Security analysis of $\varepsilon$-almost dual universal$_2$ hash functions

Masahito Hayashi

Abstract

Recently, $\varepsilon$-almost dual universal$_2$ hash functions has been proposed as a new and wider class of hash functions. This class well works even when the random seeds of hash function are subject to non-uniform distribution. This paper evaluates the security performance when we apply this kind of hash functions. We evaluate the security in several kinds of setting based on the $L_1$ distinguishability criterion and the modified mutual information criterion. The obtained evaluation is based on smoothing of Rényi entropy of order 2 and/or min entropy. Further, we apply this analysis to the secret key generation with error correction.

Index Terms

$\varepsilon$-almost dual universal$_2$ hash function, secret key generation, exponential decreasing rate, single-shot setting, equivocation rate

I. INTRODUCTION

A. Overview

Secure key generation is an important problem in information theoretic security. When a part of keys are leaked to a third party, we cannot use the key. In this case, we need to apply a hash function to the keys. Bennett et al. [4] and Håstad et al. [15] proposed to use universal$_2$ hash functions for privacy amplification and derived two universal hashing lemma, which provides an upper bound for leaked information based on Rényi entropy of order 2. Recently, Tsurumaru et al. [14] proposed to use $\varepsilon$-almost dual universal$_2$ hash functions, which is a generalization of liner universal$_2$ hash functions, and obtained a different version of two universal hashing lemma for this class of hash functions. Further, the recent paper [34] shows that the $\varepsilon$-almost dual universal$_2$ well works even when the random seeds of the hash function are subject to a non-uniform distribution while the $\varepsilon$-almost universal$_2$ does not well work in this case. The paper [34] gives concrete examples of $\varepsilon$-almost universal$_2$ hash functions that have a smaller calculation amount and a smaller number of random variables than the concatenation of Toeplitz matrix and the identity matrix, which is a typical example of universal$_2$ hash functions. Since it is quite difficult to realize the perfect random seeds, it is required from a practical viewpoint to evaluate the security based on the $\varepsilon$-almost dual universal$_2$.

Two universal hashing lemma can guarantee the security only when the length of the generated keys is less than Rényi entropy of order 2. In order to resolve this drawback, Renner [16] attached the smoothing to min entropy, which is a lower bound of conditional Rényi entropy of order 2. That is, he proposed to maximize the min-entropy among the sub-distribution whose variational distance to the true distribution is less than a given threshold. He also employs the variational distance between the true distribution and the ideal distribution as the security criterion because it satisfies the universal composability. We call this criterion the $L_1$ distinguishability criterion. Then, he derived lower bound of the extractable key length with finite-length under the $L_1$ distinguishability criterion for universal$_2$ hash functions. In other word, when we fix the size of keys, he derived a lower bound of leaked information. However, it is not easy to find the maximizing sub-distribution. he did not give the rigorous maximization of min entropy under this condition. That is, he did not give a computable lower bound when the block size is sufficiently large. Instead of the rigorous maximization of min entropy under this condition, we can consider a lower bound of the maximum of min entropy. In the following, we say that this type lower bound or the method based on this type lower bound is an approximate smoothing of min entropy. In contrast with an approximate smoothing, we say that the tight value of min entropy under the given condition or the method based on the tight value is the rigorous smoothing of min entropy. It has been believed that the rigorous smoothing of min entropy yields a good upper bound of the $L_1$ distinguishability criterion.

Instead of min entropy, the previous paper [12] applied an approximate smoothing of Rényi entropy of order 2, and derived a lower bound of leaked information explicitly for universal$_2$ hash functions under the $L_1$ distinguishability criterion. The calculation amount of the obtained bound [12] does not depend on the block size. In the independent and identical distributed case, when the key generation rate is fixed, the bound yields a lower bound of the exponential decreasing rate of leaked information. The tightness is also shown in the recent paper [38, Theorem 30]. Currently, it has not been clarified the difference between the maximization of min entropy and the maximization of of Rényi entropy of order 2. This paper focuses this kind of difference in several settings.

M. Hayashi is with Graduate School of Mathematics, Nagoya University, Furocho, Chikusaku, Nagoya, 464-8602, Japan, and Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542. (e-mail: masahito@math.nagoya-u.ac.jp)
On the other hand, many papers \cite{21, 53, 7, 6, 54, 35, 24, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 25} employ the mutual information as the security criterion. When we apply universal$_2$ hash functions, the previous paper \cite{13} derived an upper bound of leaked information in this criterion by generalizing two universal hashing lemma. Then, it shows the security when the key generation rate is less than the conditional entropy in the independent and identical distributed case. This paper gives exponential decreasing rate of mutual information. However, the mutual information does not reflect the uniformity while it reflects the independence. In order to address with the uniformity as well as leaked information, we need the modification of mutual information, which is called the modified mutual information criterion and is explained in Subsection II-C. Indeed, as is shown in Appendix A when we assume several natural conditions for our security criterion, the security criterion is restricted to the modified mutual information criterion. Hence, it is natural to employ this criterion.

When one of two security criteria goes to zero exponentially, the other also goes to zero exponentially due to the relations given in Subsection II-C. Hence, the asymptotic key generation rate does not depend on the choice of the security criterion. However, the relations given in Subsection II-C cannot decide one of their exponential decreasing rates from the other exponent. Hence, we need to consider both exponents separately.

Here, we also notice the method of information spectrum, which is a powerful and general tool for information theory. Information spectrum has been established by Han and Verdú in their seminal papers \cite{58, 59, 60, 61, 26} and the book \cite{23}. This method can derive asymptotically tight bounds of the optimal performances of various information processings. While the relation between the rigorous smoothing of min entropy and the information spectrum has been pointed out with the fidelity distance in \cite{17}, it has not been discussed based on the variational distance. This paper also addresses this relation.

### B. Contributions

This paper shows the following five results. First, combining an approximate smoothing and the result in \cite{14}, we derive a lower bound of leaked information in the $L_1$ distinguishability criterion even when we apply $\epsilon$-almost dual universal$_2$ hash functions. Then, we derive the exponential decreasing rate of leaked information under the $L_1$ distinguishability criterion under an application of $\epsilon$-almost dual universal$_2$ hash functions. We also consider the second order asymptotics \cite{18, 19, 22, 20, 21}. Under this formulation, we show that the optimality can be realized even when $\epsilon$-almost dual universal$_2$ hash functions are applied.

Second, we derive a lower bound of leaked information in the $L_1$ distinguishability criterion even when we apply $\epsilon$-almost dual universal$_2$ hash functions. In this case, we cannot employ the method in \cite{13}. We invent another smoothing method of Rényi entropy of order 2 for the modified mutual information criterion. Then, by using the obtained bound, we derive a lower bound of the exponential decreasing rate of leaked information under the mutual information criterion under the class of $\epsilon$-almost dual universal$_2$ hash functions.

Third, under the both criteria, we analyze the difference between the exponential decreasing rates of the rigorous smoothing of min entropy and our particular rigorous smoothing of Rényi entropy of order 2. Then, we show that the exponential decreasing rates by the rigorous smoothing of min entropy are strictly worse than those by our approximate smoothing of Rényi entropy of order 2 under both criteria. This fact indicates the importance of smoothing of Rényi entropy of order 2. These three results are given in Sections IV and VI. Additionally, in Section V we clarify the relation between the information spectrum and the rigorous smoothing of min entropy in the single shot setting. This characterization is helpful for evaluation based on the rigorous smoothing of min entropy.

Fourth, in section VII we also discuss the case when the key generation rate is greater than the conditional entropy rate. In this case, the leaked information behaves linearly with respect to the block size when the modified mutual information criterion is employed. Employing our approximate smoothing of min entropy, we show that the application of $\epsilon$-almost dual universal$_2$ hash functions yields asymptotically optimal rate of leaked information. That is, only the exponential decreasing rate has the difference between the smoothing of Rényi entropy of order 2 and the smoothing of min entropy. We show that the other settings have no difference between two methods in the asymptotic setting.

Finally, in Section VII we apply our result to the case with error correction. Then, we obtain upper bounds of both security criteria when we apply an error correction and $\epsilon$-almost dual universal$_2$ hash functions. This process requires several formulas among information quantities given in Section II. These obtained results are summarized as Table I.

One might consider why we discuss the exponential decreasing rate as well as the second order asymptotics. The recent paper \cite{29} numerically showed that the superiority between the finite bounds based on the exponential decreasing rate and the second order asymptotics depends on the parameters, e.g., the number of block size and the required level of the $L_1$ distinguishability criterion. When the required level of the $L_1$ distinguishability criterion is small and the number of block size is not so large, the bounds based on the exponential decreasing rate gives a better bound. Conversely, when the required level of the $L_1$ distinguishability criterion is not so small and the number of block size is sufficiently large, the bounds based on the second order asymptotics gives a better bound.

Here, we should remark the relation with the result with the quantum case. The paper \cite{57} derived lower bounds of exponential decreasing rates of both criteria under the $\epsilon$-almost dual universal$_2$ hash functions. However, as is explained in Remark 4, their bounds are strictly worse than our results. The paper \cite{57} did not discuss the exponential decreasing rates of the rigorous
TABLE 1
SUMMARY OF OBTAINED RESULTS.

| task                           | setting        | single-shot  | asymptotic | $L_1$ | MMI |
|-------------------------------|----------------|--------------|------------|-------|-----|
| PV exponent (Rényi 2)         | single-shot    | (70) in Theorem 18 | (94) in Theorem 26 | (72) in Theorem 18 | (95) in Theorem 26 |
|                               | asymptotic     | (103) in Theorem 25 | (112) in Theorem 29 | (104) in Theorem 25 | (113) in Theorem 29 |
| PV exponent (min)             | single-shot    | (83) in Theorem 21 | (90) in Theorem 23 | –     | –   |
|                               | asymptotic     | (102) in Theorem 28 | (112) in Theorem 29 | –     | –   |
| second order (min)            | single-shot    | –             | (138) in Theorem 38 | –     | –   |
|                               | asymptotic     | –             | (139) in Theorem 38 | –     | –   |
| PV & fixed EC                 | exponent (Rényi 2) | single-shot | (130) in Theorem 36 | (131) in Theorem 36 | (133) in Theorem 37 |
|                               | asymptotic     | (138) in Theorem 38 | (139) in Theorem 38 | (141) in Theorem 39 | (142) in Theorem 39 |
| PV & randomized EC            | exponent (Rényi 2) | single-shot | no improvement | (133) in Theorem 37 | (133) in Theorem 37 |

PV is privacy amplification. EC is error correction. $L_1$ is the $L_1$ distinguishability criterion. MMI is the modified mutual information criterion. (min) means the result derived by rigorous smoothing of min entropy. (Rényi 2) means the results derived by our approximate smoothing of Rényi entropy of order 2.

smoothing of min entropy because of its difficulty. Further, the paper [57] did not address with the second order analysis. Hence, our results cannot be contained in the paper [57].

C. Organization

The remaining part of this paper is organized as follows. Now, we give the outline of the preliminary parts. In Section II we introduce the information quantities for evaluating the security and derive several useful inequalities for the quantum case. We also give a clear definition for security criteria. In Section III we introduce several class of hash functions (universal$_2$ hash functions and $\epsilon$-almost dual universal$_2$ hash functions). We clarify the relation between $\epsilon$-almost dual universal$_2$ hash functions and $\delta$-biased ensemble. We also derive an $\epsilon$-almost dual universal$_2$ version of two universal hashing lemma based on Lemma for $\delta$-biased ensemble given by Dodis et al [9]. The latter preliminary parts are more technical and used for proofs of the main results. In Section IV under the $\epsilon$-almost dual universal$_2$ condition, we evaluate the $L_1$ distinguishability criterion and the modified mutual information based on the rigorous smoothing of min entropy and Rényi entropy of order 2. These parts give the definitions for concepts and quantities describing the main results. These parts are almost included in the papers [14], [57]. However, these papers are written in the quantum terminologies. For readers’ convenience, we describe these parts without quantum notations.

Next, we outline the main results. In Section V using the tail probability of a proper event, we evaluate upper bounds given by the rigorous smoothing of min entropy in Section IV with the single-shot setting. This tail probability plays a central role in information spectrum. The bounds obtained in this section have smaller complexity for calculation than those given in Section IV. In Section V using the information quantities given in Section III we evaluate upper bounds given in Section IV. The bounds obtained in this section have smaller complexity for calculation than those given in Sections V and IV. In Section VII we derive an exponential decreasing rate for both criteria when we simply apply hash functions and there is no error between Alice’s and Bob’s information. In Section VII we also discuss the case when the key generation rate is greater than the conditional entropy rate.

In Section IX we proceed to the secret key generation with error correction. In this case, we need error correction as well as the privacy amplification. We derive Gallager bound for the error probability in this setting. We also derive upper bounds for the $L_1$ distinguishability criterion and the modified mutual information for a given sacrifice rate. Based on these upper bounds, we derive the exponential decreasing rates for both criteria. In Section X we apply our result to the simplest case.

II. Preparation

A. Information quantities for single system

1) Case of sub-distributions: In order to discuss the security problem, we prepare several information quantities for sub-distributions $P_A Q_A$ on a space $A$. That is, these are assumed to satisfy the conditions $P_A(a) \geq 0$ and $\sum_a P_A(a) \leq 1$. 
Shannon entropy, Rényi entropy of order $1 + s$, and min entropy are given as

$$H(A|P_A) := - \sum_a P_A(a) \log P_A(a)$$

$$H_{1+s}(A|P_A) := - \frac{1}{s} \log \sum_a P_A(a)^{1+s}$$

$$H_{\min}(A|P_A) := - \log \max_a P_A(a)$$

with $s \in \mathbb{R} \setminus \{0\}$. Then, the function $s \mapsto s H_{1+s}(A|P_A)$ is concave. Since $\sum_a P_A(a)^{1+s} \leq \max_a P_A(a)^s$ for $s > 0$, we have

$$H_{1+s}(A|P_A) \geq H_{\min}(A|P_A). \quad (1)$$

Taking the limit, we obtain the equality

$$\lim_{s \to +\infty} H_{1+s}(A|P_A) = H_{\min}(A|P_A). \quad (2)$$

Now, we introduce two information quantities.

$$D(P_A\|Q_A) := \sum_{a \in \mathcal{A}} P_A(a) \log \frac{P_A(a)}{Q_A(a)} \quad (3)$$

$$\psi(s|P_A\|Q_A) := \log \sum_{a \in \mathcal{A}} P_A(a)^{1+s} Q_A(a)^{-s}. \quad (4)$$

Then, we can show that the map $s \mapsto \psi(s|P_A\|Q_A)$ is convex. When we apply a stochastic matrix $\Lambda$ on $\mathcal{A}$, the information processing inequalities

$$D(\Lambda(P_A)\|\Lambda(Q_A)) \leq D(P_A\|Q_A), \quad \psi(s|\Lambda(P_A)\|\Lambda(Q_A)) \leq \psi(s|P_A\|Q_A) \quad (5)$$

hold for $s \in (0, 1]$. This quantity satisfies the following property.

**Lemma 1:** $\frac{1}{2} \psi(s|P_A\|Q_A)$ is monotonically increasing for $s$ in $(0, \infty)$ and $(-\infty, 0)$.

**Proof:** For $s_1 > s_2 > 0$, the convexity yields that

$$\psi(s_2|P_A\|Q_A) \leq \frac{s_1 - s_2}{s_1} \psi(0|P_A\|Q_A) + \frac{s_2}{s_1} \psi(s_1|P_A\|Q_A). \quad (6)$$

Since $P_A$ is a sub-distribution, $\psi(0|P_A\|Q_A) \leq 0$. Hence,

$$\frac{1}{s_2} \psi(s_2|P_A\|Q_A) \leq \frac{1}{s_1} \psi(s_1|P_A\|Q_A). \quad (7)$$

For $s_2 < s_1 < 0$, the convexity yields that

$$\psi(s_1|P_A\|Q_A) \leq \frac{s_1 - s_2}{-s_2} \psi(0|P_A\|Q_A) + \frac{-s_1}{-s_2} \psi(s_2|P_A\|Q_A). \quad (8)$$

Since $P_A$ is a sub-distribution, $\psi(0|P_A\|Q_A) \leq 0$. Hence,

$$\psi(s_1|P_A\|Q_A) \leq \frac{-s_1}{-s_2} \psi(s_2|P_A\|Q_A), \quad (9)$$

which implies that

$$\frac{1}{s_2} \psi(s_2|P_A\|Q_A) \leq \frac{1}{s_1} \psi(s_1|P_A\|Q_A). \quad (10)$$

Therefore, we obtain the desired argument.

In the following, $P_{\text{mix}, \mathcal{A}}$ expresses the uniform distribution on the set $\mathcal{A}$. Since $H_{1+s}(A|P_A) = \log |\mathcal{A}| - \frac{1}{s} \psi(s|P_A\|P_{\text{mix}, \mathcal{A}})$, applying Lemma 1 we obtain the following lemma.

**Lemma 2:** The quantity $H_{1+s}(A|P_A)$ is monotonically decreasing for $s$ in $(-\infty, 0)$ and $(0, \infty)$. 

2) Case of normalized distributions: When $P_A$ and $Q_A$ are normalized distributions, the following useful properties hold. In this case, since $\lim_{s \to 0} H_{1+s}(A|P_A) = 0$, we have $\lim_{s \to 0} H_{1+s}(A|P_A,\psi) = H(A|P_A)$. Hence, we denote $H(A|P_A,\psi)$ by $H_1(A|P_A)$.

Further, we have $\psi(0|P_A)Q_A = 0$. Hence, the concavity of $s \mapsto \psi(s|P_A)Q_A$ implies $\lim_{s \to 0} \frac{1}{s} \psi(s|P_A)Q_A = D(P_A|Q_A)$. Then, Lemma 3 yields the following lemma.

Lemma 3: When $P_A$ and $Q_A$ are normalized distributions,
\[-\psi(-s|P_A)Q_A \leq sD(P_A|Q_A) \leq \psi(s|P_A)Q_A\]
for $s > 0$.

Applying Lemma 3 we obtain the following lemma.

Lemma 4: When $P_A$ is a normalized distribution,
\[H_{1-s}(A|P_A) \geq H(A|P_A) \geq H_{1+s}(A|P_A)\]
for $s > 0$.

B. Information quantities for composite system

1) Case of joint sub-distribution: Next, we prepare several information quantities for a joint sub-distribution $P_A,\psi$ on subsets $A$ and $\psi$. In the following discussion, the sub-distribution $P_A$ and $P_A,\psi$ is not necessarily normalized, and is assumed to satisfy the condition $\sum_a P_A(a) \leq 1$ or $\sum_{a,e} P_A(a,e) \leq 1$. For the sub-distributions $P_A$ and $P_A,\psi$, we define the normalized distributions $P_A,\psi_{normal}$ and $P_A,\psi_{normal}$ by $P_A,\psi_{normal}(a) := P_A(a)/\sum_a P_A(a)$ and $P_A,\psi_{normal}(a,e) := P_A,\psi(a,e)/\sum_{a,e} P_A,\psi(a,e)$. For a sub-distribution $P_A,\psi$, we define the marginal sub-distribution $P_A$ on $A$ by $P_A(a) := \sum_{e \in E} P_A,\psi(a,e)$. Then, we define the conditional sub-distribution $P_A|\psi$ on $A$ by $P_A|\psi(a) := P_A,\psi(a,e)/P_{\psi_{normal}}(e)$.

The conditional entropy, the conditional Rényi entropies, and the conditional min entropy are given as
\[H(A|E|P_A,\psi) := H(A,E|P_A,\psi) - H(E|P_{\psi_{normal}})\]
\[H_{1+s}(A|E|P_A,\psi) := -\frac{1}{s} \log \sum_{e} P_{\psi_{normal}}(e) \sum_a P_A,\psi(a,e)^{1+s}\]
\[H_{\text{min}}(A|E|P_A,\psi) := -\log \max_{(a,e):P_{\psi_{normal}}(e) > 0} P_A|\psi(a)e\]
with $s \in \mathbb{R} \setminus \{0\}$. Then, the function $s \mapsto sH_{1+s}(A|E|P_A,\psi)$ is concave. Since $\sum_e P_{\psi_{normal}}(e) \sum_a P_A|\psi(a)e^{1+s} \leq \max_{a,e:P_{\psi}(e) > 0} P_A|\psi(a)e^s$ for $s > 0$, we have
\[H_{1+s}(A|E|P_A,\psi) \geq H_{\text{min}}(A|E|P_A,\psi)\]
(13)

Taking the limit, we obtain the equality
\[\lim_{s \to +\infty} H_{1+s}(A|E|P_A,\psi) = H_{\text{min}}(A|E|P_A,\psi)\]
(14)

The conditional Shannon entropy and the conditional Rényi entropy: can be described as follows.
\[H(A|E|P_A,\psi) = \log |A| - D(P_A,\psi|P_{\psi_{normal}})\]
(15)
\[H_{1+s}(A|E|P_A,\psi) = \log |A| - \frac{1}{s} \psi(s|P_A,\psi|P_{\psi_{normal}})\]
(16)
where $P_{\psi_{normal}}$ is the uniform distribution on the set that the random variable $A$ takes values in. When we replace $P_{\psi_{normal}}$
by another normalized distribution $Q_E$ on $E$, we can generalize the above quantities.

$$H(A|E|P_{A,E}||Q_E) := \log |A| - D(P_{A,E}||P_{\text{mix},A} \times Q_E)$$

$$= - \sum_{a,e} P_{A,E}(a,e) \log \frac{P_{A,E}(a,e)}{Q_E(e)}$$

$$= H(A|E|P_{A,E}) + D(P_E||Q_E) \geq H(A|E|P_{A,E}).$$

(17)

$$H_{1+s}(A|E|P_{A,E}||Q_E) := \log |A| - \frac{1}{s} \psi(s)P_{A,E}||P_{\text{mix},A} \times Q_E)$$

$$= \frac{1}{s} \log \sum_{a,e} P_{A,E}(a,e)^{1+s}Q_E(e)^{-s},$$

(18)

$$H_{\min}(A|E|P_{A,E}||Q_E) := -\log \max_{(a,e); Q_E(e) > 0} \frac{P_{A,E}(a,e)}{Q_E(e)}.$$

The quantity $H_{1+s}(A|E|P_{A,E}||Q_E)$ can be regarded as a generalization of $H_2(A|E|P_{A,E}||Q_E)$ by Renner [16]. Further, similar to Lemma 2 applying Lemma 1, we obtain the following lemma.

