Quantum wiretap channel with non-uniform random number and its exponent and equivocation rate of leaked information

Masahito Hayashi

Abstract

A usual code for quantum wiretap channel requires an auxiliary random variable subject to the perfect uniform distribution. However, it is difficult to prepare such an auxiliary random variable. We propose a code that requires only an auxiliary random variable subject to a non-uniform distribution instead of the perfect uniform distribution. Further, we evaluate the exponential decreasing rate of leaked information and derive its equivocation rate. For practical constructions, we also discuss the security when our code consists of a linear error correcting code.

I. INTRODUCTION

Secure communication is one of important topics in quantum information. Quantum wiretap channel model is one of most adequate formulations of this problem. In this model, the authorized sender Alice sends her message to the authorized receiver Bob and keeps its secrecy for the eavesdropper Eve. The sender is often assumed to be able to use an auxiliary random variable subject to the perfect uniform distribution. This assumption is usual in the quantum setting [7], [8] as well as in the classical setting [6], [2], [11], [20]. However, in some case, it is not necessarily easy to prepare a perfect uniform random number with low cost. In order to resolve it, it is needed to construct a code that properly works with an auxiliary non-uniform random number.

Further, the quantum wiretap channel capacity does not give an estimate for the leaked information with a finite-length code because the capacity is the asymptotic limit of securely transmitted rate. Hence, from a practical viewpoint, we need to treat the decreasing speed of leaked information rather than the quantum wiretap channel capacity. For this purpose, it is usual to focus on the exponential decreasing rate (exponent) of leaked information with a given sacrifice information rate in information theory. In this setting, we fix the sacrifice information rate, and treat the speed of convergence rate of leaked information by deriving its proper upper bounds. In fact, as another setting, similar to the second order analysis for classical channel coding [24], [25], given a fixed leaked information quantity, we can treat the asymptotic rate of coding length up to the second order. This setting essentially employs the central limit theorem. However, when the fixed leaked information quantity is quite small, the convergence of the central limit theorem is slow. Hence, the latter setting cannot yield a good bound for a finite-length code in such a case. Since the former setting gives upper bounds for leaked information, the former method can provide useful upper bounds for leaked information for a finite-length code.

In fact, Devetak [7] applied the random coding to the quantum wiretap channel model. For the evaluation for leaked information, he essentially used the quantum version of the channel resolvability. The original classical version is invented by Han and Verdú [21]. In the quantum version, we approximate the given output state by the output state of the input mixture of as a small number of input states as possible. Devetak [7] treated the output approximation when the input distribution is the perfect uniform distribution. Hence, his method does not work when the perfect uniform distribution is not available. After his achievement, the book [1] derived the exponent of leaked information, however, it also assumes that the perfect uniform distribution is available because its method is based on the quantum version of the channel resolvability. Further, there is a possibility to improve the exponent in [1] because the commutative case of the exponent in [1] is smaller than that by the previous result [20]. On the other hand, the previous paper [22] treated secure multiplex coding in the classical setting, which requires the generalization of the channel resolvability to the case when the uniform random number is not available.

In this paper, we treat the quantum wiretap channel model when the sender Alice cannot use the uniform random number. Alternatively, we assume that she knows the concrete form of the distribution of the auxiliary random number and its Rényi entropy of order 2, which can be regarded as the sacrifice information rate. In the quantum wiretap channel model, two classical-quantum channels are given. One is the channel from Alice’s classical information to Bob’s quantum state. The other is the channel from Alice’s classical information to Eve’s quantum state. In this model, we focus on the code given by random coding method. Under this protocol, we give an upper bound for leaked information with the quantum mutual information criterion. We also derive an upper bound for leaked information in terms of $L_1$ distinguishability. The both bounds go to zero exponentially when the generated key rate is less than the capacity. The exponent for the former can be regarded as the
quantum version of [20] because its commutative case coincides with that by [20]. The exponent for the latter is smaller than that by [19].

When the generated key rate is larger than the capacity, the leaked information does not go to zero. In this paper, we derive the minimum leaked information rate. That is, we calculate the maximum conditional entropy [6]. In the degraded case, we obtain its single-letterized formula. Further, in order to treat a more practical setting, we derive similar bounds when our error correction codes are restricted to linear codes. Finally, as a typical example, we treat the Pauli channel. In the classical case, more deeper analyses are given in [22].

This paper is organized as follows. In Section II we prepare quantum versions of information quantities and several fundamental inequalities for latter discussion. In Section III, we treat a non-uniform extension of quantum channel resolvability, which is a strong tool for quantum wiretap channel with an auxiliary non-uniform random number. In Section IV, we proceed to the quantum wiretap channel model and derive a lower bound of the exponent of leaked information, whose commutative version coincides with that by the previous paper [20]. In Section V, we treat the case when the sacrifice information rate is less than Eve’s mutual information. Then, we derive the equivocation rate. In Section VI, we treat the case when only liner codes are available. Finally, in Section VII, we discuss the case of Pauli channels. In Appendix D, we review a result concerning privacy amplification by universal2 hash functions, which is shown in [18].

II. INFORMATION QUANTITIES

A. Notations for distributions and states

In this section, we denote the classical probability space by the calligraphic capital letter (e.g., $\mathcal{X}$) and the corresponding random variable by the capital letter (e.g., $X$). We also denote the probability distribution on $\mathcal{X}$ by $P_X$. In order to describe the transition matrix from $\mathcal{V}$ to $\mathcal{X}$, we often use the letter $\Gamma$, in which, $\Gamma_v$ denotes the probability distribution on $\mathcal{X}$ for any $v \in \mathcal{V}$. Using these notations, we define a probability distribution on the composite system $\mathcal{X} \times \mathcal{V}$ by $(\Gamma \times P_V)(x, v) := \Gamma_v(x)P_V(x)$, and a probability distribution on the system $\mathcal{X}$ by $(\Gamma \circ P_V)(x) := \sum_{v \in \mathcal{V}} \Gamma_v(x)P_V(x)$. Since any function $f$ from $\mathcal{V}$ to $\mathcal{X}$ can be regarded as a transition matrix from $\mathcal{V}$ to $\mathcal{X}$, we can define the probability distribution $f \circ P_V$ on the composite system $\mathcal{X} \times \mathcal{V}$ and the probability distribution $f \circ P_V$ on the system $\mathcal{X}$ in the above way. When two distributions $P_X$ and $P_V$ are given, $P_X \times P_V$ expresses their independent product distribution on $\mathcal{X} \times \mathcal{V}$. For a given subset $\Omega$ of $\mathcal{X}$, we denote the uniform distribution on $\Omega$ by $P_{\text{mix},\Omega}$. For positive integers $M$ and $L$, we denote the sets $\{1, \ldots, M\}$ and $\{1, \ldots, L\}$ by $\mathcal{M}$ and $\mathcal{L}$, respectively.

Next, we introduce notations by using a classical-quantum channel $W_E : x \mapsto W_E|x$ from the classical system $\mathcal{X}$ to the quantum system $H_E$. When we consider only the probability system, we simply let it as $W : x \mapsto W_x$. For a simple treatment, we identify the state $\sum_{x \in \mathcal{X}} P_X(x)|x\rangle\langle x|$ with the distribution $P_X$. We define the state $W \times P_X := \sum_{x \in \mathcal{X}} W_x \otimes P_X(x)|x\rangle\langle x|$ on the composite system $H_E \otimes H_X$, and the state $W \circ P_X := \sum_{x \in \mathcal{X}} P_X(x)W_x$ on the system $H_E$, in which, $H_X$ is a classical system spanned by the basis $\{|x\rangle\}$. For a transition matrix $\Gamma$ from $\mathcal{V}$ to $\mathcal{X}$, we define a classical-quantum channel $W \circ \Gamma : v \mapsto \sum_{x \in \mathcal{X}} W_x \Gamma_v(x)$ from the classical system $\mathcal{V}$ to the quantum system $H_E$. Since the state $(W \circ \Gamma) \circ P_V$ coincides with $W \circ (\Gamma \circ P_V)$ as a state on $H_E$, we simply denote it by $W \circ \Gamma \circ P_V$.

B. Single system

Given a normalized state $\rho$ and a non-negative operator $\sigma$ on a single quantum system $H$, we prepare relative entropy type of information quantities

$$D(\rho\|\sigma) := \text{Tr} \rho \log \frac{\rho}{\sigma}$$

Similarly, for $s \in (-1, \infty)$, we can define the functions

$$D_s(\rho\|\sigma) := \frac{1}{s} \text{Tr} \rho s^{1+s} \sigma^{-s}$$

$$D_s^*(\rho\|\sigma) := \frac{1}{s} \text{Tr} \rho s^{-1/2} \sigma^{-s}$$

Indeed, the quantity $D_s(\rho\|\sigma)$ equals the quantity $\frac{1}{s} \psi(s|\rho\|\sigma)$ given in [28], [1], [18]. While the quantity $D_s^*(\rho\|\sigma)$ is the same as the quantity $\frac{1}{s} \psi(s|\rho\|\sigma)$ given in [18], we use the notation $D_s^*(\rho\|\sigma)$ instead of $D_s^*(\rho\|\sigma)$.

Since $s \mapsto sD_s(\rho\|\sigma)$ and $s \mapsto sD_s^*(\rho\|\sigma)$ are convex and $\lim_{s \to 0} sD_s(\rho\|\sigma) = \lim_{s \to 0} sD_s^*(\rho\|\sigma) = 0$, $D_s(\rho\|\sigma)$ and $D_s^*(\rho\|\sigma)$ are monotone increasing for $s \in \mathbb{R}$. (The convexity of $s \mapsto sD_s(\rho\|\sigma)$ is shown in [1]. We can show the convexity of $s \mapsto sD_s^*(\rho\|\sigma)$ in the same way by calculating the second derivative.) Further, since $\lim_{s \to 0} D_s(\rho\|\sigma) = D(\rho\|\sigma)$ and $\lim_{s \to 0} D_s^*(\rho\|\sigma) = D(\rho\|\sigma)$,

