y) := D(P_{XY}\|P_X P_Y) = \sum_{x,y} P_Y(x) P_{Y|X}(y|x) \log \frac{P_{Y|X}(y|x)}{P_Y(y)} = D(P_{Y|X}\|P_Y|P_X)$$

Note that $I(X;Y) = I(Y;X) = D(P_{X|Y}\|P_X|P_Y)$ and it is easy to check that the conditional divergence of order α satisfies that $D_\alpha(P_{X|Y}\|P_X|P_Y) = D_\alpha(P_{Y|X}\|P_Y|P_X)$. On the other hand, it is obvious that $D_\alpha(P_{X|Y}\|P_X|P_{Y_\alpha}) = D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha})$ does not hold generally, and hence, $R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y) = R_\alpha(Y) - R_\alpha^{\mathsf{H}}(Y|X)$ also does not hold for general α.

Hence, it is natural to define a *mutual information of order* α by

$$I_\alpha^{\mathsf{H}}(X;Y) := R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y), \tag{13}$$

which is similar to the Arimoto's mutual information of order α defined by

$$I_\alpha^{\mathsf{A}}(X;Y) := R_\alpha(X) - R_\alpha^{\mathsf{A}}(X|Y), \tag{14}$$

in the context of describing channel coding theorem in a general setting [17] .

Remark 1 *Note that $I_\alpha^{\mathsf{H}}(X;Y)$ and $I_\alpha^{\mathsf{A}}(X;Y)$ are not symmetric, i.e., $I_\alpha^{\mathsf{H}}(X;Y) \neq I_\alpha^{\mathsf{H}}(Y;X)$ and $I_\alpha^{\mathsf{A}}(X;Y) \neq I_\alpha^{\mathsf{A}}(Y;X)$ in general. In addition, it is seen that $I_\alpha^{\mathsf{A}}(X;Y) \leq I_\alpha^{\mathsf{H}}(X;Y)$ in general, since $R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha^{\mathsf{A}}(X|Y)$. As we will see immediately, $R_\alpha^{\mathsf{H}}(X|Y)$ satisfies CRE as well as $R_\alpha^{\mathsf{A}}(X|Y)$, it is easy to see that both of $I_\alpha^{\mathsf{A}}(X;Y)$ and $I_\alpha^{\mathsf{H}}(X;Y)$ are non-negative, and they are equal to zero* if *X and Y are statistically independent.*

From Theorem 3, the following relation follows quite easily. Recall that $D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha}) = 0$ if X and Y are statistically independent.

Theorem 4 (Conditioning reduces entropy) *Let X and Y be random variables taking values on $\mathcal{X}$ and $\mathcal{Y}$, respectively. For all $\alpha \geq 0$, it holds that*

$$R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha(X), \tag{15}$$

where the equality holds if X and Y are statistically independent.

Remark 2 *Although Theorem 4 is a direct consequence of Theorem 3, this theorem itself directly follows from Jensen's inequality as shown below:*

Alternative Proof of Theorem 4. From Jensen's inequality, in the case of $0 \leq \alpha < 1$, we have

$$\mathbb{E}_Y\left[\sum_{x\in\mathcal{X}} P_{X|Y}(x|Y)^\alpha\right] \leq \sum_{x\in\mathcal{X}} \mathbb{E}_Y\left[P_{X|Y}(x|Y)\right]^\alpha = \sum_{x\in\mathcal{X}} P_X(x)^\alpha.$$

Similarly, it holds that $\mathbb{E}_Y\left[\sum_{x\in\mathcal{X}} P_{X|Y}(x|Y)^\alpha\right] \geq \sum_{x\in\mathcal{X}} P_X(x)^\alpha$ in the case of $\alpha \geq 1$. □

3.2 Data Processing Inequality

If random variables X, Y, and Z, taking values on finite sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively, satisfy

$$P_{XZ|Y}(x,z|y) = P_{X|Y}(x|y)P_{Z|Y}(z|y), \quad \text{for all } x \in \mathcal{X}, y \in \mathcal{Y}, \text{and } \; z \in \mathcal{Z} \tag{16}$$

we say that X, Y, and Z form a *Markov chain*, in symbols $X \leftrightarrow Y \leftrightarrow Z$. The data processing inequality (DPI, [11]) tells us that $I(X;Y) \geq I(X;Z)$ holds if $X \leftrightarrow Y \leftrightarrow Z$. We can extend Theorem 4, in the following way:

Theorem 5 (Data processing inequality) *Let X, Y, and Z be random variables taking on finite sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively, and assume that $X \leftrightarrow Y \leftrightarrow Z$. Then it holds that $I_\alpha^{\mathsf{A}}(X;Y) \geq I_\alpha^{\mathsf{A}}(X;Z)$ and $I_\alpha^{\mathsf{H}}(X;Y) \geq I_\alpha^{\mathsf{H}}(X;Z)$ for arbitrary $\alpha \geq 0$. The equality holds if and only if there exists a surjective mapping $f : \mathcal{Y} \to \mathcal{Z}$.*

Proof. Without loss of generality, we can write $Z = g(Y,R)$ where $g : \mathcal{Y} \times \mathcal{R} \to \mathcal{Z}$ is a deterministic mapping, and R is a random variable taking values on a finite set and is independent of X. Then, we have both of $R_\alpha^{\mathsf{A}}(X|g(Y,R),Y,R) = R_\alpha^{\mathsf{A}}(X|YR)$ and $R_\alpha^{\mathsf{H}}(X|g(Y,R),Y,R) = R_\alpha^{\mathsf{H}}(X|YR)$ since g is deterministic. Noticing that $X \leftrightarrow Y \leftrightarrow R$, it holds that

$$P_{X|YR}(x|y,r) = \frac{P_{XR|Y}(x,r|y)}{P_{R|Y}(r|y)} = \frac{P_{X|Y}(x|y)P_{R|Y}(r|y)}{P_{R|Y}(r|y)} = P_{X|Y}(x|y) \tag{17}$$

for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$, where the second equality is validated by the Markov chain. Hence, we have $R_\alpha^{\mathsf{A}}(X|YR) = R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|YR) = R_\alpha^{\mathsf{H}}(X|Y)$. Since conditioning reduces entropy, we obtain $R_\alpha^{\mathsf{A}}(X|Y) \leq R_\alpha^{\mathsf{A}}(X|g(Y,R))$ and $R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha^{\mathsf{H}}(X|g(Y,R))$, which completes the proof. □

Remark 3 *DPI is very useful since it implies that the quality of information degenerates by processing the information. It is worth noting that DPI generally holds only if we use $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ since DPI is extension of CRE.*

3.3 Conditioned Monotonicity

It is well known that Shannon entropy satisfies *monotonicity* and *subadditivity*, i.e., $H(X) \leq H(XY)$ and $H(XY) \leq H(X) + H(Y)$, respectively, for random variables X and Y. However, since $R_\alpha^{\mathsf{JA}}(X|Y)$ does not satisfy CRE in general [8], it is easy to see that Rényi entropy only satisfies monotonicity. Here, we show an extended monotonicity for conditional Rényi entropy, which is also useful in cryptographic applications. In the case of Shannon entropy, this results is easily verified by subadditivity, while this fact is presented in [13, (13.9) in Lemma 13.6].

Theorem 6 *Let X, Y, and Z be random variables taking values on finite sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. Then, for $\mathsf{N} \in \{\mathsf{A}, \mathsf{H}\}$, we have:*

(i) $R_\alpha^{\mathsf{N}}(X|Z) \leq R_\alpha^{\mathsf{N}}(XY|Z)$,
(ii) $R_\alpha^{\mathsf{N}}(X|Z) = R_\alpha^{\mathsf{N}}(XY|Z)$ *if and only if* $Y = f(X,Z)$ *for some (deterministic) mapping* f.

Proof. Although (i) for $R_\alpha^{\mathsf{A}}(X|Y)$ is proved in [29, Proposition 2], we will prove this claim for both conditional Rényi entropies simultaneously. For any α with $0 \leq \alpha < 1$ and arbitrary $z \in \mathcal{Z}$, it holds that

$$\sum_{x,y} P_{XY|Z}(x,y|z)^\alpha = \sum_x P_{X|Z}(x|z)^\alpha \sum_y P_{Y|XZ}(y|x,z)^\alpha \geq \sum_x P_{X|Z}(x|z)^\alpha. \tag{18}$$

Hence, we have

$$\sum_z P_Z(z) \left(\sum_{x,y} P_{XY|Z}(x,y|z)^\alpha \right)^{1/\alpha} \geq \sum_z P_Z(z) \left(\sum_x P_{X|Z}(x|z)^\alpha \right)^{1/\alpha}, \tag{19}$$

$$\sum_z P_Z(z) \sum_{x,y} P_{XY|Z}(x,y|z)^\alpha \geq \sum_z P_Z(z) \sum_x P_{X|Z}(x|z)^\alpha, \tag{20}$$

which result in $R_\alpha^{\mathsf{A}}(X|Z) \leq R_\alpha^{\mathsf{A}}(XY|Z)$ and $R_\alpha^{\mathsf{H}}(X|Z) \leq R_\alpha^{\mathsf{H}}(XY|Z)$, respectively. Equalities of $R_\alpha^{\mathsf{N}}(X|Z) = R_\alpha^{\mathsf{N}}(XY|Z)$ holds if and only if the equality of (18) holds, i.e.,

$$\sum_y P_{Y|XZ}(y|x,z)^\alpha = 1. \tag{21}$$

holds for all x, z with $P_{XZ}(x,z) > 0$. For any x, z with $P_{XZ}(x,z) > 0$, (21) holds if x and z uniquely determine $y = f(x,z)$. Coversely, for any x, z with $P_{XZ}(x,z) > 0$, if (21) holds, we can define a deterministic mapping f by $y := f(x,z)$ such that $P_{Y|XZ}(y|x,z) = 1$. Therefore, $R_\alpha^{\mathsf{N}}(X|Z) = R_\alpha^{\mathsf{N}}(XY|Z)$ is equivalent to the condition (21), and it is also equivalent to $Y = f(X,Z)$ for some deterministic mapping f.

The case of $\alpha > 1$ can be similarly discussed, and we omit it. In addition, the statement in the case $\alpha = 1$ is true, since it means the case of Shannon entropy. □

3.4 Fano's Inequality

In this section, we derive upper-bounds for $R_\alpha^{\mathsf{H}}(X|Y)$, and they can be seen as extension of Fano's inequality (see Remark 4).

Theorem 7 *Let X and Y be random variables taking values in a finite set $\mathcal{X}$. Also, let $P_e := \Pr\{X \neq Y\}$ and $\bar{P}_e := 1 - P_e$. Then, for $\alpha \geq 0$, we have the following inequalities.*

(i) If $0 \leq \alpha \leq 1$ and $P_e \geq 1 - \frac{1}{|\mathcal{X}|}$, or $\alpha \geq 1$ and $0 \leq P_e \leq 1 - \frac{1}{|\mathcal{X}|}$, it holds that

$$R_\alpha^{\mathsf{H}}(X|Y) \leq \frac{1}{1-\alpha} \log \left[(|\mathcal{X}| - 1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha \right].$$

(ii) If $0 \leq \alpha \leq 1$ and $0 \leq P_e \leq 1 - \frac{1}{|\mathcal{X}|}$, or $\alpha \geq 1$ and $P_e \geq 1 - \frac{1}{|\mathcal{X}|}$, it holds that

$$R_\alpha^{\mathsf{H}}(X|Y) \leq \frac{1}{1-\alpha} \log \left[(|\mathcal{X}| - 1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e \right].$$

Here, in the above inequalities the case $\alpha = 1$ is meant to take the limits at $\alpha = 1$, and the case $P_e = 0$ is meant to take the limits at $P_e = 0$.