Lemma 5: The quantity $H_{1+s}(A|E|P_{A,E}||Q_E)$ are monotonically decreasing for $s$ in $(0, \infty)$ and $(-\infty, 0)$.

Similar to the case of $H_{1+s}(A|P_A)$, since $\sum_e P_{E,\text{normal}}(e) \sum_a P_{A}(a|e)P_{A}(a,e)^sQ_E(e)^{-s} \leq \max_{a,e}P_{E}(e) > 0 P_{A,E}(a,e)^sQ_E(e)^{-s}$ for $s > 0$, we have

$$H_{1+s}(A|E|P_{A,E}||Q_E) \geq H_{\min}(A|E|P_{A,E}||Q_E).$$

(19)

Taking the limit, we obtain the equality

$$\lim_{s \to \pm \infty} H_{1+s}(A|E|P_{A,E}||Q_E) = H_{\min}(A|E|P_{A,E}||Q_E).$$

(20)

Due to 5, when we apply an operation $\Lambda$ on $E$, it does not act on the system $A$. Then,

$$H(A|E|\Lambda(P_{A,E}||\Lambda(Q_E)) \geq H(A|E|P_{A,E}||Q_E)$$

$$H_{1+s}(A|E|\Lambda(P_{A,E}||\Lambda(Q_E)) \geq H_{1+s}(A|E|P_{A,E}||Q_E).$$

(21)

(22)

In particular, the inequalities

$$H(A|E|\Lambda(P_{A,E})) \geq H(A|E|P_{A,E})$$

$$H_{1+s}(A|E|\Lambda(P_{A,E})) \geq H_{1+s}(A|E|P_{A,E})$$

hold. Conversely, when we apply the function $f$ to the random number $a \in A$, we have

$$H(f(A)|E|P_{A,E}) \leq H(A|E|P_{A,E}).$$

(23)

Now, we introduce another kind of conditional Rényi entropy for a joint normalized distribution $P_{A,E}$ as

$$H_{1+s}^G(A|E|P_{A,E}) := -\frac{1}{s} \log \sum_e (\sum_a P_{A,E}(a,e)^{1+s})^{1/s}$$

This quantity can be expressed as

$$H_{1+s}^G(A|E|P_{A,E}) = -\frac{1}{s} \phi\left(\frac{s}{1+s}\right) A|E|P_{A,E}$$

by using the Gallager-type function [12]:

$$\phi(s|A|E|P_{A,E}) := \log \sum_e \left(\sum_a P_{A,E}(a,e)^{1/(1-s)}\right)^{1-s}$$

$$= \log \sum_e P_E(e) \left(\sum_a P_{A|E}(a|e)^{1/(1-s)}\right)^{1-s}.$$

Lemma 6: For $s \in [-1, 1] \setminus \{0\}$, a joint sub-distribution $P_{A,E}$ satisfies the relation

$$H_{1+s}(A|E|P_{A,E}) \geq H_{1+s}^G(A|E|P_{A,E}).$$

(25)
The equality holds only when $P_{A|E=e}$ is uniform distribution for all $e \in \mathcal{E}$.

The proof of Lemma 7 is given in Appendix B. The opposite type inequality also holds as follows.

**Lemma 7:** A joint sub-distribution $P_{A,E}$ satisfies the relation

$$\max_{Q_E} H_{1+s}(A|E)|_{P_{A,E}} = H^G_{1+s}(A|E)|_{P_{A,E}}.$$  \hspace{1cm} (26)

for $s \in [-1, \infty) \backslash \{0\}$. The maximum of LHS can be realized when $Q_E(e) = \left( \sum_a P_{A,E}(a,e)^{1+s} \right)^{1/(1+s)} / \sum_e \left( \sum_a P_{A,E}(a,e)^{1+s} \right)^{1/(1+s)}$. The proof of Lemma 7 is given in Appendix C.

**Remark 1:** Iwamoto and Shikata [63] discussed conditional Rényi entropies in the different notations. They denote $H_{1+s}(A|E)|_{P_{A,E}}$ by $R^H_{1+s}(A|E)$ and $H^G_{1+s}(A|E)|_{P_{A,E}}$ by $R^G_{1+s}(A|E)$. They also compare these with other conditional Rényi entropies. Muller-Lennert et al [64] denoted $H^G_{1+s}(A|E)|_{P_{A,E}}$ by $H_{1+s}(A|E)$ in the quantum setting. Iwamoto and Shikata [63] pointed out that these quantities do not satisfy the chain rule. Instead, Muller-Lennert et al [64, Proposition 7] showed the inequality $H^G_{1+s}(A|E, E')|_{P_{A,E,E'}} \geq H_{1+s}(A, E'|E)|_{P_{A,E,E'}} - \log |\mathcal{E}'|$ for $s \in (-1, \infty)$. Also, the paper [65, Corollary 77] shows the inequality $H_{1+s(1-s)}(A|E)|_{P_{A,E,E'}} \geq H_{1+s}(A, E'|E)|_{P_{A,E,E'}} - \log |\mathcal{E}'|$ for $s \in [0, 1)$.

**C. Criteria for secret random numbers**

1) **Case of joint sub-distribution:** Next, we introduce criteria for the amount of the information leaked from the secret random number $A$ to $E$ for joint sub-distribution $P_{A,E}$. Using the $\ell_1$ norm, we can evaluate the secrecy for the state $P_{A,E}$ as follows:

$$d_1(A|E)|_{P_{A,E}} := \|P_{A,E} - P_A \times P_E\|_1.$$  \hspace{1cm} (31)

Taking into account the randomness, Renner [16] employed the $L_1$ distinguishability criteria for security of the secret random number $A$:

$$d_1'(A|E)|_{P_{A,E}} := \|P_{A,E} - P_{\text{mix},A} \times P_E\|_1.$$  \hspace{1cm} (32)

which can be regarded as the difference between the true sub-distribution $P_{A,E}$ and the ideal sub-distribution $P_{\text{mix},A} \times P_E$. It is known that the quantity is universally composable [28].
Renner\cite{16} defined the conditional $L_2$-distance from uniform of $P_{A,E}$ relative to a distribution $Q_E$ on $\mathcal{E}$:

\[
d_2(A|E|P_{A,E}\|Q_E) = \sum_{a,e} (P_{A,E}(a,e) - P_{\text{mix},A}(a)P_E(e))^2 Q_E(e)^{-1}
\]

\[
= \sum_{a,e} P_{A,E}(a,e)^2 Q_E(e)^{-1} - \frac{1}{|A|} \sum_{e} P_E(e)^2 Q_E(e)^{-1}
\]

\[
= e^{-H_2(A|E|P_{A,E}\|Q_E)} - \frac{1}{|A|} e^{-\psi(1)P_A\|Q_E)}.
\]

Using this value and a normalized distribution $Q_E$, we can evaluate $d'_1(A|E|P_{A,E})$ as follows \cite{16} Lemma 5.2.3):

\[
d'_1(A|E|P_{A,E}) \leq \sqrt{|A|}d_2(A|E|P_{A,E}\|Q_E).
\]

2) Case of joint normalized distribution: In the remaining part of this subsection, we assume that $P_{A,E}$ is a normalized distribution. The correlation between $A$ and $E$ can be evaluated by the mutual information

\[
I(A : E|P_{A,E}) := D(P_{A,E}\|P_A \times P_E).
\]

By using the uniform distribution $P_{\text{mix},A}$ on $\mathcal{A}$, the mutual information can be modified to

\[
I'(A|E|P_{A,E}) := D(P_{A,E}\|P_{\text{mix},A} \times P_E),
\]

which is called the modified mutual information and satisfies

\[
I'(A|E|P_{A,E}) = I(A : E|P_{A,E}) + D(P_A\|P_{\text{mix},A})
\]

and

\[
H(A|E|P_{A,E}) = -I'(A|E|P_{A,E}) + \log |A|.
\]

Indeed, the quantity $I(A : E|P_{A,E})$ represents the amount of information leaked by $E$, and the remaining quantity $D(P_A\|P_{\text{mix},A})$ describes the difference of the random number $A$ from the uniform random number. So, if the quantity $I'(A|E|P_{A,E})$ is small, we can conclude that the random number $A$ has less correlation with $E$ and is close to the uniform random number. As shown in Appendix A when we assume several natural constraints for the security criterion, it is restricted to the modified mutual information $I'(A|E|P_{A,E})$.

In particular, if the quantity $I'(A|E|P_{A,E})$ goes to zero, the mutual information $I(A : E|P_{A,E})$ goes to zero, and the marginal distribution $P_A$ goes to the uniform distribution. Hence, we can adopt the quantity $I'(A|E|P_{A,E})$ as a criterion for qualifying the secret random number.

Using Pinsker inequality, we obtain

\[
d_1(A|E|P_{A,E})^2 \leq 2I(A|E|P_{A,E})
\]

\[
d'_1(A|E|P_{A,E})^2 \leq 2I'(A|E|P_{A,E}).
\]

Conversely, we can evaluate $I(A : E|P_{A,E})$ and $I'(A|E|P_{A,E})$ by using $d_1(A|E|P_{A,E})$ and $d'_1(A|E|P_{A,E})$ in the following way. Applying the Fannes inequality, we obtain

\[
0 \leq I(A : E|P_{A,E}) = H(A|P_A) + H(E|P_E) - H(A, E|P_{A,E})
\]

\[
= H(A, E|P_A \times P_E) - H(A, E|P_{A,E})
\]

\[
= \sum_a P_A(a)H(E|P_E) - H(E|P_{E|A=a})
\]

\[
\leq \sum_a P_A(a)\eta(\|P_{E|A=a} - P_E\|_1, \log |E|)
\]

\[
= \eta(\|P_{E,A} - P_A \times P_E\|_1, \log |E|)
\]

\[
= \eta(d_1(A|E|P_{A,E}), \log |E|),
\]
where \( \eta(x, y) := -x \log x + xy \). Similarly, we obtain
\[
0 \leq I'(A|E|P_{A,E}) = H(A|P_{\text{mix},A}) + H(E|P_{E}) - H(A, E|P_{A,E})
\]
\[
= H(A, E|P_{\text{mix},A} \times P_{E}) - H(A, E|P_{A,E})
\]
\[
= \sum_{e} P_{E}(e)(H(A|P_{\text{mix},A}) - H(A|P_{A|E=e}))
\]
\[
\leq \sum_{e} P_{E}(e)(\|P_{\text{mix},A} - H(A|P_{A|E=e})\|_1, \log |A|)
\]
\[
\leq \eta(\|P_{\text{mix},A} \times P_{E} - P_{A,E}\|_1, \log |A|)
\]
\[
= \eta(d'_1(A|E|P_{A,E}), \log |A|).
\] (41)

### III. Random Hash functions

#### A. General random hash functions

In this section, we focus on a random function \( f_X \) from \( A \) to \( B \), where \( X \) is a random variable identifying the function \( f_X \). In this case, the total information of Eve’s system is written as \( (E, X) \). Then, by using \( P_{f_X(A),E}(b,e,x) := \sum_{a \in f^{-1}_X(b)} P_{A,E}(a,e) P_X(x) \), the \( L_1 \) distinguishability criterion is written as
\[
d'_1( f_X(A)|E, X|P_{f_X(A),E})
\]
\[
= \|P_{f_X(A),E} - P_{\text{mix},B \times P_{E}}\|_1
\]
\[
= \sum_{x} P_{X}(x)\|P_{f_X=A}(E,E) - P_{\text{mix},B \times P_{E}}\|_1
\]
\[
= \text{Ex} \|P_{f_X(A),E} - P_{\text{mix},B \times P_{E}}\|_1.
\] (42)

Also, the modified mutual information is written as
\[
I'(f_X(A)|E, X|P_{f_X(A),E})
\]
\[
= D(P_{f_X(A),E}, P_{\text{mix},B \times P_{E}})
\]
\[
= \sum_{x} P_{X}(x)D(P_{f_X=x(A),E,E}||P_{\text{mix},B \times P_{E}})
\]
\[
= \text{Ex} D(P_{f_X(A),E,}\|P_{\text{mix},B \times P_{E}}).
\] (43)

We say that a random function \( f_X \) is \( \varepsilon \)-almost universal [1], [2], [14], if, for any pair of different inputs \( a_1, a_2 \), the collision probability of their outputs is upper bounded as
\[
\text{Pr}[f_X(a_1) = f_X(a_2)] \leq \frac{\varepsilon}{|B|}.
\] (44)

The parameter \( \varepsilon \) appearing in (44) is shown to be confined in the region
\[
\varepsilon \geq \frac{|A| - |B|}{|A| - 1},
\] (45)

and in particular, a random function \( f_X \) with \( \varepsilon = 1 \) is simply called a universal function.

Two important examples of universal hash function are the Toeplitz matrices (see, e.g., [3]), and multiplications over a finite field (see, e.g., [1], [4]). A modified form of the Toeplitz matrices is also shown to be universal, which is given by a concatenation \( (X, I) \) of the Toeplitz matrix \( X \) and the identity matrix \( I \) [13]. The (modified) Toeplitz matrices are particularly useful in practice, because there exists an efficient multiplication algorithm using the fast Fourier transform algorithm with complexity \( O(n \log n) \) (see, e.g., [5]).

The following proposition holds for any universal function.

**Proposition 10 (Renner [17] Lemma 5.4.3):** Given any joint sub-distribution \( P_{A,E} \) on \( A \times E \) and any normalized distribution \( Q_E \) on \( E \), any universal hash function \( f_X \) from \( A \) to \( M := \{1, \ldots, M\} \) satisfies
\[
\text{Ex} d_2(f_X(A)|E|P_{A,E}) \leq e^{-H_2(A|E|P_{A,E})}. \quad (46)
\]

More precisely, the inequality
\[
\text{Ex} e^{-H_2(f_X(A)|E|P_{A,E})} \leq e^{-H_2(A|E|P_{A,E})} + \frac{1}{M} e^{\psi(1)|P_{E}|Q_E}
\] (47)
holds.
B. Ensemble of linear hash functions

Tsurumaru and Hayashi[14] focus on linear functions over the finite field $\mathbb{F}_2$. Now, we treat the case of linear functions over a finite field $\mathbb{F}_q$, where $q$ is a power of a prime number $p$. We assume that sets $A$, $B$ are $\mathbb{F}_q^n$, $\mathbb{F}_q^m$ respectively with $n \geq m$, and $f$ are linear functions over $\mathbb{F}_q$. Note that, in this case, there is a kernel $C$ corresponding to a given linear function $f$, which is a vector space of the dimension $n - m$ or more. Conversely, when given a vector subspace $C \subset \mathbb{F}_q^n$ of the dimension $n - m$ or more, we can always construct a linear function

$$f_C : \mathbb{F}_q^n \to \mathbb{F}_q^n / C \equiv \mathbb{F}_q^l, \quad l \leq m.$$  

(48)

That is, we can always identify a linear hash function $f_C$ and a code $C$.

When $C_X = \text{Ker} f_X$, the definition of $\varepsilon$-universal$_2$ function (44) takes the form

$$\forall x \in \mathbb{F}_q^n \setminus \{0\}, \quad \Pr[f_X(x) = 0] \leq q^{-m}\varepsilon,$$  

(49)

which is equivalent with

$$\forall x \in \mathbb{F}_q^n \setminus \{0\}, \quad \Pr[x \in C_X] \leq q^{-m}\varepsilon.$$  

(50)

This shows that the kernel $C_X$ contains sufficient information for determining if a random function $f_X$ is $\varepsilon$-almost universal$_2$ or not.

For a given random code $C_X$, we define its minimum (respectively, maximum) dimension as $t_{\text{min}} := \min_X \dim C_X$ (respectively, $t_{\text{max}} := \max_{x \in I} \dim C_X$). Then, we say that a linear random code $C_X$ of minimum (or maximum) dimension $t$ is an $\varepsilon$-almost universal$_2$ code if the following condition is satisfied

$$\forall x \in \mathbb{F}_q^n \setminus \{0\}, \quad \Pr[x \in C_X] \leq q^{t-n}\varepsilon.$$  

(51)

In particular, if $\varepsilon = 1$, we call $C_X$ a universal$_2$ code.

C. Dual universality of a random code

Based on Tsurumaru and Hayashi[14], we define several variations of the universality of a error-correcting random code and the linear function as follows. First, we define the dual random code $C_X^\perp$ of a given linear random code $C_X$ as the dual code of $C_X$. We also introduce the notion of dual universality as follows. We say that a random code $C_X$ in $\mathbb{F}_q^n$ is $\varepsilon$-almost dual universal$_2$ with minimum dimension $t$ (with maximum dimension $t$), if the dual random code $C_X^\perp$ is $\varepsilon$-almost universal$_2$ with maximum dimension $n - t$ (with minimum dimension $n - t$). Hence, we say that a linear random function $f_X$ from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$ is $\varepsilon$-almost dual universal$_2$, if the kernels $C_X$ of $f_X$ forms an $\varepsilon$-almost dual universal$_2$ code with minimum dimension $n - m$. This condition is equivalent with the condition that the linear space spanned by the generating matrix of $f_X$ forms an $\varepsilon$-almost universal$_2$ random code with maximum dimension $m$. An explicit example of a dual universal$_2$ function (with $\varepsilon = 1$) can be given by the modified Toeplitz matrix mentioned earlier [11], i.e., a concatenation $(X, I)$ of the Toeplitz matrix $X$ and the identity matrix $I$. The modified Toeplitz matrix requires $n - 1$ bits of random seeds $R$. This example is particularly useful in practice because it is both universal$_2$ and dual universal$_2$, and also because there exists an efficient algorithm with complexity $O(n \log n)$. When the random variable $R$ is not the uniform random number, the modified Toeplitz matrix is $q^{n-1}e^{-H_{\text{min}}(R)}$-almost dual universal$_2$, as shown in [34]. Therefore, we can evaluate the security of the modified Toeplitz matrix even with non-uniform random seeds. With these preliminaries, we can present the following theorem in the non-quantum setting of [14 Corollary 2]:

Proposition 11: An $\varepsilon$-almost universal$_2$ surjective linear random hash function $f_X$ from $\mathbb{F}_q^n$ to $\mathbb{F}_q^m$ is $q(1-q^m\varepsilon) + (\varepsilon-1)q^{n-m}$-almost dual universal$_2$ linear random hash function.

As a special case, we obtain the following.

Corollary 12: Any universal$_2$ linear random function $f_X$ over a finite filed $\mathbb{F}_q$ is a $q$-almost dual universal$_2$ function.

D. $\delta$-biased ensemble

Next, according to Dodis and Smith[9], we introduce $\delta$-biased ensemble of random variables $W_X$ on a vector space over a general finite field $\mathbb{F}_q$, where $q$ is the power of the prime $p$. First, we fix a non-degenerate bilinear form $(\cdot, \cdot)$ from $\mathbb{F}_q^2$ to $\mathbb{F}_p$. Then, we define $(x \cdot y) \in \mathbb{F}_p$ for $x, y \in \mathbb{F}_q^n$ as $(x \cdot y) := \sum_{j=1}^n x_j \cdot y_j$. For a given $\delta > 0$, an ensemble of random variables $\{W_X\}$ on $\mathbb{F}_q^n$ is called $\delta$-biased when the inequality

$$\mathbb{E} |\mathbb{E}_{W_X} \omega_p^{(x \cdot W_X)}|^2 \leq \delta^2$$  

holds for any $x \neq 0 \in \mathbb{F}_q^n$, where $\omega_p := e^{2\pi i/p}$.

We denote the random variable subject to the uniform distribution on a code $C \in \mathbb{F}_q^n$, by $W_C$. Then,

$$\mathbb{E}_{W_C} \omega_p^{(x \cdot W_C)} = \begin{cases} 0 & \text{if } x \notin C^\perp, \\ 1 & \text{if } x \in C^\perp. \end{cases}$$  

(53)
Using the above relation, as is suggested in [9] Case 2, we obtain the following lemma.