$$D(\rho\|\sigma) \leq D_s(\rho\|\sigma),$$

$$D(\rho\|\sigma) \leq D_s^*(\rho\|\sigma)$$




for $s \in (0, 1]$. The information processing inequalities
\[
D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D(\rho\|\sigma), \quad D_s(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_s(\rho\|\sigma)
\]
(7)
hold for $s \in (-1, 1]$ and a TP-CP map $\mathcal{E}$ [1] (5.30), (5.41)]. However, this kind of inequality does not hold for $D_s(\rho\|\sigma)$ or $D_s(\rho\|\sigma)$ in general. Since the inequalities (7) are natural property for information quantities, we consider that the quantities $D(\rho\|\sigma)$ and $D_s(\rho\|\sigma)$ describe the essential information quantities, and the quantities $D_s(\rho\|\sigma)$ and $D_s(\rho\|\sigma)$ are technical tools for our derivation. As is shown in the end of this subsection, the relations
\[
D_s(\rho\|\sigma) \leq D(\rho\|\sigma)
\]
(8)
\[
D_s^*(\rho\|\sigma) \leq D_s(\rho\|\sigma)
\]
(9)
hold for $s \in [-1, 1]$. When $\sigma$ is $I$, we obtain von Neumann entropy and Rényi entropy of $\rho$ as
\[
H(\rho) := -D(\rho\|I) = -D_s(\rho\|I)
\]
(10)
\[
H_{1+s}(\rho) := -D_{s}(\rho\|I) = -D_{s}^{*}(\rho\|I).
\]
(11)
When the state $\sigma$ has the spectral decomposition $\sigma = \sum_i s_i E_i$, the pinching map $\mathcal{E}_\sigma$ is defined as
\[
\mathcal{E}_\sigma(\rho) := \sum_i E_i \rho E_i.
\]
(12)
When $v$ is the number of the eigenvalues of $\sigma$, the inequality
\[
\rho \leq v\mathcal{E}_\sigma(\rho).
\]
(13)
holds [1] Lemma 3.8]. Hence, we obtain
\[
\sigma^{-1/2} \rho \sigma^{-1/2} \leq v\sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2}.
\]
(14)
Since $x \mapsto \log x$ is matrix monotone,
\[
\log \sigma^{-1/2} \rho \sigma^{-1/2} \leq \log v + \log \sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2}.
\]
(15)
Since
\[
\text{Tr} \rho \log \sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2} = \text{Tr} \mathcal{E}_\sigma(\rho) \log \sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2},
\]
we obtain
\[
D(\rho\|\sigma) \leq D(\mathcal{E}_\sigma(\rho)\|\sigma) + \log v = D(\mathcal{E}_\sigma(\rho)\|\sigma) + \log v.
\]
(17)
Proofs of (8) and (9): Here, we show Inequalities (8) and (9) by using Inequality (14). Since the method based on Inequalities (13) and (14) is very important in this paper, we put these proofs here not in Appendix.
For $s \in [0, 1]$, since $x^s$ is operator monotone, the inequality (14) implies that
\[
\text{Tr} \rho (\sigma^{-1/2} \rho \sigma^{-1/2})^s \leq v\text{Tr} \rho (\sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2})^s
\]
\[
= v\text{Tr} \mathcal{E}_\sigma(\rho) (\sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2})^s
\]
\[
= v\text{Tr} \mathcal{E}_\sigma(\rho) (\sigma^{1+s} \sigma^{-s}) \leq v\text{Tr} \rho^{1+s} \sigma^{-s},
\]
where $v$ is the number of eigenvalues of $\sigma$. In the $n$-fold setting, we obtain
\[
e^{nD_s(\rho\|\sigma)} \leq v_n e^{nD_s(\rho\|\sigma)},
\]
where $v_n$ is the number of eigenvalues of $\sigma \otimes n$. That is, the relation
\[
sD^*_s(\rho\|\sigma) \leq \frac{\log v_n}{n} + sD_s(\rho\|\sigma)
\]
(18)
holds. Taking the limit $n \to \infty$, we obtain (8) with $s \in [0, 1]$. Finally, taking the limit $s \to 0$, we obtain (9).
For $s \in [-1, 0]$, since $-x^s$ is operator monotone, the inequality (14) implies that
\[
\text{Tr} \rho (\sigma^{-1/2} \rho \sigma^{-1/2})^s \geq v\text{Tr} \rho (\sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2})^s
\]
\[
= v\text{Tr} \mathcal{E}_\sigma(\rho) (\sigma^{-1/2} \mathcal{E}_\sigma(\rho)\sigma^{-1/2})^s
\]
\[
= v\text{Tr} \mathcal{E}_\sigma(\rho) (\sigma^{1+s} \sigma^{-s}) \leq v\text{Tr} \rho^{1+s} \sigma^{-s},
\]
which implies that
\[
sD^*_s(\rho\|\sigma) \geq \frac{\log v_n}{n} + sD_s(\rho\|\sigma).
\]
(19)
Taking the limit $n \to \infty$, we obtain (8) with $s \in [-1, 0]$. ■
C. Composite system

This paper heavily employs mutual information type quantities rather than conditional entropy type quantities because the channel coding is closely linked to the mutual information rather than the conditional entropy even with the existence of the eavesdropper. Now, we introduce mutual information type quantities in a composite system $\mathcal{H}_X \otimes \mathcal{H}_E$ by using the relative entropy type quantities $D(\rho||\sigma)$, $D_s(\rho||\sigma)$, $\mathcal{L}(\rho||\sigma)$, and $\mathcal{L}_s(\rho||\sigma)$ when the composite state $\rho_{E,X}$ is given as $W \times P_X$ by using a distribution $P_X$ on $X$ and a classical-quantum channel $W : x \rightarrow W_x$ from the classical system $X$ to the quantum system $\mathcal{H}_E$. Using two kinds of relative entropies $D(\rho||\sigma)$ and $\mathcal{L}(\rho||\sigma)$, we define two kinds of quantum versions of the mutual information between $\mathcal{H}_X$ and $\mathcal{H}_E$ as

$$I(X;E|\rho_{E,X}) := D(\rho_{E,X}\|\rho_E \otimes \rho_X)$$

$$I_s(X;E|\rho_{E,X}) := D_s(\rho_{E,X}\|\rho_E \otimes \rho_X)$$

where $\rho_E$, $\rho_X$ are reduced density operators while the conditional entropy is given as

$$H(X|E|\rho_{E,X}) := -D(\rho_{E,X}\|\rho_E \otimes I).$$

In this notation, we use the random variable for identifying the classical system while we use the subscript of the quantum system for identifying the quantum system. Further, we define modified version mutual informations as

$$I_s(X;E|\rho_{E,X}) := D_s(\rho_{E,X}\|\rho_E \otimes \rho_X)$$

$$\mathcal{L}_s(X;E|\rho_{E,X}) := \mathcal{L}_s(\rho_{E,X}\|\rho_E \otimes \rho_X)$$

while the conditional Rényi entropy is defined $\mathcal{L}$ as

$$H_{1+s}(X|E|\rho_{E,X}) := -D_s(\rho_{E,X}\|\rho_E \otimes I).$$

Since $D_s(\rho||\sigma)$ is monotone increasing for $s \in \mathbb{R}$, the function $H_{1+s}(X|E|\rho_{E,X})$ is monotone decreasing for $s \in \mathbb{R}$. In particular,

$$H(X|E|\rho_{E,X}) \geq H_{1+s}(X|E|\rho_{E,X})$$

for $s \in (0,1]$ because $\lim_{s \to 0} H_{1+s}(X|E|\rho_{E,X}) = H(X|E|\rho_{E,X})$. When a TP-CP map $E$ on the system $\mathcal{H}_E$ is applied, Relations (7) imply that

$$I(X;E|\mathcal{E}(\rho_{E,X})) \leq I(X;E|\rho_{E,X}), \quad I_s(X;E|\mathcal{E}(\rho_{E,X})) \leq I_s(X;E|\rho_{E,X})$$

for $s \in (-1,1)$. However, $I(X;E|\rho_{E,X})$ and $\mathcal{L}_s(X;E|\rho_{E,X})$ do not satisfy the same property. In this paper, due to these properties, we consider that the quantities $I(X;E|\mathcal{E}(\rho_{E,X}))$ and $I_s(X;E|\mathcal{E}(\rho_{E,X}))$ describe essential information quantities. The other quantities $\mathcal{L}_s(X;E|\rho_{E,X})$ and $\mathcal{L}_s(X;E|\rho_{E,X})$ are technical tools for overcoming the difficulty caused by the non-commutativity. That is, our final result will be described by using the quantities $I(X;E|\mathcal{E}(\rho_{E,X}))$ and $I_s(X;E|\mathcal{E}(\rho_{E,X}))$.

Next, using a state $\sigma_E$ on $\mathcal{H}_E$, we extend $I(X;E|\rho_{E,X})$ and $I_s(X;E|\rho_{E,X})$ as

$$I(X;E|\rho_{E,X}\|\sigma_E) := D(\rho_{E,X}\|\sigma_E \otimes \rho_X)$$

$$I_s(X;E|\rho_{E,X}\|\sigma_E) := D_s(\rho_{E,X}\|\sigma_E \otimes \rho_X).$$

The conditional Rényi entropy $H(X|E|\rho_{E,X})$ is also generalized $\mathcal{L}$ as

$$H_{1+s}(X|E|\rho_{E,X}\|\sigma_E) := -D_s(\rho_{E,X}\|\sigma_E \otimes I).$$

Note that our definition of the quantity $H_2(X|E|\rho_{E,X}\|\sigma_E)$ is different from Renner’s definition $H_2(X|E|\rho_{E,X}\|\sigma_E) = -\mathcal{L}_s(\rho_{E,X}\|\sigma_E \otimes I)$ $\mathcal{L}$ in the non-commutative case. Generally, due to $\mathcal{L}$, our definition gives a smaller value than Renner’s definition.

The first quantity $I(X;E|\rho_{E,X}\|\sigma_E)$ satisfies

$$I(X;E|\rho_{E,X}\|\sigma_E) - I(X;E|\rho_{E,X}) = D(\rho_E\|\sigma_E) \geq 0$$

for any state $\sigma_E$ on $\mathcal{H}_E$. That is,

$$\min_{\sigma_E} I(X;E|\rho_{E,X}\|\sigma_E) = I(X;E|\rho_{E,X}).$$

For the minimization of the second quantity $I_s(X;E|\rho_{E,X}\|\sigma_E)$, we introduce another type of mutual information as

$$I^G_s(X;E|\rho_{E,X}) := \frac{1+s}{s} \log \text{Tr} \left( \sum_x P_X(x)(W_x)^{1+s} \right).$$
This quantity can be written by using a quantum extension of Gallager function \[9\]
\[
\phi(s|W, P_X) := \log \text{Tr} \left( \sum_x P_X(x)(W_x)^{1/(1-s)} \right) ^{1-s}
\]
as
\[
I_s^G(X; E|\rho_{E,X}) = \frac{1+s}{s} \phi\left( \frac{s}{1+s} |W, P_X\right).
\]
These quantities satisfy
\[
\lim_{s \to 0} I_s^G(X; E|\rho_{E,X}) = \lim_{s \to 0} I_s(X; E|\rho_{E,X}) = I(X; E|\rho_{E,X})
\]
and can be characterized by the following lemmas

**Lemma 1**: The equation
\[
\min_{\rho_H} I_s(X; E|\rho_{E,X}||\sigma_E) = I_s^G(X; E|\rho_{E,X})
\]
holds for \(s \in (-1, \infty)\), where \(\sigma\) is restricted to normalized states on \(\mathcal{H}_E\). The minimum is realized when
\[
\sigma_E = (\sum_x P_X(x)W_x^{1+s})^{1+1/(1+s)} / \text{Tr} \left( \sum_x P_X(x)W_x^{1+s} \right)^{1+1/(1+s)}.
\]
Thus, the inequality
\[
I_s(X; E|\rho_{E,X}) \geq I_s^G(X; E|\rho_{E,X})
\]
holds for \(s \in (-1, \infty)\).
The proof of Lemma 1 is given in Appendix A. Since \(D_s(\rho||\sigma)\) is monotone increasing for \(s\), \(I_s(X; E|\rho_{E,X}||\sigma_E)\) is also monotone increasing for \(s\). Due to Lemma 1, \(I_s^G(X; E|\rho_{E,X})\) is also monotone increasing for \(s\).

Conversely, as shown in [20, (16)], the following lemma holds.

**Lemma 2**: When all of the states \(W_x\) are commutative each other, the inequality
\[
I_s(X; E|\rho_{E,X}) \leq I_s^G(X; E|\rho_{E,X})
\]
holds for \(s \in [0, 1]\).

### D. Classical-quantum channel

Next, we characterize the above defined quantities by using a classical-quantum channel \(W\) from \(\mathcal{X}\) to \(\mathcal{H}_E\).

\[
I(X; E|W \times P_X) = \sum_x P_X(x)\text{Tr} W_x (\log W_x - \log(W \circ P_X))
\]
\[
\mathbb{I}(X; E|W \times P_X) = \sum_x P_X(x)^\mathbb{I} W_x \log((W \circ P_X)^{-1/2}W_x(W \circ P_X)^{-1/2})
\]
\[
I_s(X; E|W \times P_X) = \frac{1}{s} \log \text{Tr} \sum_x P_X(x)\text{Tr} W_x^{1+s}(W \circ P_X)^{-s}
\]
\[
I_s^G(X; E|W \times P_X) = \frac{1}{s} \log \text{Tr} \sum_x P_X(x)\text{Tr} W_x((W \circ P_X)^{-1/2}W_x(W \circ P_X)^{-1/2})^s.
\]

When we apply a quantum operation \(E\) on \(\mathcal{H}_E\), we can define another channel \(E|W : x \mapsto E(W_x)\) from the classical system \(\mathcal{X}\) to the quantum system \(\mathcal{H}_E\). Then, Inequalities (27) are rewritten as
\[
I(X; E|E|W \times P_X) \leq I(X; E|W \times P_X)
\]
\[
I_s(X; E|E|W \times P_X) \leq I_s(X; E|W \times P_X)
\]
hold for \(s \in (-1, 1]\). When \(v\) is the number of the eigenvalues of \(W \circ P_X\), Relation (12) yields an inequality opposite to (41):
\[
I(X; E|W \times P_X) \leq I(X; E|W \circ P_X [W] \times P_X) + \log v
\]
\[
= I(X; E|E|W \circ P_X [W] \times P_X) + \log v.
\]

Further, we obtain the following lemma.

**Lemma 3**: The function \(P_X \mapsto e^{sI_s(X; E|E|W \times P_X)} = e^{s\mathbb{I}(X; E|W \times P_X)}\) is convex with \(s \in [-1/2, 0]\), and is concave with \(s \in [0, \infty]\).

However, the quantity \(P_X \mapsto e^{sI_s(X; E|E|W \times P_X)}\) does not satisfy the convexity in general.

