Revisiting Conditional Rényi Entropies and Generalizing Shannon's Bounds in Information Theoretically Secure Encryption*

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. 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, a new optimal symmetric key encryption is proposed which achieve our lower bounds.

Keywords: Information Theoretic Cryptography, (Conditional) Rényi entropy, Error probability in guessing, Impossibility, Symmetric-key Encryption, Shannon's Bound, Shannon's impossibility

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 [2] 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 [3]. Furthermore, min-entropy [4] 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 [5], 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, [5]). Similarly, we also know that the number of key candidates can be no less than the cardinality of the set of plaintexts. Furthermore, Dodis [6] recently showed that the similar property also holds with respect to min-entropy. Namely, he showed the

* This paper is presented in part at International Conference on Information Theoretic Security (ICITS2013) [1].

bound on secret-keys, $R_\infty(K) \geq R_\infty(M)$, for symmetric-key encryption with perfect secrecy[3]. Also, Alimomeni and Safavi-Naini [8] 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 [9] 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. [10] are very useful. In [10], 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 the other different definitions of conditional Rényi entropies provided in [11, 12]. Moreover, they did not find the definitions of conditional Rényi entropies corresponding to several conditional min-entropies [13, 14] which are useful in cryptographic contexts. Finding reasonable explanation for these min-entropies in terms of conditional Rényi entropies is also an important contribution, since these relations actually bridge interesting information theoretic measures of 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, and it may open a new vista in the field while covering existing criteria. In particular, discussing min-entropy criteria is very important since the attacker will guess the plaintext with the highest probability (called *guessing secrecy*). From this viewpoint, it is plausible that the security should be measured by min-entropy instead of Shannon entropy. Although this fact was pointed out by Alimomeni and Safavi-Naini [8], an explicit construction of encryption satisfying the guessing secrecy criteria is not provided in the literature. Hence, it is a very interesting open problem to design information theoretic cryptography under Rényi entropic security criteria in general. In particular, we are interested in the constructions tightly meeting the lower bounds of key size measured by min-entropies or Rényi entropies in general.

1.2 Our Contributions and Organization of This Paper

Conditional Rényi entropies, revisited (Sections 2 and 3) In [10], Teixeira et al. analyzed the relations among exiting conditional Rényi entropies. However, their analyses should be reinforced with the following three aspects. First, they do not discuss the implications of their results from an axiomatic viewpoints in entropies. Recall that Rényi entropy was originally discovered in [9] axiomatically, and many kinds of entropies are derived from Rényi entropy as special cases. Hence, it is desirable to discuss conditional

[3] Merhav [7] studied the exponent of this kind of success probability in guessing for symmetric-key cryptography with *variable-length keys* in *asymptotic setup*. In this study, each key depends on the ciphertext, and hence, its length can be varied depending on the ciphertext.

Rényi entropies from axiomatic and/or technological viewpoints. Second, the analysis in [10] missed to include two important conditional Rényi entropies due to Arimoto [11] and Hayashi [12] 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 [10] since they cannot be obtained from the conditional Rényi entropies discussed in [10].

Based on the above motivations, we start our discussion from the postulates required for Shannon and Rényi entropies, and we summarize in Sect. 2.3 what kind of properties should be required and/or are interested. We choose several important properties such as non-negativity, conditioning reduces entropy, data processing inequality (DPI), etc., as the conditions hopefully required for conditional Rényi entropies while the chain rule is outside the scope. Actually, we will see in Sect. 2.4 that the chain rule does not hold generally in the case of (conditional) Rényi entropies. Then, we consider the relation between conditional Rényi entropies and conditional min-entropies. We clarify that the conditional min-entropies useful in cryptographic context are related to the conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$.

Sections 3.1–3.3 are devoted to show that the above properties hopefully required for conditional Rényi entropies are actually satisfied by $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$. Furthermore, we show an extension of Fano's inequality [15] 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 [2] and the recent one to show the converse of channel coding theorem in finite blocklength regime [16, 17] 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 (Sections 5 and 6) 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.

Then, we propose *constructions* of secret-key encryption meeting the lower bounds which we derived by Rényi entropy of order α. It is well-known that, in the case of $\alpha = 1$, one-time pad [18] is optimal in the sense of key size measured by Shannon entropy. Hence,

we are interested in the case of $\alpha \neq 1$. In particular, we will focus on the case of $\alpha = \infty$, i.e., the situation where the security is measured by the conditional min-entropies. Note that under conditional min-entropy security criteria, perfect secrecy is not guaranteed in general. Actually, we propose a symmetric-key encryption that does not satisfy the perfect secrecy but satisfies the security measured by (conditional) min-entropies. Furthermore, it turns out that the proposed symmetric key encryption achieves tight lower bounds of key-size.

Furthermore, in Sect. 6, we slightly extend our bounds in terms of conditional Rényi entropies to the one under a class of conditional entropy functions which is naturally characterized from axiomatic consideration discussed in Section 2.3. This part of contribution is mainly shown from a theoretical interest, and for possibility of apperance of new conditional entropy formulas except for $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$.

2 Conditional Rényi Entropies, Revisited

2.1 Preliminaries: Rényi Entropies and α-divergence

Definition 1 (Rényi entropy, [9]) *Let X be a random variable taking values in a finite set $\mathcal{X}$. For a real number $\alpha \geq 0$, the Rényi entropy of order α is defined by*[4]

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

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 known as Rényi divergence of order α or the normalized Chernoff α-divergence) is important.

Definition 2 (α-divergence) *Let X and Y be random variables taking values in 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 for $\alpha \in (0, \infty)$ is equal to 0 if and only if $P_X(\cdot) = P_Y(\cdot)$, similarly to Kullback-Leibler divergence. However, it is only sufficient in the cases of $\alpha \to 0$ and $\alpha \to \infty$.

[4] 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.

2.2 Definitions of Conditional Rényi Entropies

Similarly to the conditional version of Shannon entropy, it is natural to consider the *conditional* Rényi entropies. Actually, several definitions of conditional Rényi entropies have been proposed, e.g., [11, 12, 19–22]. In particular, relations and properties are discussed in [10] among three kinds of conditional Rényi entropies such as

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

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

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

defined in [19], [20, 21], and [22], 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 [22].

In addition to the above, two conditional Rényi entropies are introduced in [11] and [12], which are defined as

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

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

respectively. Both of these conditional Rényi entropies are outside the scope of [10].

$R_\alpha^{\mathsf{H}}(X|Y)$ is defined in [12] to derive an upper bound of leaked information in universal privacy amplification. $R_\alpha^{\mathsf{A}}(X|Y)$ is used in [11] to show that the strong converse of channel coding theorem. We also note that $R_\alpha^{\mathsf{A}}(X|Y)$ is implicitly used even in cryptographic contexts. In [23], $R_\alpha^{\mathsf{A}}(X|Y) = -((1+s)/s)\phi(s/(1+s)|X|Y)$ is used to bound an average security measure of privacy amplification, where $\phi(t|X|Y) := \log \sum_y (\sum_x P_{XY}(x,y)^{\frac{1}{1+t}})^{1-t}$.

Not only the conditional Rényi entropies discussed in [10] but also $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ is non-negative and is upper 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, where $\operatorname{supp} P_Y := \{y \in \mathcal{Y} \mid P_Y(y) > 0\}$. 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 $\mathcal{X}$. In addition, $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ are continuous with respect to $\alpha \in (0,\infty)$. The proofs are not so hard and we omit them (Proofs for $R_\alpha^{\mathsf{A}}(X|Y)$, see [11]). Note that the following fundamental relations hold with respect to $R_\alpha^{\mathsf{H}}(X|Y)$ and $R_\alpha^{\mathsf{A}}(X|Y)$.

Proposition 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*

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

Note that $R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha^{\mathsf{A}}(X|Y)$ for $\alpha > 1$ was proved in Lemma 7 of [23]. In addition, Proposition 1 means that: it holds even for $0 < \alpha < 1$ and its proof is simply shown by Jensen's inequality; and the cases of $\alpha = 0, 1$ are meant to take the limits at $\alpha = 0, 1$ (see Theorem 1 and Proposition 2).

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 independent from each definition, and hence, it is denoted by $R_\alpha(X|Y)$.

Axiomatic Consideration Recall that Rényi entropy is axiomatically obtained, namely, it is the unique quantity (up to a constant factor) that satisfies weakened postulates for Shannon entropy [9]. According to [9], 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[5], where $H(X|Y) := \sum_y P_Y(y)H(X|Y = y)$. 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 $y \in \mathcal{Y}$, as well as $\{P_Y(y)\}_{y\in\mathcal{Y}}$.
- $R_\alpha(X|Y)$ 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 $y \in \operatorname{supp} Y$.

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

Since the postulate (d) is replaced with $H(XY) = H(X) + H(Y)$ for independent random variables X and Y, it is natural to expect that Rényi entropies do not satisfy the chain rule. Actually, it is pointed out in [10, Theorem 5] that $R^{\mathsf{C}}_\alpha(X|Y)$ and $R^{\mathsf{RW}}_\alpha(X|Y)$ do not satisfy the chain rule for arbitrary $\alpha \neq 1$[6]. We will see in Section 2.4 that the chain rules do not hold for $R^{\mathsf{A}}_\alpha(X|Y)$ and $R^{\mathsf{H}}_\alpha(X|Y)$ either. Instead, we consider several fundamental properties related to chain rule.

First, note that $H(XY) \geq H(X)$ is derived from the chain rule of Shannon entropies since the conditional Shannon entropy is non-negative. The inequality $H(XY) \geq H(X)$ means that additional information Y increases the entropy of X. Hence, we call this inequality as "*Additional information Increases Entropy*," (AIE) as opposed to CRE which will be introduced below. In this paper, we will focus on the non-negativity and AIE for (conditional) Rényi entropies.