**Lemma 13:** When a random code $C_X$ in $\mathbb{F}_q^n$ is $\varepsilon$-almost dual universal with minimum dimension $t$, the ensemble of random variables $W_{C_X}$ in $\mathbb{F}_q^n$ is $\sqrt{\varepsilon q^{-t}}$-biased.

**Proof:** $C_X$ is $\varepsilon$-almost universal with maximum dimension $n-t$ in $\mathbb{F}_q^n$. Hence, for any $x \in \mathbb{F}_q^n$, the probability $\Pr\{x \in C_X^\perp\}$ is less than $\varepsilon q^{-t}$. Thus, (53) guarantees that the ensemble of random variables $W_{C_X}$ in $\mathbb{F}_q^n$ is $\sqrt{\varepsilon q^{-t}}$-biased. □

In the following, we treat the case of $A = \mathbb{F}_q^n$. Given a joint sub-distribution $P_{A,E}$ on $A \times E$ and a normalized distribution $P_W$ on $A$, we define another joint sub-distribution $P_{A,E} \ast P_W(a,e) := \sum_w P_W(w)P_{A,E}(a-w,e)$. Using these concepts, Dodis and Smith[9] evaluated the average of $d_2(A|E|P_{A,E} \ast P_W(x)|Q_E)$ as follows.

**Proposition 14 (52 Lemma 4):** For any joint sub-distribution $P_{A,E}$ on $A \times E$ and any normalized distribution $Q_E$ on $E$, a $\delta$-biased ensemble of random variables $\{W_X\}$ on $A = \mathbb{F}_q^n$ satisfies

$$E_Xd_2(A|E|P_{A,E} \ast P_W(x)|Q_E) \leq \delta^2 e^{-H_2(A|E|P_{A,E}|Q_E)}.$$  

(54)

More precisely,

$$E_Xd_2(A|E|P_{A,E} \ast P_W(x)|Q_E) \leq \delta^2 d_2(A|E|P_{A,E}|Q_E).$$  

(55)

The original proof by Dodis and Smith[9] discussed in the case with $q = 2$. Fehr and Schaffner[10] extended this lemma to the quantum setting in the case with $q = 2$. Their proof is based on Fourier analysis and easy to understand. The proof with a general prime power $q$ is given in Appendix [3] by generalizing the idea by Fehr and Schaffner[10]. Dodis and Smith[9] Lemma 6] also considered the case with a general prime power $q$. They did not explicitly give Proposition[14] and the definition [52] with the general case.

**Lemma 15:** Given a joint sub-distribution $P_{A,E}$ on $A \times E$ and a normalized distribution $Q_E$ on $E$. When $C_X$ is an $\varepsilon$-almost dual universal$\varepsilon_2$ code with minimum dimension $t$, the random hash function $f_{C_X}$ satisfies

$$E_Xd_2(f_{C_X}(A)|E|P_{A,E} \ast P_W(x)|Q_E) \leq \varepsilon e^{-H_2(A|E|P_{A,E}|Q_E)}.$$  

(56)

More precisely,

$$E_Xe^{-H_2(f_{C_X}(A)|E|P_{A,E}|Q_E)} \leq \varepsilon e^{-H_2(A|E|P_{A,E}|Q_E)} + \frac{1}{q^{n-t}}e^{H_1(P_E|Q_E)}.$$  

(57)

In other words, an $\varepsilon$-almost dual universal$\varepsilon_2$ function $f_X$ from $\mathbb{F}_2^n$ to $\mathbb{F}_2^{n-t}$ satisfies (56) and (57).

**Lemma 15** essentially coincides with Proposition[14] However, the concept “$\delta$-biased” does not concern a linear random hash function while the concept “$\varepsilon$-almost dual universality$\varepsilon_2$” does because the former is defined for the ensemble of random variables. That is, the latter is a generalization of a universal$\varepsilon_2$ linear hash function while the former does not. Hence, Proposition[14] cannot directly provide the performance of a linear random hash function. In contrast, Lemma[15] gives how the privacy amplification by a linear hash function decreases the leaked information. Therefore, in the following section, using Lemma[15] we treat the exponential decreasing rate when we apply the privacy amplification by an $\varepsilon$-almost dual universal$\varepsilon_2$ linear hash function.

**Proof:** Due to Lemma[13] and (54), we obtain

$$E_Xd_2(A|E|P_{A,E} \ast P_W(x)|Q_E) \leq \varepsilon q^{-t}e^{-H_2(A|E|P_{A,E}|Q_E)}.$$  

(58)

Denoting the quotient class with respect to the subspace $C$ with the representative $a \in A$ by $[a]$, we obtain

$$P_{A,E} \ast P_{W_C}(a,e) = \sum_{w \in C} q^{-t}P_{A,E}(a-w,e) = q^{-t}P_{A,E}([a],e).$$

Now, we focus on the relation $A \cong A/C \times C \cong f_{C}(A) \times C$. Then,

$$P_{A,E} \ast P_{W_{C_X}}(b,w,e) = q^{-t}P_{f_{C}(A),E}(b,e).$$

Thus,

$$d_2(A|E|P_{A,E} \ast P_W(x)|Q_E) = q^{-t}d_2(f_{C}(A)|E|P_{f_{C}(A),E}|Q_E) = q^{-t}d_2(f_{C}(A)|E|P_{A,E}|Q_E).$$  

(59)
Therefore, \((68)\) implies
\[
E_X q^{-f} d_2(f_{C_X}(A)|E|P_{A,E}\|Q_E)
\leq e^{-\varepsilon f} e^{-H_2(A|E)P_{A,E}\|Q_E)}
\]
which implies \((56)\).

Similarly, Lemma 13 \((55)\), and \((59)\) imply that
\[
E_X q^{-f} d_2(f_{C_X}(A)|E|P_{A,E}\|Q_E)
\leq e^{-\varepsilon f} e^{-H_2(A|E)P_{A,E}\|Q_E)}.
\]
Since \(E_X d_2(f_{C_X}(A)|E|P_{A,E}\|Q_E) = E_X e^{-H_2(f_{C_X}(A)|E|P_{A,E}\|Q_E) - \frac{1}{q_0} e^{\psi(1)P_E\|Q_E)}\), we have \((57)\).

IV. SECURITY BOUNDS WITH RÉNYI ENTROPY OF ORDER 2 AND MIN ENTROPY

Firstly, we consider the secure key generation problem from a common random number \(A \in A\) which has been partially eavesdropped as an information by Eve. For this problem, it is assumed that Alice and Bob share a common random number \(A \in A\), and Eve has a random number \(E\) correlated with the random number \(A\), whose distribution is \(P_E\). The task is to extract a common random number \(f(A)\) from the random number \(A \in A\), which is almost independent of Eve’s quantum state. Here, Alice and Bob are only allowed to apply the same function \(f\) to the common random number \(A \in A\). Now, we focus on the random function \(f_X\) from \(A\) to \(M = \{1, \ldots, M\}\), where \(X\) denotes a random variable describing the stochastic behavior of the function \(f_X\).

Renner\([\text{16}]\) Lemma 5.2.3 essentially evaluated \(E_X d'_1(f_X(A)|E|P_{A,E})\) by using \(E_X d_2(f_X(A)|E|P_{A,E}\|Q_E)\) as follows.

**Lemma 16:** When a state \(Q_E\) is a normalized distribution on \(E\), any random hash function \(f_X\) from \(A\) to \(\{1, \ldots, M\}\) satisfies
\[
E_X d'_1(f_X(A)|E|P_{A,E}) \leq M^{\frac{1}{2}} \sqrt{E_X d_2(f_X(A)|E|P_{A,E}\|Q_E)}.
\]
Further, the inequalities used in proof of Renner\([\text{16}]\) Corollary 5.6.1 imply that
\[
E_X d'_1(f_X(A)|E|P_{A,E}) \leq 2\|P_{A,E} - P'_{A,E}\|_1 + E_X d'_1(f_X(A)|E|P'_{A,E})
\leq 2\|P_{A,E} - P'_{A,E}\|_1 + M^{\frac{1}{2}} \sqrt{E_X d_2(f_X(A)|E|P'_{A,E}\|Q_E)}.
\]

Applying the same discussion to Shannon entropy, we can evaluate the average of the modified mutual information criterion by using \(E_X d_2(f_X(A)|E|P_{A,E}\|Q_E)\) as follows.

**Lemma 17:** Assume that \(P_{A,E}\) is a normalized distribution on \(A \times E\). Any random hash function \(f_X\) from \(A\) to \(M = \{1, \ldots, M\}\) satisfies
\[
E_X I'(f_X(A)|E|P_{A,E})
\leq \log(1 + ME_X d_2(f_X(A)|E|P_{A,E})\|P_E) \quad (60)
\]
\[
\leq ME_X d_2(f_X(A)|E|P_{A,E}\|P_E) \quad (61)
\]
Further, when a sub-distribution \(P'_{A,E}\) satisfies \(P'_{E}(e) \leq P_{E}(e)\) for any \(e \in E\) (we simplify this condition to \(P'_{E} \leq P_{E}\)), we obtain
\[
E_X I'(f_X(A)|E|P_{A,E})
\leq \eta(\|P_{A,E} - P'_{A,E}\|_1, \log M)
+ \log(1 + ME_X d_2(f_X(A)|E|P'_{A,E}\|P_E)) \quad (62)
\]
\[
\leq \eta(\|P_{A,E} - P'_{A,E}\|_1, \log M)
+ ME_X d_2(f_X(A)|E|P'_{A,E}\|P_E), \quad (63)
\]
where \(\eta(x, y) := xy - x \log x\).

**Proof:** The inequality \(\psi(1)P'_{E}\|P_E) \leq 0\) holds due to the condition \(P'_{E}(e) \leq P_{E}(e)\). Since
\[
d_2(f_X(A)|E|P'_{A,E}\|P_E)
= e^{-H_2(f_X(A)|E|P'_{A,E}\|P_E)} - \frac{1}{M} e^{\psi(1)P'_{E}\|P_E)}
\geq e^{-H_2(f_X(A)|E|P'_{A,E}\|P_E)} - \frac{1}{M},
\]
(64)
we have
\[ e^{-H_2(f_X(A)|E|P'_{A,E}||P_E)} \leq d_2(f_X(A)|E|P'_{A,E}||P_E) + \frac{1}{M}. \]

Taking the logarithm, we obtain
\[
\begin{align*}
- \log M + \log(1 + Md_2(f_X(A)|E|P'_{A,E}||P_E)) \\
\geq - H_2(f_X(A)|E|P'_{A,E}||P_E) \geq - H(f_X(A)|E|P'_{A,E}||P_E).
\end{align*}
\]

Substituting \( P_{A,E} \) to \( P'_{A,E} \), we obtain
\[
H(f_X(A)|E|P_{A,E}||P_E) = H(f_X(A)|E|P_{A,E})
\]
and using the concavity of functions \( X \) yield (61).

Since the function \( x \mapsto \log(1 + x) \) is concave, we obtain
\[
E_X I'(f_X(A)|E|P_{A,E})
\]
which implies (60). The inequality \( \log(1 + x) \leq x \) and (60) yield (61).

Due to Fannes inequality, the normalized distribution \( P_{A|E=e}(a) := \frac{P_{A,E}(a,e)}{P_E(e)} \) and the sub-distribution \( P'_{A|E=e}(a) := \frac{P'_{A,E}(a,e)}{P_E(e)} \) satisfy
\[
|H(f_X(A)|P_{A|E=e}) - H(f_X(A)|P'_{A|E=e})| \leq \eta(||P_{A|E=e} - P'_{A|E=e}||_1, \log M).
\]

Since \( \sum_e P_E(e)||P_{A|E=e} - P'_{A|E=e}||_1 = ||P_{A,E} - P'_{A,E}||_1 \), taking the average under the distribution \( P_E \), we obtain
\[
\begin{align*}
|H(f_X(A)|E|P_{A,E}||P_E) - H(f_X(A)|E|P'_{A,E}||P_E)| \\
= |\sum_e P_E(e)(H(f_X(A)|P_{A|E=e}) - H(f_X(A)|P'_{A|E=e}))|
\end{align*}
\]
\[
\leq \sum_e P_E(e)||H(f_X(A)|P_{A|E=e}) - H(f_X(A)|P'_{A|E=e})||
\]
\[
\leq \eta(\sum_e P_E(e)||P_{A|E=e} - P'_{A|E=e}||_1, \log M)
\]
\[
= \eta(||P_{A,E} - P'_{A,E}||_1, \log M).
\]

Therefore, using (67) and (65), we obtain
\[
I'(f_X(A)|E|P_{A,E})
\]
\[
= \log M - H(f_X(A)|E|P_{A,E}||P_E)
\]
\[
\leq \eta(||P_{A,E} - P'_{A,E}||_1, \log M)
\]
\[
+ \log M - H(f_X(A)|E|P'_{A,E}||P_E)
\]
\[
\leq \eta(||P_{A,E} - P'_{A,E}||_1, \log M)
\]
\[
+ \log(1 + Md_2(f_X(A)|E|P'_{A,E}||P_E)).
\]

Taking the expectation of \( X \) and using the concavity of functions \( x \mapsto \eta(x, \log M) \) and \( x \mapsto \log(1 + x) \), we obtain (62). The inequality \( \log(1 + x) \leq x \) yields (63). In this proof, the condition \( P_E(e)' \leq P_E(e) \) is crucial because Inequality (64) cannot be shown without this condition.
Now, we evaluate the security by combining Lemmas 15, 16, and 17. For this purpose, we introduce the quantities:

\[\Delta_{d,2}(M, \varepsilon | P_{A,E}) := \min_{Q_E} \min_{P_{A,E}} 2\|P_{A,E} - P'_{A,E}\|_1 + \sqrt{e}\max_{\eta} H_2(A|E|P_{A,E}^\prime|Q_E)}\]

\[= \min_{Q_E} \min_{\varepsilon > 0} 2\varepsilon + \sqrt{e}\max_{\eta} H_2(A|E|P_{A,E}^\prime|Q_E)}\]

\[= \min_{Q_E} \min_{\varepsilon > 0} 2\varepsilon + \sqrt{e}\max_{\eta} H_2(A|E|P_{A,E}^\prime|Q_E)}\]

\[\Delta_{I,2}(M, \varepsilon | P_{A,E}) := \min_{P'_{A,E}} \eta(\|P_{A,E} - P'_{A,E}\|_1, \log M) + \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

\[= \min_{\varepsilon > 0} \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

\[= \min_{\varepsilon > 0} \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

where

\[H_2^\varepsilon(A|E|P_{A,E}) := \max_{P_{A,E}} H_2(A|E|P_{A,E}^\prime|Q_E)\]

\[H_2^\varepsilon(A|E|P_{A,E}) := \max_{P_{A,E}} H_2(A|E|P_{A,E}^\prime|Q_E)\]

Note that \(H_2^\varepsilon(A|E|P_{A,E})\) is different from \(H_2(A|E|P_{A,E})\) because the definition of \(H_2^\varepsilon(A|E|P_{A,E})\) has additional constraints for \(P'_{A,E}\). Then, we can evaluate the averages of both security criteria under the \(\varepsilon\)-almost dual universal condition.

**Theorem 18:** Assume that \(Q_E\) is a normalized distribution on \(E\), \(P_{A,E}\) is a sub-distribution on \(A \times E\), and a linear random hash function \(f_X\) from \(A\) to \(M = \{1, \ldots, M\}\) is \(\varepsilon\)-almost dual universal. Then, the random hash function \(f_X\) satisfies

\[E_X d_1^\varepsilon(f_X(A)|E|P_{A,E}) \leq \sqrt{e}\max_{\eta} H_2(A|E|P_{A,E}^\prime|Q_E)}\]

\[E_X I^\varepsilon(f_X(A)|E|P_{A,E}) \leq \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

When \(P_{A,E}\) is a normalized joint distribution, it satisfies

\[E_X I^\varepsilon(f_X(A)|E|P_{A,E}) \leq \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

While the same evaluations for the \(L_1\) distinguishability criterion under the universal condition has been shown in Renner [16, Corollary 5.6.1], those for the modified mutual information criterion have not been shown even under the universal condition. All of the above evaluations under the \(\varepsilon\)-almost dual universal condition have not been discussed in Renner.

Since the function \(x \mapsto \eta(x, y)\) is concave, combining Inequality (31), we obtain the following corollary.

**Corollary 19:** When a linear random hash function \(f_X\) from \(A\) to \(M = \{1, \ldots, M\}\) is \(\varepsilon\)-almost dual universal, any joint sub-distribution \(P_{A,E}\) on \(A \times E\) satisfies

\[E_X I^\varepsilon(f_X(A)|E|P_{A,E}) \leq \varepsilon \max_{\eta} H_2(A|E|P_{A,E}^\prime|P_E)\]

for \(s \in (0, 1/2]\).

Since the function \(x \mapsto \sqrt{x}\) is concave, combining Inequality (50), we obtain the following corollary.

**Corollary 20:** When a linear random hash function \(f_X\) from \(A\) to \(M = \{1, \ldots, M\}\) is \(\varepsilon\)-almost dual universal, any joint normalized distribution \(P_{A,E}\) on \(A \times E\) satisfies

\[E_X I^\varepsilon(f_X(A)|E|P_{A,E}) \leq \sqrt{2\Delta_{I,2}(M, \varepsilon | P_{A,E})}\]

for \(s \in (0, 1/2]\).

Further, in the case of the universal condition, Renner [16, Corollary 5.6.1] proposed to replace \(H_2(A|E|P_{A,E}^\prime|Q_E)\) by the min entropy \(H_{\min}(A|E|P_{A,E}^\prime|Q_E)\) because \(H_2(A|E|P_{A,E}^\prime|Q_E) \geq H_{\min}(A|E|P_{A,E}^\prime|Q_E)\). Based on \(H_{\min}(A|E|P_{A,E}^\prime|Q_E)\), Renner [16] introduced \(\varepsilon\)-smooth min entropy as

\[H_{\min}^\varepsilon(A|E|P_{A,E}) := \max_{\|P_{A,E} - P_{A,E}^\prime\|_1 \leq \varepsilon} H_{\min}(A|E|P_{A,E}^\prime|Q_E)\]

For the evaluation of \(E_X I^\varepsilon(f_X(A)|E|P_{A,E})\), adding the condition \(P'_{E} \leq P_E\), we define

\[H_{\min}^\varepsilon(A|E|P_{A,E}) := \max_{\|P_{A,E} - P_{A,E}^\prime\|_1 \leq \varepsilon, P'_{E} \leq P_E} H_{\min}(A|E|P_{A,E}^\prime|P_E)\]

(76)
As is shown in Lemma \text{22}, $H^*_{\min}(A|E|P_{A,E})$ equals $H^*_{\min}(A|E|P_{A,E}||P_{E})$ while the former has an additional constraint. Defining the quantities

$$
\Delta_{d,\min}(M, \varepsilon|P_{A,E}) := \min_{Q_E} \min_{P'_{A,E}} \mathbb{E}_{P_{A,E}} \left[ 2 \left\| P_{A,E} - P'_{A,E} \right\|_1 + \sqrt{M \varepsilon} e^{-\frac{1}{2}H_{\min}(A|E|P'_{A,E}||Q_E)} \right]
$$

(77)

$$
= \min_{Q_E} \min_{\varepsilon > 0} \left\{ 2 \varepsilon + \sqrt{M \varepsilon} e^{-\frac{1}{2}H_{\min}(A|E|P_{A,E}||Q_E)} \right\}
$$

(78)

$$
= \min_{Q_E} \min_{\varepsilon > 0} \mathbb{E}_{P_{A,E}:H_{\min}(A|E|P_{A,E}||Q_E) \geq R} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1 + \sqrt{M \varepsilon} e^{-\frac{1}{2}R} \right],
$$

(79)

$$
\Delta_{I,\min}(M, \varepsilon|P_{A,E}) := \min_{Q_E} \min_{P'_{A,E}:P_{A,E} \leq Q_E} \eta(\|P_{A,E} - P'_{A,E}\|_1, \log M) + \varepsilon Me^{-H_{\min}(A|E|P'_{A,E}||P_E)}
$$

(80)

$$
= \min_{\varepsilon > 0} \eta(\varepsilon, \log M) + \varepsilon Me^{-\frac{1}{2}H_{\min}(A|E|P_{A,E})}
$$

(81)

$$
= \min_{R} \eta(\mathbb{E}_{P_{A,E}:P'_{A,E} \leq P_{E},H_{\min}(A|E|P'_{A,E}||P_E) \geq R} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1, \log M \right], \varepsilon Me^{-R},
$$

(82)

we obtain the following lemma.