**Proof**: The convexity and concavity of \(P_X \mapsto e^{s\mathbb{I}(X; E|W \times P_X)} = \text{Tr} \left( \sum_x P_X(x)(W_x)^{1+s} \right)^{1+1/(1+s)}\) follow from the operator convexity and operator concavity of \((x^{1+s})\) for the respective parameter \(s\).
E. Security criteria

Next, in order to give the security criteria, we characterize the leaked information when an information $X$ is transmitted via the classical-quantum channel $W$. When the information $X$ is subjected to the uniform distribution $P_{\text{mix},X}$ on $X$, the leaked information is given as

$$I_{\text{mix}}(X;E|W) := I(X;E|W \times P_{\text{mix},X}) = D(W \times P_{\text{mix},X}||(W \circ P_{\text{mix},X}) \otimes P_{\text{mix},X})$$

$$= \sum_{x \in X} P_{\text{mix},X}(x)D(W_x||W \circ P_{\text{mix},X}) = \sum_{x \in X} P_{\text{mix},X}(x)(H(W \circ P_{\text{mix},X}) - H(W_x)).$$

When we do not know the distribution of the information $X$, we adopt the following value as the criterion of the leaked information:

$$I_{\text{max}}(X;E|W) := \max_{P_X} I(X;E|W \times P_X) = \max_{P_X} D(W \times P_X||(W \circ P_X) \otimes P_X)$$

$$= \max_{P_X} \sum_{x \in X} P_X(x)D(W_x||W \circ P_X) = \max_{P_X} \sum_{x \in X} P_X(x)(H(W \circ P_X) - H(W_x)).$$

Next, we consider the leaked information by using the trace norm instead of the mutual information. When the state is given as $\rho_{E,X} = \sum_x P_X(x)W_x \otimes |x\rangle \langle x|$, the leaked information is characterized as

$$d_1(X;E|\rho_{E,X}) := \|\rho_{E,X} - \rho_E \otimes P_X\|_1. \tag{44}$$

When we take into account the uniformity as well as the independence, we employ the following quantity \[10]\:

$$d'_1(X;E|\rho_{E,X}) := \|\rho_{E,X} - \rho_E \otimes P_{\text{mix},X}\|_1. \tag{45}$$

In this notation, when a function $f : \mathcal{X} \mapsto \mathcal{Y}$ is given, $d'_1(f(X);E|\rho_{E,X})$ expresses the following quantity:

$$d'_1(f(X);E|\rho_{E,X}) = \left\| \sum_y \left( \sum_{x \in f^{-1}(y)} P_X(x)W_x \otimes |y\rangle \langle y| - \rho_E \otimes P_{\text{mix},Y} \right) \right\|_1. \tag{46}$$

Now, we consider the case when the eavesdropper’s system consists of the quantum system $\mathcal{H}_E$ and the classical system $\mathcal{Y}$. As shown in Appendix \[8]\ when $X$ and $Y$ obey the uniform distribution independently, this quantity satisfies

$$d'_1(X;Y|\rho_{X,Y,E}) \leq 2d'_1(X,Y;E|\rho_{X,Y,E}). \tag{47}$$

Using these quantities, we define the leaked information for the channel $W$. When the input information $X$ is subjected to the uniform distribution $P_{\text{mix},X}$ on $X$, the leaked information is given as

$$d_{1,\text{mix}}(X;E|W) := d_1(X;E|W \times P_{\text{mix},X}) = d'_1(X;E|W \times P_{\text{mix},X}) = \sum_{x \in X} P_{\text{mix},X}(x)||W_x - W \circ P_X\|_1.$$

When we do not know the distribution of the information $X$, we adopt the following value as the criterion of the leaked information:

$$d_{1,\text{max}}(X;E|W) := \max_{P_X} d_1(X;E|W \times P_X) = \max_{P_X} \sum_{x \in X} P_X(x)||W_x - W \circ P_X\|_1.$$

In this paper, we employ the quantities $I_{\text{mix}}(X;E|W)$, $I_{\text{max}}(X;E|W)$, $d_{1,\text{mix}}(X;E|W)$, and $d_{1,\text{max}}(X;E|W)$ as security criteria. Since $I_{\text{max}}(X;E|W)$ and $d_{1,\text{max}}(X;E|W)$ do not depend on the distribution on the input messages, the results based on these are called source universal \[27].

Using the quantum version of Pinsker inequality, we obtain

$$d_{1,\text{mix}}(X;E|W)^2 \leq 2I_{\text{mix}}(X;E|W) \tag{48}$$

$$d_{1,\text{max}}(X;E|W)^2 \leq 2I_{\text{max}}(X;E|W). \tag{49}$$
Conversely, we can bound $I_{\text{mix}}(W)$ and $I_{\text{max}}(W)$ by using $d_{1,\text{mix}}(W)$ and $d_{1,\text{max}}(W)$ in the following way. Applying Fannes's inequality, we obtain

$$0 \leq I_{\text{mix}}(X;E|W) = \sum_{x \in \mathcal{X}} P_{\text{mix},\mathcal{X}}(x)(\log d - H(W \circ P_{\text{mix},\mathcal{X}}) - H(W_x))$$

$$\leq \sum_{x \in \mathcal{X}} P_{\text{mix},\mathcal{X}}(x)\eta(\|W \circ P_{\text{mix},\mathcal{X}} - W_x\|_1, \log d_E)$$

$$\leq \eta(\sum_{x \in \mathcal{X}} P_{\text{mix},\mathcal{X}}(x)\|W \circ P_{\text{mix},\mathcal{X}} - W_x\|_1, \log d_E)$$

$$= \eta(d_{1,\text{mix}}(X;E|W), \log d_E),$$

$$0 \leq I_{\text{max}}(X;E|W) = \max_{P_X} \sum_{x \in \mathcal{X}} P_X(x)(\log d - H(W \circ P_X) - H(W_x))$$

$$\leq \max_{P_X} \sum_{x \in \mathcal{X}} P_X(x)\eta(\|W \circ P_X - W_x\|_1, \log d_E)$$

$$\leq \eta(\max_{P_X} \sum_{x \in \mathcal{X}} P_X(x)\|W \circ P_X - W_x\|_1, \log d_E)$$

$$= \eta(d_{1,\text{max}}(X;E|W), \log d_E),$$

(50)

where $d_E$ is the dimension of $\mathcal{H}_E$ and $\eta(x,y) := -x \log x + xy$.

Therefore, if the quantity $I_{\text{mix}}(X;E|W)$ or $I_{\text{max}}(X;E|W)$ goes to zero, the quantity $d_{1,\text{mix}}(X;E|W)$ or $d_{1,\text{max}}(X;E|W)$ goes to zero. The converse is also true when the quantity $d_{1,\text{mix}}(X;E|W)$ or $d_{1,\text{max}}(X;E|W)$ exponentially goes to zero and the dimension $d_E$ grows linearly. However, their speeds of both convergence do not coincide with each other. Hence, we consider both quantities.

Next, we consider the relation between $I_{\text{mix}}(X;E|W)$ and $I_{\text{max}}(X;E|W)$ in a special class of channels. A channel $W$ from the set $\mathcal{X}$ to the quantum system $\mathcal{H}_E$ is called additive when the set $\mathcal{X}$ has a structure of module and there exist a state $\rho$ on the system $\mathcal{H}_E$ and a projective representation $U$ of $\mathcal{X}$ such that $W_x = U_x \rho^{\dagger} U_x$. In this case, as shown in Appendix C, the relation

$$I_{\text{mix}}(X;E|W) = I_{\text{max}}(X;E|W)$$

holds. This equation is useful in the latter discussions.

### III. Generalization of channel resolvability

In order to treat the quantum wiretap channel model, we treat the quantum channel resolvability problem for a given classical-quantum channel $W$ from the classical system $\mathcal{X}$ to the quantum system $\mathcal{H}$, in which, the output quantum state $W_x$ on $\mathcal{H}$ is given for an element $x \in \mathcal{X}$. In the channel resolvability, we treat the approximation of the given output state with the output average state of the auxiliary input random variable when the auxiliary input random variable is subject to the uniform distribution of the subset $\mathcal{X}_0$ of the input system $\mathcal{X}$ of the given channel. That is, the purpose is minimizing the cardinality of the input subset $\mathcal{X}_0$ when $W \circ P_{\text{mix},\mathcal{X}_0} = \sum_{x \in \mathcal{X}_0} W_x$, approximates a given output state $\rho$.

Now, we generalize this problem to the case when the uniform distribution on $\mathcal{X}_0$ is not available. We assume that the auxiliary random variable is subject to a given distribution $P_A$ on the set $A$. Choosing a map $f$ from $A$ to $\mathcal{X}$, we approximate a given state $\rho$ by $W \circ f \circ P_A = \sum_{a \in A} P_A(a)W_{f(a)}$.

Now, we apply the random coding to the alphabet $X$ with the probability distribution $P_X$. The map $\Phi$ from $A$ to $\mathcal{X}$ is randomly chosen in the following way. For each $a \in A$, $\Phi(a)$ is the random variable subject to the distribution $P_X$ on $\mathcal{X}$. For $a \neq a' \in A$, $\Phi(a)$ is independent of $\Phi(a')$. Then, for a distribution $P_A$ on $A$, we can define the distribution $\Phi \circ P_A(x)$ on $\mathcal{X}$.

Then, we have the following lemma:

**Lemma 4:** For $s \in (0, 1]$, we obtain

$$E_{\Phi} D(W \circ \Phi \circ P_A||W \circ P_X) \leq \frac{\nu^s e^{-\nu s H_1(P_A) + s I(s;X|W \times P_X)}}{s},$$

(53)

where $\nu$ is the number of eigenvalues of $W \circ P_X$. Similarly,

$$E_{\Phi} e^{\frac{s}{2}(W \circ \Phi \circ P_A||W \circ P_X)} \leq E_{\Phi} e^{\frac{s}{2}(W \circ \Phi \circ P_A||W \circ P_X)}$$

$$\leq 1 + e^{-s H_1(P_A) + s I(s;X|W \times P_X)},$$

(54)

$$E_{\Phi} e^{\frac{s}{2}(W \circ \Phi \circ P_A||W \circ P_X)} \leq e^{-s I(s;X|W \times P_X)}.$$
Using (54) in Lemma 4 we obtain

\[ E_x e^{sD(W \circ \Phi \circ P_A || W \circ P_X)} \leq e^{sH_{1+}(P_A)} e^{sL^*(X; E||W \circ P_X)} \]

\[ \leq e^s (1 + e^{-sH_{1+}(P_A)} e^{sL^*(X; E||W \circ P_X)}) \]

\[ = e^s (1 + e^{-sH_{1+}(P_A)} e^{sI_e(X; E||W \circ P_X)}) \]

\[ \leq e^s (1 + e^{-sH_{1+}(P_A)} e^{sI_e(X; E||W \circ P_X)}) \]  \hspace{1cm} (55)

Next, we consider this kind of approximation when the space \( X \) has the structure of a module. That is, when a submodule \( C \subseteq X \) and an element \( y \in X \) are given, using the uniform distribution \( P_{\text{mix}, C+y} \) on the subset \( C + y := \{x+y|x \in C\} \), we approximate the output state \( W \circ P_{\text{mix}, C+y} \) with the uniform input distribution. In this case, we evaluate \( D(W \circ P_{\text{mix}, C+y} || W \circ P_{\text{mix}, X}) \). Now, we consider the condition for ensemble of submodules \( \{C[Z]\} \) of \( X \) when the submodule \( C[Z] \) is decided by a random variable \( Z \).

\textbf{Condition 1}: The cardinality of \( C[Z] \) is a constant \( L \). Any element \( x \neq 0 \in X \) is included in \( C[Z] \) with probability at most \( \frac{1}{|Z|} \).