Proof. See Appendix A.4 □

Remark 4 In Theorem 2 it is shown that $\lim_{\alpha \to 1} R_\alpha^{\mathsf{H}}(X|Y) = H(X|Y)$. On the other hand, by applying the L'Hospital's rule to the right hands of inequalities in Theorem 7, we obtain the following finite limits at $\alpha = 1$:

(i) $\displaystyle\lim_{\alpha \to 1} \frac{1}{1-\alpha} \log \left[(|\mathcal{X}| - 1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha \right] = P_e \log(|\mathcal{X}| - 1) + h(P_e)$,

(ii) $\displaystyle\lim_{\alpha \to 1} \frac{1}{1-\alpha} \log\left[(|\mathcal{X}| - 1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e \right] = P_e \log(|\mathcal{X}| - 1) + h(P_e)$,

where $h(\cdot)$ is the binary entropy function. Therefore, by taking the limit at $\alpha = 1$ for each of inequalities in Theorem 7, we obtain Fano's inequality as a special case. In this sense, our inequalities in Theorem 7 can be considered as extension of Fano's inequality.

Remark 5 Note that Fano's inequality implies $H(X|Y) \to 0$ as $P_e \to 0$. Theorem 7 implies that, for any $\alpha \geq 0$, $R_\alpha^{\mathsf{H}}(X|Y) \to 0$ as $P_e \to 0$, as we would expect.

4 Security Criteria Based on Conditional Rényi Entropies

4.1 Motivation and Significance

Our motivation and significance for considering security criteria based on conditional Rényi entropies lies in two points.

The first point lies in realistic significance which is deeply related to guessing probability by adversaries. In Section 4.3, we show that (conditional) Rényi entropies play an important role to derive a lower bound on failure of guessing by adversaries, and it turns out that our security criteria is a sufficient condition to make it reasonably large enough. Our way of thinking is also related to the recent elegant approach in information theory in order to show the converse of channel coding theorem in finite blocklength regime [30, 31].

The second point lies in mathematical importance for generalizing Shannon's impossibility (or Shannon's bounds) in information-theoretic cryptography. The purpose is to extend and unify existing notions and techniques by considering (conditional) Rényi entropies which cover various kinds of entropies such as the (conditional) Shannon entropy, collision entropy, and min-entropy. Specifically, for symmetric-key encryption protocols, there exist several known bounds on secret-keys including the Shannon's bounds (see Section 4.2). And, our purpose is to extend those bounds in a generic and unified manner by using security criteria based on conditional Rényi entropies.

4.2 Existing Lower Bounds on Secret-keys

We describe well-known Shannon's bound [4] for symmetric-key encryption and its extensions (or variants) by Dodis [5], and Alimomeni and Safavi-Naini [6]. To discribe the bounds, we use the following notation: let K, M, and C be random variables which take values in finite sets $\mathcal{K}$, $\mathcal{M}$, and $\mathcal{C}$ of secret-keys, plaintexts, and ciphertexts, respectively. Informally, a symmetric-key encryption is said to meet *perfect correctness* if it has no decryption-errors; a symmetric-key encryption is said to meet *perfect secrecy* if it reveals no information about plaintexts from ciphertexts, which is formalized by $H(M|C) = H(M)$ (see Section 5 for the formal model of encryption protocols and its explanation).

Proposition 4 (Shannon's bound: [4]) *Let Π be a symmetric-key encryption such that both encryption and decryption algorithms are deterministic. If Π satisfies perfect correctness and perfect secrecy, we have $H(K) \geq H(M)$ and $|\mathcal{K}| \geq |\mathcal{M}|$.*

Proposition 5 (Dodis's bound: Th.3 in [5]) *Let Π be a symmetric-key encryption. If Π satisfies perfect correctness and perfect secrecy, we have $R_\infty(K) \geq R_\infty(M)$.*

Proposition 6 (Alimomeni and Safavi-Naini's bound: Th.2 in [6]) *Let Π be a symmetric-key encryption such that both encryption and decryption algorithms are deterministic. If Π satisfies both $R_\infty(M) = R_\infty^{\mathsf{avg}}(M|C)$ and perfect correctness, we have $R_\infty(K) \geq R_\infty(M)$.*

4.3 Lower Bounds on Failure Probability of Adversary's Guessing

We show that lower bounds on failure probability of adversary's guessing are given by conditional Rényi entropies, $R_\alpha^{\mathsf{H}}(M|C)$ or $R_\alpha^{\mathsf{A}}(M|C)$, in general.

Let $\alpha > 1$. Suppose that an adversary obtains a ciphertext C by observing a channel, and he chooses an arbitrary function g. Let $\hat{M} := g(C)$, $P_e := \Pr\{M \neq \hat{M}\}$, and $\bar{P}_e := 1 - P_e$. The purpose of the adversary is to maximize $\Pr\{M = \hat{M}\} = \bar{P}_e$ (or equivalently, to minimize P_e) by taking a guessing strategy g. Without loss of generality, we assume $\bar{P}_e \geq 1/|\mathcal{M}|$.

First, we derive a lower bound on P_e by using $I_\alpha^{\mathsf{H}}(M;C)$. By the inequalities

$$\begin{aligned} R_\alpha(M) &= I_\alpha^{\mathsf{H}}(M;C) + R_\alpha^{\mathsf{H}}(M|C) \\ &\leq I_\alpha^{\mathsf{H}}(M;C) + R_\alpha^{\mathsf{H}}(M|\hat{M}) \qquad (22) \\ &\leq I_\alpha^{\mathsf{H}}(M;C) + \frac{1}{1-\alpha} \log\left[(|\mathcal{M}|-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha\right], \qquad (23) \end{aligned}$$

where (22) follows from DPI for $R_\alpha^{\mathsf{H}}(X|Y)$ and (23) follows from our extension of Fano's inequality (i.e., Theorem 7), we have

$$\begin{aligned} \exp\left\{(1-\alpha)[R_\alpha(M) - I_\alpha^{\mathsf{H}}(M;C)]\right\} &\geq (|\mathcal{M}|-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha \\ &\geq (1-P_e)^\alpha. \qquad (24) \end{aligned}$$

By (24), we obtain

$$P_e \geq 1 - \exp\left\{\frac{1-\alpha}{\alpha}[R_\alpha(M) - I_\alpha^{\mathsf{H}}(M;C)]\right\}.$$

Therefore, we obtain the following result.

Theorem 8 *The failure probability of adversary's guessing is lower-bounded by*

$$P_e \geq 1 - \exp\left\{\frac{1-\alpha}{\alpha}R_\alpha(M)\right\}\exp\left\{\frac{\alpha-1}{\alpha}I_\alpha^{\mathsf{H}}(M;C)\right\}. \tag{25}$$

In particular, if P_M is the uniform distribution, we have

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}}\exp\left\{\frac{\alpha-1}{\alpha}I_\alpha^{\mathsf{H}}(M;C)\right\}. \tag{26}$$

If we impose security criteria $I_\alpha^{\mathsf{H}}(M;C) \leq \epsilon$ for small ϵ (say, $\epsilon = 0$) for an encryption protocol (note that any other quantity $R_\alpha(M)$, $|\mathcal{M}|$ is independent of security of the protocol), the above lower bound can be large, and hence the adversary cannot guess a target plaintext from a ciphertext with reasonable probability even if he chooses a powerful guessing strategy g.

Remark 6 *The bound (25) is tight for $\alpha = 2$ and $\alpha = \infty$ in the following sense.*

- *Case of $\alpha = 2$: Consider the case that $I_2^{\mathsf{H}}(M;C) = 0$ and P_M is the uniform distribution. Then, (26) implies that $P_e \geq 1 - \exp(-\frac{1}{2}R_2(M)) = 1 - \frac{1}{\sqrt{|\mathcal{M}|}}$, or equivalently $\bar{P}_e \leq \frac{1}{\sqrt{|\mathcal{M}|}}$. The equality of this bound is achievable, since it is the collision probability (i.e., an adversary can take a strategy which selects a plaintext according to P_M).*
- *Case of $\alpha = \infty$: Consider the case $I_\infty^{\mathsf{H}}(M;C) = 0$. Then, (25) implies that $P_e \geq 1 - \exp(-R_\infty(M)) = 1 - \max_m P_M(m)$, or equivalently $\bar{P}_e \leq \max_m P_M(m)$. The equality of this bound is achievable, since an adversary can take a strategy $g(C) = \arg\max_m P_M(m)$.*

Second, we discuss a lower bound on P_e by using $I_\alpha^{\mathsf{A}}(M;C)$. Before discussion, we note the following previous results.

Definition 5 ([32]) *For random variables X, Y, and a real number $\rho \neq 1$, the Gallager's function is defined by*

$$E_0(\rho, P_X, P_{Y|X}) = -\log\sum_y\left(\sum_x P_X(x)P_{Y|X}(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}.$$

Proposition 7 ([17]) *For random variables X and Y, it holds that*

$$I_\alpha^{\mathsf{A}}(X;Y) = \frac{\alpha}{1-\alpha}E_0(\alpha^{-1}-1, P_{X_\alpha}, P_{Y|X})$$

where P_{X_α} is given by $P_{X_\alpha}(x) = \frac{P_X(x)^\alpha}{\sum_{\tilde{x}}P_X(\tilde{x})^\alpha}$. Conversely, for random variables X and Y, we have

$$\frac{\alpha}{1-\alpha}E_0(\alpha^{-1}-1, P_X, P_{Y|X}) = I_\alpha^{\mathsf{A}}(X_{1/\alpha};Y),$$

where $P_{X_{1/\alpha}}$ is given by $P_{X_{1/\alpha}}(x) = \frac{P_X(x)^{1/\alpha}}{\sum_{\tilde{x}}P_X(\tilde{x})^{1/\alpha}}$.

Proposition 8 ([30]) *For a real number $\alpha > 0$, and for distributions P_X, $P_{\hat{X}}$ over $\mathcal{X}$ such that $\epsilon := \Pr\{X \neq \hat{X}\} \leq 1 - \frac{1}{|\mathcal{X}|}$, it holds*

$$d_\alpha(1 - \epsilon \parallel 1/|\mathcal{X}|) \leq \frac{\alpha}{1-\alpha} E_0(\alpha^{-1} - 1, P_X, P_{\hat{X}|X}).$$

In particular, we have

$$\frac{\alpha}{\alpha - 1} \log(1 - \epsilon) + \log |\mathcal{X}| \leq \frac{\alpha}{1-\alpha} E_0(\alpha^{-1} - 1, P_X, P_{\hat{X}|X}).$$

Now, let's be back to our discussion. We use the same notation as in the case of $I^{\mathsf{H}}_\alpha(M;C)$. By combining the above propositions, we have

$$\begin{aligned}\frac{\alpha}{\alpha - 1} \log(1 - P_e) + \log m &\leq \frac{\alpha}{1-\alpha} E_0(\alpha^{-1} - 1, P_M, P_{\hat{M}|M}) \\ &= I^{\mathsf{A}}_\alpha(M_{1/\alpha}; \hat{M}) \\ &\leq I^{\mathsf{A}}_\alpha(M_{1/\alpha}; C),\end{aligned}$$

where $\hat{M} = g(C)$, $P_{M_{1/\alpha}}(m) = \frac{P_M(m)^{1/\alpha}}{\sum_{\tilde{m}} P_M(\tilde{m})^{1/\alpha}}$, and the last inequality follows from DPI for $R^{\mathsf{A}}_\alpha(X|Y)$. From the inequality, we obtain the following result.