Theorem 21: Assume that $Q_E$ is a normalized distribution on $E$, $P_{A,E}$ is a sub-distribution on $A \times E$, and a linear random hash function $f_X$ from $A$ to $M = \{1, \ldots, M\}$ is $\varepsilon$-almost dual universal. Then, the random hash function $f_X$ satisfies

$$
\mathbb{E}_{X} d_1(f_X(A)|E|P_{A,E}) \leq \Delta_{d,\min}(M, \varepsilon|P_{A,E}),
$$

(83)

$$
\mathbb{E}_{X} I'(f_X(A)|E|P_{A,E}) \leq \Delta_{I,\min}(M, \varepsilon|P_{A,E}).
$$

(84)

That is, $\Delta_{d,\min}(M, \varepsilon|P_{A,E})$ and $\Delta_{I,\min}(M, \varepsilon|P_{A,E})$ are upper bounds for leaked information in the respective criteria when the rigorous smoothing of min entropy is applied.

V. RELATION WITH INFORMATION SPECTRUM

Information spectrum can derive asymptotically tight bounds of the optimal performances of various information processing by using only the asymptotic behavior of the tail probability, e.g., $P_{A,E}\{(a, e)|P_{A}(a|e) \geq e^{-R}\}$. Hence, it can be applied without any assumption for information sources. While information spectrum originally addresses the asymptotic setting, we bound the performances in the single-shot setting by using the tail probability. We call these upper and lower bounds single-shot information spectrum bounds.

In this section, we clarify the relation between the rigorous smoothing of min entropy and single-shot information spectrum bounds. While the rigorous smoothing of min entropy employs the smooth min entropy $H^*_{\min}(A|E|P_{A,E})$, we consider the bounds $\Delta_{d,\min}(M, \varepsilon|P_{A,E})$ and $\Delta_{I,\min}(M, \varepsilon|P_{A,E})$ as functions of $\min_{P'_{A,E}:H_{\min}(A|E|P'_{A,E}||Q_E) \geq R} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1 \right]$ or $\min_{P'_{A,E}:P_{E} \leq P_{E},H_{\min}(A|E|P'_{A,E}||P_E) \geq R} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1 \right]$. That is, we employ the formulas (79) and (82) rather than (78) and (81). Then, we give their relations with the tail probability, e.g., $P_{A,E}\{(a, e)|P_{A}(a|e) \geq e^{-R}\}$ as follows.

Lemma 22:

$$
\min_{P'_{A,E}:H_{\min}(A|E|P'_{A,E}||Q_E) \geq R} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1 \right] = \min_{P'_{A:E}:H_{\min}(A|E|P'_{A,E}||Q_E) \geq R, P_{A,E} \leq P'_{A,E}} \left[ \left\| P_{A,E} - P'_{A,E} \right\|_1 \right]
$$

$$
= P_{A,E}\{(a, e)|P_{A,E}(a|e) > e^{-R}Q_E(e)\} - e^{-R}A|P_{\text{mix},A} \times Q_E{(a, e)|P_{A,E}(a, e) > e^{-R}Q_E(e)}.
$$

(85)

and

$$
(1 - \frac{1}{C})P_{A,E}\{(a, e)|P_{A,E}(a|e) > C e^{-R}Q_E(e)\} \leq P_{A,E}\{(a, e)|P_{A,E}(a|e) > e^{-R}Q_E(e)\} - e^{-R}A|P_{\text{mix},A} \times Q_E{(a, e)|P_{A,E}(a|e) > e^{-R}Q_E(e)}
$$

$$
\leq P_{A,E}\{(a, e)|P_{A,E}(a|e) > e^{-R}Q_E(e)\}
$$

(86)

for $c > 1$ and $R$.

Since the condition $P'_{A,E} \leq P_{A,E}$ is more restrictive than $P'_{A,E} \leq P_{A}$, we see that $H^*_{\min}(A|E|P_{A,E}) = H^*_{\min}(A|E|P_{A,E}||P_{E})$.

Proof: The optimal sub-distribution $P'_{A,E}$ in the first line of (85) is given as

$$
P'_{A,E}(a|e) = \begin{cases} e^{-R}Q_E(e) & \text{if } P_{A,E}(a|e) > e^{-R}Q_E(e) \\ P_{A,E}(a|e) & \text{if } P_{A,E}(a|e) \leq e^{-R}Q_E(e) \end{cases}
$$

(87)

The sub-distribution is the optimal sub-distribution in the second line of (85). Substituting the above sub-distribution in to the first line, we obtain the third line of (85).
Next, we show (86). Since \(cP_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\} \geq e^{-R}|A|P_{\text{mix},A \times Q_E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\}\), we have

\[
(1 - \frac{1}{c})P_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\} = P_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\} - eP_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\}
\]

\[
\leq P_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\} - e^{-R}|A|P_{\text{mix},A \times Q_E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\}
\]

\[
\leq P_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\} - e^{-R}|A|P_{\text{mix},A \times Q_E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\}
\]

\[
\leq P_{A,E}\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\},
\]

where the inequality (85) follows from the fact that the maximum \(\max_{\Omega} P_{A,E}(\Omega) - e^{-R}|A|P_{\text{mix},A \times Q_E}(\Omega)\) can be realized by the set \(\{(a,e)|P_{A,E}(a,e) > ce^{-R}Q_E(e)\}\).

Therefore, using the formulas (79) and (82), we obtain the following theorem.

**Theorem 23:** The upper bounds \(\Delta_{d,\min}(M, \varepsilon|P_{A,E})\) and \(\Delta_{I,\min}(M, \varepsilon|P_{A,E})\) of leaked information by the rigorous smoothing of min entropy can be evaluated as follows.

\[
2(1 - \frac{1}{c})\min_{Q_E} \min_{R'} P_{A,E}\{(a,e)|P_{A,E}(a,e)Q_E(e) > ce^{-R'}\} + \sqrt{\varepsilon}M^\frac{3}{2}e^{-\frac{1}{2}R'}
\]

(89)

\[
\leq \Delta_{d,\min}(M, \varepsilon|P_{A,E}) \leq \min_{Q_E} \min_{R'} 2P_{A,E}\{(a,e)|P_{A,E}(a,e)Q_E(e) > ce^{-R'}\} + \sqrt{\varepsilon}M^\frac{3}{2}e^{-\frac{1}{2}R'},
\]

(90)

\[
(1 - \frac{1}{c})\min_{R'} \eta(P_{A,E}\{(a,e)|P_{A,E}(a,e)Q_E(e) > ce^{-R'}\} \geq \log M) + \varepsilon M e^{-R'}
\]

(91)

\[
\leq \Delta_{I,\min}(M, \varepsilon|P_{A,E}) \leq \min_{R'} \eta(P_{A,E}\{(a,e)|P_{A,E}(a,e)Q_E(e) > ce^{-R'}\} \geq \log M) + \varepsilon M e^{-R'}
\]

(92)

for \(c > 1\).

Theorem 23 explains that the bounds \(\Delta_{d,\min}(M, \varepsilon|P_{A,E})\) and \(\Delta_{I,\min}(M, \varepsilon|P_{A,E})\) by the rigorous smoothing of min entropy have almost the same values as the single-shot information spectrum bounds. Using this characterization, we evaluate the bounds \(\Delta_{d,\min}(M, \varepsilon|P_{A,E})\) and \(\Delta_{I,\min}(M, \varepsilon|P_{A,E})\) in the latter sections. However, the bounds by the rigorous smoothing of Rényi entropy of order 2 cannot be characterized in the same way. This fact seems to indicate the possibility of the smoothing of Rényi entropy of order 2 beyond the rigorous smoothing of min entropy.

**VI. SECRET KEY GENERATION WITH NO ERROR: SINGLE-SHOT CASE**

In order to obtain useful upper bounds, we need to calculate or evaluate the quantities \(\Delta_{d,2}(M, \varepsilon|P_{A,E})^{1/2}\), \(\Delta_{I,2}(M, \varepsilon|P_{A,E})^{1/2}\), \(\Delta_{d,\max}(M, \varepsilon|P_{A,E})^{1/2}\), and \(\Delta_{I,\max}(M, \varepsilon|P_{A,E})^{1/2}\). We say that their exact value is the rigorous smoothing and upper bounds by non-optimal choice \(P'\) are approximate smoothing. The paper [12] gave a suitable approximate smoothing of Rényi entropy of order 2 and derived the following proposition.

**Proposition 24:** The inequality

\[
\Delta_{d,2}(M, 1|P_{A,E}) \leq 3M^s e^{-sH_0^{\frac{1}{1-s}}(A|E|P_{A,E})}
\]

(93)

holds for \(s \in (0, 1/2]\).

Using the same approximate smoothing, we obtain the following evaluation.

**Lemma 25:** The inequality

\[
\Delta_{d,2}(M, \varepsilon|P_{A,E}) \leq (2 + \sqrt{\varepsilon})M e^{-sH_0^{\frac{1}{1-s}}(A|E|P_{A,E})}
\]

(94)

holds for \(s \in (0, 1/2]\).

Applying a similar approximate smoothing to Theorem 13, we obtain an upper bound for \(\Delta_{I,2}(M, \varepsilon|P_{A,E})\).

**Theorem 26:** The inequality

\[
\Delta_{I,2}(M, \varepsilon|P_{A,E}) \leq \eta(M e^{-sH_{1+s}(A|E|P_{A,E})}, \varepsilon + \log M)
\]

(95)

holds for \(s \in (0, 1]\).

**Proof:** For any integer \(M\), we choose the subset \(\Omega_M := \{P_{A,E}(a)e > M^{-1}\}\), and define the sub-distribution \(P_{A,E:M}\) by

\[
P_{A,E:M}(a,e) := \begin{cases}
0 & \text{if } (a,e) \in \Omega_M \\
P_{A,E}(a,e) & \text{otherwise.}
\end{cases}
\]
For $0 \leq s \leq 1$, we can evaluate $e^{-H_2(A|E|P_{A,E} M \parallel P_E)}$ and $d_1(P_{A,E}, P_{A,E} M)$ as
\[
e^{-H_2(A|E|P_{A,E} M \parallel P_E)} = \sum_{(a,e) \in \Omega_m^a} P_{A,E}(a,e)^2 (P_E(e))^{-1}
\]
\[
\leq \sum_{(a,e) \in \Omega_m^a} P_{A,E}(a,e)^{1+s} (P_E(e))^{-s} M^{-(1-s)}
\]
\[
\leq \sum_{(a,e) \in \Omega_m^a} P_{A,E}(a,e)^{1+s} (P_E(e))^{-s} M^{-(1-s)}
\]
\[
= e^{-sH_{1+s}(A|E|P_{A,E})} M^{-(1-s)}
\]
\[
\|P_{A,E} - P_{A,E} M\|_1 = \sum_{(a,e) \in \Omega_m} P_{A,E}(a,e)
\]
\[
\leq \sum_{(a,e) \in \Omega_m} (P_{A,E}(a,e))^{1+s} M^s (P_E(e))^{-s}
\]
\[
\leq \sum_{(a,e) \in \Omega_m} (P_{A,E}(a,e))^{1+s} M^s (P_E(e))^{-s}
\]
\[
= M^s e^{-sH_{1+s}(A|E|P_{A,E})}.
\]

Substituting (96) and (97) into (72), we obtain (74) because
\[
\eta(M^s e^{-sH_{1+s}(A|E|P_{A,E})}, \varepsilon + \log M)
\]
\[
= \eta(M^s e^{-sH_{1+s}(A|E|P_{A,E})}, \log M) + \varepsilon M^s e^{-sH_{1+s}(A|E|P_{A,E})}.
\]

In the above proof, we choose $P_{A,E}'$ to be $P_{A,E} M(a,e)$, we call the approximate smoothing by this particular choice the information-spectrum-smoothing because this type of smoothing is used to derive the entropic information spectrum in [17]. Indeed, the paper [12] also employed the information-spectrum-smoothing to derive Proposition 24.

Further, $\Delta_{d,\text{min}}(M, \varepsilon | P_{A,E})$ and $\Delta_{I,\text{min}}(M, \varepsilon | P_{A,E})$ can be evaluated as follows.

**Theorem 27:** The upper bounds $\Delta_{d,\text{min}}(M, \varepsilon | P_{A,E})$ and $\Delta_{I,\text{min}}(M, \varepsilon | P_{A,E})$ of leaked information by the rigorous smoothing of min entropy can be evaluated as follows.

\[
\Delta_{d,\text{min}}(M, \varepsilon | P_{A,E}) \leq (2 + \sqrt{\varepsilon}) \min_{0 \leq s} e^{-sH_{1+s}(A|E|P_{A,E}) + sR} (1 + 2s)
\]
\[
\Delta_{I,\text{min}}(M, \varepsilon | P_{A,E}) \leq \eta(\min_{0 \leq s} e^{-sH_{1+s}(A|E|P_{A,E}) + sR}, \varepsilon + \log M).
\]

Theorem 27 gives upper bounds on $\Delta_{d,\text{min}}(M, \varepsilon | P_{A,E})$ and $\Delta_{I,\text{min}}(M, \varepsilon | P_{A,E})$. The combination of Theorems 23 and 27 shows the performance of the rigorous smoothing of min entropy. Using these bounds, we can show the tight exponential decreasing rates of $\Delta_{d,\text{min}}(M, \varepsilon | P_{A,E})$ and $\Delta_{I,\text{min}}(M, \varepsilon | P_{A,E})$.

**Proof:** Since
\[
P_{A,E}\left\{(a,e) \mid \frac{P_{A,E}(a,e)}{Q_E(e)} > e^{-R'} \right\}
\]
\[
= \sum_{(a,e) : \frac{P_{A,E}(a,e)}{Q_E(e)} > e^{-R'}} P_{A,E}(a,e)
\]
\[
\leq \sum_{(a,e) : \frac{P_{A,E}(a,e)}{Q_E(e)} > e^{-R'}} P_{A,E}(a,e) \left( \frac{P_{A,E}(a,e)}{Q_E(e)} \right)^s e^{R'}
\]
\[
\leq \sum_{(a,e)} P_{A,E}(a,e) \left( \frac{P_{A,E}(a,e)}{Q_E(e)} \right)^s e^{R'}
\]
\[
= e^{-sH_{1+s}(A|E|P_{A,E} Q_E)} + sR',
\]

(100)
choosing $R' = \frac{\log M + 2 s P_{A,E}(A|E) P_{A,E}(Q_E)}{1 + 2 s}$, we have

$$2 P_{A,E} \left\{ (a,e) \mid \frac{P_{A,E}(a,e)}{Q_E(e)} > e^{-R'} \right\} + \sqrt{\frac{\log M}{2}} e^{-\frac{1}{2} R'}$$

$$\leq 2 e^{-s H_{1+s}(A|E) P_{A,E}(Q_E) + s R'} + \sqrt{\frac{\log M}{2}} e^{-\frac{1}{2} R'}$$

$$\leq (2 + \sqrt{\epsilon}) e^{-\frac{(1 + s) H_{1+s}(A|E) P_{A,E}(Q_E) + s R}{1 + 2 s}}.$$ 

Since the above inequality holds for $s \geq 0$, Lemma 7 yields that

$$\min_{Q_E} \min_{R'} 2 P_{A,E} \left\{ (a,e) \mid \frac{P_{A,E}(a,e)}{Q_E(e)} > e^{-R'} \right\} + \sqrt{\frac{\log M}{2}} e^{-\frac{1}{2} R'}$$

$$\leq \min_{0 \leq s} (2 + \sqrt{\epsilon}) e^{-\frac{(1 + s) H_{1+s}(A|E) P_{A,E}(Q_E) + s R}{1 + 2 s}}$$

$$= (2 + \sqrt{\epsilon}) \min_{0 \leq s} e^{-\frac{R}{1 + 2 s} H_{1+s}(A|E) P_{A,E}(Q_E) + \frac{R}{1 + 2 s}}.$$ 

Hence, combining (20), we obtain (29).

Choosing $R' = \frac{\log M + s P_{A,E}(A|E) P_{A,E}(Q_E)}{1 + s}$, we have

$$\eta \left( P_{A,E} \left\{ (a,e) \mid P_{A,I}(a|e) > e^{-R'} \right\} \right), \log M \leq \epsilon M e^{-R'}$$

$$\leq \eta \left( e^{-s H_{1+s}(A|E) P_{A,E} + s R'}, \log M \leq \epsilon M e^{-R'} \right)$$

$$\leq \eta \left( e^{-s H_{1+s}(A|E) P_{A,E} + s R'}, \log M \leq \epsilon e^{-s H_{1+s}(A|E) P_{A,E} + s R} \right)$$

$$\leq \eta \left( e^{-s H_{1+s}(A|E) P_{A,E} + s R}, \epsilon + \log M \right).$$ 

Since the above inequality holds for $s \geq 0$, we have

$$\min_{R'} \eta \left( P_{A,E} \left\{ (a,e) \mid P_{A,I}(a|e) > e^{-R'} \right\} \right), \log M \leq \epsilon M e^{-R'}$$

$$\leq \min_{0 \leq s} e^{-s H_{1+s}(A|E) P_{A,E} + s R}, \epsilon + \log M$$

$$= \eta \left( \min_{0 \leq s} e^{-s H_{1+s}(A|E) P_{A,E} + s R}, \epsilon + \log M \right).$$

Hence, combining (22), we obtain (29). 

Remark 2: Here, we compare the calculation amount of obtained bounds in Sections IV, VI, and VII. In order to calculate the bounds $\Delta_{d,2}(M, \epsilon \mid P_{A,E})$, $\Delta_{d,2}(M, \epsilon \mid P_{A,E})$, $\Delta_{d,\min}(M, \epsilon \mid P_{A,E})$, and $\Delta_{I,\min}(M, \epsilon \mid P_{A,E})$ based on rigorous smoothing, we need calculate the smooth entropies, which contains several optimization. Hence, the calculation of these bounds requires at least double optimization process. Then, they need higher calculation amounts. In particular, if the block size becomes larger, their calculation amounts increase heavily.

The bounds given in Section VII are calculated from the tail probability. For example, the tail probability $P_{A,E} \left\{ (a,e) \mid P_{A,I}(a|e) > e^{-R'} \right\}$ can be characterized as the tail probability with respect to the random variable log $P_{A,E}(a|e)$ because $P_{A,E} \left\{ (a,e) \mid P_{A,I}(a|e) > e^{-R'} \right\} = P_{A,E} \left\{ (a,e) \mid \log P_{A,I}(a|e) > -R' \right\}$. Hence, in the i.i.d. case, this probability can be calculated by using statistical packages. While the calculation amount increases with a rise in the block size, it is not as large as the above cases because statistical packages can be used.

The calculation amounts of the bounds given in Section VII are quite small. In particular, in the i.i.d. case, the calculation amounts do not depend on the block size. These bounds have great advantages with respect to their calculation amounts.

VII. SECRET KEY GENERATION WITH NO ERROR: ASYMPTOTIC CASE

Next, consider the case when the information source is given by the $n$-fold independent and identical distribution $P_{A,E}$ of $P_{A,E}$, i.e., $P_{A,E} = P_{A,E}$. In this case, Ahlswede and Csiszár [7] showed that the optimal generation rate

$$G(P_{A,E}) := \sup_{(f_n,M_n)} \left\{ \lim_{n \to \infty} \frac{\log M_n}{n} \right\}$$

$$= \eta \left( \frac{d'_1(f_n(A_n)\mid E_n | P_{A,E})}{n}, \epsilon \right).$$ 

equals the conditional entropy $H(A|E)$, where $f_n$ is a function from $A^n$ to $\{1, \ldots, M_n\}$. That is, when the generation rate

$$R = \lim_{n \to \infty} \frac{\log M_n}{n}$$

is smaller than $H(A|E)$, the quantity $d'_1(f_n(A_n)|E_n|P_{A,E})$ goes to zero. In order to treat the speed of this convergence, we use the supremum of the exponential rate of decrease (exponent) for $d'_1(f_n(A_n)|E_n|P_{A,E})$ and

$$I'(f_n(A_n)|E_n|P_{A,E}) = I(f_n(A_n) : E_n|P_{A,E}) + D(P_{f_n(A_n)}(F_{min,f_n(A_n)}))$$

for a given $R$. 

Due to \((41)\), when \(d'_1(f_{C_n}(A_n)|E_n|P_{A,E}^n)\) goes to zero, \(I'(f_{C_n}(A_n)|E_n|P_{A,E}^n)\) goes to zero. Conversely, due to \(49)\), when \(I'(f_{C_n}(A_n)|E_n|P_{A,E}^n)\) goes to zero, \(d'_1(f_{C_n}(A_n)|E_n|P_{A,E}^n)\) goes to zero. So, even if we replace the security criterion by \(I'(f_{C_n}(A_n)|E_n|P_{A,E}^n)\), the optimal generation rate does not change.

Now, we consider the case when the length of generated keys behaves as \(nH(A|E|P)+\sqrt{nR}\). It is known in [29] Subsection II-D] that

\[
\lim_{n \to \infty} \min_{f} d'_1(f(A_n)|E_n|P_{A,E}^n) = 2 \int_{-\infty}^{R/\sqrt{V(P)}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx. 
\]  

(101)

Then, using Theorem 27 we obtain the following theorem.