Then, we have the following lemma:

\textbf{Lemma 5}: When the random variable \( \tilde{X} \in X \) obeys the uniform distribution on \( X \) and is independent of the choice of \( C[Z] \), we obtain

\[ E_{Z, \tilde{X}} D(W \circ P_{\text{mix}, C[Z] + \tilde{X}} || W \circ P_{\text{mix}, X}) \leq \frac{e^{sI_e(X; E||W \circ P_{\text{mix}, X})}}{s} \] \hspace{1cm} (56)

for \( s \in (0, 1] \), where \( v \) is the number of eigenvalues of \( W \circ P_{\text{mix}, X} \). Similarly,

\[ E_{Z, \tilde{X}} e^{sD(W \circ P_{\text{mix}, C[Z] + \tilde{X}} || W \circ P_{\text{mix}, X})} \leq \frac{e^{sI_e(X; E||W \circ P_{\text{mix}, X})}}{s} \] \hspace{1cm} (57)

Using (57) in Lemma 5 we obtain

\[ E_{Z, \tilde{X}} e^{sD(W \circ P_{\text{mix}, C[Z] + \tilde{X}} || W \circ P_{\text{mix}, X})} \leq e^{sI_e(X; E||W \circ P_{\text{mix}, X})} \]

\[ \leq e^s (1 + \frac{1}{L^*} e^{sL^*(X; E||W \circ P_{\text{mix}, X})}) \]

\[ \leq e^s (1 + \frac{1}{L^*} e^{sI_e(X; E||W \circ P_{\text{mix}, X})}) \] \hspace{1cm} (58)

Next, we consider the case when the uniform distribution on \( C[Z] \) is not available and only a distribution \( P_A \) on another module \( A \) is available. Now, we assume the following condition for the ensemble for injective homomorphisms \( f_Z \) from \( A \) to \( X \).

\textbf{Condition 2}: Any elements \( x \neq 0 \in X \) and \( a \neq 0 \in C \), the relation \( f_Z(a) = x \) holds with probability at most \( \frac{1}{|X|} \).

When \( X \) and \( A \) are vector spaces of a finite field \( \mathbb{F}_q \), the set of all injective morphisms from \( A \) to \( X \) satisfies Condition 2.

We choose the random variable \( \tilde{X} \in X \) that obeys the uniform distribution on \( X \) that is independent of the choice of \( f_Z \). Then, we define a map \( f_{Z,S}(a) := f_Z(a) + \tilde{X} \) and have the following lemma:

\textbf{Lemma 6}: Under the above choice, we obtain

\[ E_{Z, \tilde{X}} D(W \circ f_{Z,S} || W \circ P_{\text{mix}, X}) \leq \frac{e^{sI_e(X; E||W \circ P_{\text{mix}, X})}}{s} \] \hspace{1cm} (59)

for \( s \in (0, 1] \), where \( v \) is the number of eigenvalues of \( W \circ P_{\text{mix}, X} \). Similarly,

\[ E_{Z, \tilde{X}} e^{sD(W \circ f_{Z,S} || W \circ P_{\text{mix}, X})} \leq e^{sI_e(X; E||W \circ P_{\text{mix}, X})} \]

\[ \leq e^s (1 + \frac{1}{L^*} e^{sI_e(X; E||W \circ P_{\text{mix}, X})}) \] \hspace{1cm} (60)
Using (60) in Lemma 6, we obtain

\[
E_{Z, \tilde{X}} e^{\lambda D(W \circ f_Z; P \circ P_{\text{mix}, X})} \\
\leq e^{\lambda E_{Z, \tilde{X}} e^{\lambda D(W \circ f_Z; P \circ P_{\text{mix}, X})}} \\
\leq e^{\lambda \left(1 + e^{-s H_{1, \tilde{A}}(P_{\text{mix}})} e^{\lambda L(x; E \circ f_Z; P \circ P_{\text{mix}, X})}\right)} \\
= e^{\lambda \left(1 + e^{-s H_{1, \tilde{A}}(P_{\text{mix}})} e^{\lambda L(x; E \circ f_Z; P \circ P_{\text{mix}, X})}\right)},
\]

(61)

Here, we construct an ensemble of submodules \(\{C[Z]\}\) satisfying Condition 1 and an ensemble of injective isomorphisms \(\{f_Z\}\) satisfying Condition 2 when \(X'\) and \(A\) are given as vector spaces \(\mathbb{F}_q^k\) and \(\mathbb{F}_q^l\) of a finite field \(\mathbb{F}_q\) \((k \geq l)\). Let \(Z\) be the Toeppliz matrix of the size \((k-l) \times l\), which contains \(k-1\) random variables taking values in the finite field \(\mathbb{F}_q\), and \(Z'\) be the random matrix taking values in the set of invertible matrices of the size \(l \times l\) with the uniform distribution. When \(f_Z\) is given by the multiplication of the random matrix \((I, Z)^T\), and \(C[Z]\) is given by \(f_Z(A)\), \(\{C[Z]\}\) satisfies Condition 1. When \(f_Z\) is given by the multiplication of the random matrix \((Z', Z)^T\) with two independent random variables \(Z'\) and \(Z\), the ensemble \(\{f_Z\}\) satisfies Condition 2.

IV. QUANTUM WIRETAP CHANNEL IN A GENERAL FRAMEWORK

A. Single-shot bounds

Next, we consider the quantum wiretap channel model, in which the eavesdropper (wire-tapper) Eve and the authorized receiver Bob receive information from the authorized sender Alice. In this case, in order for Eve to have less information, Alice chooses a suitable encoding. This problem is formulated as follows. Let \(H_B\) and \(H_E\) be the quantum systems of Bob and Eve, and \(X\) be the alphabet sent by Alice. Then, the main quantum channel from Alice to Bob is described by \(W_B : x \mapsto W_B[x]\), and the wire-tapper quantum channel from Alice to Eve is described by \(W_E : x \mapsto W_E[x]\). That is, \(W_B[x]\) \((W_E[x])\) is Bob’s (Eve’s) density matrix on the system \(H_B\) \((H_E)\). In this setting, Alice chooses \(M\) distributions \(\Gamma_1, \ldots, \Gamma_M\) on \(X\), and she generates \(x \in X\) subject to \(\Gamma_m\) when she wants to send the message \(m \in M = \{1, \ldots, M\}\). In this case, Bob (Eve) receives the density matrix \(W_B \circ \Gamma_m = \sum_x \Gamma_m(x) W_B[x]\) \((W_E \circ \Gamma_m) \equiv \sum_x \Gamma_m(x) W_E[x]\). Bob prepares \(M\)-outcome POVM \(\{D_1, \ldots, D_M\}\) of \(H_B\). Here, we regard \(\{\Gamma_1, \ldots, \Gamma_M\}\) as a transition matrix \(\Gamma\) from \(M\) to \(X\). Therefore, the triplet \((M, \Gamma, \{D_1, \ldots, D_M\}\) is called a code, and is described by \(\Phi\). Its performance is given by the following three quantities. The first is the size \(M\), which is denoted by \(|\Phi|\). The second is the average error probability \(\epsilon_B(\Phi)\):

\[
\epsilon_B(\Phi) := \frac{1}{M} \sum_{m=1}^M \text{Tr} W_B \circ \Gamma_m (I - D_m) - D_m,
\]

and the third is Eve’s information regarding the transmitted message. Using the channel \(W_E \circ \Gamma : m \mapsto W_E \circ \Gamma_m\), we can describe this quantity by the following ways:

\[
I_{\text{mix}}(M; E|W_E \circ \Gamma) = I(M; E|W_E \circ \Gamma \times P_{\text{mix}, M})
\]

\[
I_{\text{max}}(M; E|W_E \circ \Gamma) = \max_{P_{\text{mix}, M}} I(M; E|W_E \circ \Gamma \times P_{\text{mix}, M})
\]

\[
d_{1, \text{mix}}(M; E|W_E \circ \Gamma) = \sum_{m=1}^M \frac{1}{M} \|W_E \circ \Gamma_m - W_E \circ \Gamma \circ P_{\text{mix}, M}\|_1
\]

\[
d_{1, \text{max}}(M; E|W_E \circ \Gamma) = \max_{P_{\text{mix}, M}} \sum_{m=1}^M P_{\text{mix}, M}(m) \|W_E \circ \Gamma_m - W_E \circ \Gamma \circ P_{\text{mix}, M}\|_1.
\]

In the usual setting, the distribution \(\Gamma_i\) can be chosen to a uniform distribution. However, sometimes, it is difficult to prepare a perfect uniform distribution in a realistic setting. So, in the following, we make our code with a non-uniform distribution \(P_L\) on \(L = \{1, \ldots, L\}\).

Now, we make a code \(\Phi[Z]\) for the quantum wiretap channel based on the random coding method for given integers \(L\) and \(M\). In the following, Alice is allowed to generate a random number on \(\{1, \ldots, L\}\) with the distribution \(P_L\). First, we generate the random code \(\Phi[Z]\) with size \(LM\), which is described by the LM independent and identical random variables \(Z\) subject to the distribution \(P_X\) on \(X\). That is, all of \(\{\Phi[Z]\}_{l,m}\) are independent and obeys the distribution \(P_X\) on \(X\). For integers \(l = 1, \ldots, L\) and \(m = 1, \ldots, M\), As is guaranteed in the previous paper [13], we choose the decoder (POVM) \(\mathcal{D}_{l,m}[Z]\) of the code \(\Phi[Z]\) such that the ensemble expectation of the average error probability concerning decoding the input message \(A\) is less than \(4(ML)\epsilon e^{-s_{f_0}(X; B)W_B \times P_X}\) for \(0 \leq s \leq 1\). The proof given in [13] is valid even if the prior distribution for sent messages is not uniform. That is, when the message \((l,m)\) is sent with the probability \(P_{LM}(l,m)\), the ensemble expectation of the
average error probability concerning decoding the input message $A$ is bounded by $4(ML)^s e^{-sI_{\mathcal{F}}(X;B)|W_B \times P_X}$ for $0 \leq s \leq 1$. Using the code $\Phi[Z]$, Alice encodes her message $M = 1, \ldots, M$ in the following way. In order to send her message $m$, she generates the random number $L$ subject to $P_L$, and inputs the element $\Phi[Z]_{L,m}$ in the channel $W_B$. That is, Alice generates the input signal $X$ subject to the distribution $\Gamma[Z]_{m}$, which is defined as $\Gamma[Z]_{m}(x) := \sum_{l} P_{L}(l) \delta_{\Phi[Z]_{L,m,l}}(x)$, where

$$\delta_{x,x'} = \begin{cases} 1 & \text{if } x = x' \\ 0 & \text{if } x \neq x'. \end{cases}$$  \hspace{1cm} (62)

Bob recovers the values $l$ and $m$ by using the decoder $\{D_{l,m}[Z]\}$, and discards the value $l$. That is, Bob recovers the value $m$ by using the decoder $D_{m}[Z] := \sum_{l=1}^{L} D_{l,m}[Z]$. We denote the code $(M, \Gamma[Z], \{D_{m}[Z]\})$ by $\Phi[Z]$. Therefore, the above discussion in [13] yields that

$$\mathbb{E}\epsilon(W_B|\Phi[Z]) \leq 4 \min_{0 \leq s \leq 1} (ML)^s e^{-sI_{\mathcal{F}}(X;B)|W_B \times P_X}.$$  \hspace{1cm} (63)

Using (53), for $0 < s \leq 1$, we obtain

$$\mathbb{E}\mathcal{I}z D(W_E \circ \Gamma[Z]_m||W_E \circ P_X) \leq \frac{v^s}{s} e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)},$$  \hspace{1cm} (64)

where $v$ is the number of eigenvalues of $W_E \circ P_X$. Thus, using (30) and (64), we obtain

$$\mathbb{E}\mathcal{I}_{\text{mix}} z (M; E|W_E \circ \Gamma[Z])$$

$$= \mathbb{E}\mathcal{I}z (M; E||W_E \circ \Gamma[Z]) \times P_{\text{mix},M}$$

$$\leq \mathbb{E}\mathcal{I}z (M; E||W_E \circ \Gamma[Z]) \times P_{\text{mix},M||W_E \circ P_X}$$

$$= \mathbb{E}\mathcal{I}z \sum_{m=1}^{M} \frac{1}{M} D(W_E \circ \Gamma[Z]_m||W_E \circ P_X)$$

$$\leq \frac{v^s}{s} e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)}.$$  \hspace{1cm} (65)

Similarly, as is shown latter, we obtain

$$\mathbb{E}\mathcal{I}_{\text{mix}} d_1(z)(M; E|W_E \circ \Gamma[Z])$$

$$\leq (4 + \sqrt{v_t}) e^{-\frac{e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)}}{s}}$$  \hspace{1cm} (66)

$$\mathbb{E}\mathcal{I}_{\text{mix}} d_1(z)(M; E|W_E \circ \Gamma[Z])$$

$$\leq (4 + \sqrt{\lambda_t}) e^{-\frac{e^{-\frac{e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)}}{s}}}}$$  \hspace{1cm} (67)

for $0 < s \leq 1$, where $v_t$ is the number of eigenvalues of $\sum_x P_X(x)W_{E|X}^{1+s}$ and $\lambda_t$ is defined as the real number $\log a_1 - \log a_0$ by using the maximum eigenvalue $a_1$ and the minimum eigenvalue $a_0$ of $\sum_x P_X(x)W_{E|X}^{1+s}/\text{Tr} \sum_x P_X(x)W_{E|X}^{1+s}$.