Proposition 9 *The failure probability of adversary's guessing is lower-bounded by*

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha - 1}{\alpha} I^{\mathsf{A}}_\alpha(M_{1/\alpha}; C)\right\}. \tag{27}$$

In particular, if P_M is the uniform distribution, we have

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha - 1}{\alpha} I^{\mathsf{A}}_\alpha(M; C)\right\}. \tag{28}$$

Remark 7 *If P_M is the uniform distribution, the bound (26) is directly obtained from the bound (28) since $I^{\mathsf{A}}_\alpha(M;C) \leq I^{\mathsf{H}}_\alpha(M;C)$. However, it is not the case in general.*

Therefore, $I^{\mathsf{H}}_\alpha(M;C) \leq \epsilon$ or $I^{\mathsf{A}}_\alpha(M;C) \leq \epsilon$ for an extremely small $\epsilon \in [0,1]$ is a sufficient condition to show that the failure probability of adversary's guessing is large enough (or equivalently, the success probability of adversary's guessing is small enough). Our security criteria based on conditional Rényi entropies is $I^{\mathsf{H}}_\alpha(M;C) \leq \epsilon$ or $I^{\mathsf{A}}_\alpha(M;C) \leq \epsilon$, which is equivalent to $R_\alpha(M) - R^{\mathsf{H}}_\alpha(M|C) \leq \epsilon$ or $R_\alpha(M) - R^{\mathsf{A}}_\alpha(M|C) \leq \epsilon$, and it is natural to consider the security criteria in terms of an adversary's guessing probability.

5 Generalizing Shannon's Impossibility in Encryption

In this section, we extend the bounds in Section 4.2 in a generic and unified manner by using security criteria based on conditional Rényi entropies.

5.1 The Model and Security Definition

We explain the traditional model of (symmetric-key) encryption protocols. In the following, let $\mathcal{M}$ (resp. $\mathcal{C}$) be a finite set of plaintexts (resp. a finite set of ciphertexts). Also, let M be a random variable which takes plaintexts in $\mathcal{M}$ and P_M its distribution. C denotes a random variable which takes ciphertexts $c \in \mathcal{C}$.

Let $\mathit{\Pi} = ([P_{ED}], \pi_{enc}, \pi_{dec})$ be an *encryption* protocol as defined below:

- Let P_{ED} be a probability distribution over $\mathcal{E}\times\mathcal{D}$ which is a finite set of pairs of encryption and decryption keys. $[P_{ED}]$ is a key generation algorithm, and it outputs $(e,d)\in\mathcal{E}\times\mathcal{D}$ according to P_{ED};
- π_{enc} is an encryption algorithm. It takes an encryption key $e\in\mathcal{E}$ and a plaintext $m\in\mathcal{M}$ on input, and it outputs a ciphertext $c\leftarrow\pi_{enc}(e,m)$, which will be sent via an authenticated channel;
- π_{dec} is a decryption algorithm. It takes on input a decryption key $d\in\mathcal{D}$ and a ciphertext $c\in\mathcal{C}$, and it outputs $\tilde{m}\leftarrow\pi_{dec}(d,c)$ where $\tilde{m}\in\mathcal{M}$.

If $\Pi=([P_K],\pi_{enc},\pi_{dec})$ (i.e., $[P_{ED}]=[P_{KK}]$ and $e=d$), Π is said to be a *symmetric-key encryption*.

In this paper, we do not require that π_{enc} is deterministic, namely, π_{enc} can be randomized. Also, we assume that Π meets *perfect correctness*, namely, it satisfies $\pi_{dec}(d,\pi_{enc}(e,m))=m$ for any possible (e,d) and m. In addition, we consider the case where an encryption protocol Π is usable at most one time (i.e., the one-time model).

Let P_M be a distribution on $\mathcal{M}$, and we assume that it is fixed in the following discussion.

Definition 6 (Secrecy) For $\alpha\geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. An encryption protocol Π is said to meet *ϵ-secrecy with respect to $R_\alpha(\cdot|\cdot)$*, if it satisfies

$$R_\alpha(M)-R_\alpha(M|C)\leq\epsilon.$$

In particular, Π meets *perfect secrecy with respect to $R_\alpha(\cdot|\cdot)$*, if $\epsilon=0$ above.

Note that the traditional notion of perfect secrecy (i.e., $H(M|C)=H(M)$) is equivalent to that of perfect secrecy with respect to $H(\cdot|\cdot)=R_1^{\mathsf{H}}(\cdot|\cdot)=R_1^{\mathsf{A}}(\cdot|\cdot)$ (i.e., $\alpha=1$)[7]. Also, ϵ-secrecy with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ (resp., $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$) is equivalent to $I_\alpha^{\mathsf{H}}(M;C)\leq\epsilon$ (resp., $I_\alpha^{\mathsf{A}}(M;C)\leq\epsilon$) (see Section 4.3).

5.2 Basic Idea for Generalization of Shannon's Impossibility

By Shannon's work [4], it is well known that we have $H(K)\geq H(M)$ for symmetric-key encryption with perfect secrecy (see Proposition 4), which is often called Shannon's impossibility. It will be natural to generalize or extend it to the Rényi entropy. However, there exist some difficulties to generalize it in a technical viewpoint, since in general conditional Rényi entropies do not always have rich properties like the conditional Shannon entropy as we have seen in Sections 2 and 3. In this subsection, we briefly explain our idea of generalizing Shannon's impossibility to the Rényi entropy.

First, let's recall two proof techniques used for deriving $H(K)\geq H(M)$ below, where PS, PC, and CRE mean perfect secrecy, perfect correctness, and conditioning reduces entropy, respectively.

Proof I

$$\begin{aligned}H(M)&=H(M|C)\ \text{(by PS)}\\&=H(M|C)-H(M|KC)\ \text{(by PC)}\\&=I(M;K|C)\\&=H(K|C)-H(K|MC)\\&\leq H(K|C)\\&\leq H(K)\ \text{(by CRE)}\end{aligned}$$

Proof II

$$\begin{aligned}H(M)&=H(M|C)\ \text{(by PS)}\\&\leq H(MK|C)\ \text{(by conditioned monotonicity)}\\&=H(K|C)+H(M|KC)\ \text{(by chain rule)}\\&=H(K|C)\ \text{(by PC)}\\&\leq H(K)\ \text{(by CRE)}\end{aligned}$$

In addition to PS and PC, the property commonly used in both proofs is CRE. From this point of view, it would be reasonable to consider a class of conditional Rényi entropies $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ which satisfy CRE. In addition, in order to complete the proofs, the useful property of the mutual information (i.e., $I(X;Y)=I(Y;X)$) is used in Proof I, while the properties of

[7] This condition is equivalent to $I(M;C)=0$, or equivalently, M and C are independent (i.e., $P_{MC}=P_MP_C$).

conditioned monotonicity and chain rule are used in Proof II. At this point, one may think it hopeless to apply the technique in Proof I, since $I_\alpha^{\mathsf{H}}(X;Y) \neq I_\alpha^{\mathsf{H}}(Y;X)$ and $I_\alpha^{\mathsf{A}}(X;Y) \neq I_\alpha^{\mathsf{A}}(Y;X)$ in general; and also one may think it hopeless to apply the technique even in Proof II, since each of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ does not satisfy the (weak) chain rule in general. Nonetheless, our idea is to follow that of Proof II: our technical point is not to use the (weak) chain rule, but to successfully utilize the equality condition of conditioned monotonicity in the case of PC. Owing to our new results about conditional Rényi entropies in Sections 2 and 3, we can prove extension of Shannons's impossibility in a highly simple and unified way compared to other ways used for the proofs in the bounds in Section 4.2, as will be seen in Section 5.3.

5.3 Lower Bounds

We newly derive a family of lower bounds on secret-keys with respect to (conditional) Rényi entropies in a comprehensive way. And, it will be seen that our new bounds include all the existing bounds in Section 4.2 as special cases.

Theorem 9 *For arbitrary $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Let $\Pi = ([P_{ED}], \pi_{enc}, \pi_{dec})$ be an encryption protocol satisfying perfect correctness. Then, we have the following bounds.*

(i) (Lower bound on encryption-keys' size) If Π satisfies $R_\alpha(C) \leq R_\alpha(C|M) + \epsilon$ and π_{enc} is deterministic, we have $R_\alpha(E) \geq R_\alpha(C) - \epsilon$.

(ii) (Lower bound on decryption-keys' size) Suppose that Π satisfies $R_\alpha(M) \leq R_\alpha(M|C) + \epsilon$. Then, we have $R_\alpha(D) \geq R_\alpha(M) - \epsilon$.

(iii) (Lower bound on ciphertexts' size) It holds that $R_\alpha(C) \geq R_\alpha(M)$.

Proof. First, we can show (i) as follows.

$$R_\alpha(C) \leq R_\alpha(C|M) + \epsilon \overset{\text{(a)}}{\leq} R_\alpha(CE|M) + \epsilon \overset{\text{(b)}}{=} R_\alpha(E|M) + \epsilon \overset{\text{(c)}}{=} R_\alpha(E) + \epsilon, \tag{29}$$

where (a) follows from Theorem 6 (i), (b) follows from Theorem 6 (ii) since π_{enc} is deterministic, and (c) follows from that M and E are independent.

Secondly, we can show (ii) as follows.

$$R_\alpha(M) \leq R_\alpha(M|C) + \epsilon \overset{\text{(a)}}{\leq} R_\alpha(MD|C) + \epsilon \overset{\text{(b)}}{=} R_\alpha(D|C) + \epsilon \overset{\text{(c)}}{\leq} R_\alpha(D) + \epsilon, \tag{30}$$

where (a) follows from Theorem 6 (i), (b) follows from Theorem 6 (ii) since Π meets perfect correctness, and (c) follows from that both $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ satisfy CRE (see Theorem 4).

Finally, we show (iii). Let $\hat{K} := (E, D)$. Then, we get

$$R_\alpha(M) \overset{\text{(a)}}{=} R_\alpha(M|\hat{K}) \overset{\text{(b)}}{\leq} R_\alpha(MC|\hat{K}) \overset{\text{(c)}}{=} R_\alpha(C|\hat{K}) \overset{\text{(d)}}{\leq} R_\alpha(C), \tag{31}$$

where (a) follows from that $\hat{K}$ and M are independent, (b) follows from Theorem 6 (i), (c) also follows from Theorem 6 (ii) since Π meets perfect correctness, and (d) follows from that both $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ satisfy CRE (see Theorem 4). □

In particular, we obtain the following results for symmetric-key encryption protocols.

Corollary 1 *For arbitrary $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets perfect correctness. Then, we have the following.*

(i) If Π satisfies $R_\alpha(M) \leq R_\alpha(M|C) + \epsilon$, it holds that $R_\alpha(K) \geq R_\alpha(M) - \epsilon$.

(ii) If Π satisfies $R_\alpha(C) \leq R_\alpha(C|M) + \epsilon$ and π_{enc} is deterministic, we have $R_\alpha(K) \geq R_\alpha(C) - \epsilon$ and $R_\alpha(C) \geq R_\alpha(M)$.

Proof. Suppose $E = D = K$ in Theorem 9. The statement (i) follows from (ii) of Theorem 9. Furthermore, the statement (ii) follows from (i) and (iii) of Theorem 9. □

Corollary 2 *For arbitrary $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets perfect correctness and ϵ-secrecy with respect to $R_\alpha(\cdot|\cdot)$. Then, it holds that $R_\alpha(K) \geq R_\alpha(M) - \epsilon$.*

Interestingly, the following proposition shows that traditional perfect secrecy implies a family of lower bounds of the Rényi entropy $R_\alpha(\cdot)$ for all $\alpha \geq 0$.