**Theorem 28:** We choose a polynomial \(P(n)\). When a random linear function \(f_{X^n}\) from \(A^n\) to \(\{1, \ldots, \lfloor e^{nH(A|E|P)+\sqrt{nR}} \rfloor\}\) is \(P(n)\)-almost dual universal, the relations

\[
\lim_{n \to \infty} E_{X^n} d'_1(f_{X^n}(A_n)|E_n|P_{A,E}^n) = \lim_{n \to \infty} \min_{f} d'_1(f(A_n)|E_n|P_{A,E}^n) = 2 \int_{-\infty}^{R/\sqrt{V(P)}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx  
\]

(102)

hold, where we take the minimum under the condition that \(f\) is a function from \(A^n\) to \(\{1, \ldots, \lfloor e^{nH(A|E|P)+\sqrt{nR}} \rfloor\}\) and \(V(P) := \sum_{a \in E} P_{A,E}(a,e)(\log P_{A,E}(a,e)) - H(A|E|P)^2\).

Lemma 28 implies that any \(P(n)\)-almost dual universal hash function realizes the optimality in the sense of the second order asymptotics when we employ the \(L_1\) distinguishability criterion. This analysis is obtained from our approximate smoothing of min entropy. That is, this analysis does not require an approximate smoothing of Rényi entropy of order 2. The second order analysis with the mutual information criterion is not so easy. We do not treat this issue.

**Proof:** We applying \((90)\) in Theorem 23 with \(R' = nH(A|E|P)+\sqrt{nR}+n^{1/4}\). Then, the central limit theorem guarantees that

\[
E_{X^n} d'_1(f_{X^n}(A_n)|E_n|P_{A,E}^n) \leq \Delta_{d,\min}(e^{nH(A|E|P)+\sqrt{nR}+n^{1/4}}, P(n)|P_{A,E}^n) \\
\leq 2P_{A,E}^n \left\{ (a,e) \in |P_{A,E}(a,e)| > e^{-nH(A|E|P)-\sqrt{nR}+n^{1/4}} \right\} + \sqrt{P(n)e^{-n^{1/4}}}  \\
-2 \int_{R/\sqrt{V(P)}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx.
\]

Since \(\min_f d'_1(f(A_n)|E_n|P_{A,E}^n) \leq d'_1(f_{X^n}(A_n)|E_n|P_{A,E}^n)\), combining (101), we obtain (102).

Now, we proceed to the exponential decreasing rate when we choose the key generation rate \(R\) is greater than \(H(A|E|P)\). Since the discussion for the exponential decreasing rate is more complex, more delicate treatment is required. First, we should remark that the exponential decreasing rate depends on the choice of the security criterion. Then, we obtain the following theorem.

**Theorem 29:** We choose a polynomial \(P(n)\). When a linear random function \(f_{X^n}\) from \(A^n\) to \(\{1, \ldots, \lfloor e^{nR} \rfloor\}\) is \(P(n)\)-almost dual universal, the relations

\[
\liminf_{n \to \infty} -\frac{1}{n} \log E_{X^n} d'_1(f_{X^n}(A_n)|E_n|P_{A,E}^n) \geq \liminf_{n \to \infty} -\frac{1}{n} \log \Delta_{d,2}(e^{nR}, P(n)|P_{A,E}^n) \geq e_G(P_{A,E}|R)  \\
\liminf_{n \to \infty} -\frac{1}{n} \log E_{X^n} I'(f_{X^n}(A_n)|E_n|P_{A,E}^n) \geq \liminf_{n \to \infty} -\frac{1}{n} \log I_{d,2}(e^{nR}, P(n)|P_{A,E}^n) \geq e_H(P_{A,E}|R)
\]

(103)

(104)

hold, where

\[
e_G(P_{A,E}|R) := \max_{0 \leq t \leq 1} t(H_t^G(A|E|P_{A,E}) - R)  \\
e_H(P_{A,E}|R) := \max_{0 \leq s \leq 1} s(H_{1+s}(A|E|P_{A,E}) - R).
\]

(105)

(106)

**Proof:** (103) can be shown by Theorem 25 (104) can be shown by Theorem 26.

As is shown in Appendix E:2A the following relation between two exponents \(e_H(P_{A,E}|R)\) and \(e_G(P_{A,E}|R)\) holds.

**Lemma 30:** we obtain

\[
\frac{1}{2} e_H(P_{A,E}|R) \leq e_G(P_{A,E}|R)  \\
e_H(P_{A,E}|R) \geq e_G(P_{A,E}|R).
\]

(107)

(108)

First, we consider the tightness of Inequality (103). Corollary 20 yields the exponent \(\frac{e_H(P_{A,E}|R)}{2}\) for the \(L_1\) distinguishability criterion. Lemma 30 shows that the exponents by Theorem 25 is better than that by Corollary 20. Further, it is also shown in [38] Theorem 30 that there exists a sequence of universal functions \(f_{X^n}\) from \(A^n\) to \(\{1, \ldots, \lfloor e^{nR} \rfloor\}\) such that

\[
\limsup_{n \to \infty} -\frac{1}{n} \log E_{X^n} d'_1(f_{X^n}(A_n)|E_n|P_{A,E}^n) \leq e_G(P_{A,E}|R),
\]

(109)
where
\[\hat{e}_G(P_{A,E}|R) := \max_{0 \leq t} t \cdot \frac{H^G_{1-t}(A|E|P_{A,E}) - R}{1-t}.\]  
(110)

When the maximum \(\max_{0 \leq t} t(H^G_{1-t}(A|E|P_{A,E}) - R)\) is attained with \(t \in (0, \frac{1}{2}]\), we have \(e_G(P_{A,E}|R) = \hat{e}_G(P_{A,E}|R)\).

Assume that \(P(n) \geq 1\). Then, Since \(\Delta_{d,2}(e^{nR}, P(n)|P_{A,E}) \leq \sqrt{P(n)} \Delta_{d,2}(e^{nR}, 1|P_{A,E})\), combining (93), (103), and (109) we have
\[
\lim_{n \to \infty} -\frac{1}{n} \log \Delta_{d,2}(e^{nR}, P(n)|P_{A,E}) = \lim_{n \to \infty} -\frac{1}{n} \log \Delta_{d,2}(e^{nR}, 1|P_{A,E}) = e_G(P_{A,E}|R).
\]  
(111)

That is, our evaluation (103) for \(\Delta_{d,2}(e^{nR}, P(n)|P_{A,E})\) is sufficiently tight in the large deviation sense.

Next, we consider the tightness of Inequality (104). Corollary [19] yields the exponent \(e_G(P_{A,E}|R)\) for the modified mutual information criterion. Lemma [50] shows that the exponent by Theorem [26] is better than that by Corollary [19]. Further, the lower bound of the exponent \(e_G(P_{A,E}|R)\) is the same as that given in the previous paper [13] under the universal condition. Since the bound given in [13] is the best lower bound of the exponent, our evaluation (104) for \(\Delta_{d,2}(e^{nR}, P(n)|P_{A,E})\) is as good as the existing evaluation [13] in the large deviation sense.

From the above discussion, we find that the exponents directly obtained by our approximate smoothing of Rényi entropy of order 2 are better than the exponents derived from the combination of Inequality (39) and (41) and the exponent of the other criterion. This fact indicates that we need to choose our approximate smoothing dependently on the security criterion.

**Remark 3:** Now, we consider the relation with the recent paper [27] discussing the quantum case as including the non-quantum case. When \(A = \mathbb{F}_q\), we focus on a \(1 + P(n)g^{-n+|nR|}\)-almost universal surjective linear function \(f_{X^n}\) over the field \(\mathbb{F}_q\) from \(\mathbb{F}_q^n\) to \(\mathbb{F}_q^{|nR|}\). Thanks to Proposition [11] the surjective linear random function \(f_{X^n}\) over the field \(\mathbb{F}_q\) is \(q+P(n)\)-almost dual universal. Hence, we obtain (103), which can recover a part of the result by [27] with the case of linear functions in the non-quantum case. The paper [27] showed the security with an \(\epsilon_{n}\)-almost universal hash function when \(\epsilon_{n}\) approaches to 1. Since we assume the surjectivity, our method cannot recover the result by [27] with the linear hash function perfectly.

Now, we clarify how better our smoothing of Rényi entropy of order 2 is than the rigorous smoothing of min entropy. As is shown in Appendix [E] we obtain the following theorem.

**Theorem 31:** The relations
\[
\lim_{n \to \infty} -\frac{1}{n} \log \Delta_{d,\min}(e^{nR}, \epsilon|P^n_{A,E}) = \hat{e}_G(P_{A,E}|R) := \max_{0 \leq s} \frac{s(H^G_{1+s}(A|E|P_{A,E}) - R)}{1 + 2s}
\]  
(112)

\[
\lim_{n \to \infty} -\frac{1}{n} \log \Delta_{I,\min}(e^{nR}, \epsilon|P^n_{A,E}) = \hat{e}_H(P_{A,E}|R) := \max_{0 \leq s} \frac{sH^F_{1+s}(A|E|P_{A,E}) - sR}{1 + s}
\]  
(113)

hold.

For comparison the exponents by the rigorous smoothing of min entropy and our approximate smoothing of Rényi entropy of order 2, as is shown in Appendix [E], we have the following lemma by using Theorem [27].

**Lemma 32:** The inequalities
\[
e_G(P_{A,E}|R) > \hat{e}_G(P_{A,E}|R)
\]  
(114)

\[
e_H(P_{A,E}|R) > \hat{e}_H(P_{A,E}|R)
\]  
(115)

hold when \(P_{A|E^c} = 0\) is not a uniform distribution for an element \(e \in E\). The equalities \(e_G(P_{A,E}|R) = \hat{e}_G(P_{A,E}|R)\) and \(e_H(P_{A,E}|R) = \hat{e}_H(P_{A,E}|R)\) hold when \(P_{A|E^c} = 0\) is a uniform distribution for any element \(e \in E\).

Theorem 31 and Lemma 32 show that the rigorous smoothing of min entropy cannot attain the exponents \(e_G(P_{A,E}|R)\) and \(e_H(P_{A,E}|R)\). That is, the bounds \(\Delta_{d,2}(e^{nR}, |P^n_{A,E})\) and \(\Delta_{I,2}(e^{nR}, |P^n_{A,E})\) by the rigorous smoothing of Rényi entropy of order 2 are strictly better than the bounds \(\Delta_{d,\min}(e^{nR}, \epsilon|P^n_{A,E})\) and \(\Delta_{I,\min}(e^{nR}, \epsilon|P^n_{A,E})\) by the rigorous smoothing of min entropy in the sense of large deviation. This fact indicates the importance of smoothing of Rényi entropy of order 2.

In summary, while a smoothing of min entropy yields the tight bound in the sense of the second order asymptotics, the rigorous smoothing of min entropy cannot yield the tight bound in the sense of the exponential decreasing rate.

**Remark 4:** Here, we give the relation with the results in the quantum case [57]. The paper [57] showed that
\[
\frac{1}{n} \log \Delta_{d,2}(e^{nR}, P(n)|P^n_{A,E}) \geq \max_{0 \leq t \leq \frac{1}{2}} \frac{t}{2(1-t)} (H^G_{1-t}(A|E|P_{A,E}) - R)
\]  
(116)

\[
\frac{1}{n} \log \Delta_{I,2}(e^{nR}, P(n)|P^n_{A,E}) \geq \max_{0 \leq s \leq 1} \frac{s}{2 - s} (H^F_{1+s}(A|E|P_{A,E}) - R).
\]  
(117)
The RHSs of (116) and (117) are smaller than \( e_G(P_{A,E}|R) \) and \( e_H(P_{A,E}|R) \), respectively. Hence, our result is better in the non-quantum case.

VIII. EQUIVOCATION RATE OF SECRET KEY GENERATION WITHOUT ERROR CORRECTION

When the key generation rate \( R \) is larger than the conditional entropy \( H(A|E|P_{A,E}) \), the leaked information does not go to zero. In this case, it is natural to consider the rate of the conditional entropy rate of generated keys or the rate of the modified mutual information \( H(A|E|P_{A,E}) \). The former rate is called the equivocation rate, and is known to be less than the conditional entropy \( H(A|E|P_{A,E}) \). That is, the rate of the modified mutual information is larger than \( R - H(A|E|P_{A,E}) \). Now, we show that the minimum rate of the modified mutual information \( R - H(A|E|P_{A,E}) \) can be achieved by an \( \varepsilon \)-almost dual universal hash function. For this purpose, we employ (63) instead of (62). Then, we obtain a slightly stronger evaluation than Theorem 21.

**Theorem 33:** Assume that \( Q_E \) is a normalized distribution on \( E \), \( P_{A,E} \) is a sub-distribution on \( A \times E \), and a linear random hash function \( f_X \) from \( A \) to \( M = \{1, \ldots, M\} \) is \( \varepsilon \)-almost dual universal. Then, the random hash function \( f_X \) satisfies

\[
\mathbb{E}_X I'(f_X(A)|E|P_{A,E}) \leq \Delta_{I,\min}(M, \varepsilon|P_{A,E}),
\]

where

\[
\Delta_{I,\min}(M, \varepsilon|P_{A,E}) := \min_{Q_E} \min_{P_{A,E}: P_E \leq Q_E} \eta(\|P_{A,E} - P'_{A,E}\|_1, \log M) + \log(1 + \varepsilon M e^{-H_{\min}(A|E)|P'_{A,E}|P_E})
\]

\[
= \min_{\varepsilon, \ell > 0} \eta(\ell, \log M) + \log(1 + \varepsilon Me^{-H_{\min}(A|E)|P_{A,E}}),
\]

\[
= \min_{R'} \eta(\min_{P_{A,E}: P_E \leq P_{E,H_{\min}(A|E)|P_{A,E}|P_E} \geq R} \|P_{A,E} - P'_{A,E}\|_1, \log M) + \log(1 + \varepsilon Me^{-R'})
\]

Further, by using similar discussions as Sections V and VI, the upper bound \( \Delta_{I,\min}(M, \varepsilon|P_{A,E}|P_{A,E}) \) can be evaluated as follows.

**Theorem 34:** Any polynomial \( P(n) \) satisfies

\[
\lim_{n \to \infty} \frac{1}{n} \Delta_{I,\min}(\lfloor e^{nR} \rfloor, P(n)|P_{A,E}^n) = R - H(A|E|P_{A,E})
\]

for \( R \geq H(A|E|P_{A,E}) \).

**Proof:** Inequality (122) follows from Lemma 33 and (121). Inequality (123) follows from (100) with \( Q_E = P_E \).

**Theorem 35:** Any polynomial \( P(n) \) satisfies

\[
\lim_{n \to \infty} \frac{1}{n} \Delta_{I,\min}(\lfloor e^{nR} \rfloor, P(n)|P_{A,E}^n) = R - H(A|E|P_{A,E})
\]

for \( R > H(A|E|P_{A,E}) \).

**Proof:** It is known by (31) that any sequence of hash function from \( A \) to \( \{1, \ldots, \lfloor e^{nR} \rfloor \} \) satisfies

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E}_{X,n} I'(f_{X,n}(A)|E|P_{A,E}) \geq R - H(A|E|P_{A,E}).
\]

Hence, it is enough to show that

\[
\lim_{n \to \infty} \frac{1}{n} \Delta_{I,\min}(\lfloor e^{nR} \rfloor, P(n)|P_{A,E}^n) \leq R - H(A|E|P_{A,E}).
\]

We choose \( R' < H(A|E|P_{A,E}) \). Relation (123) implies that

\[
\frac{1}{n} \Delta_{I,\min}(\lfloor e^{nR} \rfloor, P(n)|P_{A,E}^n) \leq \frac{1}{n} \eta(\min_{s \geq 0} e^{sn(R' - H_{\min}(A|E)|P_{A,E}|nR)} + \frac{1}{n} \log(1 + P(n)e^{n(R' - R')})
\]

Since \( R' < H(A|E|P_{A,E}) \), the value \( \min_{s \geq 0} e^{sn(R' - H_{\min}(A|E)|P_{A,E}|nR)} \) goes to zero exponentially. Hence, the term \( \frac{1}{n} \eta(\min_{s \geq 0} e^{sn(R' - H_{\min}(A|E)|P_{A,E}|nR)} + \frac{1}{n} \log(1 + P(n)e^{n(R' - R')}) \) goes to zero. Since \( \frac{1}{n} \log(1 + P(n)e^{n(R' - R')}) \leq R - R' + \frac{1}{n} \log(1 + P(n)) \to R - R' \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \Delta_{I,\min}(\lfloor e^{nR} \rfloor, P(n)|P_{A,E}^n) \leq R - R'.
\]

Since \( R' \) is an arbitrary real number satisfying \( R' < H(A|E|P_{A,E}) \), we obtain (126).
IX. SECRET KEY GENERATION WITH ERROR CORRECTION

A. Protocol

Next, we apply the above discussions to secret key generation with public communication. Alice is assumed to have an initial random variable $a \in A$, which generates with the probability $p_a$, and Bob and Eve are assumed to have their random variables $B \in B$ and $E \in E$, respectively (or initial quantum states $\rho_{B|a}$ and $\rho_{E|a}$ on their quantum systems $\mathcal{H}_B$ and $\mathcal{H}_E$, respectively.) The task for Alice and Bob is to share a common random variable almost independent of Eve’s quantum state by using a public communication. The quality is evaluated by three quantities: the size of the final common random variable, the probability of the disagreement of their final variables (error probability), and the information leaked to Eve, which can be quantified by the $L_1$ distinguishability criterion or the modified mutual information criterion between Alice’s final variables and Eve’s random variable.

In order to construct a protocol for this task, we assume that the set $A$ is a vector space on a finite field $\mathbb{F}_q$. Indeed, even if the cardinality $|A|$ is not a prime power, it become a prime power by adding elements with zero probability. Hence, we can assume that the cardinality $|A|$ is a prime power $q$ without loss of generality. Then, the secret key agreement can be realized by the following two steps: The first is the error correction, and the second is the privacy amplification. In the error correction, Alice and Bob prepare a linear subspace $C_1 \subset A$ and the representatives $a(x)$ of all cosets $x \in A/C_1$. Alice sends the coset information $[A] \in A/C_1$ to Bob in stead of her random variable $A \in A$, and Bob obtains his estimate $A$ of $A \in A$ from his random variable $B \in B$ (or his quantum state on $\mathcal{H}_B$) and $[A] \in A/C_1$. Alice obtains her random variable $A_1 := A - a([A]) \in C_1$, and Bob obtains his random variable $A'_1 := A - a([B]) \in C_1$. In the privacy amplification, Alice and Bob prepare a common hash function $f$ on $C_1$. Then, applying the hash function $f$ to the their variables $A_1$ and $A'$, they obtain their final random variables $f(A_1)$ and $f(A')$.

Indeed, the above protocol depends on the choice of estimator that gives the estimate $\hat{A}$ from $[A] \in A/C_1$ and his random variable $B \in B$ (or his quantum state on $\mathcal{H}_B$). In the remaining part of this section, we give the estimator depending on the setting and discuss the performance of this protocol.

B. Error probability

In the following, we give the concrete form of the estimator and evaluate the error probability. In this case, we apply the Bayesian decoder, which is given as

$$\hat{A}([A], B) := \arg\max_{A' \in [A]} P(A', B) = \arg\max_{A' \in [A]} P(A' | B).$$

In this case, the error probability is characterized as follows.

$$P_e[P_{A,B}, C_1] := \sum_a P_A(a) \sum_b P_B|A|(b|a) \Xi_{a,b}(C_1),$$

where

$$\Xi_{a,b}(C_1) := \begin{cases} 1 & \exists a' \in C_1 + a \setminus \{a\}, \frac{P_{A,B}(a', b)}{P_{A,B}(a, b)} \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

For any $s \in (0, 1]$, the quantity $\Xi_{a,b}(C_1)$ satisfies

$$\Xi_{a,b}(C_1) \leq (\Xi_{a,b}(C_1))^s$$

$$\Xi_{a,b}(C_1) \leq \sum_{a' \in C_1 + a \setminus \{a\}} \left( \frac{P_{A,B}(a', b)}{P_{A,B}(a, b)} \right)^s.$$

Thus, the error probability $P_e[P_{A,B}, C_1]$ can be evaluated as

$$P_e[P_{A,B}, C_1] \leq \sum_a \sum_b P_{A,B}(a, b) \left( \sum_{a' \in C_1 + a \setminus \{a\}} \left( \frac{P_{A,B}(a', b)}{P_{A,B}(a, b)} \right)^s \right)^s$$

$$= \sum_b \sum_a P_{A,B}(a, b) \left( \sum_{a' \in C_1 + a \setminus \{a\}} P_{A,B}(a', b) \right)^s.$$
Now, we randomly choose the code $C_1$ as an $\varepsilon$-almost universal code $C_\chi$ with the dimension $t$. Then,

$$
E_X P_{s|A,B,C\chi} \leq E_X \sum_b \sum_a P_{A,B}(a,b) \sum_{a' \in C_\chi + a} P_{A,B}(a',b) \varepsilon
$$

$$
\leq \sum_b \sum_a P_{A,B}(a,b) \varepsilon \sum_{a' \in C_\chi + a} P_{A,B}(a',b) \varepsilon
$$

$$
\leq \sum_b \sum_a P_{A,B}(a,b) \varepsilon \sum_{a' \neq a} P_{A,B}(a',b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) |A| \sum_b \sum_a P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) |A| \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

$$
= \varepsilon (\frac{q^t}{|A|}) ^s \sum P_{A,B}(a,b) \varepsilon
$$

\[ (129) \]

C. Leaked information with fixed error correction code

As is mentioned in the previous sections, we have two criteria for quality of secret random variables. Given a code $C_1 \subset A$ and a hash function $f$, the first criterion is $d_1'(f(A_1)||[A],E|P_{A,E})$, and the second criterion is $I'(f(A_1)||[A],E|P_{A,E})$. Note that the random variable $A$ can be written by the pair of $A_1$ and $[A]$ given in Subsection IX-A. Then, we obtain the following theorem.