Finally, we consider what code is derived from the above random coding discussion. Using the Markov inequality, we obtain

$$P_Z \{ \epsilon(W_B|\Phi[Z]) \geq 3 \mathbb{E}\mathcal{I}_z (W_B|\Phi[Z]) \} \leq \frac{1}{3}$$

$$P_Z \{ I_{\text{mix}} (M; E|W_E \circ \Gamma[Z]) \geq 3 \mathbb{E}\mathcal{I}_{\text{mix}} (M; E|W_E \circ \Gamma[Z]) \} \leq \frac{1}{3}$$

$$P_Z \{ d_{1,\text{mix}} (M; E|W_E \circ \Gamma[Z]) \geq 3 \mathbb{E}\mathcal{I}_{d_{1,\text{mix}}} (M; E|W_E \circ \Gamma[Z]) \} \leq \frac{1}{3}.$$  \hspace{1cm} (68)

Therefore, the existence of a good code is guaranteed in the following way. That is, we give the concrete performance of a code whose existence is shown in the above random coding method.

**Theorem 1:** Assume that a random variable subject to the distribution $P_L$ on $\{1, \ldots, L\}$ is available. For any integer $M$ and any probability distribution $P_X$ on $\mathcal{X}$, there exists a code $\Phi$ with an encoder $\Gamma$ such that the code $\Phi$ only uses the distribution $P_L$ for mixing the input alphabet and

$$|\Phi| = M$$

$$\epsilon(W_B|\Phi) \leq 12 \min_{0 \leq s \leq 1} (ML)^s e^{-sI_{\mathcal{F}}(X;B)|W_B \times P_X}$$  \hspace{1cm} (69)

$$I_{\text{mix}} (M; E|W_E \circ \Gamma) \leq 3 \min_{0 \leq s \leq 1} \frac{v^s}{s} e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)}$$  \hspace{1cm} (70)

$$d_{1,\text{mix}} (M; E|W_E \circ \Gamma) \leq 3 \min_{0 \leq s \leq 1} \mu_s e^{-\frac{e^{-sH_{1+s}(P_L)+sI_s(X;E|W_E \times P_X)}}{s}}$$  \hspace{1cm} (71)

where $\mu_s := \min\{4 + \sqrt{v_t}, (4 + \sqrt{\lambda_t})e^s\}$. 


Corollary 1: Assume that a random variable \( L \) subject to the distribution \( P_L \) on \( \{1, \ldots, L\} \) is available. Then, for any integer \( M \) and any probability distribution \( P_X \) on \( \mathcal{X} \), there exists a code \( \Phi \) with an encoder \( \Gamma \) such that the code \( \Phi \) only uses the distribution \( P_L \) for mixing the input alphabet and

\[
|\Phi| = M/4
\]

\[
\epsilon_{\max}(W_B|\Phi) \leq 48 \min_{0 \leq s \leq 1} (ML)^s e^{-sI_{L-s}(X;B|W_B \times P_X)}
\]

\[
I_{\max}(M;E|W_E \circ \Gamma) \leq \frac{12}{s} \min_{0 \leq s \leq 1} e^{sH_{1+s}(P_L)+sI_{L}(X;E|W_E \times P_X)}
\]

\[
d_{1,\max}(M;E|W_E \circ \Gamma) \leq 24 \min_{0 \leq s \leq 1} \mu_s e^{sH_{1+s}(P_L)+sI_{L}(X;E|W_E \times P_X)}.
\]

**Proof**: Now, we prove Corollary 1 using code \( \Phi \) given in Theorem 1. When \( M \) is regarded as a random variable obeying the uniform distribution on \( \mathcal{M} \), Markov inequality guarantees that \( \text{Tr} W_B \circ \Gamma_M (I - \mathcal{D}_M) \geq 4\epsilon(B|\Phi) \) holds at most probability 1/4. Similarly, \( D(W_E \circ \Gamma_M || W_E \circ \Gamma \circ P_{\text{mix},M}) \geq 4I_{\text{mix}}(M;E|W_E \circ \Gamma) \) and \( \|W_E \circ \Gamma_M - W_E \circ \Gamma \circ P_{\text{mix},M}\|_1 \geq 4d_{1,\text{mix}}(M;E|W_E \circ \Gamma) \) at most probability 1/4, respectively. So, the random variable \( M \) satisfies the three relations \( \text{Tr} W_B \circ \Gamma_M (I - \mathcal{D}_M) \leq 4\epsilon(B|\Phi) \), \( D(W_E \circ \Gamma_M || W_E \circ \Gamma \circ P_{\text{mix},M}) \leq 4I_{\text{mix}}(M;E|W_E \circ \Gamma) \), and \( \|W_E \circ \Gamma_M - W_E \circ \Gamma \circ P_{\text{mix},M}\|_1 \leq 4d_{1,\text{mix}}(M;E|W_E \circ \Gamma) \) at least probability 1/4, respectively. In other words, there exist at least \( M/4 \) elements \( m_1, \ldots, m_{M/4} \) such that \( \text{Tr} W_B \circ \Gamma_{m_l}(I - \mathcal{D}_{m_l}) \leq 4\epsilon(B|\Phi) \), \( D(W_E \circ \Gamma_{m_l} || W_E \circ \Gamma \circ P_{\text{mix},M}) \leq 4I_{\text{mix}}(M;E|W_E \circ \Gamma) \), and \( \|W_E \circ \Gamma_{m_l} - W_E \circ \Gamma \circ P_{\text{mix},M}\|_1 \leq 4d_{1,\text{mix}}(M;E|W_E \circ \Gamma) \) for \( l = 1, \ldots, M/4 \).

So, we define the code \( \Phi := \{M/4\}, \{\Gamma_{m_1}, \ldots, \Gamma_{m_{M/4}}\}, \{\mathcal{D}_{m_1}, \ldots, \mathcal{D}_{m_{M/4}}\} \) with \( \tilde{\Gamma}_l := \Gamma_{m_l} \) and \( \tilde{\mathcal{D}}_l := \mathcal{D}_{m_l} \). Then, \( \epsilon_{\max}(W_B|\Phi) \leq 4\epsilon(B|\Phi) \). Consider a distribution \( P_{M'} \) on \( \mathcal{M}' := \{1, \ldots, M/4\} \). Then, using (30), we obtain

\[
I_{\max}(M';E|W_E \circ \tilde{\Gamma}) = \max_{P_{M'}} I(M';E|W_E \circ \tilde{\Gamma} \times P_{M'}) \leq \max_{P_{M'}} I(M';E|W_E \circ \tilde{\Gamma} \times P_{M'} || W_E \circ \Gamma \circ P_{\text{mix},M}) \]

\[
= \frac{1}{M/4} \sum_{l=1}^{M/4} P_{M'}(l)D(W_E \circ \Gamma_{m_l} || W_E \circ \Gamma \circ P_{\text{mix},M}) \leq 4I_{\text{mix}}(M;E|W_E \circ \Gamma).
\]

Since any distribution \( P_{M'} \) on \( \mathcal{M}' \) satisfies \( \|W_E \circ \tilde{\Gamma} \circ P_{M'} - W_E \circ \Gamma \circ P_{\text{mix},M}\|_1 \leq \max_{l=1, \ldots, M/4} \|W_E \circ \tilde{\Gamma}_l \circ P_{\text{mix},M}\|_1 \),

\[
d_{1,\max}(M';E|W_E \circ \tilde{\Gamma}) = \max_{P_{M'}} \sum_{l=1}^{M/4} P_{M'}(l)\|W_E \circ \tilde{\Gamma}_l - W_E \circ \Gamma \circ P_{M'}\|_1 \leq 2 \max_{l=1, \ldots, M/4} \|W_E \circ \tilde{\Gamma}_l - W_E \circ \Gamma \circ P_{\text{mix},M}\|_1 \leq 8d_{1,\text{mix}}(M;E|W_E \circ \Gamma).
\]

So, we obtain the desired argument.

**Proofs of (66) and (67):** In order to show (66), we consider another protocol generating a secret random number. Alice and Bob prepare \( l \) random permutations \( g = (g_1, \ldots, g_l) \) among \( \{1, \ldots, M\} \). First, Alice sends Bob the message \( l \) and \( m \) based on the code \( \Phi[Z'] \). Second, Alice and Bob apply the \( l \) random permutations to their message in the way \((m, l) \mapsto (g_l(m), l)\). Finally, Alice and Bob discard \( l \) and obtain \( g_l(m) \). That is, Alice and Bob apply a hash function \( f_g : (m, l) \mapsto g_l(m) \). Alice and Bob apply the encoder \( \Gamma[Z, g]_m(x) := \sum_{l=1}^l P_l(l)\delta_{g_l(m_{(l)(i)})}x \). Bob recovers the value \( m \) by using the decoder \( D_m[Z, g] := \sum_{l=1}^l D_{g_l(m)}[Z] \). We denote the code \((M, \Gamma[Z, g], \{D_m[Z, g]\})\) by \( \Phi[Z, g] \). Due to the construction, the hash function \( f_g \) satisfies Condition 3 in Appendix D. Applying (177), for any density \( \sigma \) on the system \( \mathcal{H}_E \), we obtain

\[
E_{g[Z]} d_{1,\text{mix}}(M;E|W_E \circ \Gamma[Z, g])
\]

\[
= \sum_{g[Z]} d_{1,\text{mix}}(f_g(M, L); E|W_E \circ \Phi[Z] \times P_{\text{mix},M} \times \mathcal{L}) \leq (4 + \sqrt{\sigma}) \sqrt{\text{Tr} \sum_{m=1}^M \sum_{l=1}^L \left( \frac{P_l(l)}{M} \right)^{1+s} W_{E|\Phi[Z]_m} \sigma^{-s})^{1/2}}
\]

\[
= (4 + \sqrt{\sigma}) \left( \sum_{m=1}^M \frac{1}{M} \sum_{l=1}^L P_l(l)^{1+s} \text{Tr} W_{E|\Phi[Z]_m} \sigma^{-s})^{1/2} \right).
\]
where \( v \) is the number of eigenvalues of \( \sigma \). Taking the average concerning the variable \( Z \), we obtain

\[
E_Z E_{\mathbf{g}|Z} d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z, \mathbf{g}])
\]

\[
\leq (4 + \sqrt{v}) E_Z \left( \sum_{m=1}^{M} \sum_{l=1}^{L} P_L(l)^{1+s} \text{Tr} W_{E|\Phi[Z]|_{l,m}}^{1+s} \right)^{1/2}
\]

\[
\leq (4 + \sqrt{v}) (\sum_{m=1}^{M} \sum_{l=1}^{L} P_L(l)^{1+s} E_Z \text{Tr} W_{E|\Phi[Z]|_{l,m}}^{1+s})^{1/2}
\]

\[
= (4 + \sqrt{v}) (\sum_{l=1}^{L} P_L(l)^{1+s} \text{Tr} \sum_{x} P_X(x)(W_{E|x})^{1+s})^{1/2}.
\] (76)

According to Lemma 1, we choose \( \sigma \) to be \( c(\sum_{x} P_X(x)(W_{E|x})^{1+s})^{1/(1+s)} \) with the normalizing constant \( c \) and \( v \) to be the number of eigenvalues of \( \sum_{x} P_X(x)(W_{E|x})^{1+s} \). Then, we obtain

\[
E_Z E_{\mathbf{g}|Z} d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z, \mathbf{g}])
\]

\[
\leq (4 + \sqrt{v}) (\sum_{l=1}^{L} P_L(l)^{1+s} \text{Tr} (\sum_{x} P_X(x)W_{E|x})^{1+s})^{1/2}
\]

\[
= (4 + \sqrt{v}) e^{-\frac{s}{2} H_{l+s}(P_L) + \frac{s}{2} I_s^G(X; E|W_E \times P_X)}.
\] (77)

Finally, when \( g \) is fixed, the statistical behavior of \( \Gamma[Z, g] \) is the same as that of \( \Gamma[Z] \). That is,

\[
E_Z d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z, \mathbf{g}]) = E_Z d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z])
\] (78)

for \( g \), which implies (66). Similarly, we can show (67) by using (176) instead of (175).