Corollary 3 *Let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets both perfect correctness and perfect secrecy. Then, for any $\alpha \geq 0$, it holds that $R_\alpha(K) \geq R_\alpha(M)$. In particular, if π_{enc} is deterministic, we have $R_\alpha(K) \geq R_\alpha(C) \geq R_\alpha(M)$.*

Proof. For arbitrary $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ or $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. If Π meets perfect secrecy, or equivalently, M and C are independent, it holds that $R_\alpha(M|C) = R_\alpha(M)$ and $R_\alpha(C|M) = R_\alpha(C)$. Then, from Corollary 1 and by applying $\epsilon = 0$, the proof is completed. □

Remark 8 *Note that the Shannon's bounds (i.e., Proposition 4) are special cases of Corollary 3, since they are obtained by applying $\alpha = 0, 1$ in Corollary 3*[8]*. Also, Dodis's bound (i.e., Proposition 5) is a special case of Corollary 3, since it is obtained by applying $\alpha = \infty$ in Corollary 3. Furthermore, Alimomeni and Safavi-Naini's bound (i.e., Proposition 6) is a special case of Corollary 2, since it is obtained by applying $\epsilon = 0$ and $R_\infty^{\mathsf{avg}}(\cdot|\cdot) = \lim_{\alpha\to\infty} R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ in Corollary 2*[9]*. Therefore, since Corollaries 2 and 3 are special cases of Theorem 9, all the bounds are special cases of ours in Theorem 9.*

5.4 Construction

We note that $H(M|C) = H(M)$ implies $R_\alpha(M|C) = R_\alpha(M)$ for all $\alpha \geq 0$, where $R_\alpha(\cdot|\cdot)$ is $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ or $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Therefore, in this sense security criteria based on the Shannon entropy implies security criteria based on the Rényi entropy. However, the converse is not true in general. Actually, security criteria based on the min-entropy is strictly weaker than that of the Shannon entropy. Although in [6] it is not shown that the lower bound in Proposition 6 is tight for symmetric-key encryption protocols which do not meet perfect security, we can show that it is tight by considering the following simple construction.

Suppose $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0, 1\}$ and $P_K(0) = P_M(0) = p$ with $1/2 < p < 1$. We consider the one-time pad for 1-bit encryption $\Pi_1 = ([P_K], \pi_{enc}, \pi_{dec})$, where $\pi_{enc}(k, m) = k \oplus m$ and $\pi_{dec}(k, c) = k \oplus c$.

Proposition 10 *The above protocol Π_1 does not meet perfect secrecy, and Π_1 satisfies perfect secrecy with respect to $R_\infty^{\mathsf{avg}}(\cdot|\cdot)$, or equivalently $I_\infty^{\mathsf{A}}(M; C) = 0$. Furthermore, it holds that $R_\infty(K) = R_\infty(M)$ in Π_1.*

Proof. For the above protocol Π_1, it holds that

$$P_{M|C}(m|1) = \frac{1}{2} \text{ for any } m \in \{0, 1\}, \quad P_{M|C}(0|0) = \frac{p^2}{p^2 + (1-p)^2}, \quad P_{M|C}(1|0) = \frac{(1-p)^2}{p^2 + (1-p)^2}.$$

Hence, it is clear that Π_1 does not meet perfect secrecy. On the other hand, we have

$$\begin{aligned} R_\infty^{\mathsf{avg}}(M|C) &= -\log\left(\sum_c P_C(c) \max_m P_{M|C}(m|c)\right) = -\log\left(P_C(0) \cdot \frac{p^2}{p^2 + (1-p)^2} + P_C(1) \cdot \frac{1}{2}\right) \\ &= -\log\left(p^2 + p(1-p)\right) = -\log p = R_\infty(M). \end{aligned}$$

[8] Strictly speaking, the bounds are slightly more general than Shannon's ones, since we have removed the assumption that π_{enc} and π_{dec} are deterministic

[9] Strictly speaking, the bound is slightly more general than Alimomeni and Safavi-Naini's one, since we do not assume that π_{enc} and π_{dec} are deterministic.

In addition, it is obvious that $R_\infty(K) = R_\infty(M) = -\log p$. Therefore, the proof is completed. □

Remark 9 *In the above construction Π_1, we note that $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(M|C) = R_\infty^{\mathsf{wst}}(M|C) < R_\infty(M)$. Therefore, Π_1 does not meet perfect secrecy with respect to $R_\infty^{\mathsf{wst}}(\cdot|\cdot)$. Also, we note that $R_\infty^{\mathsf{wst}}(C|M) < R_\infty(C)$, and Π_1 illustrates $I_\infty^{\mathsf{A}}(M;C) \neq I_\infty^{\mathsf{A}}(C;M)$ for the random variables M and C, while Π_1 meets $I_\infty^{\mathsf{H}}(M;C) = I_\infty^{\mathsf{H}}(C;M)(\neq 0)$.*

In general, for any sufficiently large $\alpha \geq 0$, the following construction shows that the lower bound in Corollary 2 for symmetric-key encryption protocols is tight in an asymptotical sense.

Suppose $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}$ and $P_M(0) = p$ and $P_K(0) = q$ such that $p = \frac{1}{2}(1+\delta_1)$, $q = p + \delta_2$, and $0 < \delta_i$ and $\delta_i = o(1/\alpha)$ for $i = 1, 2$. We consider the one-time pad for 1-bit encryption $\Pi_2 = ([P_K], \pi_{enc}, \pi_{dec})$, where $\pi_{enc}(k,m) = k \oplus m$ and $\pi_{dec}(k,c) = k \oplus c$.

Proposition 11 *For a sufficiently large $\alpha \geq 0$, the above protocol Π_2 does not meet perfect secrecy, and Π_2 meets ϵ-secrecy with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$, or equivalently $I_\alpha^{\mathsf{H}}(M;C) = \epsilon$, with $\epsilon = o(1/\alpha)$. Furthermore, it holds that $R_\alpha(K) = R_\alpha(M) - o(1/\alpha)$ in Π_2.*

Proof. See Appendix A.5. □

Remark 10 *Note that the above construction Π_2 meets ϵ-secrecy with respect to $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$, or equivalently $I_\alpha^{\mathsf{A}}(M;C) = \epsilon$, with $\epsilon = o(1/\alpha)$. This fact directly follows from Proposition 11 and the inequality $I_\alpha^{\mathsf{A}}(M;C) \leq I_\alpha^{\mathsf{H}}(M;C)$. Also, by calculation (see Appendix A.5), we can see that Π_2 illustrates $I_\alpha^{\mathsf{H}}(M;C) \neq I_\alpha^{\mathsf{H}}(C;M)$ for the random variables M and C.*

6 Generalizing Shannon's Impossibility in Secret Sharing

As in the case of encryption in Section 5, we can also consider and show similar results for secret sharing protocols. In this section, we use the following notation: for any finite set $\mathcal{Z}$, let $\mathcal{P}(\mathcal{Z}) := \{Z \subset \mathcal{Z}\}$ be the family of all subsets of $\mathcal{Z}$. Also, for any finite set $\mathcal{Z}$ and any non-negative integer z, let $\mathcal{P}(\mathcal{Z}, z) := \{Z \subset \mathcal{Z} \mid |Z| \leq z\}$ be the family of all subsets of $\mathcal{Z}$ whose cardinality is less than or equal to z.

6.1 The Model and Security Definition

Let $\{1, 2, \ldots, n\}$ be a finite set of IDs of n users. Also, for every $i \in \{1, 2, \ldots, n\}$, let $\mathcal{V}_i$ be a finite set of shares of the user i, and P_{V_i} is its associated distribution on $\mathcal{V}_i$. In addition, let $\mathcal{S}$ be a finite set of secret information and P_S its associated distribution.

Let $\Pi = ([P_S], \pi_{share}, \pi_{comb})$ be a (t,n)-*secret sharing protocol*, where $1 \leq t \leq n$, as defined below:

- $[P_S]$ is a sampling algorithm for secret information, and it outputs a secret $s \in \mathcal{S}$ according to a probability distribution P_S;
- π_{share} is a randomized algorithm for generating shares for all users, and it is executed by a honest entity called *dealer*. It takes a secret $s \in \mathcal{S}$ on input and outputs $(v_1, v_2, \ldots, v_n) \in \prod_{i=1}^n \mathcal{V}_i$; and
- π_{comb} is an algorithm for recovering a secret. It takes t shares on input and outputs a secret $s \in \mathcal{S}$.

In this paper, we assume that Π meets *perfect correctness*: for any possible secret $s \in \mathcal{S}$, and for all possible shares $(v_1, v_2, \ldots, v_n) \leftarrow \pi_{share}(s)$, it holds that $\pi_{comb}(v_{i_1}, v_{i_2}, \ldots, v_{i_t}) = s$ for any subset $\{i_1, i_2, \ldots, i_t\} \subset \{1, 2, \ldots, n\}$.

In the following, for any subset $U := \{i_1, i_2, \ldots, i_u\} \subset \{1, 2, \ldots, n\}$, we use the notation $V_U := (V_{i_1}, V_{i_2}, \ldots, V_{i_u})$. Then, we give security formalization for a (t,n)-secret sharing protocol as follows.

Definition 7 (Security) For $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Let Π be a (t,n)-secret sharing protocol. Then, Π is said to meet *ϵ-security with respect to* $R_\alpha(\cdot|\cdot)$, if for any colluding group $\mathcal{W} \in \mathcal{P}(\{1,2,\ldots,n\}, t-1)$ it satisfies

$$R_\alpha(S) - R_\alpha(S|V_{\mathcal{W}}) \leq \epsilon.$$

In particular, Π meets *perfect security with respect to* $R_\alpha(\cdot|\cdot)$ if $\epsilon = 0$ above.

Note that the traditional (t,n)-threshold secret sharing protocol is equivalent to the (t,n)-secret sharing protocol with perfect security with respect to $R_1(\cdot|\cdot) = H(\cdot|\cdot)$[10].

As in Section 4.3, the security criteria in Definition 7 is explained in terms of failure probability of guessing by colluding groups: Let $\mathcal{W} \in \mathcal{P}(\{1,2,\ldots,n\}, t-1)$ be any colluding group, and suppose that $\mathcal{W}$ takes arbitrary strategy g for guessing a secret S. Let $\hat{S} := g(V_{\mathcal{W}})$ and $P_e := \Pr\{S \neq \hat{S}\}$. Without loss of generality, we assume $P_e \leq 1 - 1/|\mathcal{S}|$. Then, we have the following propositions, and the proofs are shown in a similar way in Section 4.3.

Theorem 10 *The failure probability of guessing by colluders $\mathcal{W}$ is lower-bounded by*

$$P_e \geq 1 - \exp\left\{\frac{1-\alpha}{\alpha} R_\alpha(S)\right\} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{H}}(S; V_{\mathcal{W}})\right\}.$$

In particular, if P_S is the uniform distribution, we have

$$P_e \geq 1 - |\mathcal{S}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{H}}(S; V_{\mathcal{W}})\right\}.$$

Proposition 12 *The failure probability of guessing by colluders $\mathcal{W}$ is lower-bounded by*

$$P_e \geq 1 - |\mathcal{S}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{A}}(S_{1/\alpha}; V_{\mathcal{W}})\right\}.$$

In particular, if P_S is the uniform distribution, we have

$$P_e \geq 1 - |\mathcal{S}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{A}}(S; V_{\mathcal{W}})\right\}.$$

6.2 Lower Bounds

First, the following tight lower bounds for (t,n)-threshold secret sharing protocols are well known.

Proposition 13 ([33]) *Let Π be a (t,n)-threshold secret sharing protocol. Then, for every $i \in \{1,2,\ldots,n\}$, it holds that $H(V_i) \geq H(S)$ and $|\mathcal{V}_i| \geq |\mathcal{S}|$.*

We next show a new lower bound in a general setting, and we will see that Proposition 13 is a special case of ours (see Remark 11).