Second, it is also known that Rényi entropies *do not satisfy* the *subadditivity* since only the additivity for independent random variables is required for Rényi entropies instead of the postulate (d). Subadditivity for Shannon entropy is written as $H(XY) \leq H(X) + H(Y)$, which is equivalent to $H(X|Y) \leq H(X)$ due to the chain rule. This inequality is called as "*Conditioning reduces entropy*" [24], CRE for short. Even if the chain rule does not hold for Rényi entropies, we have possibilities that CRE holds for Rényi entropies.

Summarizing so far, instead of the chain rules, three fundamental properties discussed above are formally given in a form of

$$0 \overset{\text{(8a)}}{\leq} R_\alpha(X|Y) \overset{\text{(8b)}}{\leq} R_\alpha(X) \overset{\text{(8c)}}{\leq} R_\alpha(WX) \tag{8}$$

[5] This form of the chain rule is inductively obtained by using the postulate (d) in [9, p. 547].

[6] In the case of $\alpha = 1$, conditional Rényi entropies coincide with conditional Shannon entropy, and hence, chain rule is obviously 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.

for arbitrary $\alpha \geq 0$, and random variables W, X, and Y, which gives an upper and a lower bounds of $R_\alpha(X)$. The inequality (8a) means that the conditional Rényi entropies is non-negative. In addition, CRE (8b) states that the entropy of random variable X decreases if some information Y related to X is revealed. On the other hand, AIE (8c) implies that the entropy of X increases if some information is added. We will investigate the inequalities with respect to (8a) and (8b) in this paper, while (8c) itself is not directly investigated. Instead, (8c) is proved as a special case of (9c) which is introduced later in (9)[7].

The equalities of (8a) and (8b) hold under the following conditions:

- (8a): $R_\alpha(X|Y) = 0$ holds if $X = f(Y)$ for a certain (deterministic) mapping $f : \mathcal{Y} \to \mathcal{X}$.
- (8b): $R_\alpha(X) = R_\alpha(X|Y)$ holds if X and Y are independent.

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

Inequality (8) shows an upper and a lower bounds of $R_\alpha(X)$. Since we are interested in conditional Rényi entropies, it might be natural to require its upper and lower bounds in a similar manner to (8), namely,

$$0 \overset{\text{(9a)}}{\leq} R_\alpha(X|YZ) \overset{\text{(9b)}}{\leq} R_\alpha(X|Z) \overset{\text{(9c)}}{\leq} R_\alpha(WX|Z). \tag{9}$$

The inequality (9a) is essentially the same with (8a). Hence, we are concerned in this paper with the inequalities (9b) and (9c). As a natural extension of Shannon entropies, inequalities (9b) and (9c) hold with equalities under the following conditions, which will be proved in later sections:

- (9b): $R_\alpha(X|YZ) = R_\alpha(X|Z)$ holds if[8] $X \leftrightarrow Z \leftrightarrow Y$.
- (9c): $R_\alpha(X|Z) = R_\alpha(WX|Z)$ holds if $W = f(X, Z)$ for a certain (deterministic) mapping $f : \mathcal{X} \times \mathcal{Z} \to \mathcal{W}$.

In the case of conditional Shannon entropies, i.e., $\alpha \to 1$, (9b) combining with chain rule results in $H(XYZ) + H(Z) \leq H(XZ) + H(YZ)$. This inequality implies that Shannon entropy is *polymatroid* [26], which is one of the most important properties of Shannon entropy. While Rényi entropy is not[9] polymatroid for general $\alpha \geq 0$, the lower bound of (9) is not only a simple extension of CRE but also has an important connection to the structure of information measures.

The inequality (9c) is not only a mathematical extension of (8c) but also useful in proving information theoretic inequality. Actually, the equality case of (9c) plays a crucial role in this paper, see Proof II in Section 5.2. In addition, in the case of (conditional) Shannon entropies, i.e., $\alpha \to 1$, (9c) is introduced in [25, (13.9) in Lemma 13.6] as an important property.

Finally, we note that stronger inequality than $H(X|YZ) \leq H(X|Z)$, which corresponds to (9b), is known for Shannon entropy if $X \leftrightarrow Z \leftrightarrow Y$. In this case, it holds that $H(X|Y) \leq H(X|Z)$, which is equivalent to $I(X;Y) \geq I(X;Z)$, called *Data Processing Inequality* (DPI). In this paper, we will study DPI for conditional Rényi entropies, i.e.,

$$R_\alpha(X|Z) \geq R_\alpha(X|Y) \quad \text{if} \quad X \leftrightarrow Z \leftrightarrow Y, \tag{10}$$

where the equality holds under the following condition:

[7] Note that (8c) is equivalent to $R_\alpha^{\mathsf{JA}}(W|X) \geq 0$, which obviously holds if (8a) holds. However, (9c) does not follow from $R_\alpha^{\mathsf{JA}}(W|X) \geq 0$, and hence, it is necessary to be investigated.

[8] If random variables X, Y, and Z, taking values in finite sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively, satisfy $P_{XZ|Y}(x, z|y) = P_{X|Y}(x|y)P_{Z|Y}(z|y)$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$, we say that X, Y, and Z form a *Markov chain* in this order, in symbols $X \leftrightarrow Y \leftrightarrow Z$.

[9] In the case of Rényi entropy, the corresponding inequality, i.e., $R_\alpha(XYZ) + R_\alpha(Z) \leq R_\alpha(XZ) + R_\alpha(YZ)$ does not hold for general α since it is equivalent to $R_\alpha^{\mathsf{JA}}(X|Z) \geq R_\alpha^{\mathsf{JA}}(X|YZ)$.

Table 1. Summary of properties of conditional Rényi entropies; AIE: Additional Information Increases Entropy, CRE: Conditioning Reduces Entropy, DPI: Data Processing Inequality

Eq.		(2): $R_\alpha^{\mathsf{C}}(\cdot\|\cdot)$	(3): $R_\alpha^{\mathsf{JA}}(\cdot\|\cdot)$	(4): $R_\alpha^{\mathsf{RW}}(\cdot\|\cdot)$	(5): $R_\alpha^{\mathsf{H}}(\cdot\|\cdot)$	(6): $R_\alpha^{\mathsf{A}}(\cdot\|\cdot)$
	Chain Rule	No (obvious)	✓ (obvious)	No (obvious)	No (obvious)	No (obvious)
	Weak Chain Rule	No ([10])	✓ (obvious)	✓ ([10])	No (Prop. 4)	No (Prop. 4)
(8a):	Non-negetivity	✓ ([10])	✓ ([10])	✓ ([10])	✓	✓
(8b):	CRE	No ([10])	No ([10])	✓ ($\alpha \geq 1$)[10]	✓ (Thm. 2)	✓ ([11, 31])
(9b):	CRE conditioned by Z	No ([10])	No ([10])	✓ ($\alpha \geq 1$)[10]	✓ (Thm. 4)	✓ (Thm. 4)
(9c):	Conditioned AIE	✓ (Thm. 5)	✓ (Thm. 5)	✓ (Thm. 5)	✓ (Thm. 5)	✓ (Thm. 5)
(10):	DPI	No (obvious)	No (obvious)	No (obvious)	✓ (Thm. 6)	✓ (Thm. 6)

– (10): the equality holds if there exists a surjective mapping $f : \mathcal{Z} \to \mathcal{Y}$.

Whether each conditional Rényi entropy given by (2)–(6) satisfies each property discussed above is summarized in Table 1, which will be proved in later sections.

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 particularly 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

$$R_\infty^{\mathsf{avg}}(X|Y) := -\log \mathbb{E}_Y \left[\max_x P_{X|Y}(x|Y)\right] \tag{11}$$

proposed in [14] is important from a cryptographic viewpoint, e.g., [14, 27–30]. We can also find the *worst case* conditional min-entropy (e.g., in the analysis of physically unclonable functions (PUFs), see [13]).

$$R_\infty^{\mathsf{wst}}(X|Y) := -\log \max_{\substack{x\in\mathcal{X} \\ y\in \mathrm{supp}\, P_Y}} P_{X|Y}(x|y). \tag{12}$$

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 [10] 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 (11) is not sufficiently analyzed in [10] 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 [10] actually bridge the conditional Shannon entropy and the conditional min-entropy appeared in cryptography as shown below. Hence, in the forthcoming discussion, we will mainly focus on the properties of conditional Rényi entropies $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$.

[10] In [10], a counterexample of CRE in the case of $\alpha > 1$. However, it is easy to show that CRE holds for $0 \leq \alpha \leq 1$.

Theorem 1 *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 [11]. For the rest of the proofs, see Appendix A.1. □

Finally, we consider the limits of conditional Rényi entropies as $\alpha \to 0$, which can be considered as conditional *Hartley entropies*.

Proposition 2 *For random variables X and Y, the following relations hold:*

$$\lim_{\alpha\to 0} R_\alpha^{\mathsf{C}}(X|Y) = \mathbb{E}_Y\left[\log|\mathrm{supp}P_{XY}|\right] \tag{13}$$

$$\lim_{\alpha\to 0} R_\alpha^{\mathsf{JA}}(X|Y) = \log|\mathrm{supp}P_{XY}| - \log|\mathrm{supp}P_Y| \tag{14}$$

$$\lim_{\alpha\to 0} R_\alpha^{\mathsf{RW}}(X|Y) = \log\max_{y\in\mathcal{Y}}\left\{\left|\mathrm{supp}P_{X|Y=y}\right|\right\} \tag{15}$$

$$\lim_{\alpha\to 0} R_\alpha^{\mathsf{A}}(X|Y) = \log\max_{y\in\mathcal{Y}}\left\{\left|\mathrm{supp}P_{X|Y=y}\right|\right\} \tag{16}$$

$$\lim_{\alpha\to 0} R_\alpha^{\mathsf{H}}(X|Y) = \log\mathbb{E}_Y\left[\left|\mathrm{supp}P_{X|Y}\right|\right] \tag{17}$$

Proof. See Appendix A.2. □

It is interesting to see that average and the worst cases of conditional *min-entropies* correspond to $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$, respectively, while average and the worst cases of conditional *Hartley entropies* correspond to $R_\alpha^{\mathsf{H}}(X|Y)$ and $R_\alpha^{\mathsf{A}}(X|Y)$, respectively.