Remark 1: One might consider that Inequality (64) could be derived from a kind of privacy amplification lemma \([23, \text{Theorem 1]} \) similar to (66) from \([20, (12)] \). However, the strategy cannot yield Inequality (64) due to the following reason. In order to show an inequality corresponding to (76), we need to show

\[
E_Z \sum_{m,l} \frac{P_L(l)^{1+s}}{M} \text{Tr} (W_{E \circ \Gamma[Z]|_{l,m}}^{1+s}) \leq \left( \sum_{m,l} \frac{P_L(l)}{M} W_{E \circ \Gamma[Z]|_{l,m}}^{1+s} \right)^{-s}
\]

\[
\leq \left( \sum_{l} P_L(l)^{1+s} \text{Tr} (\sum_{x} P_X(x)W_{E|x})^{1+s} \right)^{-s}.
\] (79)

However, only the opposite inequality holds, in general. Hence, this method cannot be applied to the proof of (64).

Further, using (48) and (50), we can obtain other type bounds. Combining (48) and (64), we obtain an alternative bound of \( E_Z d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z]) \) as

\[
E_Z d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z]) \leq \sqrt{E_Z d_{1,\text{mix}}(M; E|W_E \circ \Gamma[Z])^2}
\]

\[
\leq \sqrt{\frac{2}{N^2/2}} e^{-\frac{s}{2} H_{l+s}(P_L) + \frac{s}{2} I_s^G(X; E|W_E \times P_X)}.
\] (80)

Combining (50), (66), and (67), we obtain

\[
E_Z D(W_E \circ \Gamma[Z]|_{m} || W_E \circ P_X)
\]

\[
\leq (\log d_E - \log \mu_s - \frac{s}{2} H_{l+s}(P_L) + \frac{s}{2} I_s^G(X; E|W_E \times P_X))
\]

\[
\cdot \mu_s e^{-\frac{s}{2} H_{l+s}(P_L) + \frac{s}{2} I_s^G(X; E|W_E \times P_X)}.
\]

Hence, instead of (64), we obtain an alternative bound of \( E_Z I_{\text{mix}}(M; E|W_E \circ \Gamma[Z]) \) as

\[
E_Z I_{\text{mix}}(M; E|W_E \circ \Gamma[Z])
\]

\[
\leq (\log d_E - \log \mu_s - \frac{s}{2} H_{l+s}(P_L) - \frac{s}{2} I_s^G(X; E|W_E \times P_X))
\]

\[
\cdot \mu_s e^{-\frac{s}{2} H_{l+s}(P_L) + \frac{s}{2} I_s^G(X; E|W_E \times P_X)}.
\] (81)

As is shown in Subsection IV-C these alternative bounds are weaker than the bounds (65), (66), and (67) in the asymptotic setting.
B. Asymptotic analysis

In the following, we focus on the n-fold discrete memoryless channels of the channels $W_B$ and $W_E$, which are written as $W_B^{(n)}$ and $W_E^{(n)}$. The $n$-independent and identical distribution $P^n_X$ of $P_X$ satisfies the additive equation $I_s(X; B|W_B^{(n)} \times P^n_X) = nI_s(X; B|W_B \times P_X)$. In this case, we assume that a random variable subject to the distribution $P_{L_n}$ on $L_n := \{1, \ldots, L_n\}$ is available. Thus, there exists a code $\Phi_n[P_X]$ with an encoder $\Gamma_n[P_X]$ for any integer $M_n$, and any probability distribution $P_X$ on $\mathcal{X}$ such that the code $\Phi_n$ only uses the distribution $P_{L_n}$ for mixing the input alphabet and

$$|\Phi_n| = M_n/4$$

$$\epsilon_{\max}(W_B^{(n)}|\Phi_n[P_X]) \leq \frac{48}{n} \min_{0 \leq s \leq 1} (M_n L_n)^s e^{-sI_{-s}(X; B|W_B \times P_X)} \quad (82)$$

$$I_{\max}(W_E^{(n)} \circ \Gamma_n[P_X]) \leq 12v_n \min_{0 \leq s \leq 1} \frac{e^{-sH_{1+s}(P_{L_n}) + usI_s(X; E|W_E \times P_X)}}{s} \quad (83)$$

and

$$d_{1,\max}(W_E^{(n)} \circ \Gamma_n[P_X]) \leq 24 \min_{0 \leq s \leq 1} \mu_{s,n} e^{-sH_{1+s}(P_{L_n}) + n\Phi_s(X; E|W_E \times P_X)}, \quad (84)$$

$$\mu_{s,n} := \min\{4 + \sqrt{v_n}, 4 + \sqrt{\max\{n, 1\}}\}, \quad (85)$$

where $v_n$ and $v_{s,n}$ are the numbers of eigenvalues of $(W_E \circ P_X)^{\otimes n}$ and $(\sum_{x} P(x)(W_E(x))^{1+s})^{\otimes n}$. The numbers $v_n$ and $v_{s,n}$ are bounded by $(\dim H) \cdot n + 1$. Hence, due to (33), when $M_n, L_n \geq e^{nI(X; B|W_B \times P_X)}$ and $e^{-sH_{1+s}(P_{L_n})} \leq L_n^s \leq e^{n\Phi_s(X; E|W_E \times P_X)}$ for the values (82), (83), and (84) go to zero. That is, the rate $\max P_X I(X; B|W_B \times P_X) - I(X; E|W_E \times P_X)$ can be asymptotically attained, as shown by Devetak [7]. Now, we focus on the exponential decreasing rates of our upper bounds. We assume that $H_2(P_{L_n}) \geq nR$, which implies that $H_{1+s}(P_{L_n}) \geq H_2(P_{L_n}) \geq nR$ for $s \in [0, 1]$. In fact, $H_2(P_{L_n})$ can be regarded as the sacrifice information.

Now, we denote the ensemble of codes and encoders given in Subsection [IV-A] with the $n$-i.i.d. distribution $P^n_X$ by $\Phi[Z_n, P^n_X]$ and $\Gamma[Z_n, P^n_X]$, respectively. Using the ensemble, we define two kinds of the decreasing rates under the above code:

$$e_1(R|W_E, P_X) := \lim_{n \to \infty} -\frac{1}{n} \log E_{Z_n} I_{\max}(M_n; E|W_E^{(n)} \circ \Gamma[Z_n, P^n_X])$$

$$e_d(R|W_E, P_X) := \lim_{n \to \infty} -\frac{1}{n} \log E_{Z_n} d_{1,\max}(M_n; E|W_E^{(n)} \circ \Gamma[Z_n, P^n_X]).$$

Inequality (83) yields

$$e_1(R|W_E, P_X) \geq e_0(R|W_E, P_X) := \max_{0 \leq s \leq 1} sR - sI_s(X; E|W_E \times P_X),$$

and Inequality (84) yields

$$e_d(R|W_E, P_X) \geq G(R|W_E, P_X) := \max_{0 \leq s \leq 1} \frac{s}{2} R - \frac{s}{2} \Phi_s(X; E|W_E \times P_X).$$

Note that the exponent $e_0(R|W_E, P_X)$ with the commutative case is the same as that by the previous paper [20]. However, the exponent $e_G(R|W_E, P_X)$ with the commutative case is smaller than that by the previous paper [19].

In the quantum wiretap channel model, it is known that a pre noisy processing $\Gamma : \mathcal{V} \to \mathcal{X}$ may improve the capacity in the classical case, where $\Gamma$ is a stochastic matrix from $\mathcal{V}$ to $\mathcal{X}$. When we apply the pre noisy processing $\Gamma$, the rate $\max_{P_{\Gamma}} I(V; B|W_B \circ \Gamma \times P_V) - I(V; E|W_E \circ \Gamma \times P_V)$ can be attained asymptotically, where $P_{\Gamma}$ is the distribution on $\mathcal{V}$. Applying our method to the pair of channels $W_B \circ \Gamma$ and $W_E \circ \Gamma$, we obtain an upper bound for error probability and leaked information, which goes to zero exponentially and can attain the rate $\max_{P_{\Gamma}, \Gamma} I(V; B|W_B \circ \Gamma \times P_V) - I(V; E|W_E \circ \Gamma \times P_V)$. Further, by choosing $\Gamma$ as a stochastic matrix from $\mathcal{V}$ to $\mathcal{X}'$, the rate $\lim_{n \to \infty} \frac{1}{n} \max_{P_{\Gamma}, \Gamma} I(V; B|W_B^{(n)} \circ \Gamma \times P_V) - I(V; E|W_E^{(n)} \circ \Gamma \times P_V)$ can be attained asymptotically.

Now, we define the capacity $C_{W_B, W_E}$

$$C_{W_B, W_E} := \sup_{\{\Phi_n, \Gamma_n\}} \left\{ \lim_{n \to \infty} \frac{\log |\Phi_n|}{n} \left| \lim_{n \to \infty} I_{\max}(M_n; E|W_E \circ \Gamma_n) = 0, \lim_{n \to \infty} \epsilon_B(\Phi_n) = 0 \right. \right\}.$$
Proposition 1:

\[
C_{W_B, W_E} = \lim_{n \to \infty} \frac{1}{n} \max_{P_{x, i}} I(V; B|W_B^{(n)} \circ \Gamma \times P_V) - I(V; E|W_E^{(n)} \circ \Gamma \times P_V).
\]  
(87)

If there is a quantum channel \( C \) such that \( C(W_{B|x}) = W_{E|x} \), the quantum wiretap channel \( W_B, W_E \) is called degraded. It is known that the degraded channel \( W_B, W_E \) satisfies that\[14\] (9.62)

\[
\max_{P_x} I(X; B|W_B \times P_X) - I(X; E|W_E \times P_X)
= \frac{1}{n} \max_{P_{x, i}} I(V; B|W_B^{(n)} \circ \Gamma \times P_V) - I(V; E|W_E^{(n)} \circ \Gamma \times P_V).
\]

That is,

\[
C_{W_B, W_E} = \max_{P_x} I(X; B|W_B \times P_X) - I(X; E|W_E \times P_X).
\]

The detail property of \( C_{W_B, W_E} \) has been studied by \[26\] in this case.

In the classical case, Csiszár et al \[2\] showed

\[
\max_{P_{x, i}} I(V; B|W_B \circ \Gamma \times P_V) - I(V; E|W_E \circ \Gamma \times P_V)
= \frac{1}{n} \max_{P_{x, i}} I(V; B|W_B^{(n)} \circ \Gamma \times P_V) - I(V; E|W_E^{(n)} \circ \Gamma \times P_V).
\]

That is,

\[
C_{W_B, W_E} = \max_{P_{x, i}} I(V; B|W_B \circ \Gamma \times P_V) - I(V; E|W_E \circ \Gamma \times P_V).
\]

C. Comparison

When we replace the role of the pair of \(66\) and \(67\) by that of \(80\), we obtain another bound

\[
e_d(R|W_E, P_X) \geq \frac{1}{2} e_R(R|W_E, P_X).
\]

Similarly, replacing the role of \(64\) by that of \(81\), we obtain

\[
e_f(R|W_E, P_X) \geq e_G(R|W_E, P_X).
\]

For a comparison between these inequalities and Inequalities \(64\) and \(83\), we have the following lemma.

Lemma 7: In the general case, we have

\[
\frac{1}{2} e_R(R|W_E, P_X) \leq e_G(R|W_E, P_X).
\]  
(88)

When \( W_x \) are commutative each other, we have

\[
e_G(R|W_E, P_X) \leq e_R(R|W_E, P_X).
\]  
(89)

Hence, Inequality \(84\) provides a better bound for \(e_d(R|W_E, P_X)\). That is, Inequality \(64\) is a better evaluation for a sufficiently large number \( n \). In the commutative case, Inequality \(83\) provides a better bound for \(e_f(R|W_E, P_X)\). That is, Inequalities \(66\) and \(67\) are better evaluations for a sufficiently large number \( n \). These numerical comparisons for a non-commutative example will be given in Section VII.

Proof: Inequality \(85\) of Lemma 1 yields that

\[
\frac{1}{2} e_R(R|W_E, P_X) = \max_{0 \leq s \leq 1} \frac{s}{2} R - \frac{s}{2} I_s(X; E|W_E \times P_X)
\leq \max_{0 \leq s \leq 1} \frac{s}{2} R - \frac{s}{2} I_s^G(X; E|W_E \times P_X) = e_G(R|W_E, P_X),
\]

which implies \(88\).