Theorem 11 *For arbitrary $\alpha \geq 0$, let $R_\alpha(\cdot|\cdot)$ be any of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Let Π be a (t,n)-secret sharing scheme which meets ϵ-security with respect to $R_\alpha(\cdot|\cdot)$. Then, we have $R_\alpha(V_i) \geq R_\alpha(S) - \epsilon$ for every $i \in \{1,2,\ldots,n\}$.*

Proof. For any $i \in \{1,2,\ldots,n\}$, we take $\mathcal{W} \in \mathcal{P}(\{1,2,\ldots,n\}, t-1)$ such that $i \notin \mathcal{W}$ and $|\mathcal{W}| = t-1$. Then, we have

$$R_\alpha(S) \leq R_\alpha(S|V_{\mathcal{W}}) + \epsilon \overset{\text{(a)}}{\leq} R_\alpha(S, V_i|V_{\mathcal{W}}) + \epsilon \overset{\text{(b)}}{=} R_\alpha(V_i|V_{\mathcal{W}}) + \epsilon \overset{\text{(c)}}{\leq} R_\alpha(V_i) + \epsilon, \tag{32}$$

where (a) follows from Theorem 6 (i), (b) follows from Theorem 6 (ii) since Π meets perfect correctness, and (c) follows from that both $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ satisfy CRE (see Theorem 4). □

As in the case of encryption (i.e., Corollary 3), Theorem 11 implies the following result. The proof is shown in the same way as that of Corollary 3.

[10] This condition is equivalent to $I(S; V_{\mathcal{W}}) = 0$, or equivalently, S and $V_{\mathcal{W}}$ are independent (i.e., $P_{SV_{\mathcal{W}}} = P_S P_{V_{\mathcal{W}}}$) for all $\mathcal{W} \in \mathcal{P}(\{1,2,\ldots,n\}, t-1)$.

Corollary 4 *Let Π be a (t, n)-threshold secret sharing protocol. Then, for any $\alpha \geq 0$, it holds that $R_\alpha(V_i) \geq R_\alpha(S)$ for every $i \in \{1, 2, \ldots, n\}$.*

Remark 11 *Note that Proposition 13 is a special case of Corollary 4, since the bounds in Proposition 13 are obtained by applying $\alpha = 0, 1$ in Corollary 4.*

By applying $\alpha = \infty$ in Corollary 4, we obtain the bound $R_\infty(V_i) \geq R_\infty(S)$ for every $i \in \{1, 2, \ldots, n\}$, which can be considered as an analogue of Dodis's bound (i.e., Proposition 5) in the context of secret sharing protocols.

By Theorem 11 and the fact that $R_\infty^{\mathsf{avg}}(\cdot|\cdot) = \lim_{\alpha \to \infty} R_\alpha^{\mathsf{A}}(\cdot|\cdot)$, an analogue of Alimomeni and Safavi-Naini's bound (i.e., Proposition 6) in secret sharing can also be obtained. To the best of authors' knowledge, this kind of security is first proposed in the context of secret sharing protocols in this paper.

Corollary 5 *Let Π be a (t, n)-secret sharing protocol which meets perfect security with respect to $R_\infty^{\mathsf{avg}}(\cdot|\cdot)$. Then, we have $R_\infty(V_i) \geq R_\infty(S)$ for every $i \in \{1, 2, \ldots, n\}$.*

6.3 Construction

The lower bound in Corollary 5 is almost tight, since there exits the following construction Π_3 which is an analogue of the construction Π_1 in Section 5.

Let S, and $V_1, V_2, \ldots, V_n$ be binary random variables. Assume that S and $V_1, V_2, \ldots, V_{n-1}$ are independent and they satisfy $P_S(0) = P_{V_1}(0) = \cdots = P_{V_n-1}(0) = p$, for $1/2 < p < 1$. Then, we generate V_n by $V_n := S \oplus V_1 \oplus V_2 \oplus \cdots \oplus V_{n-1}$.

Proposition 14 *The construction Π_3 realizes an (n, n)-secret sharing protocol which meets perfect security with respect to $R_\infty^{\mathsf{avg}}(\cdot|\cdot)$. In addition, the share sizes in Π_3 are almost optimal (or the lower bound in Corollary 5 is almost tight) in the sense that $R_\infty(V_i) = R_\infty(S)$ for $i = 1, 2, \ldots, n-1$, but $R_\infty(V_n) > R_\infty(S)$.*

Proof. See Appendix A.6. □

Remark 12 *In the construction Π_3, S and $V_{\mathcal{W}}$ are not statistically independent if $n \in \mathcal{W}$, but are independent if $n \notin \mathcal{W}$. Therefore, Π_3 is not an (n, n)-threshold secret sharing protocol.*

7 Further Extension of Our Results

In Sections 5 and 6, we have derived lower bounds in a generic and unified manner by using security criteria based on conditional Rényi entropies (i.e., by using $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$). In this section, from a theoretical interest, we further extend the results to a wide class of conditional entropies which includes $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$.

7.1 A Class of Pairs of Entropies and Conditional Entropies under Consideration

In the proof of our bound in Theorem 9, we note that it is crucial to use the properties of CRE and conditioned monotonicity of $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$. Therefore, in order to further extend Theorem 9 in a generic way, we consider a wide class of entropies and conditional entropies satisfying several properties including CRE and conditioned monotonicity. In addition to the above consideration, we take into account the axiomatic consideration in Section 2.3 for conditional entropies. From the aspect above, we define the following class of pairs of entropy and conditional entropy functions.

Definition 8 *Let Σ be a class of pairs of entropy and conditional entropy functions such that, for any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$, it satisfies the following conditions.*

1. *(Unconditioning implies entropy) If Y is the random variable taking a constant (i.e., Y is deterministic), a conditional entropy function implies an entropy function $F(\cdot|Y) = F(\cdot)$, namely $F(X|Y) = F(X)$ for any random variable X.*
2. *(Symmetricity) $F(X|Y)$ is symmetric with respect to $\{P_{X|Y}(x|y)\}_{x\in\mathcal{X}}$ for each $y \in \mathcal{Y}$, and $\{P_Y(y)\}_{y\in\mathcal{Y}}$.*
3. *(Continuity) $F(X|Y)$ is a continuous function with respect to P_{XY}.*
4. *(Uniformity implies maximum) $F(X|Y) = 1$ if a binary random variable X is uniformly distributed for given Y.*
5. *(Non-negativity) $F(X|Y) \geq 0$ for all random variables X and Y.*
6. *(Conditioned monotonicity) (i) $F(X|Z) \leq F(XY|Z)$ for all random variables X, Y, and Z; and in particular , (ii) $F(X|Z) = F(XY|Z)$ if $Y = f(X,Z)$ for some (deterministic) mapping f.*
7. *(CRE) $F(X|Y) \leq F(X)$ for all random variables X and Y, where equality holds if X and Y is independent.*

Note that all the properties in Definition 8 are focused on and discussed in Section 2.3, and more importantly, we have explained why we consider all the properties as important and reasonable ones for conditional entropies. As we have seen, the class Σ actually contains $(R_\alpha(\cdot), R_\alpha^{\mathsf{H}}(\cdot|\cdot))$ and $(R_\alpha(\cdot), R_\alpha^{\mathsf{A}}(\cdot|\cdot))$ for all $\alpha \geq 0$. In addition, Σ contains $(R_\alpha(\cdot), R_\alpha^{\mathsf{RW}}(\cdot|\cdot))$ for any $\alpha > 1$, and its proof is straightforward from [8, 21]. Therefore, we have the following proposition.

Proposition 15 *The class Σ in Definition 8 contains*

(i) $(R_\alpha(\cdot), R_\alpha^{\mathsf{H}}(\cdot|\cdot))$ *for any* $\alpha \geq 0$*;*
(ii) $(R_\alpha(\cdot), R_\alpha^{\mathsf{A}}(\cdot|\cdot))$ *for any* $\alpha \geq 0$*; and*
(iii) $(R_\alpha(\cdot), R_\alpha^{\mathsf{RW}}(\cdot|\cdot))$ *for any* $\alpha > 1$*.*

By using the class Σ, we further extend our results in Sections 5 and 6, as will be seen in the following sections.

7.2 Encryption

The model of encryption protocols is the same as that in Section 5.1. However, we consider the following security definition instead of Definition 6.

Definition 9 (Secrecy) Let Π be an encryption protocol. Then, for any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 8, Π is said to meet *ϵ-secrecy with respect to* $(F(\cdot), F(\cdot|\cdot))$, if it satisfies

$$F(M) - F(M|C) \leq \epsilon.$$

Then, we derive a family of lower bounds on secret-keys for all entropy and conditional entropy functions in Σ in Definition 8 in a comprehensive way. Theorem 12, and Corollaries 6, 7 and 8 below are extension of Theorem 9, and Corollaries 1, 2 and 3, respectively. Their proofs can be shown in the same way as those in Section 5.3, and we omit them here.

Theorem 12 *Let $\Pi = ([P_{ED}], \pi_{enc}, \pi_{dec})$ be an encryption protocol satisfying perfect correctness. Then, for any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 8, we have the following.*

(i) (Lower bound on encryption-keys' size) If Π satisfies $F(C) \leq F(C|M) + \epsilon$ and π_{enc} is deterministic, we have $F(E) \geq F(C) - \epsilon$.
(ii) (Lower bound on decryption-keys' size) Suppose that Π satisfies $F(M) \leq F(M|C) + \epsilon$. Then, we have $F(D) \geq F(M) - \epsilon$.
(iii) (Lower bound on ciphertexts' size) It holds that $F(C) \geq F(M)$.

Corollary 6 *Let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets perfect correctness. For any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 8, we have the following.*

(i) If Π satisfies $F(M) \leq F(M|C) + \epsilon$, it holds that $F(K) \geq F(M) - \epsilon$.
(ii) If Π satisfies $F(C) \leq F(C|M) + \epsilon$ and π_{enc} is deterministic, we have $F(K) \geq F(C) - \epsilon$ and $F(C) \geq F(M)$.

Corollary 7 *Let $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 8, and let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets perfect correctness and ϵ-secrecy with respect to $(F(\cdot), F(\cdot|\cdot))$. Then, it holds that $F(K) \geq F(M) - \epsilon$.*

Corollary 8 *Let $\Pi = ([P_K], \pi_{enc}, \pi_{dec})$ be a symmetric-key encryption protocol which meets both perfect correctness and perfect secrecy. Then, for any entropy function $F(\cdot)$ appearing in Σ in Definition 8, it holds that $F(K) \geq F(M)$. In particular, if π_{enc} is deterministic, we have $F(K) \geq F(C) \geq F(M)$.*

7.3 Secret Sharing

The model of secret sharing protocols is the same as that in Section 6.1. We consider the following security definition instead of Definition 7.

Definition 10 (Security) Let Π be a (t, n)-secret sharing protocol. Then, for any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 8, Π is said to meet *ϵ-security with respect to* $(F(\cdot), F(\cdot|\cdot))$, if for all $\mathcal{W} \in \mathcal{P}(\{1, 2, \ldots, n\}, t-1)$ it satisfies $F(S) - F(S|V_{\mathcal{W}}) \leq \epsilon$.

Then, as in the case of encryption, we can derive a family of lower bounds on shares for all entropy and conditional entropy functions in Σ in Definition 8 in a comprehensive way. Theorem 13 and Corollary 9 below are extension of Theorem 11 and Corollary 4, respectively. The proof of Theorem 13 is given in the same way as that of Theorem 11, and we omit it here.