**Theorem 36:** Assume that $f_X$ is an $\varepsilon$-almost dual universal hash function from $A/C_1$ to $M = \{1, \ldots, M\}$. The relations

$$
E_X d_1'(f_X(A_1||[A],E|P_{A,E})
$$

$$
\leq (2 + \sqrt{\varepsilon}) (|A|/L)^s e^s H_{1+\varepsilon}^G(A||[A],E|P_{A,E})
$$

$$
E_X I'(f_X(A_1||[A],E|P_{A,E})
$$

$$
\leq \eta(\eta(M^s e^s H_{1+\varepsilon}^G(A||[A],E|P_{A,E}), \varepsilon + \log M).
$$

hold for $s \in (0,1/2]$ and $s' \in (0,1]$, where $L$ is the the amount of sacrifice information $|C_1|/M$.

**Proof:** Proposition $24$ and $27$ with $s = \frac{1}{1+s}$ guarantee that

$$
E_X d_1'(f_X(A_1||[A],E|P_{A,E})
$$

$$
\leq (2 + \sqrt{\varepsilon}) M^s e^s H_{1+\varepsilon}^G(A_1||[A],E|P_{A,E})
$$

$$
\leq (2 + \sqrt{\varepsilon}) M^s (|A|/|C_1|)^s e^s H_{1+\varepsilon}^G(A_1||[A],E|P_{A,E})
$$

$$
= (2 + \sqrt{\varepsilon}) (M^s |A|/|C_1|)^s e^s H_{1+\varepsilon}^G(A||[A],E|P_{A,E}),
$$

which implies $130$. Proposition $24$ Lemma $6$ and $27$ with $s' = \frac{1}{1+s'}$ guarantee that

$$
E_X I'(f_X(A_1||[A],E|P_{A,E})
$$

$$
\leq \eta(M^s e^s H_{1+\varepsilon}^G(A_1||[A],E|P_{A,E}), \varepsilon + \log M)
$$

$$
\leq \eta(M^s e^s H_{1+\varepsilon}^G(A_1||[A],E|P_{A,E}), \varepsilon + \log M)
$$

$$
\leq \eta(M^s (|A|/|C_1|)^s e^s H_{1+\varepsilon}^G(A_1||[A],E|P_{A,E}), \varepsilon + \log M)
$$

$$
= \eta((|A|/L)^s \varepsilon e^s H_{1+\varepsilon}^G(A||[A],E|P_{A,E}), \varepsilon + \log M),
$$

which implies $131$. □
D. Leaked information with randomized error correction code

Next, we evaluate leaked information when the error correcting code $C_1$ is chosen as an $\epsilon_1$-almost universal $2$ code. In this case, the evaluation for the average of the modified mutual information criterion can be improved to the following way.

**Theorem 37:** We choose the code $C_1$ as an $\epsilon_1$-almost universal $2$ code $C_X$ with the dimension $t$. Assume that $f_Y$ is an $\epsilon_2$-almost dual universal $2$ hash function from $A/C_X$ to $M = \{1, \ldots, M\}$, the random variables $X$ and $Y$ are independent, and $\epsilon_2 \geq 1$.

$$E_{X,Y} T'(f_Y(A_1)|A|C_X, E|P_{A,E}) \leq \eta'(\frac{|A|}{q^t}) e^{-sH_{1+}(A|E|P_{A,E})} \log M + \frac{\epsilon_2}{\epsilon_1} \log \epsilon_1. \quad (133)$$

for $s \in (0, 1]$.

**Proof:** We choose a joint sub-distribution $P_{A,E}'(a, e)$ such that $P_{A,E}'(a, e) \leq P_{A,E}(a, e)$. Due to (57), we obtain

$$E_Y e^{-H_2(f_Y(A_1)|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} \leq \epsilon_2 e^{-H_2(f_Y(A_1)|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} + \frac{1}{|A|} \epsilon_2 e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)}$$

$$\leq \epsilon_2 e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} + \frac{1}{|A|} e^{\psi(1)(P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} e^{-H_2(A|E|P_{A,E}'|P_E)}$$

$$= \epsilon_2 e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} + \frac{1}{|A|} \sum_{a} \sum_{e} P_{A,E}'(a, e) e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)}$$

Since

$$e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} = \sum_{a} \sum_{e} P_{A,E}'(a, e) e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)}$$

and

$$E_X \sum_{a} \sum_{e} P_{A,E}'(a, e) \left( \sum_{a' \in C_X + a \setminus \{a\}} P_{A,E}'(a', e) \right) (P_E(e))^{-1} \leq \epsilon_2 \sum_{a} \sum_{e} P_{A,E}'(a, e) (P_E(e))^{-1}$$

we have

$$E_X e^{-H_2(1\|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} \leq \epsilon_1 \frac{q^t}{|A|} e^{\psi(1)(P_{\text{mix}}|A|C_X \times P_E)} \leq \epsilon_1 \frac{q^t}{|A|} \quad (134)$$

where the first inequality follows from $\epsilon_2 \geq 1$.

Hence, we obtain

$$E_{X,Y} e^{-H_2(f_Y(A_1)|A|C_X, E|P_{A,E}'|P_{\text{mix}}|A|C_X \times P_E)} \leq \epsilon_2 \frac{|A|}{q^t} e^{-H_2(A|E|P_{A,E}'|P_E)} + \frac{1}{M} \epsilon_1$$

$$= \frac{1}{M} \epsilon_1 (1 + \frac{\epsilon_2 |A|}{q^t} e^{-H_2(A|E|P_{A,E}'|P_E)}).$$
Applying Jensen’s inequality to \( x \mapsto \log x \), we obtain
\[
\begin{align*}
\mathbb{E}_{X,Y} - H_2(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}'|P_{\text{mix},|A|C_X} \times P_E) \\
\leq \log M + \log \epsilon_1 \\
+ \log \left(1 + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}\right).
\end{align*}
\]
Using (66), (17), and Lemma 8, we obtain
\[
\begin{align*}
I'(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}) \\
= \log M - H(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}) \\
\leq \eta(\|P_{A,E} - P_{A,E}'\|_1, \log M) \\
+ \log \epsilon_1 + \log \left(1 + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}\right) \\
\leq \eta\|P_{A,E} - P_{A,E}'\|_1, \log M \\
+ \log \epsilon_1 + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}.
\end{align*}
\]
Applying the same discussion as the proof of Theorem 26, we obtain
\[
\begin{align*}
\mathbb{E}_{X,Y} I'(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}') \\
\leq \eta\|P_{A,E} - P_{A,E}'\|_1, \log M + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}. \\
\end{align*}
\]
Hence, we obtain
\[
\begin{align*}
\mathbb{E}_{X,Y} I'(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}) \\
\leq \eta(\|P_{A,E} - P_{A,E}'\|_1, \log M) \\
+ \log \epsilon_1 + \log \left(1 + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}\right) \\
\leq \eta\|P_{A,E} - P_{A,E}'\|_1, \log M \\
+ \log \epsilon_1 + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}.
\end{align*}
\]
Applying the same discussion as the proof of Theorem 26 we obtain
\[
\begin{align*}
\mathbb{E}_{X,Y} I'(f_Y(A_1)[|A|]_{C_X}, E|P_{A,E}') \\
\leq \eta\|P_{A,E} - P_{A,E}'\|_1, \log M + \frac{\epsilon_2 |A|}{\epsilon_1 q} M e^{-H_2(A|E|P_{A,E}'|P_E)}. \\
\end{align*}
\]
\[\text{(135)}\]

E. Asymptotic analysis

Next, we consider the case when the joint distribution is given as the \( n \)-fold independent and identical distribution \( P_{A,B,E}^n \) of a distribution \( P_{A,B,E} \), where \( A \) is \( \mathbb{F}_q \). In this setting, we can treat the error probability and the leaked information separately. Now, we fix a code \( C_{1,n} \in \mathbb{F}_q^n \) with the dimension \( \frac{n - \log q}{\log n} \).

Theorem 38: Let \( P(n) \) be an arbitrary polynomial. When \( f_X \) is a \( P(n) \)-almost dual universal2 hash function from \( \mathbb{F}_q^n/C_{1,n} \) to \( \mathbb{F}_q^{n - \log q} \), the relations
\[
\begin{align*}
\liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{E} d'_1(f_X(A_1,n)[|A_n|, E_n P_{A,E}^n) \\
\geq \max_{0 \leq s \leq 1/2} s\log q + sH_{1/2}(A|E|P_{A,E}) \\
eq e_{G}(P_{A,E}|\log q - R_2), \\
\liminf_{n \to \infty} -\frac{1}{n} \log \mathbb{E} I'(f_X(A_1,n)[|A_n|, E_n P_{A,E}^n) \\
\geq \max_{0 \leq s \leq 1/2} s\log q + sH_{1/2}(\phi(s|A|E|P_{A,E}) \\
eq e_{G}(P_{A,E}|\log q - R_2),
\end{align*}
\]
hold.

Proof: 130 and (131) yield the above inequality.

Hence, due to \( 29 \), when \( R_1 \leq \log q - H(A|B|P_{A,B}) \), the error probability goes to zero exponentially. Similarly, when \( R_2 \geq \log q - H(A|E|P_{A,E}) \), the leaked information goes to zero exponentially in both criteria. In the above case, the key
generation rate $R_1 - R_2$ is less than $H(A|E|P_{A,E}) - H(A|B|P_{A,B})$. This value is already obtained by Ahlswede & Csiszár\cite{7}, Maurer\cite{6}.

Next, we consider the case when the error correcting code is chosen randomly. In this case, the exponential decreasing rate for $I'(f_X(A_{1,n})||A_n, E_n|P^*_n)$ can be improved as follows.

**Theorem 39:** For an arbitrary polynomial $P(n)$ and the independent random variables $X, Y$, we assume that the random code $C_X$ with the dimension $[n R_1 / \log q]$ is universal 2 and $f_Y$ is a $P(n)$-almost dual universal 2 hash function from $\mathbb{F}_q^n / C_X$ to $\mathbb{F}_q^{[n R_1 / \log q]}$, the relations (138) and

\[
\lim_{n \to \infty} \frac{1}{n} \log E_{X_n, Y} P^*_n[P_{A,B}, C_{X_n}]
\geq \max_{0 \leq s \leq 1} s(\log q - R) - s H^G_{1/s}(A|B|P_{A,B})
\]

(140)

\[
\lim_{n \to \infty} \frac{1}{n} \log E_{X_n, Y} I'(f_Y(A_{1,n})||A_n|C_X, E_n|P^*_n)
\geq \max_{0 \leq s \leq 1} s(R_2 - \log q) + s H_{1+s}(A|E|P_{A,E})
= e_H(P_{A,E}|\log q - R_2)
\]

(141)

hold.

**Proof:** Theorem 37 implies (141). Since $C_{X_n}$ is a universal 2 code in $\mathbb{F}_q^n$ with the dimension $[n R_1 / \log q]$, due to (129), the error probability can be bounded as

\[
E_{X_n, Y} P^*_n[P_{A,B}, C_{X_n}]
\leq e^{n(s(R_1 - \log q) + s H^G_{1/s}(A|B|P_{A,B}))}
\]

for $s \in [0, 1]$, which implies (140).

**Remark 5:** The RHS of (138) in Theorem 38 is the same as the exponent of (12) (66)). The RHS of (141) in Theorem 39 is the same as the exponent of (13) (28)). However, the codes of Theorems 38 and 39 are essentially different from those given in [13], [12] because the condition of hash functions in Theorems 38 and 39 is weaker than the universal 2 condition, which is essentially employed in [13], [12]. That is, Theorems 38 and 39 require only $P(n)$-almost dual universal 2 hash function.

Further, the RHS of (141) in Theorem 39 is better than the RHS of (139). However, the protocol considered in (141) in Theorem 39 is different from that in (139) in Theorem 38. We have to randomize the code $C_1$ for (141) in Theorem 39 while the bound (139) in Theorem 38 is obtained with a fixed code $C_1$.

**X. SIMPLE EXAMPLE**

As a simple example, we assume that $A = B = E = \mathbb{F}_p$ and for two distributions $P_X$ and $P'_X$ are given on $\mathcal{X} = \mathbb{F}_p$, the joint distribution is given as

\[
P_{A,B,E}(a, b, e) = \frac{1}{p} P'_X(b - a) P_X(e - a).
\]

(142)

Then,

\[
e^{-s H^G_{1/s}(A|E|P_{A,E})} = \sum_{x} \left( \frac{1}{p} \sum_{a} P_X(b - a)^{1/(1-s)} \right)^{1-s}
\]

\[
= \sum_{x} P_X(x)^{1/(1-s)}
\]

\[
e^{-s H_{1+s}(X|P_X)} = e^{-s H_{1+s}(X|P_X)}
\]

(143)

and

\[
e^{-s H_{1+s}(A|E|P_{A,E})} = e^{-s H_{1+s}(X|P_X)}
\]

Hence, $e_G(P_{A,E}|R)$ and $e_H(P_{A,E}|R)$ are simplified to

\[
e_G(P_{A,E}|R) = \max_{0 \leq s \leq 1/2} s(H_{1+s}(X|P_X) - R)
\]

(144)

\[
e_H(P_{A,E}|R) = \max_{0 \leq s \leq 1} s(H_{1+s}(X|P_X) - R).
\]

(145)

Similarly, we obtain

\[
H^G_{1/s}(A|E|P_{A,B}) = H_{1/s}(X|P'_X).
\]

(146)
Now, we choose the rate $R_1$ of size of code $C_1$. When $C_{X_n}$ is the $P(n)$-almost universal$_2$ code in $F^n_p$ with the dimension $\left\lceil n \frac{R_2}{\log p} \right\rceil$, due to (129), the error probability can be bounded as

$$E_{X_n} P_e[P^n_{A,B}, C_{X_n}] \leq P(n)e^{-n(s(R_1 - \log p) - sH_{\frac{1}{n}}(X|P'_X))}.$$ 

That is,

$$\liminf_{n \to \infty} -\frac{1}{n} \log E_{X_n} P_e[P^n_{A,B}, C_{X_n}] \geq \max_{0 \leq s \leq 1} s(\log p - R_1) + sH_{\frac{1}{n}}(X|P'_X).$$

On the other hand, since $e_H(P_{A,E}|R) > e_G(P_{A,E}|R)$ due to (143) and (145), the randomization of error correcting code improves the evaluation of the quantity $I'(f_X(A_{1,n})|[A_n], E_n|P^n_{A,E})$. The difference between $e_H(P_{A,E}|R)$ and $e_G(P_{A,E}|R)$ is numerically evaluated in Fig 1.

![Fig. 1. Lower bounds of exponent. Normal line: $e_H(P_{A,E}|R)$, Thick line: $e_G(P_{A,E}|R)$ with $p = 2$, $P_X(0) = 0.9$, $P_X(1) = 0.1$.](image)

XI. Conclusion

We have derived upper bounds for the leaked information in the modified mutual information criterion and the $L_1$ distinguishability criterion when we apply an $\epsilon$-almost dual universal$_2$ hash function for privacy amplification. (Theorems 26 and 25 in Section VI). Then, we have derived lower bounds on their exponential decreasing rates in the i.i.d. setting. (Theorem 29 in Section VII). We have also applied our result to the case when we need error correction. In this case, we apply the privacy amplification after error correction as given in Subsection IX-A. Then, we have derived upper bounds for the information leaked with respect to the final keys in the respective criteria as well as upper bounds for the probability for disagreement in the final keys (Theorems 36 and 57 in Section IX). Applying them to the i.i.d. setting, we have derived lower bounds on their exponential decreasing rates. (Theorems 38 and 59 in Section IX).

We have rigorously compared the exponents by the rigorous smoothing of min-entropy and our approximate smoothing of Rénnyi entropy of order 2. That is, we have clarified the upper bounds of leaked information via the rigorous smoothing of min-entropy in the both criteria. That is, we have compared $\Delta_{d,2}(M, \epsilon|P_{A,E})$ and $\Delta_{d,\min}(M, \epsilon|P_{A,E})$ for Rénnyi entropy of order 2, and have done $\Delta_{I,2}(M, \epsilon|P_{A,E})$ and $\Delta_{I,\min}(M, \epsilon|P_{A,E})$ for modified mutual information criterion. We have derived the exponents of the upper bounds (Theorem 31 in Section VI), and have shown that the exponents are strictly worse than the exponents by our approximate smoothing of Rénnya entropy of order 2 (Lemma 32 in Section VI). This fact shows the importance of an approximate smoothing of Rénnyi entropy of order 2. The obtained exponents are summarized in Table II.

Due to Pinsker inequality and Inequality (41), the exponential convergence of one criterion yields the exponential convergence of the other criterion. However, we have shown that better exponential decreasing rates can be obtained by separate derivations. For example, our approximate smoothing of Rénnyi entropy of order 2 yields the exponent $e_G(P_{A,E}|R)$ for the $L_1$ distinguishability criterion, which yields the exponent $e_H(P_{A,E}|R)$ for the modified mutual information criterion by using Pinsker inequality. Similarly, our approximate smoothing of Rénnyi entropy of order 2 yields the exponent $e_G(P_{A,E}|R)$ for the modified mutual information criterion, which yields the exponent $e_H(P_{A,E}|R)$ for the $L_1$ distinguishability criterion by Inequality (41). Since $e_G(P_{A,E}|R) \geq e_H(P_{A,E}|R)$ and $e_H(P_{A,E}|R) \geq e_G(P_{A,E}|R)$, the exponents directly derived by our approximate smoothing of Rénnyi entropy of order 2 are better than the exponents derived from the combination of the exponent for the other criterion and the inequality.

We have also shown that the application of $\epsilon$-almost dual universal hash function attains the asymptotically optimal performance in the sense of the second order asymptotics as well as in that of the asymptotic equivocation rate. These
facts have been shown by using the approximate smoothing of min entropy. We can conclude that ε-almost dual universal hash functions are very a useful class of hash functions. Further, these discussions show that the approximate smoothing of min entropy is sufficiently powerful except for the exponential decreasing rate. That is, the exponential decreasing rate requires more delicate evaluation than other settings.

ACKNOWLEDGMENTS

The author is grateful to Dr. Toyohiro Tsurumaru, Dr. Shun Watanabe, Dr. Marco Tomamichel, Dr. Mario Berta, Dr. William Henry Rosgen, Dr. Li Ke, and Dr. Markus Grassl for a helpful comments. He is also grateful to the referee of the first version of [14] for informing the literatures [9], [10]. He is partially supported by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071 and the National Institute of Information and Communication Technology (NICT), Japan. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

APPENDIX A

MODIFIED MUTUAL INFORMATION CRITERION

It is natural to adopt a quantity expressing the difference between the true state and the ideal state $P_{\text{mix},A} \times P_E$ as a security criterion. However, there are several quantities expressing the difference between two states. Both $d'_1(A|E|P)$ and $I'(A|E|P)$ are characterized in this way. Here, we show that the modified mutual information $I'(A|E|P)$ can be derived in a natural way.

It is natural to assume the following condition for the security criterion $C(A;E|P)$ as well as the the permutation invariance on $A$ and $E$.

$\text{C1 Chain rule}$ $C(A,B|E|P) = C(B|E|P) + C(A|B,E|P)$.

$\text{C2 Linearity}$ When the supports of two marginal distributions $P_{E,1}$ and $P_{E,2}$ are disjoint as subsets of $E$, $C(A|E|\lambda P_1 + (1-\lambda)P_2) = \lambda C(A|E|P_1) + (1-\lambda)C(A|E|P_2)$.

$\text{C3 Range}$ $\log |A| \geq C(A|E|P) \geq 0$.

$\text{C4 Ideal case}$ $C(A|E|P_{\text{mix},A} \otimes P_E) = 0$.

$\text{C5 Normalization}$ $C(A|E||a|) = \log |A|$.

Unfortunately, the $L_1$ distinguishability does not satisfies C1 Chain rule. However, we have the following theorem.

Theorem 40: $C(A|E|P)$ satisfies all of the above properties if and only if $C(A|E|P)$ coincides with the modified mutual information criterion $I'(A|E|P) = \log |A| - H(A|E|P)$.

Hence, it is natural to adopt the modified mutual information criterion $I'(A|E|P)$ as a security criterion. In particular, if one emphasizes C1 Chain rule rather than the universal composability, it is better employ the modified mutual information criterion $I'(A|E|P)$.