Inequality \(89\) is shown in the following way. Since the map \( s \mapsto I_s(X; E|W_E \times P_X) \) is monotonically increasing, Lemma 2 yields

\[
\frac{s}{2} I_s^G(X; E|W_E \times P_X) \geq \frac{s}{2} I_{\frac{s}{2}}(X; E|W_E \times P_X)
\geq \frac{s}{2} I_{1/2}(X; E|W_E \times P_X).
\]
Then,
\[
\epsilon_G(R|W_E, P_X) = \max_{0 \leq s \leq 1} \frac{s}{2} R - \frac{s}{2} I_s^G(X; E|W_E \times P_X)
\]
\[
\leq \max_{0 \leq s \leq 1} \frac{s}{2} R - \frac{s}{2} I_{s/2}(X; E|W_E \times P_X) \leq \epsilon_R(R|W_E, P_X),
\]
which implies (89).

On the other hand, combining (9.79) and (9.53) in the book [1], we obtain
\[
e_{\ell}(R|W_E, P_X) \geq \max_{0 \leq s \leq 1} \frac{sR - sI_s(X; E|W_E \times P_X)}{2(1 + s)}.
\]
\[
e_d(R|W_E, P_X) \geq \max_{0 \leq s \leq 1} \frac{sR - sI_s(X; E|W_E \times P_X)}{2(1 + s)}.
\]

Our lower bounds of exponents \(\epsilon_G(R|W_E, P_X)\) and \(\epsilon_{\ell}(R|W_E, P_X)\) improve them.

Next, we compare the evaluations (64) and (81) for \(d_E = \dim \mathcal{H}_E\) is two, \(v\) increases only linearly. However, when \(d_E \geq 3\), they increase with the order \(n^{d_E - 1}\). This factor might be not negligible when \(n\) is not sufficiently large. However, when \(d_E\) is finite, \(\lambda\) increases linearly. In the evaluation (81), we have other factors, i.e., \(\log d_E\) and the logarithm of the upper bound given in (66) or (67). These factors also increase linearly. Hence, in the evaluation (81), all of the polynomial factors increase in the order \(n^{3/2}\) while the polynomial factor in the evaluation (64) increases in the order \(n^{s(d_E - 1)}\). Now, we fix the optimal \(s\) in the evaluation (64). Then, when \(d_E > \frac{3}{2} + 1\), the polynomial factor in the evaluation (81) is smaller than that in the evaluation (64). Hence, when \(n\) is not sufficiently large, the evaluation (81) might be better than the evaluation (64).

Finally, we compare these evaluations when the channel \(W^{(n)}\) is not stationary memoryless and the distribution \(P_{L_n}\) is not independent and identical. In this case, the speeds of increase of \(v\) and \(v_s\) are not polynomial in general. Hence, even though \(n\) is sufficiently large, the factor \(v\) and \(v_s\) are not negligible. However, \(\lambda_s\) increases linearly when the logarithm of the minimum eigenvalue of \((\sum_x P_X(x)^{(n)}(W^{(n)})^{1+s})^{-1/3} / \text{Tr}(\sum_x P_X(x)^{(n)}(W^{(n)})^{1+s})^{-1} \to \infty\) behaves linearly. Hence, the evaluations (67) and (81) work well under the above weak assumption.

V. Equivocation rate

When the coding rate \(R\) is larger than the capacity, the Eve’s information does not go to zero. In this case, it is usual in the classical setting to evaluate the limit of \(\frac{I_{\max}(M; E|W_E \circ \Gamma_{\alpha})}{15}\) with a sequence of encoders \(\Gamma_{\alpha}\). In the same construction as Section IV by using (33) and the convexity of \(x \mapsto e^x\), (53) yields the inequalities
\[
F_{\mathcal{Z}} e^{sI_{\max}(M; E|W_E \circ \Gamma_{\alpha})}
\]
\[
\geq F_{\mathcal{Z}} e^{sI_{\max}(M; E|W_E \circ \Gamma_{\alpha})} - F_{\mathcal{Z}} e^{sI_{\max}(M; E|W_E \circ \Gamma_{\alpha})} + sD(W_E \circ \Gamma_{\alpha} || W_E \circ P_X)
\]
\[
= F_{\mathcal{Z}} e^{sD(W_E \circ \Gamma_{\alpha} || W_E \circ P_X)}
\]
\[
\leq F_{\mathcal{Z}} e^{sD(W_E \circ \Gamma_{\alpha} || W_E \circ P_X)} (1 + e^{-sH_{1+s}(P_{L_n})} e^{sI_s(X; E|W_E \times P_X)})
\]
\[
= e^{s(1 + e^{-sH_{1+s}(P_{L_n})} e^{sI_s(X; E|W_E \times P_X)})}
\]
for \(0 < s \leq 1\), where \(v\) is the number of eigenvalues of \(W_E \circ P_X\). Here, (90), (91), and (92) follow from (50), the convexity of \(x \mapsto e^x\), and (53), respectively.

Since the Markov inequality guarantees the inequality
\[
P_{\mathcal{Z}} \{ e(W_B|\Phi[Z]) > 2E_{\mathcal{Z}} e(W_B|\Phi[Z]) \} < \frac{1}{2}
\]
\[
P_{\mathcal{Z}} \{ e^{sI_{\max}(M; E|W_E \circ \Gamma_{\alpha})} > 2E_{\mathcal{Z}} e^{sI_{\max}(M; E|W_E \circ \Gamma_{\alpha})} \} < \frac{1}{2},
\]
there exists a code $\Phi$ with an encoder $\Gamma$ such that $\epsilon(W_B|\Phi) \leq 2E\epsilon(W_B|\Phi[Z])$ and $e^{sI_{\text{mix}}(M;E|W_E \circ \Gamma)} \leq 2E\epsilon^{sI_{\text{mix}}(M;E|W_E \circ \Gamma[Z])}$.

Since the relations

$$\frac{1}{s} \log 2 v^s (1 + e^{-H_{1+\eta}(P_L)} e^{sI_s(X;E|W_E \times P_X)})$$

$$\leq \log v + \frac{1}{s} (\log 4 + [\log e^{-sH_{1+\eta}(P_L)} e^{sI_s(X;E|W_E \times P_X)}]_+)$$

$$= \log v + \frac{1}{s} (\log 4 + [sI_s(X;E|W_E \times P_X) - sH_{1+\eta}(P_L)]_+)$$

hold, the existence of a good code is guaranteed in the following way.

**Theorem 2:** Assume that a random variable $L$ subject to the distribution $P_L$ on $\{1, \ldots, L\}$ is available for an auxiliary random number. There exists a code $\Phi$ with an encoder $\Gamma$ for any integer $M$, and any probability distribution $P_X$ on $\mathcal{X}$ such that the encoder $\Gamma$ uses only the distribution $P_L$ for mixing the input alphabet and

$$|\Phi| = M$$

$$\epsilon(W_B|\Phi) \leq 8 \min_{0 \leq s \leq 1} (ML)^s e^{-sI_{\text{mix}}(X;B|W_B \times P_X)}$$

$$I_{\text{mix}}(M;E|W_E \circ \Gamma) \leq \log v + \frac{1}{ns} (\log 4 + [nsI_s(X;E|W_E \times P_X) - sH_{1+\eta}(P_L)]_+),$$

for $0 < s \leq 1$, where $v$ is the number of eigenvalues of $W_E \circ P_X$.

In the $n$-fold discrete memoryless channels $W_B^{(n)}$ and $W_E^{(n)}$ of the channels $W_B$ and $W_E$, the additive equation $I_s(X;B|W_B^{(n)} \times P_X^0) = nI_s(X;B|W_B \times P_X)$ holds. Assume that a random variable subject to the distribution $P_L$ on $\{1, \ldots, L\}$ is available. Thus, for any integer $M_n$ and any probability distribution $P_X$ on $\mathcal{X}$, there exists a code $\Phi_n$ with an encoder $\Gamma_n$ such that the encoder $\Gamma_n$ only uses the distribution $P_L$ for mixing the input alphabet and

$$|\Phi_n| = M_n$$

$$\epsilon_B(\Phi_n) \leq 8 \min_{0 \leq s \leq 1} (M_nL_n)^s e^{-nsI_{\text{mix}}(X;B|W_B \times P_X)}$$

$$I_{\text{mix}}(M_n;E|W_E^{(n)} \circ \Gamma_n) \leq \log v_n + \frac{1}{ns} (\log 4 + [nsI_s(X;E|W_E \times P_X) - sH_{1+\eta}(P_L)]_+),$$

for $0 < s \leq 1$, where $v_n$ is the number of eigenvalues of $W_E \circ P_X^\otimes n$.

When the sacrifice information rate is $R_0$, i.e., $sH_{1+\eta}(P_L) \leq sR_0$, the above code $\Phi_n$ satisfies

$$\lim_{n \to \infty} I_{\text{mix}}(M_n;E|W_E^{(n)} \circ \Gamma_n) \leq [I_s(X;E|W_E \times P_X) - R_0]_+$$

for $0 < s \leq 1$. Since the function $s \mapsto I_s(X;E|W_E \times P_X)$ is monotone increasing, $\inf_{0 < s \leq 1} I_s(X;E|W_E \times P_X) = \lim_{s \to 0} I_s(X;E|W_E \times P_X) = I(X;E|W_E \times P_X)$. Therefore, we obtain

$$\lim_{n \to \infty} I_{\text{mix}}(M_n;E|W_E^{(n)} \circ \Gamma_n) \leq I(X;E|W_E \times P_X) - R_0.$$ (98)

Now, we define the leaked information rate:

$$I_{W_B,W_E}(R) := \inf_{\{\Phi_n\}} \left\{ \lim_{n \to \infty} I_{\text{mix}}(M_n;E|W_E^{(n)} \circ \Gamma_n) \left| \epsilon_B(\Phi_n) \to 0, \frac{\log M_n}{n} \to R \right. \right\}.$$ (97)

We assume $M_n = e^{nR - \sqrt{n}}$, $L_n = e^{nR_0}$ and $I(X;B|W_B \times P_X) = R + R_0$, the error probability $\epsilon_B(\Phi_n)$ goes to zero in the above construction. Then,

$$\lim_{n \to \infty} I_{\text{mix}}(M_n;E|W_E^{(n)} \circ \Gamma_n) \leq I(X;E|W_E \times P_X) - (I(X;B|W_B \times P_X) - R)$$

$$= R - (I(X;B|W_B \times P_X) - I(X;E|W_E \times P_X)),$$

which implies

$$I_{W_B,W_E}(R) \leq R - (I(X;B|W_B \times P_X) - I(X;E|W_E \times P_X)).$$ (99)

Define

$$H(R) := \lim_{n \to \infty} \max_{\{P_{V_n},\Gamma_n\}} \frac{I(V_n;B|W_B^{(n)} \circ \Gamma_n \times P_{V_n}) - I(V;E|W_E^{(n)} \circ \Gamma_n \times P_{V_n})}{n},$$ (100)
where we take the maximum under the condition $R \leq \frac{I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n})}{n}$. Note that the limit in RHS of (100) equals the limit infimum in RHS of (100). Now, we choose a pair $(P_{V_n}, \Gamma_n)$ such that $R \leq \frac{I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n})}{n}$. Applying (59) to the channel pair $(W_B^{(n)} \circ \Gamma_n, W_E^{(n)} \circ \Gamma_n)$ with the distribution $P_{V_n}$, we obtain

$$I_{W_B,W_E}(R) \leq R - \frac{I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n}) - I(V_n; E|W_E^{(n)} \circ \Gamma_n \times P_{V_n})}{n},$$

which implies that

$$I_{W_B,W_E}(R) \leq R - H(R).$$

(101)

In fact, we obtain the following theorem.

**Theorem 3:**

$$I_{W_B,W_E}(R) = R - H(R).$$

(102)

The proof of this theorem will be given in the end of this subsection. According to [6], we define the equivocation rate:

$$H_{W_B,W_E}(R) := \inf_{\{P_{V_n}, \Gamma_n\}} \lim_{n \to \infty} \frac{H(M_n|E|W_E^{(n)} \circ \Gamma_n \times P_{mix,M_n})}{n} \left\{ \epsilon_B(\Phi_n) \to 0 \right\}$$

where $M_n$ is the random variable to be sent and $\Phi_n = (M_n, \{\Gamma_n\}, \{D_n\})$. Since $H(M_n|E|W_E^{(n)} \circ \Gamma_n \times P_{mix,M_n}) = \log M_n - I_{mix}(M; E|W_E^{(n)} \circ \Gamma_n)$, we obtain

$$H_{W_B,W_E}(R) = H(R).$$

(103)

We define the critical rate

$$R^* := \sup_{\{(P_{V_n}, \Gamma_n)\}} I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n}),$$

(104)

where we take the supremum under the condition $\lim_{n \to \infty} \frac{I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n}) - I(V_n; E|W_E^{(n)} \circ \Gamma_n \times P_{V_n})}{n} = C_{W_B,W_E}$. Then, we have $H(R) = C_{W_B,W_E}$ for $R \leq R^*$. For $R > R^*$, $H(R)$ is smaller than the capacity $C_{W_B,W_E}$. We also have the following lemma.