Theorem 13 *For any $(F(\cdot), F(\cdot|\cdot))$ in Σ in Definition 8 and any (t, n)-secret sharing scheme Π which meets ϵ-security with respect to $(F(\cdot), F(\cdot|\cdot))$, it holds that $F(V_i) \geq F(S) - \epsilon$ for every $i \in \{1, 2, \ldots, n\}$.*

Corollary 9 *Let Π be a (t, n)-threshold secret sharing protocol. Then, for any entropy function $F(\cdot)$ appearing in Σ in Definition 8, it holds that $F(V_i) \geq F(S)$ for every $i \in \{1, 2, \ldots, n\}$.*

References

1. Shannon, C.: A mathematical theory of communication. Bell Systems Technical Journal **27** (July and Oct. 1948) 379–423
2. Hartley, R.V.L.: Transmission of information. Bell System Technical Journal **7**(3) (July 1928) 535–563
3. Haståd, J., Impagliazzo, R., Levin, L.A., Luby, M.: A pseudorandom generator from one-way function. SIAM Journal of Computing (1994) 1364–1396
4. Shannon, C.E.: Communication theory of secrecy systems. Bell Tech. J. **28** (Oct. 1949) 656–715
5. Dodis, Y.: Shannon impossibility, revisited. Proc. of the 6th International Conference on Information Theoretic Security (ICITS 2012), LNCS7412, Springer-Verlag (August 2012) 100–110 IACR Cryptology ePrint Archive (preliminary short version): `http://eprint.iacr.org/2012/053`.
6. Alimomeni, M., Safavi-Naini, R.: Guessing secrecy. Proc. of the 6th International Conference on Information Theoretic Security (ICITS 2012), LNCS7412, Springer-Verlag (August 2012) 1–13
7. Rényi, A.: On measures of information and entropy. Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960 (1961) 547–561
8. Teixeira, A., Matos, A., Antunes, L.: Conditional Rényi entropies. IEEE Trans. Information Theory **58**(7) (July 2012) 4273–4277
9. Hayashi, M.: Exponential decreasing rate of leaked information in universal random privacy amplification. IEEE Trans. Information Theory **57**(6) (2011) 3989–4001
10. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In: Advances in Cryptology - EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004, Proceedings. Volume 3027 of Lecture Notes in Computer Science., Springer (2004) 523–540
11. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Second edn. Wiley and Interscience (2006)

12. Fano, R.M.: Class notes for transmission of information (course 6.574). Technical report, MIT, Cambridge, MA (1952)
13. Stinson, D.R.: Cryptography: Theory and Practice. Third edn. Chapman & Hall/CRC (2005)
14. Shamir, A.: How to share a secret. Communications of the ACM **22**(11) (1979) 612–613
15. Blakley, G.R.: Safeguarding cryptographic keys. AFIPS 1979 National Computer Conference **48** (1979) 313–317
16. Vernam, G.S.: Cipher printing telegraph systems for secret wire and radio telegraphic communications. J. of American Institute for Electrical Engineering **45** (1926) 109–115
17. Arimoto, S.: Information measures and capacity of order α for discrete memoryless channels. Colloquia Mathematica Societatis János Bolyai, 16. Topics in Information Theory (1975) 41–52
18. Cachin, C.: Entropy Measures and Unconditional Security in Cryptography. PhD thesis, Swiss Federal Institute of Technology, Zürich, Switzerland (1997)
19. Jizba, P., Arimitsu, T.: Generalized statistics: Yet another generalization. Physica A **340** (2004) 110–116
20. Jizba, P., Arimitsu, T.: The world according to Rényi: Thermodynamics of multifractal systems. Annals of Physics **312** (2004) 17–59
21. Renner, R., Wolf, S.: Simple and tight bounds for information reconciliation and privacy amplification. Advances in Cryptology–ASIACRYPT2005, LNCS4515, Springer-Verlag (2005) 199–216
22. Fujishige, S.: Polymatroidal dependence structure of a set of random variables. Information and Control **39** (1978) 55–72
23. Dodis, Y., Katz, J., Reyzin, L., Smith, A.: Robust fuzzy extractors and authenticated key agreement from close secrets. Advances in Cryptology–CRYPTO2006, LNCS4117, Springer-Verlag (2006) 232–250
24. Dodis, Y., Katz, J., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. SIAM Journal on Computing **38**(1) (2008) 97–139
25. Barak, B., Dodis, Y., Krawcayk, H., Pereira, O., Pietrzak, K., Standaert, F.X., Yu, Y.: Leftover hash lemma, revisited. Advances in Cryptology–CRYPTO2011, LNCS7417, Springer-Verlag (2011) 1–20
26. Dodis, Y., Kanukurthi, B., Katz, J., Reyzin, L., Smith, A.: Robust fuzzy extractors and authenticated key agreement from close secrets. IEEE Trans. on Information Theory (2012) 6207–6222
27. Dodis, Y., Yu, Y.: Overcoming weak expectations. Tenth Workshop on Theory of Cryptography–TCC2013 (2013) to appear.
28. Katzenbeisser, S., Kocabaş, Ü., Rožić, V., Sadeghi, A.R., Verbauwhede, I., Wachsmann, C.: PUFs: Myth, fact or busted? a security evaluation of physically unclonable functions (PUFs) cast in silicon. Proc. of CHES2012, LNCS7248 (2012) 283–301
29. Arikan, E.: An inequality on guessing and its application to sequential decoding. IEEE Trans. Information Theory **42**(1) (1996) 99–105
30. Polyanskiy, Y., Verdú, S.: Arimoto channel coding converse and Rényi divergence. Forty-Eighth Annual Allerton Conference (2010) 1327–1333
31. Polyanskiy, Y., Poor, V., Verdú, S.: Channel coding rate in the finite blocklength regime. IEEE Trans. Inform. Theory **56**(5) (2010) 2307–2359
32. Gallager, R.G.: A simple derivation of the coding theorem and some applications. IEEE Trans. Inform. Theory **11**(1) (1965) 41–52
33. Karnin, E.D., Greene, J.W., Hellman, M.E.: On secret sharing systems. IEEE Trans. Inform. Theory **29**(1) (1983) 35–41

A Technical Proofs

A.1 Proof of Theorem 1

Proof: Note that $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ is written as

$$\begin{aligned}
R_\alpha^{\mathsf{A}}(X|Y) &:= \frac{\alpha}{1-\alpha}\log\sum_{y\in\mathcal{Y}}P_Y(y)\left(\sum_{x\in\mathcal{X}}P_{X|Y}(x|y)^\alpha\right)^{1/\alpha} \\
&= \frac{\alpha}{1-\alpha}\log\mathbb{E}_Y\left[\left(\sum_{x\in\mathcal{X}}P_{X|Y}(x|Y)^\alpha\right)^{1/\alpha}\right], \qquad (33)
\end{aligned}$$

$$\begin{aligned}
R_\alpha^{\mathsf{H}}(X|Y) &:= \frac{\alpha}{1-\alpha}\log\left(\sum_{y\in\mathcal{Y}}P_Y(y)\sum_{x\in\mathcal{X}}P_{X|Y}(x|y)^\alpha\right)^{1/\alpha} \\
&= \frac{\alpha}{1-\alpha}\log\left(\mathbb{E}_Y\left[\sum_{x\in\mathcal{X}}P_{X|Y}(x|Y)^\alpha\right]\right)^{1/\alpha}. \qquad (34)
\end{aligned}$$

Due to Jensen's inequality, in the case of $0 \le \alpha \le 1$, it holds that

$$\left(\mathbb{E}_Y\left[\sum_{x\in\mathcal{X}} P_{X|Y}(x|Y)^\alpha\right]\right)^{1/\alpha} \le \mathbb{E}_Y\left[\left(\sum_{x\in\mathcal{X}} P_{X|Y}(x|Y)^\alpha\right)^{1/\alpha}\right]. \tag{35}$$

Combining (33)–(35), we can conclude that $R_\alpha^{\mathsf{H}}(X|Y) \le R_\alpha^{\mathsf{A}}(X|Y)$.

Similar arguments can be applied to the case of $\alpha \ge 1$, which completes the proof. □

A.2 Proof of Theorem 2

(i) The equality $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{A}}(X|Y) = H(X|Y)$ is proved in [17]. $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(X|Y) = H(X|Y)$ is easily verified by the L'Hospital's rule. Namely, we have

$$\begin{aligned}\lim_{\alpha\to 1} R_\alpha^{\mathsf{H}}(X|Y) &= -\lim_{\alpha\to 1}\frac{\mathrm{d}}{\mathrm{d}\alpha}\log\mathbb{E}_Y\left[\sum_{x\in X} P_{X|Y}(x|Y)^\alpha\right] = -\lim_{\alpha\to 1}\mathbb{E}_Y\left[\sum_{x\in X}\frac{\mathrm{d}}{\mathrm{d}\alpha}P_{X|Y}(x|Y)^\alpha\right]\\ &= -\mathbb{E}_Y\left[\sum_{x\in X} P_{X|Y}(x|Y)\log P_{X|Y}(x|Y)\right] = \sum_{y\in\mathcal{Y}} P_Y(y)H(X|Y=y).\end{aligned} \tag{36}$$

(ii) We first prove $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{A}}(X|Y) = R_\infty^{\mathsf{avg}}(X|Y)$. Observing that

$$\max_x P_{X|Y}(x|y) \le \left\{\sum_x P_{X|Y}(x|y)^\alpha\right\}^{1/\alpha} \le |\mathcal{X}|^{1/\alpha}\max_x P_{X|Y}(x|y) \tag{37}$$

holds for arbitrarily fixed $y \in \mathcal{Y}$, it holds that

$$\lim_{\alpha\to\infty}\frac{\alpha}{1-\alpha}\log\sum_y P_Y(y)\left\{\sum_x P_{X|Y}(x|y)^\alpha\right\}^{1/\alpha} = -\log\sum_y P_Y(y)\max_x P_{X|Y}(x|y), \tag{38}$$

which means that $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{A}}(X|Y) = R_\infty^{\mathsf{avg}}(X|Y)$ holds.

Then, we prove $\lim_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(X|Y) = R_\infty^{\mathsf{wst}}(X|Y)$. We can check that for every fixed $y \in \mathcal{Y}$

$$\max_x P_{X|Y}(x|y)^\alpha \le \sum_x P_{X|Y}(x|y)^\alpha \le |\mathcal{X}|\max_x P_{X|Y}(x|y)^\alpha. \tag{39}$$

The expectations of the upper and the lower bounds in (39) with respect to Y can be further bounded as

$$\sum_y P_Y(y)|\mathcal{X}|\max_x P_{X|Y}(x|y)^\alpha \le |\mathcal{X}| \max_{\substack{x\in\mathcal{X}\\ y\in\operatorname{supp}P_Y}} P_{X|Y}(x|y)^\alpha \tag{40}$$

and

$$P_Y(y^*)\max_x P_{X|Y}(x|y^*)^\alpha \le \sum_y P_Y(y)\max_x P_{X|Y}(x|y)^\alpha \tag{41}$$

respectively, where we define that $y^* \in \operatorname{supp}P_Y$ attains the maximum of $P_{X|Y}(x|y)$ over the set $\mathcal{X}\times\operatorname{supp}P_Y$.

Now, we can assume that α is sufficiently large, say $\alpha > 1$. Then, noticing that $1/(1-\alpha) < 0$ and from (39)–(41), we have

$$R_\alpha^{\mathsf{H}}(X|Y) \ge \frac{1}{1-\alpha}\log\left\{|\mathcal{X}| \max_{\substack{x\in\mathcal{X}\\ y\in\operatorname{supp}P_Y}} P_{X|Y}(x|y)^\alpha\right\}$$

and

$$R_\alpha^{\mathsf{H}}(X|Y) \le \frac{1}{1-\alpha}\log\left\{P_Y(y^*) \max_{\substack{x\in\mathcal{X}\\ y\in\operatorname{supp}P_Y}} P_{X|Y}(x|y)^\alpha\right\}.$$

Hence, we have $\liminf_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(X|Y) \ge -\log\max_{x,y} P_{X|Y}(x|y)$, and $\limsup_{\alpha\to\infty} R_\alpha^{\mathsf{H}}(X|Y) \le -\log\max_{x,y} P_{X|Y}(x|y)$, since $|\mathcal{X}|$ is finite, which imply the claim of the proposition.