Proof of Theorem 40: First, we show that the modified mutual information criterion $I'(A|E|P) = \log |A| - H(A|E|P)$ satisfies all of the above conditions. We can trivially check the conditions C4 Ideal case and C5 Normalization. We show other conditions. C1 Chain rule can be shown as follows.

$I'(A,B|E|P) = \log |A| + \log |B| - H(A,B,E|P) + H(E|P)$

$= \log |A| + \log |B| - H(B,E|P) + H(E|P) - H(A,B,E|P) + H(B,E|P)$

$= \log |A| + \log |B| - H(B,E|P) - H(A,B|E|P) = I'(A|B,E|P) + I'(B|E|P)$.

When two marginal distributions $P_{E,1}$ and $P_{E,2}$ are distinguishable on $E$,

$I'(A|E|\lambda P_1 + (1-\lambda)P_2) = \log |A| - H(A,E|\lambda P_1 + (1-\lambda)P_2) + H(E|\lambda P_1 + (1-\lambda)P_2)$

$= \log |A| - \lambda H(A,E|P_1) - (1-\lambda)H(A,E|P_2) - h(\lambda) + \lambda H(E|P_1) + (1-\lambda)H(E|P_2) + h(\lambda)$

$= \log |A| - \lambda H(A,E|P_1) - (1-\lambda)H(A,E|P_2) + \lambda H(E|P_1) + (1-\lambda)H(E|P_2)$

$= \lambda I'(A|E|P_1) + (1-\lambda)I'(A|E|P_2)$.
which implies C2 Linearity. $I'(A|E|P) = D(P||P_{a|x,A} \otimes P_E) \geq 0$. Since $H(A,E|P) \geq 0$, $I'(A|E|P)$ satisfies C3 Range. Thus, $I'(A|E|P)$ satisfies all of the above properties.

Next, we show that an quantity satisfying all of the above properties is the modified mutual information criterion $I'(A|E|P) = \log|A| - H(A|E|P)$. For this purpose, we focus on $\tilde{H}(A|E|P) := \log|A| - C(A|E|P)$. Due to C1 Linearity, we have

$$\tilde{H}(A|E|P) = \sum_e P_E(e) \tilde{H}(A|E|P_{A|E=e}).$$

Further, we see that the quantity $\tilde{H}(A|E|P_{A|E=e})$ satisfies Khinchin’s axioms [56] for entropy because of the remaining properties. Hence, we find that $\tilde{H}(A|E|P_{A|E=e}) = H(P_{A|E=e})$. Thus, $\tilde{H}(A|E|P)$ is equal to the conditional entropy $H(A|E|P)$. Hence, $C(A|E|P) = I'(A|E|P)$.

**APPENDIX B**

**PROOF OF LEMMA 6**

For $s \in (0, 1]$ and two functions $X(a)$ and $Y(a)$, the Hölder inequality

$$\sum_a X(a)Y(a) \leq \left( \sum_a |X(a)|^{1/(1-s)} \right)^{1-s} \left( \sum_a |Y(a)|^{1/s} \right)^{s}$$

holds. The equality holds only when $X(a)$ is a constant times of $Y(a)$. Substituting $P_{A,E}(a,e)$ and $(P_{A,E}(a,e)/P_E(e))$ to $X(a)$ and $Y(a)$, we obtain

$$e^{-sH_{1+s}(A|E|P_{A,E})} = \sum_e \sum_a P_{A,E}(a,e) \left( \frac{P_{A,E}(a,e)}{P_E(e)} \right)^s$$

$$\leq \sum_e \left( \sum_a P_{A,E}(a,e) \right)^{1/(1-s)} \left( \sum_a \frac{P_{A,E}(a,e)}{P_{E,normal}(e)} \right)^s$$

$$= \sum_e \left( \sum_a P_{A,E}(a,e) \right)^{1/(1-s)} \left( \sum_a \frac{P_{A,E}(a,e)}{P_{E,normal}(e)} \right)^s$$

$$= e^{-sH_{1+s}(A|E|P_{A,E})}$$

for $s \in (0, 1]$ because $\sum_a P_{A,E}(a,e)/P_{E,normal}(e) = \frac{P_E(e)}{P_{E,normal}(e)} \leq 1$. The equality condition holds only when $P_{A|E=e}$ is uniform distribution for all $e \in E$.

For $s \in [-1, 0)$ and two functions $X(a)$ and $Y(a)$, the reverse Hölder inequality [35]

$$\sum_a X(a)Y(a) \geq \left( \sum_a |X(a)|^{1/(1-s)} \right)^{1-s} \left( \sum_a |Y(a)|^{1/s} \right)^{s}$$

holds. The same substitution yields

$$e^{-sH_{1+s}(A|E|P_{A,E})} \geq e^{-sH^{(1+s)}_{1+s}(A|E|P_{A,E})}$$

for $s \in [-1, 0)$ because $\sum_a P_{A,E}(a,e)/P_{E,normal}(e) = (\frac{P_E(e)}{P_{E,normal}(e)})^s \geq 1$. The equality condition holds only when $P_{A|E=e}$ is uniform distribution for all $e \in E$.

**APPENDIX C**

**PROOF OF LEMMA 7**

For two non-negative functions $X(e)$ and $Y(e)$, the reverse Hölder inequality [35]

$$\sum_e X(e)Y(e) \geq \left( \sum_e X(e)^{1/(1+s)} \right)^{1+s} \left( \sum_e Y(e)^{-1/s} \right)^{-s}$$

holds for $s \in (0, \infty]$. Substituting $\sum_a P_{A,E}(a,e)^{1+s}$ and $Q_E(e)^{-s}$ to $X(e)$ and $Y(e)$, we obtain

$$e^{-sH_{1+s}(A|E|P_{A,E},Q_E)} = \sum_e \sum_a P_{A,E}(a,e)^{1+s}Q_E(e)^{-s}$$

$$\geq \left( \sum_e \sum_a P_{A,E}(a,e)^{1+s} \right)^{1+s} \left( \sum_e Q_E(e)^{-s} \right)^{-s}$$

$$= \left( \sum_e \sum_a P_{A,E}(a,e)^{1+s} \right)^{1+s} \left( \sum_e Q_E(e)^{-s} \right)^{-s}$$

$$= e^{-sH^{(1+s)}_{1+s}(A|E|P_{A,E})}$$
for $s \in (0, \infty]$. Since the equality holds when $Q_E(e) = (\sum_a P_{A,E}(a,e)^{1+s})^{1/(1+s)} / \sum_a (\sum_a P_{A,E}(a,e)^{1+s})^{1/(1+s)}$, we obtain

$$\min_{Q_E} e^{-sH_{1+s}(A|E)P_{A,E}(Q_E)} = e^{-sH_{1+s}^G(A|E)P_{A,E}},$$

which implies (26) with $s \in (0, \infty]$.

For two non-negative functions $X(e)$ and $Y(e)$, the Hölder inequality

$$\sum_e X(e)Y(e) \leq (\sum_e X(e)^{1/(1+s)})^{1+s}(\sum_e Y(e)^{-1/s})^{-s}$$

holds for $s \in [-1, 0)$. The same substitution yields

$$e^{-sH_{1+s}(A|E)P_{A,E}(Q_E)} \leq e^{-sH_{1+s}^G(A|E)P_{A,E}}$$

for $s \in [-1, 0)$. Hence, similarly we obtain (26) with $s \in [-1, 0)$.

**APPENDIX D**

**PROOF OF LEMMA 9**

**Lemma 41:**

$$-\frac{d}{ds}H_{1+s}^G(A|E)P_{A,E}$$

$$= \sum_{a,e} P_{A,E:s}(a,e) \left( \log P_{A|E}(a|e) - \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) + \phi\left(\frac{s}{1+s} |A|E P_{A,E}\right),$$

$$\left(1 + s\right) \sum_{a,e} P_{A,E:s}(a,e) \left( \frac{1}{1+s} \log P_{A|E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right)^2$$

$$- \left(1 + s\right) \left( \sum_{a,e} P_{A,E:s}(a,e) \left( \frac{1}{1+s} \log P_{A|E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) \right)^2.$$ (147)

Hence, when we regard $H_{1+s}^G(A|E)P_{A,E}$ as $H(A|E)P_{A,E}$ and $P_{A|E=e}$ is not a uniform distribution for an element $e \in \mathcal{E}$, the function $s \mapsto -sH_{1+s}^G(A|E)P_{A,E}$ is strictly convex in $(-1, \infty)$. That is, the map $s \mapsto sH_{1+s}^G(A|E)P_{A,E}$ is strictly concave and then the map $s \mapsto H_{1+s}^G(A|E)P_{A,E}$ is strictly monotonically decreasing for $s \in (-1, \infty)$.

**Proof:** We define

$$\varphi(s) := \sum_e P_E(e)(\sum_a P_{A|E}(a|e)^{1+s})^{1/s}.$$

Then,

$$\frac{d\varphi(s)}{ds}$$

$$= \sum_{a,e} P_{A,E:s}(a,e) \left( \frac{1}{1+s} \log P_{A|E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right)$$

$$= \varphi(s) \sum_{a,e} P_{A,E:s}(a,e) \left( \frac{1}{1+s} \log P_{A|E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right).$$

Since

$$-\frac{d}{ds}sH_{1+s}^G(A|E)P_{A,E}$$

$$= \phi\left(\frac{s}{1+s} |A|E P_{A,E}\right) + (1 + s) \frac{d\varphi(s)}{ds} \varphi(s)^{-1},$$

we obtain (147).
Next, we show (148). Since

\[
\begin{align*}
    \frac{d^2 \varphi(s)}{ds^2} & = \sum_{a,e} \frac{P_{A,E}(a|e)^{1+s} P_E(e)}{(\sum_a P_{A,E}(a|e)^{1+s})^{1+s}} \left( \frac{1}{1+s} \log P_{A,E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A,E}(a|e)^{1+s}) \right)^2 \\
    & \quad + \sum_{a,e} \frac{P_{A,E}(a|e)^{1+s} P_E(e)}{(\sum_a P_{A,E}(a|e)^{1+s})^{1+s}} \left( - \frac{2}{(1+s)^2} \log P_{A,E}(a|e) + \frac{2}{(1+s)^3} \log(\sum_a P_{A,E}(a|e)^{1+s}) \right) \\
    & = \varphi(s) \left( \sum_{a,e} P_{A,E,s}(a,e) \left( \frac{1}{1+s} \log P_{A,E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A,E}(a|e)^{1+s}) \right)^2 \\
    & \quad - \frac{2}{(1+s)} \frac{d \varphi(s)}{ds} \right),
\end{align*}
\]

We have

\[
\begin{align*}
    \frac{d^2}{ds^2} & (1+s) \phi \left( \frac{s}{1+s} |A| E |P_{A,E} \right) \\
    & = (1+s) \frac{d^2}{ds^2} \phi \left( \frac{s}{1+s} |A| E |P_{A,E} \right) + 2 \frac{d}{ds} \phi \left( \frac{s}{1+s} |A| E |P_{A,E} \right) \\
    & = (1+s) \varphi(s) \frac{d^2 \varphi(s)}{ds^2} \varphi(s)^2 + 2 \frac{d \varphi(s)}{ds} \varphi(s) \\
    & = (1+s) \sum_{a,e} P_{A,E,s}(a,e) \left( \frac{1}{1+s} \log P_{A,E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A,E}(a|e)^{1+s}) \right)^2 \\
    & \quad - (1+s) \left( \sum_{a,e} P_{A,E,s}(a,e) \left( \frac{1}{1+s} \log P_{A,E}(a|e) - \frac{1}{(1+s)^2} \log(\sum_a P_{A,E}(a|e)^{1+s}) \right) \right)^2,
\end{align*}
\]

which implies (148).

**APPENDIX E**

PROOFS OF COMPARISONS OF EXPONENTS

A. Proof of Lemma 30

Inequality (108) can be shown from (25). Lemma 7 yields that

\[
\begin{align*}
    \frac{1}{2} e_G(P_{A,E}|R) \\
    & = \max_{0 \leq s \leq 1} \frac{s}{2} H_{1+s}(A|E|P_{A,E}) - \frac{s}{2} R \\
    & \leq \max_{0 \leq s \leq 1} \frac{s}{2} H_{1+s}^G(A|E|P_{A,E}) - \frac{s}{2} R \\
    & = \max_{0 \leq t \leq 1/2} \frac{t}{2(1-t)} \left( H_{1+s}^G(A|E|P_{A,E}) - R \right) \\
    & \leq \max_{0 \leq t \leq 1/2} t(H_{1+s}^G(A|E|P_{A,E}) - R) \\
    & = e_G(P_{A,E}|R),
\end{align*}
\]

where \( t = \frac{s}{1+s} \), i.e., \( s = \frac{t}{1-t} \). Inequality (149) follows from the non-negativity of the RHS of (149) and the inequality \( \frac{1}{2(1-t)} \leq 1 \).

B. Proof of Lemma 32

Lemma 9 implies that

\[
H_{1+s}^G(A|E|P_{A,E}) < H_{1+s}^G(A|E|P_{A,E})
\]
Choosing \( t = \frac{s}{1+s} \), we have
\[
\max_{0 \leq s} \left( \frac{sH_{1+t}^G(A|E|P_{A,E}) - sR}{1 + 2s} \right) = \max_{0 \leq s \leq 1} \left( \frac{t(H_{1+t}^G(A|E|P_{A,E}) - sR)}{1 + t} \right) < \max_{0 \leq t \leq 1} \left( \frac{t(H_{1+t}^G(A|E|P_{A,E}) - sR)}{1 + t} \right),
\]
which implies (114). Similarly, since \( H_{1+t}(A|E|P_{A,E}) \) is strictly monotonically increasing with respect to \( t \),
\[
\max_{0 \leq s} \left( \frac{sH_{1+s}(A|E|P_{A,E}) - sR}{1 + s} \right) = \max_{0 \leq s \leq 1} \left( \frac{tH_{1+s}(A|E|P_{A,E}) - tR}{1 + t} \right) < \max_{0 \leq t \leq 1} \left( \frac{tH_{1+t}(A|E|P_{A,E}) - tR}{1 + t} \right),
\]
which implies (115).

When \( P_{A|E=e} \) is a uniform distribution for any element \( e \in \mathcal{E} \), \( H_{1+t}(A|E|P_{A,E}) \) and \( H_{1+t}^G(A|E|P_{A,E}) \) do not depend on \( t \).

Hence, we obtain
\[
\max_{0 \leq s} \left( \frac{sH_{1+t}(A|E|P_{A,E}) - sR}{1 + 2s} \right) = \max_{0 \leq t \leq 1} \left( \frac{tH_{1+t}(A|E|P_{A,E}) - tR}{1 + t} \right) = \frac{\max_{0 \leq t \leq 1} [tH_{1+t}(A|E|P_{A,E}) - tR]}{1 + t} = H(A|E|P_{A,E}) - R,
\]
which imply the equalities \( e_G(P_{A,E}|R) = e_G(P_{A,E}|R) \) and \( e_H(P_{A,E}|R) = e_H(P_{A,E}|R) \).

### Appendix F

**Rigorous smoothing of min entropy**

**A. Proof of (113) of Theorem 51**

First, \( \Delta_{t,\min}(e^{nR},\varepsilon|P_{A,E}^n) \) is the upper bound by the rigorous smoothing of min entropy in the modified mutual information criterion as is mentioned in (84). Using the relation (29) in Theorem 27, we obtain
\[
\lim \inf_{n \to \infty} \frac{-1}{n} \log \Delta_{t,\min}(e^{nR},\varepsilon|P_{A,E}^n) \geq \max_{0 \leq s} \left( \frac{sH_{1+s}(A|E|P_{A,E}) - sR}{1 + s} \right). \tag{150}
\]

Now, we show the opposite inequality. Applying the Cramér Theorem (56), we obtain
\[
\lim_{n \to \infty} \frac{-1}{n} \log P_{A,E} \{ (a,c) \in \mathcal{A} \times \mathcal{E} | P_{A,E}^n(a|e) \geq 2e^{-nR'} \} = \max_{0 \leq s} sH_{1+s}(A|E|P_{A,E}) - sR'. \tag{151}
\]

Since \( sH_{1+s}(A|E|P_{A,E}) - sR' \) is monotone decreasing with respect to \( R' \) and \( R' - R \) is monotone increasing with respect to \( R' \), we have
\[
\max_{R'} \min_{s} \left( \frac{sH_{1+s}(A|E|P_{A,E}) - sR', R' - R} {1 + s} \right) \geq \frac{sH_{1+s}(A|E|P_{A,E}) - sR}{1 + s} \geq \frac{sH_{1+s}(A|E|P_{A,E}) - sR}{1 + s}. \tag{152}
\]

Using the lower bound (29) in Theorem 23, with \( c = 2 \), (151), and (152), we have
\[
\lim_{n \to \infty} \frac{-1}{n} \log \min_{0 < \eta < \infty} \left( \eta, nR + e^{nR - H_{\min}(A|E)|P_{A,E}^n}) \right) \leq \frac{-1}{n} \log \min_{0 < \eta < \infty} \left( \eta, 2nR + e^{nR - H_{\min}(A|E)|P_{A,E}^n}) \right) \leq \max_{0 \leq s} \frac{sH_{1+s}(A|E|P_{A,E}) - sR', R' - R} {1 + s} \geq \frac{sH_{1+s}(A|E|P_{A,E}) - sR}{1 + s}. \tag{153}
\]

Hence, we obtain (113).
B. Proof of (112) of Theorem 31

The quantity \( \Delta_{d, \min}(e^n R, \varepsilon[P^n_{A,E}]) \) is the upper bound by rigorous smoothing of min entropy in the \( L_1 \) distinguishability criterion as is mentioned in (33). Using the relation (98) in Theorem 27 we obtain

\[
\liminf_{n \to \infty} \frac{-1}{n} \log \Delta_{d, \min}(e^n R, \varepsilon[P^n_{A,E}]) \geq \max_{0 \leq s} \frac{s H_{1+s}^G(A|E) P_{A,E} - s R}{1 + 2s}.
\]

(154)

We show the opposite inequality in (112) by using the following lemma. The proof of Lemma 42 will be shown latter.

Lemma 42: The following inequality

\[
\lim_{n \to \infty} \frac{-1}{n} \log \min_{Q_{E,n}} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n}(e)} \geq 2e^{-n R'} \} \leq \max_{0 \leq s} s H_{1+s}^G(A|E)P_{A,E} - s R'.
\]

(155)

Using (155) in Lemma 42 and the lower bound (89) in Theorem 25 with \( c = 2 \), we obtain

\[
\lim_{n \to \infty} \frac{-1}{n} \log(\min_{\varepsilon_2 > 0} 2\varepsilon_1 + e^{\varepsilon_2 n R} e^{-\varepsilon_2 H_{1+\varepsilon_2}^G(A|P^n_{A,E})})
\]

\[
\leq \lim_{n \to \infty} \frac{-1}{n} \log(\min_{R'} \min_{Q_{E,n}} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n}(e)} \geq 2e^{-n R'} \} + e^{\varepsilon_2 n (R - R')})
\]

\[
= \max_{R'} \lim_{n \to \infty} \frac{-1}{n} \log(\min_{Q_{E,n}} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n}(e)} \geq 2e^{-n R'} \} + e^{\varepsilon_2 n (R - R')})
\]

\[
= \max_{R'} \min_{0 \leq s} \left( \lim_{n \to \infty} \frac{-1}{n} \log(\min_{Q_{E,n}} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n}(e)} \geq 2e^{-n R'} \}), \frac{R' - R}{2} \right)
\]

\[
\leq \max_{R'} \min_{0 \leq s} \left( \max \min_{s} s H_{1+s}^G(A|E)P_{A,E} - s R', \frac{R' - R}{2} \right)
\]

\[
= \max_{R'} \min_{0 \leq s} \left( \max \min_{s} s H_{1+s}^G(A|E)P_{A,E} - s R', \frac{R' - R}{2} \right)
\]

\[
= \max_{R'} \min_{0 \leq s} \left( s H_{1+s}^G(A|E)P_{A,E} - s R', \frac{R' - R}{2} \right).
\]

(156)

Further, \( s H_{1+s}^G(A|E)P_{A,E} - s R' \) is monotone increasing with respect to \( R' \) and \( \frac{R' - R}{2} \) is monotone decreasing with respect to \( R' \). Solving the equation \( s H_{1+s}^G(A|E)P_{A,E} - s R' = \frac{R' - R}{2} \) with respect to \( R' \), we have \( R' = \frac{2s H_{1+s}^G(A|E)P_{A,E} + R}{1 + 2s} \), which implies that

\[
\max_{R'} \min_{0 \leq s} \left( s H_{1+s}^G(A|E)P_{A,E} - s R', \frac{R' - R}{2} \right) = \frac{s H_{1+s}^G(A|E)P_{A,E} - s R}{1 + 2s}.
\]

Thus,

\[
\max_{R'} \min_{0 \leq s} \left( s H_{1+s}^G(A|E)P_{A,E} - s R', \frac{R' - R}{2} \right) = \frac{s H_{1+s}^G(A|E)P_{A,E} - s R}{1 + 2s}.
\]

Hence, we obtain (112).