2.4 Chain Rule and Weak Chain Rule for Conditional Rényi Entropies

Based on the discussion of the previous section, it is hard to expect that the conditional Rényi entropies satisfy the *chain rule*. According to [10], we can readily know that the chain rule will not hold *with equality* if the conditional Rényi entropies satisfy CRE since, by defining the conditional Rényi entropy as $R_\alpha^{\mathsf{JA}}(X|Y) := R_\alpha(XY) - R_\alpha(Y)$ [20, 21], 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 3 ([10]) *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 3 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 neither $R_\alpha^{\mathsf{A}}(X|Y)$ nor $R_\alpha^{\mathsf{H}}(X|Y)$ satisfy the chain rule even in a weak sense.

Proposition 4 *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 $\varphi^{\mathsf{N}}(\alpha) := R_\alpha(XY) - R^{\mathsf{N}}_\alpha(X|Y) - R_\alpha(Y)$ *for* $\mathsf{N} \in \{\mathsf{A}, \mathsf{H}\}$ *are depicted in Fig. 1–(a),(b) in Appendix B, which means that* $R_\alpha(XY) > R^{\mathsf{N}}_\alpha(X|Y) + R_\alpha(Y)$ *holds only when* $\alpha \in (0,1)$ *with Case I, but* $R_\alpha(XY) < R^{\mathsf{N}}_\alpha(X|Y) + R_\alpha(Y)$ *holds only when* $\alpha \in (0,1)$ *with Case II. 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 Conditional Rényi Entropies

As is pointed out in Theorem 1, the conditional Rényi entropies $R^{\mathsf{A}}_\alpha(X|Y)$ and $R^{\mathsf{H}}_\alpha(X|Y)$ 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, [24]), 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 [10] that $R^{\mathsf{C}}_\alpha(X|Y)$, $R^{\mathsf{JA}}_\alpha(X|Y)$, and $R^{\mathsf{RW}}_\alpha(X|Y)$ given by (2)–(4), respectively, *do not satisfy* CRE in general[11]. Fortunately, however, we will point out in this section that $R^{\mathsf{A}}_\alpha(X|Y)$ and $R^{\mathsf{H}}_\alpha(X|Y)$, which are outside the scope of [10], satisfy CRE in general.

Theorem 2 (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*

$$R^{\mathsf{H}}_\alpha(X|Y) \leq R_\alpha(X), \tag{18}$$

where the equality holds if and only if X *and* Y *are statistically independent in the case of* $\alpha \in (0, \infty)$. *However, independency between* X *and* Y *is not necessary but sufficient for the equality of* (18) *when* $\alpha \to 0$ *and* $\alpha \to \infty$.

Remark 1 *The same result also holds for* $R^{\mathsf{A}}_\alpha(X|Y)$, *which was proved in [11, 31][12]. Then,* (18) *is immediately obtained by recalling CRE for* $R^{\mathsf{A}}_\alpha(X|Y)$ *and the relation given by* (7) *in Proposition 1. Namely, it holds that* $R^{\mathsf{H}}_\alpha(X|Y) \leq R^{\mathsf{A}}_\alpha(X|Y) \leq R_\alpha(X)$.

Due to CRE for $R^{\mathsf{A}}_\alpha(X|Y)$ *and* $R^{\mathsf{H}}_\alpha(X|Y)$, *it is immediately seen that* $R^{\mathsf{avg}}_\infty(X|Y)$ *and* $R^{\mathsf{wst}}_\infty(X|Y)$ *also satisfy CRE, though it is possible to show it directly.*

Proof. In the case of $\alpha \to 1$, Theorem 2 holds from the properties of Shannon entropies. From Jensen's inequality, in the case of $0 < \alpha < 1$, we have

$$\mathbb{E}_Y \left[\sum_{x \in \mathcal{X}} P_{X|Y}(x|Y)^\alpha \right] \leq \sum_{x \in \mathcal{X}} \mathbb{E}_Y \left[P_{X|Y}(x|Y) \right]^\alpha = \sum_{x \in \mathcal{X}} P_X(x)^\alpha, \tag{19}$$

[11] We can show that CRE is satisfied by $R^{\mathsf{RW}}_\alpha(X|Y)$ in the case of $\alpha > 1$. See Proposition 14 of Section 6.
[12] However, [11, 31] did not discuss the condition for equality. See Remark 2.

where the equality holds $P_{X|Y}(x|Y) = P_X(x)$ with probability 1 for all $x \in \mathcal{X}$, i.e., X and Y are statistically independent. 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 $\alpha \geq 1$, and the equality holds if and only if X and Y are statistically independent.

In the cases of $\alpha \to 0$ and $\alpha \to \infty$, (19) is also valid from the continuity of $R_\alpha^{\mathsf{H}}(X|Y)$ and $R_\alpha(X)$.

Finally, we show that the independency between X and Y is not necessary but sufficient for the equality of (18) when $\alpha \to 0$ and $\alpha \to \infty$.

To see this in the case of $\alpha \to 0$, we consider the arbitrarily correlated random variables X and Y satisfying $\mathrm{supp}P_{XY} = \mathcal{X}\times\mathcal{Y}$. Then, it holds that $\mathrm{supp}P_{X|Y=y} = \mathcal{X}$ for all $y \in \mathcal{Y}$, and hence, we have $\lim_{\alpha\to 0} R_\alpha^{\mathsf{H}}(X|Y) = \log \mathbb{E}_Y\left[|\mathrm{supp}P_{X|Y}|\right] = \log|\mathcal{X}| = \log|\mathrm{supp}P_X|$. Hence, the equality of (18) holds even if X and Y are statistically correlated.

In the case of $\alpha \to \infty$, we introduce the following probability transition from Y to X with $\mathcal{X} = \{0,1,2\}$ and $\mathcal{Y} = \{0,1\}$: $P_{X|Y}(0|0) = 2/3$, $P_{X|Y}(1|0) = 1/12$, $P_{X|Y}(2|0) = 1/4$, $P_{X|Y}(0|1) = 2/3$, $P_{X|Y}(1|1) = 1/4$, $P_{X|Y}(2|1) = 1/12$, which makes X and Y statistically correlated if Y is not uniform. In the case of $(P_Y(0), P_Y(1)) = (1/3, 2/3)$, we have $(P_X(0), P_X(1), P_X(2)) = (2/3, 7/36, 5/36)$, and hence, it follows that $R_\infty^{\mathsf{wst}}(X|Y) = R_\infty(X) = -\log(2/3)$. □

Remark 2 *As is pointed in Remark 1, CRE for $R_\alpha^{\mathsf{A}}(X|Y)$, i.e., $R_\alpha^{\mathsf{A}}(X|Y) \leq R_\alpha(X)$ is proved in [11, 31] and the condition for equality is given in the case of $\alpha \in (0,\infty)$ while it is not discussed in the cases of $\alpha \to 0$ and $\alpha \to \infty$. By observing the relation $R_\alpha^{\mathsf{H}}(X|Y) \leq R_\alpha^{\mathsf{A}}(X|Y) \leq R_\alpha(X)$, the examples provided in the proof of Theorem 2 also satisfy $R_\alpha^{\mathsf{A}}(X|Y) = R_\alpha(X)$ when $\alpha \to 0$ and $\alpha \to \infty$. Actually, these examples also results in $\lim_{\alpha\to 0} R_\alpha^{\mathsf{A}}(X|Y) = \log\max_{y\in\mathcal{Y}}|\mathrm{supp}P_{X|Y=y}| = \log|\mathcal{X}| = \log|\mathrm{supp}P_X|$, and $R_\infty^{\mathsf{avg}}(X|Y) = R_\infty(X) = -\log(2/3)$.*

We also note that another example of $R_\alpha^{\mathsf{A}}(X|Y) = R_\alpha(X)$ for correlated X and Y in the case of $\alpha \to \infty$ will be given in Sect. 5.4 in the context of a symmetric key encryption.

Although Proposition 2 can be proved directly as shown above, this proof does not tell us the difference between both sides of (18). To see this gap, we introduce a conditional α-divergence defined by the same idea with $R_\alpha^{\mathsf{H}}(X|Y)$ in the following form.

Definition 4 ([17]) *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 $P_{X_1Y}(\cdot,y) = W(\cdot|y)Q(y)$, $P_{X_2Y}(\cdot,y) = V(\cdot|y)Q(y)$ 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

$$D_\alpha(X_1\|X_2|Y) := \frac{1}{\alpha-1}\log\sum_{x,y}\frac{W(x|y)^\alpha}{V(x|y)^{\alpha-1}}Q(y) \tag{20}$$

which is also written as $D_\alpha(W\|V|Q)$ depending on the context.

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 5 *Let X_1, X_2, and Y be random variables following the probability distributions $P_{X_1Y}(\cdot,y) = W(\cdot|y)Q(y)$, $P_{X_2Y}(\cdot,y) = V(\cdot|y)Q(y)$ for all $y \in \mathcal{Y}$. Then, the following properties are satisfied:*

(i) $\lim_{\alpha\to 1} D_\alpha(X_1\|X_2|Y) = D(X_1\|X_2|Y) := \sum_{x,y} Q(y)W(x|y)\log(W(x|y)/V(x|y))$.
(ii) $D_\alpha(X_1\|X_2|Y) \geq 0$ *for all* $\alpha \geq 0$, *where the equality holds if and only if* $W(\cdot|y) = V(\cdot|y)$ *for all* $y \in \operatorname{supp} Q$ *in the case of* $\alpha \in (0,\infty)$.

Proof. The property (i) is pointed out in [17] without proof. We provide the formal proofs for (i) and (ii) in Appendix A.3 for readers' convenience. □

Then, the following relation holds, which can be seen as an alternative proof for Theorem 2 owing to $D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha}) \geq 0$. Moreover, the condition for the equality of (18) is easily derived by recalling that $D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha}) = 0$ if and only if X and Y are statistically independent for the case of $\alpha \in (0,\infty)$.