A.3 Proof of Proposition 3

(i) For $0 \leq \alpha < 1$, it follows that

$$\begin{aligned}\sum_{x,y} \frac{W(x|y)^{\alpha}}{V(x|y)^{\alpha-1}} Q(y) &= \sum_{x,y} \left(\frac{W(x|y)}{V(x|y)} \right)^{\alpha} P_{YZ}(x,y) = \mathbb{E}_{YZ} \left[\left(\frac{W(Y|Z)}{V(Y|Z)} \right)^{\alpha} \right] \\ &\leq \left\{ \mathbb{E}_{YZ} \left[\frac{W(Y|Z)}{V(Y|Z)} \right] \right\}^{\alpha} = \left\{ \sum_{x,y} \frac{W(x|y)}{V(x|y)} V(x|y) Q(y) \right\}^{\alpha} = 1 \qquad (42)\end{aligned}$$

where the inequality follows from Jensen's inequality. The quality holds if $W(Y|Z)/V(Y|Z)$ is constant with probability 1, which implies $W(\cdot|y) = V(\cdot|y)$ for $y \in \operatorname{supp} Q$.

Similarly, we have

$$\sum_{x,y} \frac{W(x|y)^{\alpha}}{V(x|y)^{\alpha-1}} Q(y) \geq 1$$

for $\alpha \geq 1$. Hence, we have $D_{\alpha}(X_1 \| X_2 | Y) \geq 0$ for all $\alpha \geq 0$.

(ii) The claim directly follows from the L'Hospital's rule:

$$\begin{aligned}\lim_{\alpha \to 1} D_{\alpha}(X_1 \| X_2 | Y) &= \frac{\mathrm{d}}{\mathrm{d}\alpha} \log \sum_{x,y} \left\{ \frac{W(x|y)}{V(x|y)} \right\}^{\alpha} V(x|y) Q(y) \bigg|_{\alpha=1} \\ &= \sum_{x,y} Q(y) W(x|y) \log \frac{W(x|y)}{V(x|y)} = D(W \| V | Q) = D(X_1 \| X_2 | Y). \qquad (43)\end{aligned}$$

□

A.4 Proof of Theorem 7

Let $m := |\mathcal{X}|$. We define a random variable Z and its associated distribution P_Z over $\mathcal{X} \times \mathcal{X}$ as follows. For $(i,j) \in \mathcal{X} \times \mathcal{X}$, we define

$$P_Z(i,j) := \begin{cases} \dfrac{\bar{P}_e}{m} & \text{if } i = j, \\ \dfrac{P_e}{m(m-1)} & \text{if } i \neq j. \end{cases}$$

Also, for any fixed $j \in \mathcal{X}$, we define a distribution $P_{Z_1}(\cdot|j)$ over $\mathcal{X}$ by

$$P_{Z_1}(i|j) := \begin{cases} \bar{P}_e & \text{if } i = j, \\ \dfrac{P_e}{m-1} & \text{if } i \neq j. \end{cases}$$

Note that $P_{Z_1}(i|j) = m P_Z(i,j)$ for $(i,j) \in \mathcal{X} \times \mathcal{X}$. Then, by non-negativity of the conditional α-divergence we have

$$\begin{aligned}0 \leq D_{\alpha}(XY \| Z | Y) &= \frac{1}{\alpha - 1} \log \left[\sum_j P_Y(j) \sum_i \left(\frac{P_{XY}(i,j)}{P_Y(j)} \right)^{\alpha} P_{Z_1}(i|j)^{1-\alpha} \right] \\ &= \frac{1}{\alpha - 1} \log \left[m^{1-\alpha} \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} P_Z(i,j)^{1-\alpha} \right]. \qquad (44)\end{aligned}$$

On the other hand, we get

$$\begin{aligned}
&\sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} P_Z(i,j)^{1-\alpha} \\
&\quad = \sum_{i \neq j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} P_Z(i,j)^{1-\alpha} + \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} P_Z(i,i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \sum_{i \neq j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} + \left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \left(\sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} - \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \\
&\qquad\quad + \left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} \\
&\qquad\quad + \left(\sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \left[\left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} - \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \right]. \qquad (45)
\end{aligned}$$

Therefore, by (44) and (45) we obtain

$$\begin{aligned}
0 \geq \frac{1}{1-\alpha} \log \Bigg\{ & \left(\frac{P_e}{m-1}\right)^{1-\alpha} \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} \\
& + \left(\sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \left[\bar{P}_e^{1-\alpha} - \left(\frac{P_e}{m-1}\right)^{1-\alpha} \right] \Bigg\} \qquad (46)
\end{aligned}$$

For simplicity, we set

$$\begin{aligned}
r &:= \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha}, \qquad & s &:= \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha}, \\
a &:= \left(\frac{P_e}{m-1}\right)^{1-\alpha}, \qquad & b &:= \bar{P}_e^{1-\alpha} - \left(\frac{P_e}{m-1}\right)^{1-\alpha},
\end{aligned}$$

and then (46) is written in the form:

$$\frac{1}{1-\alpha} \log(ar + sb) \leq 0. \qquad (47)$$

Suppose that $0 \leq \alpha < 1$ and $P_e \neq 0$ (i.e., $a > 0$). Then, (47) implies

$$r \leq a^{-1}(1 - sb) = (m-1)^{1-\alpha} P_e^{\alpha-1} + s(1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}). \qquad (48)$$

Here, we note that $1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha} \geq 0$ (resp., ≤ 0) if $P_e \geq 1 - \frac{1}{m}$ (resp., $P_e \leq 1 - \frac{1}{m}$).

Now, we need the following lemma.

Lemma 1 *For a real number $\alpha \geq 0$, it holds that:*

(i) $\bar{P}_e \leq s \leq \bar{P}_e^{\alpha}$ *if* $0 \leq \alpha \leq 1$*;*
(ii) $\bar{P}_e^{\alpha} \leq s \leq \bar{P}_e$ *if* $\alpha \geq 1$*.*

Proof. It is trivial that the statement is true for $\alpha = 0, 1$. Thus, we consider the case of $\alpha \neq 0, 1$.

First, we show (i). Suppose $0 < \alpha < 1$. Then, we have

$$s = \sum_i P_{XY}(i,i)^\alpha P_Y(i)^{1-\alpha} \geq \sum_i P_{XY}(i,i)^\alpha P_{XY}(i,i)^{1-\alpha} = \sum_i P_{XY}(i,i) = \bar{P}_e.$$

On the other hand, we consider a function

$$f(x_1, \ldots, x_m, y_1, \ldots, y_m) = \sum_{i=1}^m x_i^\alpha y_i^{1-\alpha} \quad (0 \leq x_i \leq y_i)$$

subject to the constraints $\sum_{i=1}^m x_i = \bar{P}_e$ and $\sum_{i=1}^m y_i = 1$. For arbitrary $(x_1, \ldots, x_m, y_1, \ldots, y_m)$ satisfying the above condition, we define a random variable W by $\Pr(W = x_i/y_i) = y_i$ for $i = 1, 2, \ldots, m$. Then, since $g(w) := w^\alpha$ is a concave function, it holds that

$$\mathbb{E}_W\left[g(W)\right] \leq g(\mathbb{E}_W\left[W\right])$$

by Jensen's inequality. Therefore, we have

$$f(x_1, \ldots, x_m, y_1, \ldots, y_m) \leq \bar{P}_e^\alpha,$$

and hence $s \leq \bar{P}_e^\alpha$ (Note that this inequality can also be shown by using Lagrange multipliers).

Next, suppose that $\alpha > 1$. In this case, we can similarly show $s \leq \bar{P}_e$. In addition, by using the similar discussion in the case $0 < \alpha < 1$, we can also prove $s \geq \bar{P}_e^\alpha$. □

If $0 \leq \alpha < 1$ and $P_e \geq 1 - \frac{1}{m}$, from (48) and (i) in Lemma 1 it follows that

$$\begin{aligned} r &\leq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e^\alpha (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha. \end{aligned} \tag{49}$$

If $0 \leq \alpha < 1$ and $0 < P_e \leq 1 - \frac{1}{m}$, from (48) and (i) in Lemma 1 it follows that

$$\begin{aligned} r &\leq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e. \end{aligned} \tag{50}$$

Next, suppose that $\alpha > 1$ and $P_e \neq 0$. Then, (47) implies

$$\begin{aligned} r &\geq a^{-1}(1 - sb) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} + s(1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}). \end{aligned} \tag{51}$$

Here, we note that $1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha} \geq 0$ (resp., ≤ 0) if $P_e \leq 1 - \frac{1}{m}$ (resp., $P_e \geq 1 - \frac{1}{m}$).

If $\alpha > 1$ and $P_e \geq 1 - \frac{1}{m}$, from (51) and (ii) in Lemma 1 it follows that

$$\begin{aligned} r &\geq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e. \end{aligned} \tag{52}$$

If $\alpha > 1$ and $0 < P_e \leq 1 - \frac{1}{m}$, from (51) and (ii) in Lemma 1 it follows that

$$\begin{aligned} r &\geq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e^\alpha (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha. \end{aligned} \tag{53}$$

Therefore, from (49), (50), (52) and (53), it holds that

$$\begin{aligned} R_\alpha^{\mathsf{H}}(X|Y) &= \frac{1}{1-\alpha} \log r \\ &\leq \begin{cases} \frac{1}{1-\alpha} \log\left[(m-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha\right] \\ \quad \text{if } 0 \leq \alpha < 1 \text{ and } P_e \geq 1 - \frac{1}{m}, \text{ or } \alpha > 1 \text{ and } 0 < P_e \leq 1 - \frac{1}{m}, \\ \frac{1}{1-\alpha} \log\left[(m-1)^{1-\alpha} P_e^{\alpha-1}(1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e\right] \\ \quad \text{if } 0 \leq \alpha < 1 \text{ and } 0 < P_e \leq 1 - \frac{1}{m}, \text{ or } \alpha > 1 \text{ and } P_e \geq 1 - \frac{1}{m} \end{cases} \end{aligned} \tag{54}$$

For the case $\alpha = 1$, the left hand of (54) implies $\lim_{\alpha\to 1} R_\alpha(X|Y) = H(X|Y)$ by Theorem 2-(ii). In addition, the right hands of (54) have a finite limit at $\alpha = 1$, and it is equal to Fano's inequality (see Remark 4). Therefore, (54) holds even for $\alpha = 1$.