Proof of Lemma 42: We show Lemma 42 by using Lemmas 43 and 43 which will be given latter. For any distribution \( Q_{E,n} \), we define the permutation invariant distribution \( Q_{E,n, \text{inv}} \) by

\[
Q_{E,n, \text{inv}}(e) := \sum_{g \in S_n} \frac{1}{n!} Q_{E,n}(g(e)),
\]

where \( S_n \) is the \( n \)-th permutation group and \( g(e) \) is the element permuted from \( e \in \mathcal{E}^n \) by \( g \in S_n \). Then, we have

\[
P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n}(e)} \geq 2e^{-n R'} \}
\]

\[
= P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | P^n_{A,E}(a,e) \geq 2e^{-n R'} Q_{E,n}(e) \}
\]

\[
\geq \frac{1}{2} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | P^n_{A,E}(a,e) \geq 4e^{-n R'} Q_{E,n, \text{inv}}(e) \}
\]

\[
= \frac{1}{2} P^n_{A,E} \{(a,e) \in \mathcal{A}^n \times \mathcal{E}^n | \frac{P^n_{A,E}(a,e)}{Q_{E,n, \text{inv}}(e)} \geq 4e^{-n R'} \},
\]
where the inequality follows from Lemma 43. Here, we denote the set of types of \( E \) by \( T_{n,E} \). For any element \( Q_E \in T_{n,E} \), we denote the uniform distribution over the subset of elements whose type is \( Q_E \) by \( \bar{Q}_E \). Now, we define the distribution

\[
Q_{E,n,\text{inv,mix}}(e) := \frac{1}{|T_{n,E}|} \sum_{Q_E \in T_{n,E}} \hat{Q}_E(e).
\]

Since \( Q_{E,n,\text{inv}} \leq |T_{n,E}|Q_{E,n,\text{inv,mix}}(e) \), we have

\[
\frac{1}{2} P^n_{A,E}(\{(a, e) \in A^n \times E^n | P^n_{A,E}(a, e) \geq 4e^{-nR'}Q_{E,n,\text{inv}}(e)\})  \\
\geq \frac{1}{2} P^n_{A,E}(\{(a, e) \in A^n \times E^n | P^n_{A,E}(a, e) \geq 4|T_{n,E}|e^{-nR'}Q_{E,n,\text{inv,mix}}(e)\}).
\]

For given sequence \( (a, e) \in A \times E \), we denote the type of \( (a, e) \) by \( P^n_{A,E} \) and its marginal distribution over \( E \) of \( P^n_{A,E} \) by \( P^n_{E} \). Then, \( P^n_{A,E}(a, e) = e^{-n(D(P^n_{A,E}\|P^T_{A,E})+H(P^n_{A,E}))} \) and \( |T_{n,E}|Q_{E,n,\text{inv,mix}}(e) = e^{-nH(P^n_{E})} \). That is, the condition \( P^n_{A,E}(a, e) \geq 4|T_{n,E}|e^{-nR'}Q_{E,n,\text{inv,mix}}(e) \) is equivalent to the condition \( D(P^n_{A,E}\|P_{A,E}) + H(P^n_{A,E}) \leq \frac{\log 4}{n} + H(P^n_{E}) + R' \). We denote the set of sequences whose types are \( P^n_{A,E} \) by \( T_{P^n_{A,E}} \). Hence,

\[
\frac{1}{2} P^n_{A,E}(\{(a, e) \in A^n \times E^n | P^n_{A,E}(a, e) \geq 4|T_{n,E}|e^{-nR'}Q_{E,n,\text{inv,mix}}(e)\}) = \frac{1}{2} P^n_{A,E}(T_{P^n_{A,E}}) \]

\[
\geq \frac{1}{2} P^n_{A,E}(T_{P^n_{A,E}}).
\]

Since \( P^n_{A,E}(T_{P^n_{A,E}}) \geq e^{-nD(P^n_{A,E}\|P_{A,E})} \), taking the limit, we have

\[
\lim_{n \to \infty} \frac{-1}{n} \log \frac{1}{2} P^n_{A,E}(\{(a, e) \in A^n \times E^n | P^n_{A,E}(a, e) \geq 4|T_{n,E}|e^{-nR'}Q_{E,n,\text{inv,mix}}(e)\}) \leq \max_{P^n_{A,E}} D(P^n_{A,E}\|P_{A,E}) + D(P^n_{A,E}\|P_{A,E}) + H(P^n_{A,E}) \leq R' + \max_{P^n_{A,E}} D(P^n_{A,E}\|P_{A,E}) + H(A|E|P^n_{A,E}) \leq R'.
\]

Hence, combining Lemma 45, we obtain (155).

Lemma 43: The relation

\[
P^n_{A}\{a \in A^n | c \geq f(a)\} \geq \frac{1}{2} P_{\text{mix},A}\{a \in A^n | c \geq \frac{1}{n!} \sum_{g \in S_n} f(g(a))\}  \tag{157}
\]

holds for any function \( f \).

Proof: Lemma 43 can be shown by applying Lemma 44 to all of distributions conditioned with type.

Lemma 44: The relation

\[
P_{\text{mix},A}\{a|c \geq f(a)\} \geq \frac{1}{2} P_{\text{mix},A}\{a|c \geq \frac{1}{|A|} \sum_{a} f(a)\} \tag{158}
\]

holds for any function \( f \).

Proof: Markov inequality implies that

\[
P_{\text{mix},A}\{a|c < f(a)\} \leq \frac{1}{c} \frac{1}{|A|} \sum_{a} f(a).
\]

When \( c \geq \frac{2}{|A|} \sum_{a} f(a) \), \( 1 - \frac{1}{c} \frac{1}{|A|} \sum_{a} f(a) \) is greater than \( \frac{1}{2} \). Hence,

\[
P_{\text{mix},A}\{a|c \geq f(a)\} = 1 - P_{\text{mix},A}\{a|c < f(a)\} \geq 1 - \frac{1}{c} \frac{1}{|A|} \sum_{a} f(a) \geq \frac{1}{2} P_{\text{mix},A}\{a|c \geq \frac{2}{|A|} \sum_{a} f(a)\}.
\]

Lemma 45: The relation

\[
\max_{P^n_{A,E}} \{ D(P^n_{A,E}\|P_{A,E}) + H(A|E|P^n_{A,E}) \leq R' \}
\]

\[
= \max_{0 \leq s} s H^G_{1+s}(A|E|P_{A,E}) - s R'. \tag{159}
\]

holds.
Proof: We show Lemma 45 by using Lemma 41 which will be given latter. We employ a generalization of the method used in [62, Appendix D]. First, we define the distribution $P_{A,E,s}(a,e)$ as

$$P_{A,E,s}(a,e) := \frac{P_{A|E}(a|e)P_{E}(e)}{(\sum_a P_{A|E}(a|e)^{1+s})^{(1/s)} (\sum_e P_{E}(e)(\sum_a P_{A|E}(a|e)^{1+s})^{1/s})}.$$ 

That is, we have

$$P_{A|E,s}(a|e) = \frac{P_{A|E}(a|e)^{1+s}}{\sum_a P_{A|E}(a|e)^{1+s}},$$

$$P_{E,s}(e) = \frac{P_{E}(e)(\sum_a P_{A|E}(a|e)^{1+s})^{1/s}}{(\sum_e P_{E}(e)(\sum_a P_{A|E}(a|e)^{1+s})^{1/s})}.$$

Hence,

$$D(P_{A,E,s}||P_{A,E}) = \sum_{a,e} P_{A,E,s}(a,e) \left( s \log P_{A|E}(a|e) - \frac{s}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E}),$$

$$H(A|E|P_{A,s}) = \sum_{a,e} P_{A,E,s}(a,e) \left( -(1+s) \log P_{A|E}(a|e) + \log(\sum_a P_{A|E}(a|e)^{1+s}) \right),$$

$$D(P_{A|E,s}||P_{A,E}) + H(A|E|P_{A,s}),$$

$$= \sum_{a,e} P_{A,E,s}(a,e) \left( -\log P_{A|E}(a|e) + \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E}).$$

Given $s \geq 0$, we choose an arbitrary distribution $P'_{A,E}$ such that

$$D(P_{A,E}'||P_{A,E}) = D(P_{A,E}'||P_{A,E}).$$

Since

$$D(P'_{A,E}||P_{A,E}) = \sum_{a,e} P'_{A,E}(a,e) \left( \log P'_{A,E}(a,e) - \log P_{A,E}(a,e) \right),$$

$$D(P'_{A,E}'||P_{A,E}) = \sum_{a,e} P'_{A,E}(a,e) \left( \log P'_{A,E}(a,e) - -(1+s) \log P_{A|E}(a|e) - \log P_{E}(e) \right) + \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E}),$$

we have

$$D(P'_{A,E}||P_{A,E,s}) = D(P'_{A,E}||P_{A,E,s}) + D(P_{A,E,s}||P_{A,E}) - D(P'_{A,E}||P_{A,E}),$$

$$= \sum_{a,e} P_{A,E,s}(a,e) \left( s \log P_{A|E}(a|e) - \frac{s}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E})$$

$$- \sum_{a,e} P'_{A,E}(a,e) \left( s \log P_{A|E}(a|e) - \frac{s}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E})$$

$$= \sum_{a,e} (P_{A,E,s}(a,e) - P'_{A,E}(a,e)) \left( s \log P_{A|E}(a|e) - \frac{s}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right).$$

Hence,

$$H(A|E|P_{A,s}) - H(A|E|P'_{A,E}) + D(P'_{E}||P_{E,s})$$

$$= \sum_{a,e} P_{A,E,s}(a,e) \left( -\log P_{A|E}(a|e) + \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) + \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E})$$

$$- \sum_{a,e} P'_{A,E}(a,e) \left( -\log P_{A|E}(a|e) + \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right) - \frac{s}{1+s} H_{1+s}^G(A|E|P_{A,E})$$

$$= \sum_{a,e} (P_{A,E,s}(a,e) - P'_{A,E}(a,e)) \left( -\log P_{A|E}(a|e) + \frac{1}{1+s} \log(\sum_a P_{A|E}(a|e)^{1+s}) \right)$$

$$= -sD(P'_{A,E}||P_{A,E,s}) \leq 0.$$
Since $D(P_E^t \| P_E; s) \geq 0$, we have $H(A|E|P_{A,E,s}) \leq H(A|E|P_{A,E})$, which implies

$$H(A|E|P_{A,E,s}) + D(P_{A,E,s} \| P_{A,E}) \leq H(A|E|P_{A,E}) + D(P_{A,E}^t \| P_{A,E}).$$

Since the map $s \mapsto D(P_{A,E,s} \| P_{A,E})$ is continuous, we have

$$\min_{s \geq 0} \{D(P_{A,E} \| P_{A,E})|D(P_{A,E}^t \| P_{A,E}) + H(A|E|P_{A,E}) \leq R'\}$$

$$= \min_{s \geq 0} \{D(P_{A,E,s} \| P_{A,E})|D(P_{A,E,s} \| P_{A,E}) + H(A|E|P_{A,E,s}) \leq R'\}.$$ 

Now, we choose $s_0 \geq 0$ such that

$$D(P_{A,E}^t \| P_{A,E}) + H(A|E|P_{A,E})$$

$$= \sum_{a,e} P_{A,E}^t(a,e) \left( - \log P_{A,E}(a|e) + \frac{1}{1 + s_0} \log \left( \sum_a P_{A,E}(a|e)^{1+s_0} \right) \right) + \frac{s_0}{1 + s_0} H_{1+s_0}^G(A|E)P_{A,E}$$

$$= R',$$ 

which implies that

$$\sum_{a,e} P_{A,E}^t(a,e) \left( - \log P_{A,E}(a|e) + \frac{1}{1 + s_0} \log \left( \sum_a P_{A,E}(a|e)^{1+s_0} \right) \right) = R' - \frac{s_0}{1 + s_0} H_{1+s_0}^G(A|E)P_{A,E}.$$ 

Then,

$$\min_{s \geq 0} \{D(P_{A,E,s} \| P_{A,E})|D(P_{A,E,s} \| P_{A,E}) + H(A|E|P_{A,E,s}) \leq R'\}$$

$$= \sum_{a,e} P_{A,E}^t(a,e) \left( s_0 \log P_{A,E}(a|e) - \frac{s_0}{1 + s_0} \log \left( \sum_a P_{A,E}(a|e)^{1+s_0} \right) \right) + \frac{s_0}{1 + s_0} H_{1+s_0}^G(A|E)P_{A,E}$$

$$= - s_0 \sum_{a,e} P_{A,E}^t(a,e) \left( - \log P_{A,E}(a|e) + \frac{1}{1 + s_0} \log \left( \sum_a P_{A,E}(a|e)^{1+s_0} \right) \right) + \frac{s_0}{1 + s_0} H_{1+s_0}^G(A|E)P_{A,E}$$

$$= - s_0 (R' + \phi\left( \frac{s_0}{1 + s_0} |A|E|P_{A,E} \right)) + \frac{s_0}{1 + s_0} H_{1+s_0}^G(A|E)P_{A,E}$$

$$= - s_0 R' + s_0 H_{1+s_0}^G(A|E)P_{A,E}$$

$$= \max_{s \geq 0} - s R' + s H_{1+s}^G(A|E)P_{A,E},$$

where the reason of the equation is the following. Due to Lemma 41, the function $s \mapsto - s H_{1+s}^G(A|E)P_{A,E}$ is convex, and $- R' = - \frac{d}{ds} s H_{1+s}^G(A|E)P_{A,E}$. Then, we obtain (159). \[\Box\]

**APPENDIX G**

**PROOF OF PROPOSITION 14**

First, remember that $A$ is a vector space $\mathbb{F}_q^n$ and $\mathcal{E}$ is a general discrete set. We make preparation before our proof of Proposition 14. We define the $\ell^2$ norm over the space $L^2(A \times \mathcal{E})$ as

$$\|f\|_2^2 := \sum_{a \in A, e \in \mathcal{E}} |f(a,e)|^2, \quad \forall f \in L^2(A \times \mathcal{E}).$$

(160)

Then, we define the discrete Fourier transform $\mathcal{F}$ on $L^2(A \times \mathcal{E})$ as

$$\mathcal{F}(f)(a', e) := q^{\frac{d}{2}} \sum_{a \in A} (d'a) f(a, e), \quad \forall f \in L^2(A \times \mathcal{E}), \forall a', \forall e \in \mathcal{E},$$

(161)

which satisfies $||\mathcal{F}f||_2 = ||f||_2$. For $\forall f, g \in L^2(A \times \mathcal{E})$, the convolution $f \ast g$:

$$f \ast g(a,e) := \sum_{a' \in A} f(a - a', e)g(a', e).$$

(162)

satisfies

$$\mathcal{F}(f \ast g)(a, e) = q^{\frac{d}{2}} \mathcal{F}(f)(a, e)\mathcal{F}(g)(a, e).$$

(163)

We prepare the following lemma.

**Lemma 46:** When $f_{P_{A,E;Q_E}} \in L^2(A \times \mathcal{E})$ is defined as

$$f_{P_{A,E;Q_E}}(a, e) := P_{A,E}(a, e)Q_E(e)^{-\frac{1}{2}}.$$ 

(164)
we have
\[ \|f_{P_{A,E},Q_E}\|_2^2 = e^{-H_2(A|E|P_{A,E}\|Q_E)} \]  \tag{165} 
\[ \sum_{e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(0,e)|^2 = e^{-\psi(1|P_E\|Q_E)} \tag{166} \]
\[ \sum_{a \neq 0 \in \mathcal{A},e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(a,e)|^2 = d_2(A|E|P_{A,E}\|Q_E). \tag{167} \]

**Proof:** (165) and (166) are shown as follows.
\[ \|f_{P_{A,E},Q_E}\|_2^2 = \sum_{a,e} (P_{A,E}(a,e)Q_E(e)^{-\frac{1}{2}})^2 = e^{-H_2(A|E|P_{A,E}\|Q_E)} \]
\[ \sum_{e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(0,e)|^2 = \sum_{e} (\sum_{a} P_{A,E}(a,e)Q_E(e)^{-\frac{1}{2}})^2 = \sum_{e} (P_E(e)Q_E(e)^{-\frac{1}{2}})^2 = e^{-\psi(1|P_E\|Q_E)}. \]

(167) is shown as follows.
\[ \sum_{a \neq 0 \in \mathcal{A},e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(a,e)|^2 = \|\mathcal{F}(f_{P_{A,E},Q_E})\|_2^2 - \sum_{e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(0,e)|^2 \]
\[ = \|f_{P_{A,E},Q_E}\|_2^2 - \sum_{e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E})(0,e)|^2 \]
\[ = e^{-H_2(A|E|P_{A,E}\|Q_E)} - e^{-\psi(1|P_E\|Q_E)} = d_2(A|E|P_{A,E}\|Q_E). \]

**Proof of Proposition 12** Now, we choose \( g_X \in L^2(\mathcal{A} \times \mathcal{E}) \) as
\[ g_X(a,e) := P_{W_X}(a). \tag{168} \]
Then,
\[ f_{P_{A,E},Q_E} * g_X = f_{P_{A,E} * P_{W_X},Q_E}. \tag{169} \]
The assumption yields that
\[ \mathbb{E}_X|\mathcal{F}(g_X)(a,e)|^2 = \mathbb{E}_X|q^{-\frac{1}{2}} \sum_{a \in \mathcal{A}} \omega_p^{(a',a)} P_{W_X}(a)|^2 \leq \delta^2 q^{-n} \tag{170} \]
for \( a' \neq 0 \in \mathcal{A} \). Hence,
\[ \mathbb{E}_X d_2(A|E|P_{A,E} * P_{W_X}\|Q_E) \leq \mathbb{E}_X \sum_{a \neq 0 \in \mathcal{A},e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E} * P_{W_X},Q_E})(a,e)|^2 \tag{b} \]
\[ \mathbb{E}_X \sum_{a \neq 0 \in \mathcal{A},e \in \mathcal{E}} |\mathcal{F}(f_{P_{A,E},Q_E} * g_X)(a,e)|^2 \leq \mathbb{E}_X \sum_{a \neq 0 \in \mathcal{A},e \in \mathcal{E}} |q^{\frac{1}{2}} \mathcal{F}(f_{P_{A,E},Q_E})(a,e) \mathcal{F}(g_X)(a,e)|^2 \tag{c} \]
\[ \leq \delta^2 \mathbb{E}_X \sum_{a \neq 0,e} |\mathcal{F}(f_{P_{A,E},Q_E})(a,e)|^2 \tag{d} \]
which shows (54) and (55). Here, (a), (b), (c), (d), and (e) follow from (167), (169), (163), (170), and (167), respectively.

**References**

[1] L. Carter and M. Wegman, “Universal classes of hash functions,” *J. Comput. System Sci.*, vol. 18, No. 2, 143–154, 1979.
[2] M. N. Wegman and J. L. Carter, “New Hash Functions and Their Use in Authentication and Set Inequality,” *J. Comput. System Sci.*, vol. 22, pp.265-279 (1981).
[3] Y. Mansour, N. Nisan, and P. Tiwari, “The Computational Complexity of Universal Hashing,” in STOC ’90, Proceedings of the twenty-second annual ACM symposium on Theory of computing, pp.235-243 (1990).
[4] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, “Generalized privacy amplification,” *IEEE Trans. Inform. Theory*, vol. 41, 1915-1923 (1995).
[5] G. H. Golub, and C. F. Van Loan, *Matrix Computation*, Third Edition, The Johns Hopkins University Press, 1996.
[6] U. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inform. Theory*, vol. 39, 733–742, 1993.
[7] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography part 1: Secret sharing,” *IEEE Trans. Inform. Theory*, vol. 39, no. 4, 1121–1132, 1993.
[8] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.
[9] Y. Dodis and A. Smith, “Correcting Errors Without Leaking Partial Information,” STOC 2005.
[10] S. Fehr and C. Schaffner, “Randomness Extraction via Delta-Biased Masking in the Presence of a Quantum Attacker,” Theory of Cryptography, Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19–21, 2008, pages, 465-48.
[60] S. Vembu and S. Verdu, “Generating random bits from an arbitrary source: fundamental limits,” IEEE Trans. Inform. Theory, vol. 41, no. 5, pp. 1322-1332, Sep. 1995.
[61] T.-S. Han and O. Uchida, “Source code with cost as a nonuniform random number generator,” IEEE Trans. Inform. Theory, vol. 46, no. 2, pp. 712-717, Mar. 2000.
[62] M. Hayashi, M. Koashi, K. Matsumoto, F. Morikoshi, and A. Winter, “Error exponents for entanglement concentration,” Journal of Physics A: Mathematical and General, vol. 36, no.2, pp.527-553 (2003).
[63] M. Iwamoto and J. Shikata, “Information Theoretic Security for Encryption Based on Conditional Renyi Entropies,” [http://eprint.iacr.org/2013/440] (2013)
[64] M. Muller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Renyi entropies: a new definition and some properties,” [arXiv:1306.3142] (2013).
[65] M. Hayashi, and R. Matsumoto, “Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages,” [arXiv:1202.1332] (2013).