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

$$R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y) = D_\alpha(P_{Y|X}\|P_Y|P_{X_\alpha}) \tag{21}$$

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

Although this theorem follows from the identity introduced in [17, Equation (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

$$\begin{aligned}\left\{\sum_x P_X(x)^\alpha\right\}^{-1}\sum_{x,y} P_Y(y)P_{X|Y}(x|y)^\alpha &= \left\{\sum_x P_X(x)^\alpha\right\}^{-1}\sum_{x,y} P_{XY}(x,y)^\alpha P_Y(y)^{1-\alpha}\\ &= \sum_{x,y}\frac{P_X(x)^\alpha}{\sum_x P_X(x)^\alpha}P_{Y|X}(y|x)^\alpha P_Y(y)^{1-\alpha}. \end{aligned}\tag{22}$$

Taking the logarithms of both sides of (22) and multiplying $-1/(1-\alpha)$, we obtain (21). □

This relation (21) 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

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

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, $R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y) = R_\alpha(Y) - R_\alpha^{\mathsf{H}}(Y|X)$ does not hold for general α as well.

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

$$I_\alpha^{\mathsf{H}}(X;Y) := R_\alpha(X) - R_\alpha^{\mathsf{H}}(X|Y), \tag{23}$$

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

$$I_\alpha^{\mathsf{A}}(X;Y) := R_\alpha(X) - R_\alpha^{\mathsf{A}}(X|Y), \tag{24}$$

in the context of describing channel coding theorem in a general setting [11] .

Remark 3 *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* $I_\alpha^{\mathsf{A}}(X;Y) \leq I_\alpha^{\mathsf{H}}(X;Y)$ *generally holds from* (7) *of Proposition 1. Since* $R_\alpha^{\mathsf{H}}(X|Y)$ *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 and only if* X *and* Y *are statistically independent. in the case of* $\alpha \in (0,\infty)$. *However, it should be stressed that the independency only sufficient if* $\alpha \to \infty$.

3.2 Upper and Lower Bounds of Conditional Rényi Entropies

We discuss (9) for $R^{\mathsf{A}}(\cdot|\cdot)$ and $R^{\mathsf{H}}(\cdot|\cdot)$, which is an extended inequality of (8) as we have seen in Sect. 2.3. Inequality (9) indicates an upper and a lower bounds of the conditional Rényi entropies.

The following theorem proves the inequality (9b):

Theorem 4 *For arbitrary RVs X, Y, and Z, it holds for all $\alpha \geq 0$ that*

$$R^{\mathsf{A}}_\alpha(X|Z) \geq R^{\mathsf{A}}_\alpha(X|YZ) \tag{25}$$

$$R^{\mathsf{H}}_\alpha(X|Z) \geq R^{\mathsf{H}}_\alpha(X|YZ) \tag{26}$$

where the equalities hold if and only if $X \leftrightarrow Z \leftrightarrow Y$ in the case of $\alpha \in (0, \infty)$. However, $X \leftrightarrow Z \leftrightarrow Y$ is not necessary but sufficient for the equalities of (25) *and* (26) *when $\alpha \to 0$ and $\alpha \to \infty$.*

Proof. In the case of $\alpha \to 1$, the theorem obviously holds due to the properties of conditional Shannon entropies. Hence, we first prove[13] (25) in the case of $\alpha \in (0, 1)$. Note that

$$\begin{aligned} R^{\mathsf{A}}_\alpha(X|Z) &= \frac{\alpha}{1-\alpha} \log \mathbb{E}_Y \left[\left\{ \sum_{x \in \mathcal{X}} P_{X|Z}(x|Z)^\alpha \right\}^{1/\alpha} \right] \\ &= \frac{\alpha}{1-\alpha} \log \mathbb{E}_Z \left[\left\| P_{X|Z}(\cdot|Z) \right\|_\alpha \right], \end{aligned} \tag{27}$$

where we define an α-norm for the probability distribution P_X by $\|P_X\| := \{\sum_x P_X(x)^\alpha\}^{1/\alpha}$ for a positive real number $\alpha \geq 0$ and a random variable X taking values in a finite set $\mathcal{X}$.

Note that the α-norm $\|\boldsymbol{x}\|_\alpha := \left\{\sum_{x \in \mathcal{X}} x^\alpha\right\}^{1/\alpha}$ is a strictly concave function of $\boldsymbol{x} \in (\mathbb{R}^+)^n$ in the case of $\alpha \in (0, 1)$, which can be easily verified from Minkowski's inequality (for instance, see [32, Thm. 25, page 31]). Hence, it holds for arbitrary $y \in \mathcal{Y}$ that

$$\begin{aligned} \|P_{X|Z}(\cdot|z)\|_\alpha &= \left\| \mathbb{E}_Y \left[P_{X|YZ}(\cdot|Y, z) \right] \right\|_\alpha \\ &\geq \mathbb{E}_Y \left[\|P_{X|YZ}(\cdot|Y, z)\|_\alpha \right] \end{aligned} \tag{28}$$

from Jensen's inequality. The equality holds if and only if $P_{X|YZ}(x|Y, z)$ is constant for all $x \in \mathcal{X}$ with probability 1, i.e., $P_{X|YZ}(x|y, z) = P_{X|Z}(x|z)$ for all $y \in \mathcal{Y}$. Applying $\alpha/(1-\alpha) \log \mathbb{E}_Y[\cdot]$ to both sides, we have (25) in the case of $\alpha \in (0, 1)$. The quality holds if and only if $X \leftrightarrow Z \leftrightarrow Y$. In the case of $\alpha \in (1, \infty)$, we can prove (25) in a similar manner.

Then, we prove (26) in the case of $\alpha \in (0, 1)$. Note that

$$\begin{aligned} R^{\mathsf{H}}_\alpha(X|Z) &= \frac{1}{1-\alpha} \log \mathbb{E}_Z \left[\sum_{x \in \mathcal{X}} P_{X|Z}(x|Z)^\alpha \right] \\ &= \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} \mathbb{E}_Y \left[P_{X|Z}(x|Z)^\alpha \right]. \end{aligned} \tag{29}$$

Due to Jensen's inequality for the concave function x^α, $\alpha \in (0, 1)$, it holds for arbitrary $y \in \mathcal{Y}$ that

$$\begin{aligned} \mathbb{E}_{YZ} \left[P_{Z|YZ}(x|YZ)^\alpha \right] &\leq \mathbb{E}_Y \left[\mathbb{E}_Z \left[P_{X|YZ}(x|YZ) \right]^\alpha \right] \\ &= \mathbb{E}_Y \left[P_{X|Y}(x|Y)^\alpha \right] \end{aligned} \tag{30}$$

[13] While (25) was proved in [11, 31], we show the proof for the readers' convenience and for checking the condition for equality.

in the case of $\alpha \in (0,1)$. Hence, we obtain (26). The equality holds if and only if $X \leftrightarrow Z \leftrightarrow Y$ due to in the same discussion with the case of (28). In the case of $\alpha > 1$, we can prove (26) in a similar manner.

Finally, in the cases of $\alpha \to 0$ and $\alpha \to \infty$, inequalities (25) and (26) are valid from the continuity of $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$, respectively.

The conditions of equalities in (25) and (26) when $\alpha \to 0$ and $\alpha \to \infty$ can be discussed almost the same line with the proof of Theorem 2. □

Remark 4 *Note that the other kinds of conditional Rényi entropies such as $R_\alpha^{\mathsf{C}}(X|Z)$, $R_\alpha^{\mathsf{RW}}(X|Z)$, and $R_\alpha^{\mathsf{JA}}(X|Z)$, do not satisfy* (9b) *in general since they do not satisfy CRE, i.e.,* (8b) *in general.*

Then, we prove (9c) which gives an upper bound of conditional Rényi entropies as an extension of CRE (8c). We prove that all five conditional Rényi entropies satisfy the conditioned AIE while several inequalities such as CRE and DPI do not hold for several conditional Rényi entropies.

Theorem 5 *Let W, X, and Z be random variables taking values in finite sets $\mathcal{W}$, $\mathcal{X}$, and $\mathcal{Z}$, respectively. Then, for each of $\mathsf{N} \in \{\mathsf{C}, \mathsf{JA}, \mathsf{RW}, \mathsf{A}, \mathsf{H}\}$ and for all $\alpha \geq 0$, we have:*

(i) $R_\alpha^{\mathsf{N}}(X|Z) \leq R_\alpha^{\mathsf{N}}(WX|Z)$,
(ii) $R_\alpha^{\mathsf{N}}(X|Z) = R_\alpha^{\mathsf{N}}(WX|Z)$ *if $W = f(X,Z)$ for some (deterministic) mapping $f : \mathcal{X} \times \mathcal{Z} \to \mathcal{W}$.*

Proof. Although (i) for $R_\alpha^{\mathsf{A}}(X|Y)$ is proved in [31, Proposition 2], we will prove this claim for all remaining conditional Rényi entropies simultaneously. For any α with $\alpha \in [0,1)$ and arbitrary $z \in \mathcal{Z}$, it holds that

$$\sum_{w,x} P_{WX|Z}(w,x|z)^\alpha = \sum_x P_{X|Z}(x|z)^\alpha \sum_w P_{W|XZ}(w|x,z)^\alpha \geq \sum_x P_{X|Z}(x|z)^\alpha. \tag{31}$$

Hence, we have

$$\max_z \sum_{w,x} P_{WX|Z}(w,x|z)^\alpha \geq \max_z \sum_x P_{X|Z}(x|z)^\alpha, \tag{32}$$

$$\sum_z P_Z(z) \left(\sum_{w,x} P_{WX|Z}(w,x|z)^\alpha \right)^{1/\alpha} \geq \sum_z P_Z(z) \left(\sum_x P_{X|Z}(x|z)^\alpha \right)^{1/\alpha}, \tag{33}$$

$$\sum_z P_Z(z) \sum_{w,x} P_{WX|Z}(w,x|z)^\alpha \geq \sum_z P_Z(z) \sum_x P_{X|Z}(x|z)^\alpha, \tag{34}$$

which result in $R_\alpha^{\mathsf{RW}}(X|Z) \leq R_\alpha^{\mathsf{RW}}(WX|Z)$, $R_\alpha^{\mathsf{A}}(X|Z) \leq R_\alpha^{\mathsf{A}}(WX|Z)$ and $R_\alpha^{\mathsf{H}}(X|Z) \leq R_\alpha^{\mathsf{H}}(WX|Z)$, respectively.