For the case $P_e = 0$, the left hand of (54) implies $\lim_{P_e\to 0} R_\alpha(X|Y) = R_\alpha(X|X) = 0$, and the right hands of (54) imply

$$\begin{aligned}
&\lim_{P_e\to 0} \frac{1}{1-\alpha} \log\left[(m-1)^{1-\alpha} P_e^\alpha + \bar{P}_e^\alpha\right] = 0 \text{ (for } \alpha \geq 1),\\
&\lim_{P_e\to 0} \frac{1}{1-\alpha} \log\left[(m-1)^{1-\alpha} P_e^{\alpha-1}(1-\bar{P}_e^{2-\alpha}) + \bar{P}_e\right] = 0 \text{ (for } 0 \leq \alpha \leq 1).
\end{aligned}$$

Therefore, (54) holds for $P_e = 0$. □

A.5 Proof of Theorem 11

It is easily seen that Π_2 does not meet perfect secrecy since $q \neq 1/2$. And, it holds that:

$$\begin{aligned}
R_\alpha(M) &= \frac{1}{1-\alpha}\log(p^\alpha + (1-p)^\alpha)\\
&= \frac{1}{1-\alpha}\log\left[\left(\frac{1}{2}\right)^\alpha (1-\delta_1)^\alpha + \left(\frac{1}{2}\right)^\alpha (1+\delta_1)^\alpha\right] \qquad (55)\\
&= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha (2+o(1)), \qquad (56)\\
R_\alpha(K) &= \frac{1}{1-\alpha}\log(q^\alpha + (1-q)^\alpha) = \frac{1}{1-\alpha}\log(p^\alpha + (1-p)^\alpha + o(1))\\
&= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha (2+o(1)), \qquad (57)\\
R_\alpha(C) &= \frac{1}{1-\alpha}\log\left[\left(\frac{1}{2}\right)^\alpha (1-\delta_1^2)^\alpha + \left(\frac{1}{2}\right)^\alpha (1+\delta_1^2)^\alpha + o(1)\right] \qquad (58)\\
&= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha (2+o(1)), \qquad (59)\\
R_\alpha^{\mathsf{H}}(M|C) &= \frac{1}{1-\alpha}\log\sum_c P_C(c)\sum_m P_{M|C}(m|c)^\alpha\\
&= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha \left[(1-\delta_1^2) + \frac{1}{2}\frac{(1-\delta_1)^{2\alpha}}{(1+\delta_1^2)^{\alpha-1}} + \frac{1}{2}\frac{(1+\delta_1)^{2\alpha}}{(1+\delta_1^2)^{\alpha-1}} + o(1)\right]\\
&= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha (2+o(1)), \qquad (60)\\
R_\alpha^{\mathsf{H}}(C|M) &= \frac{1}{1-\alpha}\log\sum_m P_M(m)\sum_c P_{C|M}(c|m)^\alpha\\
&= \frac{1}{1-\alpha}\log\left[p(q^\alpha + (1-q)^\alpha) + (1-p)(q^\alpha + (1-q)^\alpha)\right]\\
&= \frac{1}{1-\alpha}\log\left(q^\alpha + (1-q)^\alpha\right) = R_\alpha(K). \qquad (61)
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
I_\alpha^{\mathsf{H}}(M;C) = R_\alpha(M) - R_\alpha^{\mathsf{H}}(M|C) &= \frac{1}{1-\alpha}\log\frac{2+o(1)}{2+o(1)} = \log(1+o(1))^{\frac{1}{\alpha-1}}\\
&= \log(1+o(1/\alpha)) = o(1/\alpha),
\end{aligned}$$

where the last equality follows from $\log(1+x) = x - o(x)$. Similarly, we also have $R_\alpha(M) - R_\alpha(K) = o(1/\alpha)$. Therefore, the proof is completed. Finally, for Remark 10 we see that $I_\alpha^{\mathsf{H}}(M;C) \neq I_\alpha^{\mathsf{H}}(C;M)$ by calculation. □

A.6 Proof of Theorem 14

Since $R_\infty^{\mathsf{avg}}(X|Y)$ satisfies "conditioning reduces entropy," it is sufficient to show $R_\infty^{\mathsf{avg}}(S|V_{\mathcal{W}}) = R_\infty(S)$ only in the case of $|\mathcal{W}| = n-1$. Let $q := 1-p$.

First, consider the case where $n \notin \mathcal{W}$. In this case, it is easy to see that

$$P_{S|V_1V_2\cdots V_{n-1}}(s|v_1, v_2, \dots, v_{n-1}) = P_S(s)$$

holds since S and $V_1, V_2, \dots, V_{n-1}$ are independent. Hence, $R_\infty^{\mathsf{avg}}(S|V_{\mathcal{W}}) = R_\infty(S)$ obviously holds in this case.

Next, we consider the case of $n \in \mathcal{W}$. From the symmetricity, it is sufficient to consider the case where we have[11] $\boldsymbol{v} = \{v_2, \dots, v_{n-1}\}$ in addition to v_n. Let $\sigma : \{0,1\}^{n-2} \to \{0,1\}$ be the mapping that computes exclusive OR of all inputs. Due to the construction and the independency among S and $V_1, V_2, \dots, V_{n-1}$, the probability $P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n)$ can be calculated in the following cases:

Case 1: $\sigma(\boldsymbol{v}) \oplus v_n = 1$, i.e., $(\sigma(\boldsymbol{v}), v_n) = (0,1)$ or $(\sigma(\boldsymbol{v}), v_n) = (1,0)$:

$$P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n) = P_{SV_1V}(1,0,\boldsymbol{v}) + P_{SV_1V}(0,1,\boldsymbol{v}) = 2pqP_{\boldsymbol{V}}(\boldsymbol{v}) \tag{62}$$

Case 2: $\sigma(\boldsymbol{v}) \oplus v_n = 0$, i.e., $(\sigma(\boldsymbol{v}), v_n) = (0,0)$ or $(\sigma(\boldsymbol{v}), v_n) = (1,1)$:

$$P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n) = P_{SV_1V}(0,0,\boldsymbol{v}) + P_{SV_1V}(1,1,\boldsymbol{v}) = (p^2+q^2)P_{\boldsymbol{V}}(\boldsymbol{v}) \tag{63}$$

Furthermore, note that the following relation:

$$P_{S|\boldsymbol{V}V_n}(s|\boldsymbol{v}, v_n) = P_{SV_1|\boldsymbol{V}V_n}(s,1|\boldsymbol{v}, v_n) + P_{SV_1|\boldsymbol{V}V_n}(s,0|\boldsymbol{v}, v_n). \tag{64}$$

Now, consider **Case 1**. In this case it is easy to see that

$$P_{SV_1|\boldsymbol{V}V_n}(0,1|\boldsymbol{v}, v_n) = P_{SV_1|\boldsymbol{V}V_n}(1,0|\boldsymbol{v}, v_n) = \frac{1}{2}, \text{ and, } P_{SV_1|\boldsymbol{V}V_n}(0,0|\boldsymbol{v}, v_n) = P_{SV_1|\boldsymbol{V}V_n}(1,1|\boldsymbol{v}, v_n) = 0. \tag{65}$$

Hence, (64) becomes $P_{S|\boldsymbol{V}V_n}(s|\boldsymbol{v}, v_n) = 1/2$, which leads to

$$P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n) \max_s P_{S|\boldsymbol{V}V_n}(s|\boldsymbol{v}, v_n) = pq \cdot P_{\boldsymbol{V}}(\boldsymbol{v}). \tag{66}$$

Next, consider **Case 2**. In this case, it is easy to see that

$$P_{SV_1|\boldsymbol{V}V_n}(0,1|\boldsymbol{v}, v_n) = P_{SV_1|\boldsymbol{V}}(1,0|\boldsymbol{v}, v_n) = 0$$

and hence, (64) becomes

$$\max_s P_{S|\boldsymbol{V}V_n}(s|\boldsymbol{v}, v_n) = \max\left\{P_{SV_1|\boldsymbol{V}V_n}(0,0|\boldsymbol{v}, v_n), P_{SV_1|\boldsymbol{V}V_n}(1,1|\boldsymbol{v}, v_n)\right\}. \tag{67}$$

Here, $P_{SV_1|\boldsymbol{V}V_n}(0,0|\boldsymbol{v}, v_n)$ can be calculated as follows:

$$P_{SV_1|\boldsymbol{V}V_n}(0,0|\boldsymbol{v}, v_n) = \frac{P_{SV_1\boldsymbol{V}V_n}(0,0,\boldsymbol{v}, v_n)}{P_{\boldsymbol{V}}(\boldsymbol{v}, v_n)} = \frac{P_{SV_1}(0,0)P_{\boldsymbol{V}}(\boldsymbol{v})}{P_{\boldsymbol{V}}(\boldsymbol{v}, v_n)} = \frac{p^2}{p^2+q^2} \tag{68}$$

Similarly, we have $P_{SV_1|\boldsymbol{V}V_n}(1,1|\boldsymbol{v}, v_n) = q^2/(p^2+q^2)$. Hence, because of $p > q$, (67) becomes

$$P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n) \max_s P_{S|\boldsymbol{V}}(s|\boldsymbol{v}) = p^2 \cdot P_{\boldsymbol{V}}(\boldsymbol{v}). \tag{69}$$

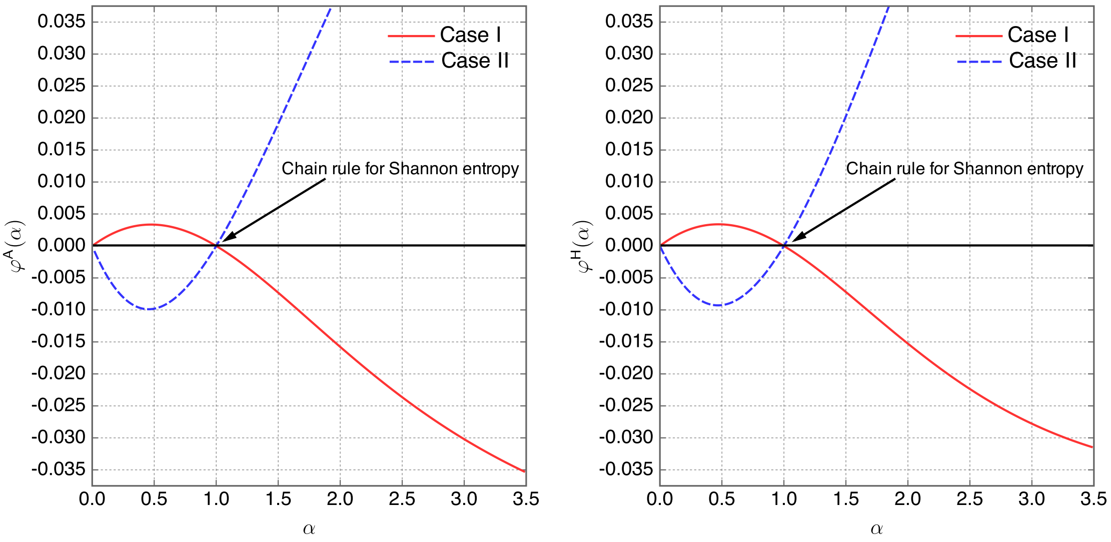
Summarizing (66) and (69), we have

$$\begin{aligned}\sum_{\boldsymbol{v}, v_n} P_{\boldsymbol{V}V_n}(\boldsymbol{v}, v_n) \max_s P_{S|\boldsymbol{V}V_n}(s|\boldsymbol{v}, v_n) &= \sum_{\substack{\boldsymbol{v}:\sigma(\boldsymbol{v})=0\\ v_n=1}} \sum_{\substack{\boldsymbol{v}:\sigma(\boldsymbol{v})=1\\ v_n=0}} pqP_{\boldsymbol{V}}(\boldsymbol{v}) + \sum_{\substack{\boldsymbol{v}:\sigma(\boldsymbol{v})=0\\ v_n=0}} \sum_{\substack{\boldsymbol{v}:\sigma(\boldsymbol{v})=1\\ v_n=1}} p^2P_{\boldsymbol{V}}(\boldsymbol{v}) \\ &= \sum_{\boldsymbol{v}:\sigma(\boldsymbol{v})=0,1} (pq+p^2)P_{\boldsymbol{V}}(\boldsymbol{v}) = p \end{aligned} \tag{70}$$

Hence, we obtain $R_\infty(S|V_{\mathcal{W}}) = R_\infty(S) = -\log p$, which completes the proof. □

B Graphs of $\varphi^{\mathsf{N}}(\alpha)$ for Cases I and II in Example 1

[11] In the case of $n = 2$, we set $\boldsymbol{v} = \emptyset$.



(a) $\varphi^{\mathsf{A}}(\alpha) = R_\alpha(X) - R_\alpha^{\mathsf{A}}(X|Y)$

(b) $\varphi^{\mathsf{H}}(\alpha) = R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y)$

Fig. 1. Graphs of $\varphi^{\mathsf{N}}(\alpha)$, $\mathsf{N} \in \{\mathsf{A}, \mathsf{H}\}$ for Cases I and II in Example 1