Furthermore, (31) also implies that $R_\alpha(WX) \geq R_\alpha(X)$ since we can consider the case of Z is constant with probability 1 in (31). Hence, the conditioned AIE $R_\alpha^{\mathsf{JA}}(X|Z) \leq R_\alpha^{\mathsf{JA}}(WX|Z)$ is obvious from the AIE of Rényi entropies since it is equivalent to $R_\alpha(XZ) \leq R_\alpha(WXZ)$ from its definition. Similarly, $R_\alpha^{\mathsf{C}}(X|Z) \leq R_\alpha^{\mathsf{C}}(WX|Z)$ also holds due to the AIE of Rényi entropies, i.e., $R_\alpha(X|Z=z) \leq R_\alpha(WX|Z=z)$ for all $z \in \mathcal{Z}$.

If there exits a deterministic mapping $f : \mathcal{X} \times \mathcal{Z} \to \mathcal{W}$, it holds that

$$\sum_w P_{W|XZ}(w|x,z)^\alpha = 1. \tag{35}$$

for $\alpha \in [0,1)$ since $P_{W|XZ}(w|x,z) \in \{0,1\}$ for all $(x,z) \in \mathrm{supp}P_{XZ}$. Then, (31) holds with equality, which also makes (32)–(34) hold with equalities in the case of $\alpha \in (0,1)$. In the case of $\alpha \to 0$, this argument is also valid (even for (33)), and hence, the existence of such f is sufficient for the equalities for all $\alpha \in [0,1)$.

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

Remark 5 *As we proved, the existence of such a deterministic map $f : \mathcal{X} \times \mathcal{Z} \to \mathcal{W}$ is sufficient condition for the equality of $R_\alpha^{\mathsf{N}}(X|Z) \leq R_\alpha^{\mathsf{N}}(WX|Z)$. However, it is easy to show that the existence of such f is* not necessary *to satisfy $R_\alpha^{\mathsf{N}}(X|Z) = R_\alpha^{\mathsf{N}}(WX|Z)$ when $\alpha \to \infty$. To see this, it is sufficient to show $R_\infty(X) = R_\infty(WX)$ holds even if such f does not exist. Such a case happens when $\mathcal{X} = \mathcal{Y} = \{0,1\}$ and $P_{WX}(0,1) = 1/2$, $P_{WX}(1,0) = P_{WX}(1,1) = 1/4$, and $P_{WX}(0,0) = 0$. In this case, it holds that $R_\infty(WX) = R_\infty(X) = \log 2$ but such a deterministic map f does not exist. Actually, in this case, we can check that $H(WX) = (3/2)\log 2 > H(X) = \log 2$.*

3.3 Data Processing Inequality

The data processing inequality (DPI, [24]) tells us that $I(X;Y) \geq I(X;Z)$ holds if $X \leftrightarrow Y \leftrightarrow Z$. We can extend Theorem 2, in the following way:

Theorem 6 (Data processing inequality) *Let X, Y, and Z be random variables taking in 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 $\alpha \geq 0$. In the case of $\alpha \in (0,\infty)$, 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 in 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

$$P_{X|YR}(x|y,r) = \frac{P_{XR|Y}(x,r|y)}{P_{R|Y}(r|y)} = \frac{P_{X|Y}(x|y)P_{R|Y}(r|y)}{P_{R|Y}(r|y)} = P_{X|Y}(x|y) \tag{36}$$

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 $R_\alpha^{\mathsf{H}}(X|YR) = R_\alpha^{\mathsf{H}}(X|Y)$. Due to CRE, 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 6 *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 for $R_\alpha^{\mathsf{A}}(X|Y)$ and $R_\alpha^{\mathsf{H}}(X|Y)$ since DPI is an extension of CRE, as summarized in Table 1.*

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 7).

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 $0 \le \alpha \le 1$ *and* $0 \le P_e \le 1 - \frac{1}{|\mathcal{X}|}$*, or* $\alpha \ge 1$ *and* $P_e \ge 1 - \frac{1}{|\mathcal{X}|}$*, it holds that*

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

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

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

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 8 Note that Fano's inequality implies $H(X|Y) \to 0$ as $P_e \to 0$. Theorem 7 implies that, for any $\alpha \ge 0$, $R_\alpha^{\mathsf{H}}(X|Y) \to 0$ 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 [16, 17].

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, Hartley 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 [5] for symmetric-key encryption and its extensions (or variants) by Dodis [6], and Alimomeni and Safavi-Naini [8]. To describe 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 6 (Shannon's bound: [5]) *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 7 (Dodis's bound: Theorem 3 in [6]) *Let Π be a symmetric-key encryption. If Π satisfies perfect correctness and perfect secrecy, we have $R_\infty(K) \geq R_\infty(M)$.*

Remark 9 *Note that a similar result with Proposition 7 is proved in [33] using information spectrum methods [34]. In [33, Theorem 5], it is clarified that the inf-spectral rate of the secret key is not less than that of the plaintext. Noticing the recent results [35] of smooth min-entropy [36], the asymptotic version of min-entropy is equivalent to the inf-spectral entropy. Hence, we can say that an asymptotic version of Proposition 7 is proved in [33]. However, [6] directly proves $R_\infty(K) \geq R_\infty(M)$ in a non-asymptotic setup.*

Proposition 8 (Alimomeni and Safavi-Naini's bound: Theorem 2 in [8]) *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

$$\begin{aligned} R_\alpha(M) &= I_\alpha^{\mathsf{H}}(M;C) + R_\alpha^{\mathsf{H}}(M|C) \\ &\overset{\text{(a)}}{\leq} I_\alpha^{\mathsf{H}}(M;C) + R_\alpha^{\mathsf{H}}(M|\hat{M}) \\ &\overset{\text{(b)}}{\leq} I_\alpha^{\mathsf{H}}(M;C) + \frac{1}{1-\alpha}\log\left[(|\mathcal{M}|-1)^{1-\alpha}P_e^\alpha + \bar{P}_e^\alpha\right], \end{aligned} \tag{37}$$

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

$$\begin{aligned} \exp\left\{(1-\alpha)[R_\alpha(M) - I_\alpha^{\mathsf{H}}(M;C)]\right\} &\geq (|\mathcal{M}|-1)^{1-\alpha}P_e^\alpha + \bar{P}_e^\alpha \\ &\geq (1-P_e)^\alpha. \end{aligned} \tag{38}$$

By (38), 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*

$$P_e \geq 1 - \exp\left\{\frac{1-\alpha}{\alpha}R_\alpha(M)\right\}\exp\left\{\frac{\alpha-1}{\alpha}I_\alpha^{\mathsf{H}}(M;C)\right\}. \tag{39}$$

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

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{H}}(M;C)\right\}. \tag{40}$$

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 10 *The bound (39) is tight for $\alpha = \infty$ in the following sense: Consider the case $I_\infty^{\mathsf{H}}(M;C) = 0$. Then, (39) 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 ([37]) *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 9 ([11]) *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 10 ([17]) *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 $I_\alpha^{\mathsf{H}}(M;C)$. By combining the above propositions, we have

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

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 $R_\alpha^{\mathsf{A}}(X|Y)$. From the inequality, we obtain the following result.

Proposition 11 *The failure probability of adversary's guessing is lower-bounded by*

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{A}}(M_{1/\alpha}; C)\right\}. \tag{41}$$

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

$$P_e \geq 1 - |\mathcal{M}|^{\frac{1-\alpha}{\alpha}} \exp\left\{\frac{\alpha-1}{\alpha} I_\alpha^{\mathsf{A}}(M; C)\right\}. \tag{42}$$

Remark 11 *If P_M is the uniform distribution, the bound (40) is directly obtained from the bound (42) since $I_\alpha^{\mathsf{A}}(M;C) \leq I_\alpha^{\mathsf{H}}(M;C)$. However, it is not the case in general.*

Therefore, $I_\alpha^{\mathsf{H}}(M;C) \leq \epsilon$ or $I_\alpha^{\mathsf{A}}(M;C) \leq \epsilon$ for 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 $I_\alpha^{\mathsf{H}}(M;C) \leq \epsilon$ or $I_\alpha^{\mathsf{A}}(M;C) \leq \epsilon$, which is equivalent to $R_\alpha(M) - R_\alpha^{\mathsf{H}}(M|C) \leq \epsilon$ or $R_\alpha(M) - R_\alpha^{\mathsf{A}}(M|C) \leq \epsilon$, 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 and P_M be a random variable which takes plaintexts in $\mathcal{M}$ and its distribution, respectively. C denotes a random variable which takes ciphertexts $c \in \mathcal{C}$.

Let $\Pi = ([P_{ED}], \pi_{enc}, \pi_{dec})$ 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., $H(M|C) = H(M)$)[14] is equivalent to that of perfect secrecy with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ or $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ for $\alpha \in (0, \infty)$ (see Theorem 2). Also, ϵ-secrecy with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ (resp., $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$) is equivalent to $I_\alpha^{\mathsf{H}}(M;C) \le \epsilon$ (resp., $I_\alpha^{\mathsf{A}}(M;C) \le \epsilon$) (see Section 4.3).

5.2 Basic Idea for Generalization of Shannon's Impossibility

By Shannon's work [5], it is well known that we have $H(K) \ge H(M)$ for symmetric-key encryption with perfect secrecy (see Proposition 6), 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 $H(K) \ge H(M)$ below, where PS, PC, and CRE mean perfect secrecy, perfect correctness, and conditioning reduces entropy, respectively.

Proof I.

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

Proof II.

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

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 conditioned AIE, i.e., (9c), and chain rule are used in Proof II. At this point, one may think it hopeless to apply the technique in Proof I, since $I_\alpha^{\mathsf{H}}(X;Y) \ne I_\alpha^{\mathsf{H}}(Y;X)$ and $I_\alpha^{\mathsf{A}}(X;Y) \ne I_\alpha^{\mathsf{A}}(Y;X)$ 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 AIE 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.

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

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 $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ and $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$. 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 the size of encryption-keys) 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 the size of decryption-keys) 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 the size of ciphertexts) It holds that $R_\alpha(C) \geq R_\alpha(M)$.

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

$$R_\alpha(C) \leq R_\alpha(C|M) + \epsilon \overset{\text{(a)}}{\leq} R_\alpha(CE|M) + \epsilon \overset{\text{(b)}}{=} R_\alpha(E|M) + \epsilon \overset{\text{(c)}}{=} R_\alpha(E) + \epsilon, \tag{43}$$

where (a) follows from Theorem 5 (i), (b) follows from Theorem 5 (ii) since π_{enc} is deterministic, and (c) follows from that M and E are independent.

Secondly, we can show (ii) as follows.

$$R_\alpha(M) \leq R_\alpha(M|C) + \epsilon \overset{\text{(a)}}{\leq} R_\alpha(MD|C) + \epsilon \overset{\text{(b)}}{=} R_\alpha(D|C) + \epsilon \overset{\text{(c)}}{\leq} R_\alpha(D) + \epsilon, \tag{44}$$

where (a) follows from Theorem 5 (i), (b) follows from Theorem 5 (ii) since Π meets perfect correctness, and (c) follows from that both $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$ and $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ satisfy CRE (see Theorem 2).

Finally, we show (iii) as follows.

$$R_\alpha(M) \overset{\text{(a)}}{=} R_\alpha(M|D) \overset{\text{(b)}}{\leq} R_\alpha(MC|D) \overset{\text{(c)}}{=} R_\alpha(C|D) \overset{\text{(d)}}{\leq} R_\alpha(C), \tag{45}$$

where (a) follows from that D and M are independent, (b) follows from Theorem 5 (i), (c) also follows from Theorem 5 (ii) since Π meets perfect correctness, and (d) follows from that both $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ and $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$ satisfy CRE (see Theorem 2). □

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 $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ and $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$. 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 $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ and $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$. 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 corollary 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 $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ or $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$. 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 12 *Note that the Shannon's bounds (i.e., Proposition 6) are special cases of Corollary 3, since they are obtained by applying $\alpha = 0, 1$ in Corollary 3*[15]*. Also, Dodis's bound (i.e., Proposition 7) is a special case of Corollary 3, since it is obtained by applying $\alpha \to \infty$ in Corollary 3. Furthermore, Alimomeni and Safavi-Naini's bound (i.e., Proposition 8) is a special case of Corollary 2, since it is obtained by applying $\epsilon = 0$ and $R^{\mathsf{avg}}_\infty(\cdot|\cdot) = \lim_{\alpha\to\infty} R^{\mathsf{A}}_\alpha(\cdot|\cdot)$ in Corollary 2*[16]*. 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 Constructions

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 $R^{\mathsf{H}}_\alpha(\cdot|\cdot)$ or $R^{\mathsf{A}}_\alpha(\cdot|\cdot)$. 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 the case of $\alpha \to \infty$. Actually, as we will see in the following, security criteria based on the min-entropy is strictly weaker than that of the Shannon entropy. Although in [8] it is not shown that the lower bound in Proposition 8 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 12 *The above protocol Π_1 does not meet perfect secrecy, and Π_1 satisfies perfect secrecy with respect to $R^{\mathsf{avg}}_\infty(\cdot|\cdot)$, or equivalently $I^{\mathsf{A}}_\infty(M; C) = 0$. Furthermore, it holds that $R_\infty(K) = R_\infty(M)$ in Π_1.*

Proof. For the above protocol Π_1, it holds that

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

[15] 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

[16] 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.

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

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

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

Remark 13 *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 the case of $\alpha \in (0,\infty)$, the perfect secrecy with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ implies the traditional perfect secrecy, namely, independency of C and M. However, in the case of ϵ-security with respect to $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ does not imply the perfect secrecy if $\epsilon > 0$. 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 asymptotic 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 13 *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 14 *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 13 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 Further Extension of Our Results

In Section 5, 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 $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$.

6.1 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 AIE 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 AIE. 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 7 *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 AIE) (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 7 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 $(R_\alpha(\cdot), R^{\mathsf{H}}_\alpha(\cdot|\cdot))$ and $(R_\alpha(\cdot), R^{\mathsf{A}}_\alpha(\cdot|\cdot))$ for all $\alpha \geq 0$. In addition, Σ contains $(R_\alpha(\cdot), R^{\mathsf{RW}}_\alpha(\cdot|\cdot))$ for any $\alpha > 1$, and its proof is straightforward from [10, 22]. Therefore, we have the following proposition.

Proposition 14 *The class Σ in Definition 7 contains*

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

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

6.2 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 8 (Secrecy) Let Π be an encryption protocol. Then, for any $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 7, 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 7 in a comprehensive way. Theorem 10, and Corollaries 4, 5 and 6 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 10 *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 7, we have the following.*

(i) (Lower bound on the size of encryption-keys) 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 the size of decryption-keys) Suppose that Π satisfies $F(M) \leq F(M|C)+\epsilon$. Then, we have $F(D) \geq F(M) - \epsilon$.
(iii) (Lower bound on the size of ciphertexts) It holds that $F(C) \geq F(M)$.

Corollary 4 *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 7, 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 5 *Let $(F(\cdot), F(\cdot|\cdot)) \in \Sigma$ in Definition 7, 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 6 *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 7, 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 Conclusion

Information theoretic cryptography was discussed based on conditional Rényi entropies. Our discussion focused not only on cryptography but also on the definitions of conditional Rényi entropies and the related information theoretic inequalities.

First, we revisited conditional Rényi entropies, and clarified what kind of properties are required and actually satisfied. Based on the axiomatic discussion for conditional Rényi entropies, we listed several hopefully required properties for condition Rényi entropies such as non-negativity, chain rule, conditioning reduces entropy (CRE), additional information increases entropy (AIE), data processing inequality (DPI), and their extended inequalities. Since five conditional Rényi entropies are proposed so far, we investigated each definition of conditional Rényi entropy actually satisfies each property. As a result, we concluded that the conditional Rényi entropies proposed by Arimoto [11] and Hayashi [12], denoted by $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$, respectively, satisfy all properties except chain rule and weak chain rule. In addition, we pointed out that $R_\alpha^{\mathsf{A}}(\cdot|\cdot)$ and $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ correspond to the conditional min-entropies $R_\infty^{\mathsf{avg}}(\cdot|\cdot)$ and $R_\infty^{\mathsf{wst}}(\cdot|\cdot)$, respectively, when $\alpha \to \infty$. Finally, we presented Fano's inequality for $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$.

Then, we proposed 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 was 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 [6], and Alimomeni and Safavi-Naini [8].

Finally, new optimal symmetric key cryptography is proposed which achieve our lower bounds. In particular, we succeeded in constructing symmetric key cryptography which is secure under the conditional min-entropy $R_\infty^{\mathsf{avg}}(\cdot|\cdot)$ and the conditional Rényi entropy $R_\alpha^{\mathsf{H}}(\cdot|\cdot)$ with (almost) tight key sizes.

Our discussion can be extended to wider classes of conditional entropies. We discuss such a framework of extended conditional entropies, and provide an extended model of symmetric key encryption. We also derive the lower bounds of key sizes, which includes the case of encryption based on conditional Rényi entropies.

Acknowledgments

The authors would like to thank the anonymous referees of ICITS2013 for their helpful comments. They also grateful Prof. Hirosuke Yamamoto with the University of Tokyo for bringing their attention to [7]. Mitsugu Iwamoto is supported by JSPS KAKENHI Grant No. 23760330. Junji Shikata is supported by JSPS KAKENHI Grant No. 23500012.

References

1. Iwamoto, M., Shikata, J.: Information theoretic security for encryption based on conditional Rényi entropies. In: International Conference on Information Theoretic Security (ICITS2013). (November 2013) 103–121
2. Shannon, C.: A mathematical theory of communication. Bell Systems Technical Journal **27** (July and Oct. 1948) 379–423
3. Hartley, R.V.L.: Transmission of information. Bell System Technical Journal **7**(3) (July 1928) 535–563
4. Haståd, J., Impagliazzo, R., Levin, L.A., Luby, M.: A pseudorandom generator from one-way function. SIAM Journal of Computing (1994) 1364–1396
5. Shannon, C.E.: Communication theory of secrecy systems. Bell Tech. J. **28** (Oct. 1949) 656–715
6. 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`.
7. Merhav, N.: A large-deviations notions of perfect secrecy. IEEE Trans. Information Theory **30**(2) (2003) 506–508
8. 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
9. Rényi, A.: On measures of information and entropy. Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960 (1961) 547–561
10. Teixeira, A., Matos, A., Antunes, L.: Conditional Rényi entropies. IEEE Trans. Information Theory **58**(7) (July 2012) 4273–4277
11. 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
12. Hayashi, M.: Exponential decreasing rate of leaked information in universal random privacy amplification. IEEE Trans. Information Theory **57**(6) (2011) 3989–4001
13. 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
14. 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
15. Fano, R.M.: Class notes for transmission of information (course 6.574). Technical report, MIT, Cambridge, MA (1952)
16. Polyanskiy, Y., Poor, V., Verdú, S.: Channel coding rate in the finite blocklength regime. IEEE Trans. Inform. Theory **56**(5) (2010) 2307–2359
17. Polyanskiy, Y., Verdú, S.: Arimoto channel coding converse and Rényi divergence. Forty-Eighth Annual Allerton Conference (2010) 1327–1333
18. 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
19. Cachin, C.: Entropy Measures and Unconditional Security in Cryptography. PhD thesis, Swiss Federal Institute of Technology, Zürich, Switzerland (1997)
20. Jizba, P., Arimitsu, T.: Generalized statistics: Yet another generalization. Physica A **340** (2004) 110–116
21. Jizba, P., Arimitsu, T.: The world according to Rényi: Thermodynamics of multifractal systems. Annals of Physics **312** (2004) 17–59
22. Renner, R., Wolf, S.: Simple and tight bounds for information reconciliation and privacy amplification. Advances in Cryptology–ASIACRYPT2005, LNCS4515, Springer-Verlag (2005) 199–216
23. Hayashi, M.: Tight exponential analysis of universally composable privacy amplification and its applications. arXiv:1010.1358 (2010)
24. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Second edn. Wiley and Interscience (2006)
25. Stinson, D.R.: Cryptography: Theory and Practice. Third edn. Chapman & Hall/CRC (2005)

26. Fujishige, S.: Polymatroidal dependence structure of a set of random variables. Information and Control **39** (1978) 55–72
27. 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
28. 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
29. 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
30. Dodis, Y., Yu, Y.: Overcoming weak expectations. Tenth Workshop on Theory of Cryptography–TCC2013, LNCS4117, Springer-Verlag (2013) 1–22
31. Arikan, E.: An inequality on guessing and its application to sequential decoding. IEEE Trans. Information Theory **42**(1) (1996) 99–105
32. Hardy, G., Littlewood, J.E., Pólya, G.: Inequalities. 2nd. edn. Cambridge mathematical library (1934)
33. Koga, H.: New coding theorems for fixed-length source coding and Shannon's cipher system with a general source. In: ISITA2008. (December 2008) 251–256
34. Han, T.S.: Information-Spectrum Methods in Information Theory. Springer-Verlag (2003)
35. Tomamichel, M., Hayashi, M.: A hierarchy of information quantities for finite block length analysis of quantum tasks. arXiv:1208.1478 (2012)
36. Renner, R., Wolf, S.: Smooth Rényi entropy and its applications. In: ISIT2004. (June–July 2004) 232
37. Gallager, R.G.: A simple derivation of the coding theorem and some applications. IEEE Trans. Inform. Theory **11**(1) (1965) 41–52

A Technical Proofs

A.1 Proof of Theorem 1

(i) The equality $\lim_{\alpha\to 1} R_\alpha^{\mathsf{A}}(X|Y) = H(X|Y)$ is proved in [11]. $\lim_{\alpha\to 1} R_\alpha^{\mathsf{H}}(X|Y) = H(X|Y)$ is easily verified by the L'Hospital's rule. Namely, we have

$$\begin{aligned}\lim_{\alpha\to 1} R_\alpha^{\mathsf{H}}(X|Y) &= -\lim_{\alpha\to 1}\frac{\mathrm{d}}{\mathrm{d}\alpha}\log \mathbb{E}_Y\left[\sum_{x\in X} P_{X|Y}(x|Y)^\alpha\right] = -\lim_{\alpha\to 1}\mathbb{E}_Y\left[\sum_{x\in X}\frac{\mathrm{d}}{\mathrm{d}\alpha}P_{X|Y}(x|Y)^\alpha\right]\\ &= -\mathbb{E}_Y\left[\sum_{x\in X} P_{X|Y}(x|Y)\log P_{X|Y}(x|Y)\right] = \sum_{y\in\mathcal{Y}} P_Y(y)H(X|Y=y).\end{aligned} \tag{46}$$

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

$$\max_x P_{X|Y}(x|y) \le \left\{\sum_x P_{X|Y}(x|y)^\alpha\right\}^{1/\alpha} \le |\mathcal{X}|^{1/\alpha}\max_x P_{X|Y}(x|y) \tag{47}$$

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

$$\lim_{\alpha\to\infty}\frac{\alpha}{1-\alpha}\log\sum_y P_Y(y)\left\{\sum_x P_{X|Y}(x|y)^\alpha\right\}^{1/\alpha} = -\log\sum_y P_Y(y)\max_x P_{X|Y}(x|y), \tag{48}$$

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 $y\in\mathcal{Y}$

$$\max_x P_{X|Y}(x|y)^\alpha \le \sum_x P_{X|Y}(x|y)^\alpha \le |\mathcal{X}|\max_x P_{X|Y}(x|y)^\alpha. \tag{49}$$

The expectations of the upper and the lower bounds in (49) with respect to Y can be further bounded as

$$\sum_y P_Y(y)|\mathcal{X}| \max_x P_{X|Y}(x|y)^\alpha \leq |\mathcal{X}| \max_{\substack{x\in\mathcal{X} \\ y\in \mathrm{supp}\, P_Y}} P_{X|Y}(x|y)^\alpha \tag{50}$$

and

$$P_Y(y^*) \max_x P_{X|Y}(x|y^*)^\alpha \leq \sum_y P_Y(y) \max_x P_{X|Y}(x|y)^\alpha \tag{51}$$

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

Now, we can assume that α is sufficiently large, say $\alpha > 1$. Then, noticing that $1/(1-\alpha) < 0$ and from (49)–(51), we have

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

and

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

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

A.2 Proof of Proposition (2)

All relations except (16) are almost obvious by noticing that $\lim_{\alpha\to 0} R_\alpha(X) = \log|\mathrm{supp} P_X|$. Equation (16) is immediately obtained from the following lemma.

Lemma 1 *For arbitrary random variable Z taking values in a finite set $\mathcal{Z}$, it holds that* $\lim_{\alpha\to 0} \mathbb{E}\left[Z^{1/\alpha}\right]^\alpha = \max_{z\in\mathcal{Z}} z$.

Proof of Lemma 1

Let $S_\alpha(Z) := \mathbb{E}_Z\left[Z^{1/\alpha}\right]^\alpha$ and $z^* := \max_{z\in\mathcal{Z}} z$. Lemma 1 is verified from the following two inequalities.

We first check that $S_\alpha(Z) \leq z^*$ for arbitrary $\alpha \geq 0$, which is immediately obtained by observing that $S_\alpha(Z) \leq \left\{\sum_{z\in\mathcal{Z}} P_Z(z) z^{*1/\alpha}\right\}^\alpha = z^*$.

Then we prove the second inequality $\lim_{\alpha\to 0} S_\alpha(Z) \geq z^*$ from the easily verified relation such that $P_Z(z^*) \leq \sum_{z\in\mathcal{Z}} P_Z(z)\left(z/z^*\right)^{1/\alpha}$. This inequality is equivalent to $z^{*1/\alpha} P_Z(z^*) \leq \mathbb{E}_Z\left[Z^{1/\alpha}\right]$, i.e., $z^* P_Z(z^*)^\alpha \leq S_\alpha(Z)$. Taking the limits $\alpha \to 0$ for both sides of $z^* P_Z(z^*)^\alpha \leq S_\alpha(Z)$, we obtain $\lim_{\alpha\to 0} S_\alpha(Z) \geq z^*$. □

Now, we can prove (16). Let $Z := \sum_{x\in\mathcal{X}} P_{X|Y}(x|Y)^\alpha$ with $\mathcal{Z} := \mathcal{Y}$. From Lemma 1 combined with the easily verified identity $\lim_{\alpha\to 0} \sum_{x\in\mathcal{X}} P_{X|Y}(x|y)^\alpha = \log\left|\mathrm{supp} P_{X|Y=y}\right|$ for arbitrary $y \in \mathcal{Y}$, (16) yields immediately. □

A.3 Proof of Proposition 5

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

$$\lim_{\alpha\to 1} D_\alpha(X_1\|X_2|Y) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\log\sum_{x,y}\left\{\frac{W(x|y)}{V(x|y)}\right\}^\alpha V(x|y)Q(y)\Bigg|_{\alpha=1}$$
$$= \sum_{x,y} Q(y)W(x|y)\log\frac{W(x|y)}{V(x|y)} = D(W\|V|Q) = D(X_1\|X_2|Y). \quad (52)$$

(ii) For $\alpha\in(0,1)$, it follows that

$$\sum_{x,y}\frac{W(x|y)^\alpha}{V(x|y)^{\alpha-1}}Q(y) = \sum_{x,y}\left(\frac{W(x|y)}{V(x|y)}\right)^\alpha P_{YZ}(x,y)$$
$$= \mathbb{E}_{YZ}\left[\left(\frac{W(Y|Z)}{V(Y|Z)}\right)^\alpha\right]$$
$$\le \left\{\mathbb{E}_{YZ}\left[\frac{W(Y|Z)}{V(Y|Z)}\right]\right\}^\alpha$$
$$= \left\{\sum_{x,y}\frac{W(x|y)}{V(x|y)}V(x|y)Q(y)\right\}^\alpha = 1 \quad (53)$$

where the inequality follows from Jensen's inequality. The quality holds if and only 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, in the case of $\alpha\in(1,\infty)$, we have $\sum_{x,y}\frac{W(x|y)^\alpha}{V(x|y)^{\alpha-1}}Q(y)\ge 1$, where the equality holds if and only if $W(\cdot|y) = V(\cdot|y)$ for $y\in\operatorname{supp}Q$.

Hence, $D_\alpha(X_1\|X_2|Y)\ge 0$ holds for $\alpha\in(0,1)\cup(1,\infty)$, but due to the continuity of conditional α-divergence, this inequality holds for all $\alpha\ge 0$. □

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

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

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

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

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

$$0\le D_\alpha(XY\|Z|Y) = \frac{1}{\alpha-1}\log\left[\sum_j P_Y(j)\sum_i\left(\frac{P_{XY}(i,j)}{P_Y(j)}\right)^\alpha P_{Z_1}(i|j)^{1-\alpha}\right]$$
$$= \frac{1}{\alpha-1}\log\left[m^{1-\alpha}\sum_{i,j}P_{XY}(i,j)^\alpha P_Y(j)^{1-\alpha}P_Z(i,j)^{1-\alpha}\right]. \quad (54)$$

On the other hand, we get

$$\begin{aligned}
&\sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} P_Z(i,j)^{1-\alpha} \\
&\quad = \sum_{i\neq j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} P_Z(i,j)^{1-\alpha} + \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} P_Z(i,i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \sum_{i\neq j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} + \left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \left(\sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} - \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \\
&\qquad\qquad + \left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} \sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \\
&\quad = \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} \\
&\qquad\qquad + \left(\sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \left[\left(\frac{\bar{P}_e}{m}\right)^{1-\alpha} - \left(\frac{P_e}{m(m-1)}\right)^{1-\alpha} \right]. \qquad (55)
\end{aligned}$$

Therefore, by (54) and (55) we obtain

$$\begin{aligned}
0 \geq \frac{1}{1-\alpha} \log \Bigg\{ &\left(\frac{P_e}{m-1}\right)^{1-\alpha} \sum_{i,j} P_{XY}(i,j)^{\alpha} P_Y(j)^{1-\alpha} \\
&+ \left(\sum_{i} P_{XY}(i,i)^{\alpha} P_Y(i)^{1-\alpha} \right) \left[\bar{P}_e^{1-\alpha} - \left(\frac{P_e}{m-1}\right)^{1-\alpha} \right] \Bigg\} \qquad (56)
\end{aligned}$$

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 (56) is written in the form:

$$\frac{1}{1-\alpha} \log(ar + sb) \leq 0. \qquad (57)$$

Suppose that $0 \leq \alpha < 1$ and $P_e \neq 0$ (i.e., $a > 0$). Then, (57) implies

$$r \leq a^{-1}(1 - sb) = (m-1)^{1-\alpha} P_e^{\alpha-1} + s(1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}). \qquad (58)$$

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 2 *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

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

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 (58) and (i) in Lemma 2 it follows that

$$\begin{aligned} r &\leq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e^{\alpha} (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha} + \bar{P}_e^{\alpha}. \end{aligned} \tag{59}$$

If $0 \leq \alpha < 1$ and $0 < P_e \leq 1 - \frac{1}{m}$, from (58) and (i) in Lemma 2 it follows that

$$\begin{aligned} r &\leq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e. \end{aligned} \tag{60}$$

Next, suppose that $\alpha > 1$ and $P_e \neq 0$. Then, (57) implies

$$\begin{aligned} r &\geq a^{-1}(1 - sb) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} + s(1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}). \end{aligned} \tag{61}$$

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 (61) and (ii) in Lemma 2 it follows that

$$\begin{aligned} r &\geq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha-1} (1 - \bar{P}_e^{2-\alpha}) + \bar{P}_e. \end{aligned} \tag{62}$$

If $\alpha > 1$ and $0 < P_e \leq 1 - \frac{1}{m}$, from (61) and (ii) in Lemma 2 it follows that

$$\begin{aligned} r &\geq (m-1)^{1-\alpha} P_e^{\alpha-1} + \bar{P}_e^{\alpha} (1 - (m-1)^{1-\alpha} P_e^{\alpha-1} \bar{P}_e^{1-\alpha}) \\ &= (m-1)^{1-\alpha} P_e^{\alpha} + \bar{P}_e^{\alpha}. \end{aligned} \tag{63}$$

Therefore, from (59), (60), (62) and (63), it holds that

$$R_\alpha^{\mathsf{H}}(X|Y) = \frac{1}{1-\alpha}\log r \leq \begin{cases} \frac{1}{1-\alpha}\log\left[(m-1)^{1-\alpha}P_e^\alpha + \bar{P}_e^\alpha\right] \\ \quad \text{if } 0 \le \alpha < 1 \text{ and } P_e \ge 1 - \frac{1}{m}, \text{ or } \alpha > 1 \text{ and } 0 < P_e \le 1 - \frac{1}{m}, \\ \frac{1}{1-\alpha}\log\left[(m-1)^{1-\alpha}P_e^{\alpha-1}(1-\bar{P}_e^{2-\alpha}) + \bar{P}_e\right] \\ \quad \text{if } 0 \le \alpha < 1 \text{ and } 0 < P_e \le 1 - \frac{1}{m}, \text{ or } \alpha > 1 \text{ and } P_e \ge 1 - \frac{1}{m} \end{cases} \tag{64}$$

For the case $\alpha = 1$, the left hand of (64) implies $\lim_{\alpha\to 1} R_\alpha(X|Y) = H(X|Y)$ by Theorem 1-(ii). In addition, the right hands of (64) have a finite limit at $\alpha = 1$, and it is equal to Fano's inequality (see Remark 7). Therefore, (64) holds even for $\alpha = 1$.

For the case $P_e = 0$, the left hand of (64) implies $\lim_{P_e\to 0} R_\alpha(X|Y) = R_\alpha(X|X) = 0$, and the right hands of (64) imply

$$\lim_{P_e\to 0} \frac{1}{1-\alpha}\log\left[(m-1)^{1-\alpha}P_e^\alpha + \bar{P}_e^\alpha\right] = 0 \text{ (for } \alpha \ge 1),$$
$$\lim_{P_e\to 0} \frac{1}{1-\alpha}\log\left[(m-1)^{1-\alpha}P_e^{\alpha-1}(1-\bar{P}_e^{2-\alpha}) + \bar{P}_e\right] = 0 \text{ (for } 0 \le \alpha \le 1).$$

Therefore, (64) holds for $P_e = 0$. □

A.5 Proof of Theorem 13

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

$$\begin{aligned} R_\alpha(M) &= \frac{1}{1-\alpha}\log(p^\alpha + (1-p)^\alpha) \\ &= \frac{1}{1-\alpha}\log\left[\left(\frac{1}{2}\right)^\alpha(1-\delta_1)^\alpha + \left(\frac{1}{2}\right)^\alpha(1+\delta_1)^\alpha\right] \\ &= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha(2+o(1)), \end{aligned} \tag{65}$$

$$\begin{aligned} R_\alpha(K) &= \frac{1}{1-\alpha}\log(q^\alpha + (1-q)^\alpha) = \frac{1}{1-\alpha}\log(p^\alpha + (1-p)^\alpha + o(1)) \\ &= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha(2+o(1)), \end{aligned} \tag{66}$$

$$\begin{aligned} R_\alpha(C) &= \frac{1}{1-\alpha}\log\left[\left(\frac{1}{2}\right)^\alpha(1-\delta_1^2)^\alpha + \left(\frac{1}{2}\right)^\alpha(1+\delta_1^2)^\alpha + o(1)\right] \\ &= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha(2+o(1)), \end{aligned} \tag{67}$$

$$\begin{aligned} R_\alpha^{\mathsf{H}}(M|C) &= \frac{1}{1-\alpha}\log\sum_c P_C(c)\sum_m P_{M|C}(m|c)^\alpha \\ &= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha\left[(1-\delta_1^2) + \frac{1}{2}\frac{(1-\delta_1)^{2\alpha}}{(1+\delta_1^2)^{\alpha-1}} + \frac{1}{2}\frac{(1+\delta_1)^{2\alpha}}{(1+\delta_1^2)^{\alpha-1}} + o(1)\right] \\ &= \frac{1}{1-\alpha}\log\left(\frac{1}{2}\right)^\alpha(2+o(1)), \end{aligned} \tag{68}$$

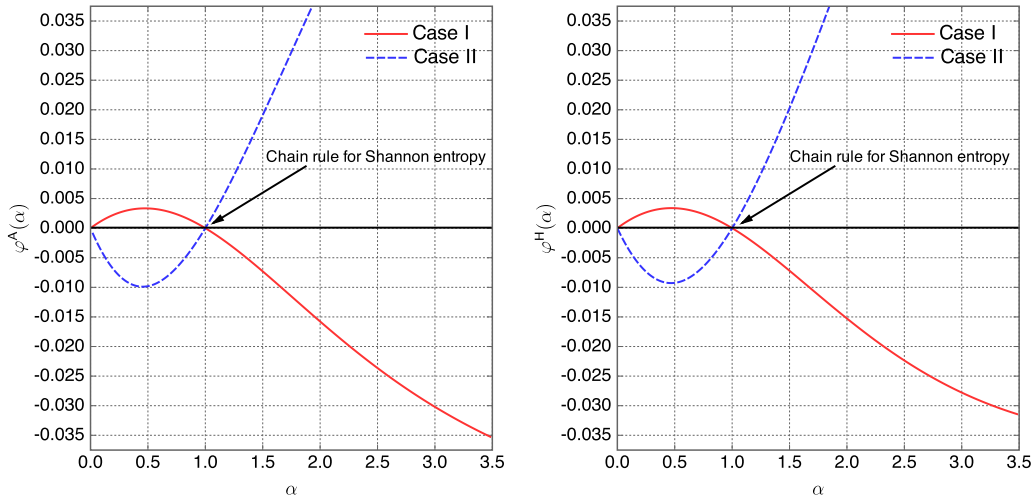
$$\begin{aligned} R_\alpha^{\mathsf{H}}(C|M) &= \frac{1}{1-\alpha}\log\sum_m P_M(m)\sum_c P_{C|M}(c|m)^\alpha \\ &= \frac{1}{1-\alpha}\log\left[p(q^\alpha + (1-q)^\alpha) + (1-p)(q^\alpha + (1-q)^\alpha)\right] \\ &= \frac{1}{1-\alpha}\log\left(q^\alpha + (1-q)^\alpha\right) = R_\alpha(K). \end{aligned} \tag{69}$$

Therefore, we get

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

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 14 we see that $I_\alpha^{\mathsf{H}}(M;C) \neq I_\alpha^{\mathsf{H}}(C;M)$ by calculation. □

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



(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