**Lemma 8:** In the case of degraded channel, $H(R)$ is calculated as follows.

$$H(R) = \max_{P_X:I(X;B|W_B \times P_X)\geq R} I(X; B|W_B \times P_X) - I(X; E|W_E \times P_X).$$

(105)

In the general classical case, $H(R)$ can be single-letterized by using two auxiliary random variables[6]. In the general quantum case, the converse part for the single-letterized formula has a crucial difficulty, and the direct part requires a quantum analogue of the superposition coding. Thus, we do not treat the single-letterization of $H(R)$.

**Proof of Lemma 5** Since the channel is degraded, the inequality

$$I(V_n; B|W_B^{(n)} \circ \Gamma_n \times P_{V_n}) - I(V_n; E|W_E^{(n)} \circ \Gamma_n \times P_{V_n})$$

$$\leq I(X^n; B|W_B^{(n)} \times \Gamma_n \circ P_{V_n}) - I(X^n; E|W_E^{(n)} \times \Gamma_n \circ P_{V_n})$$

holds [5 Exercise 9.19]. Further, for two classical-quantum channels $W_B$ and $\tilde{W}_B$ with the input classical system $\mathcal{X}$ and $\tilde{\mathcal{X}}$, we define the classical-quantum channel $W_B \otimes \tilde{W}_B$ by $(W_B \otimes \tilde{W}_B)_z := W_{B|x} \otimes \tilde{W}_{B|z}$. For a distribution $P_{X,\tilde{X}}$ on $\mathcal{X} \times \tilde{\mathcal{X}}$, we choose $P_X$ and $P_{\tilde{X}}$ to be the marginal distributions of $P_{X,\tilde{X}}$. Then, we have [5 Exercise 9.13]

$$I(X, \tilde{X}; B|W_B \otimes \tilde{W}_B \times P_{X,\tilde{X}}) - I(X, \tilde{X}; E|W_E \otimes \tilde{W}_E \times P_{X,\tilde{X}})$$

$$\leq I(X; B|W_B \times P_X) - I(X; E|W_E \times P_X) + I(\tilde{X}; B|\tilde{W}_B \times P_{\tilde{X}} - I(\tilde{X}; E|\tilde{W}_E \times P_{\tilde{X}})$$

and

$$I(X, \tilde{X}; B|W_B \otimes \tilde{W}_B \times P_{X,\tilde{X}}) \leq I(X; B|W_B \times P_X) + I(\tilde{X}; B|\tilde{W}_B \times P_{\tilde{X}}).$$
Hence, for $0 \leq k \leq n$,
\[
\max_{\langle P, \Gamma \rangle: I(V;B|W_B^{(n)} \circ \Gamma \times P_V) \geq R} I(V;B|W_B^{(n)} \circ \Gamma \times P_V) - I(V;E|W_E^{(n)} \circ \Gamma \times P_V)
\]
\[
= \max_{P_X: I(X^n;B|W_B^{(n)} \times P_X) \geq R} I(X^n;B|W_B^{(n)} \times P_X^n) - I(X;E|W_E^{(n)} \times P_X^n)
\]
\[
\geq \max_{(P_X, P_{X^n-k}): I(X^n-k;B|W_B^{(n-k)} \times P_{X^n-k}) \geq R} (I(X^k;B|W_B^{(k)} \times P_{X^k}) - I(X^k;E|W_E^{(k)} \times P_{X^k}))
\]
\[
+ (I(X^n-k;B|W_B^{(n-k)} \times P_{X^n-k}) - I(X^n-k;E|W_E^{(n-k)} \times P_{X^n-k}))
\]
\[
= \max_{(P_X, \ldots, P_{X,n})} \sum_{i=1}^{n} (I(X;B|W_B \times P_X,i) - I(X;E|W_E \times P_X,i)).
\]

Since, as is shown latter,
\[
\sum_{i=1}^{n} I(X;B|W_B \times P_X,i) \leq n I(X;B|W_B \times \sum_{i=1}^{n} \frac{1}{n} P_X,i),
\]
(106)
\[
\sum_{i=1}^{n} (I(X;B|W_B \times P_X,i) - I(X;E|W_E \times P_X,i))
\]
\[
\leq n (I(X;B|W_B \times \sum_{i=1}^{n} \frac{1}{n} P_X,i) - I(X;E|W_E \times \sum_{i=1}^{n} \frac{1}{n} P_X,i)),
\]
(107)

the maximum is realized when $P_{X,i} = \sum_{i=1}^{n} \frac{1}{n} P_X,i$. Therefore, we obtain (105).

Now, we show (106) and (107). (106) is shown from the concavity of von Neumann entropy. The proof of (107) is more difficult. It is enough to show
\[
\lambda(I(X;B|W_B \times P_X) - I(X;E|W_E \times P_X)) + (1 - \lambda)(I(X;B|W_B \times P_X') - I(X;E|W_E \times P_X'))
\]
\[
\leq I(X;B|W_B \times (\lambda P_X + (1 - \lambda) P_X')) - I(X;E|W_E \times (\lambda P_X + (1 - \lambda) P_X')).
\]

We choose a TP-CP map $\Lambda$ such that $W_E|_{\mathcal{X}} = \Lambda(W_B|_{\mathcal{X}})$. Then, we obtain
\[
D(W_E \circ P_X \| W_E \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
= D(\Lambda(W_B \circ P_X) \| \Lambda(W_B \circ (\lambda P_X + (1 - \lambda) P_X')))
\]
\[
\leq D(W_B \circ P_X \| W_B \circ (\lambda P_X + (1 - \lambda) P_X')).
\]
(108)

Using (108), we obtain
\[
I(X;B|W_B \times (\lambda P_X + (1 - \lambda) P_X')) - I(X;E|W_E \times (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
- \lambda(I(X;B|W_B \times P_X) - I(X;E|W_E \times P_X)) + (1 - \lambda)(I(X;B|W_B \times P_X') - I(X;E|W_E \times P_X'))
\]
\[
= H(W_B \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
- H(W_B \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
- \lambda H(W_B \circ P_X) - (1 - \lambda) H(W_B \circ P_X')
\]
\[
+ \lambda H(W_E \circ P_X) + (1 - \lambda) H(W_E \circ P_X')
\]
\[
= \lambda D(W_B \circ P_X \| W_B \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
+ (1 - \lambda) D(W_B \circ P_X' \| W_B \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
+ \lambda D(W_E \circ P_X \| W_E \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
+ (1 - \lambda) D(W_E \circ P_X' \| W_E \circ (\lambda P_X + (1 - \lambda) P_X'))
\]
\[
\geq 0
\]

Proof of Theorem 3 It is sufficient to show the inequality opposite to (101). Let $\Phi_n = (M_n, \{\Gamma_n\}, \{D_n\})$ be the sequence attaining $I_{W_B, W_E}(R)$. Then, when $M_n$ is the random variable to be sent and $\hat{M}_n$ is the random variable to be received, Fano inequality implies that
\[
h(\epsilon(W_B^n|\Phi_n)) + \epsilon(W_B^n|\Phi_n) \log M_n \geq H(M_n | \hat{M}_n).
\]
That is,
\[-h(\epsilon(W_B^{(n)}|\Phi_n)) + (1 - \epsilon(W_B^{(n)}|\Phi_n)) \log M_n\]
\[\leq I(M_n; \hat{M}_n) \leq I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n}).\]

Thus,
\[\frac{1}{n} \log M_n \leq \frac{I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n}) + h(\epsilon(W_B^{(n)}|\Phi_n))}{n(1 - \epsilon(W_B^{(n)}|\Phi_n))}.
\]

Taking the limit, we obtain
\[R \leq \lim_{n \to \infty} \frac{I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n}) + h(\epsilon(W_B^{(n)}|\Phi_n))}{n(1 - \epsilon(W_B^{(n)}|\Phi_n))} = \lim_{n \to \infty} \frac{I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n})}{n}.
\]

Further,
\[\lim_{n \to \infty} \frac{I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n}) - I(M_n; E|W_C^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n})}{n} \leq H(R).
\]

Since \(I_{W_B,W_E}(R) = \text{lim}_{n \to \infty} \frac{I(M_n; E|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n})}{n}\),
\[R - I_{W_B,W_E}(R) \leq \lim_{n \to \infty} \frac{I(M_n; B|W_B^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n}) - I(M_n; E|W_C^{(n)} \circ \Gamma_n \times P_{\text{mix},M_n})}{n} \leq H(R).
\]

That is
\[I_{W_B,W_E}(R) \geq R - H(R),\]
which implies \(102\).

\[\text{VI. Wire-tap Channel with Linear Coding}\]

\textbf{A. The case when uniform distribution is available}

1) General case: In a practical sense, we need to take into account the decoding time. For this purpose, we often restrict our error correcting codes to linear codes. While the constructions of codes in this section are different from those in Section [LV] the bounds obtained in this section are similar to those in Section [IV]. Hence, the evaluation in the \(n\)-fold discrete memoryless case can be derived in the same way as that in Section [LV] from the single-shot case by substituting the channel \(W_C^{(n)}\) into the channel \(W_E\). The source-universality also can be shown in the same way. Therefore, this section discusses only the single-shot case.

In the following, we consider the case where the sender’s space \(\mathcal{X}\) has the structure of a module \(\mathbb{F}_q^n\). First, we regard a submodule \(C_1 \subseteq \mathcal{X}\) as the set of transmitted messages, and focus on its decoding \(\{D_x\}_{x \in C_1}\) by the authorized receiver. In the following, for any element \(x \in \mathcal{X}\), \(x|C_2\) denotes the coset concerning the quotient by \(C_2\), and \(x|C_1\) denotes the coset concerning the quotient by \(C_1\). When the code \(C_2\) is fixed, \(x|C_1\) is simplified to \(x\). Based on a submodule \(C_2\) of \(C_1\), we construct a code for a quantum wiretap channel \(\Phi_{C_1,C_2} = (C_1/C_2, \{2^{C_1/C_2} - 1\} \times C_2, \{D_x\}_{x \in C_1/C_2})\) as follows. The encoding distribution corresponding to the message \(x \in C_1/C_2\) is given as the uniform distribution \(P_{\text{mix},x}\) on the coset \(x := x + C_2\), and the decoding \(D_x\) is given as the subset \(\sum_{x' \in x+C_2} D_{x'}\). As a generalization, for \(y|C_1 \in \mathcal{X}/C_1\), we consider a code \(\Phi_{y|C_1,C_2} = (C_1/C_2, \{2^{y|C_1,C_2} - 1\} \times y|C_1/C_2, \{D_x\}_{x \in y|C_1/C_2})\) as follows. The encoding distribution corresponding to the message \(x \in y|C_1/C_2\) is given as the uniform distribution on the coset \(x := x + C_2\), and the decoding \(D_x\) is given as the subset \(\sum_{x' \in x+C_2} D_{x'}\). Sometimes, this type code is useful.

Now, we treat the ensemble of the code pairs \(\{C_2\}_{C_1} \subseteq \{C_1\}\) with the size \(|C_1/C_2| = M\) and \(|C_2| = L\). We identify the set \([y|C_1/C_2]/C_2[/Z] with the set of messages \(\mathcal{M} = \{1, \ldots, M\}\). We also define \(v\), \(v_s\), \(\chi_s\) with \(P_X = P_{\text{mix},\mathcal{X}}\) similar to Subsection [V-L].
Lemma 9: We suppose that the coset \([Y]_{C_1[Z]} \in X/C_1[Z]\) is chosen with the uniform distribution for the given code pair \(C_2[Z] \subset C_1[Z]\) and that the ensemble \(\{C_2[Z]\}\) satisfies Condition \([1]\) as a subset of \(X\). The relations
\[
E_{Z,Y} I_{mix}(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]}) \\
\leq u^s \frac{sI_l(X; E|W_E \times P_{mix,x})}{L^s s},
\]
(109)
\[
E_{Z,Y} sI_{mix}(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]}) \\
\leq u^s (1 + \frac{e^{sI_l(X; E|W_E \times P_{mix,x})}}{L^s} )
\]
(110)
\[
E_{Z,Y} E\{I_{mix}(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]}) \} \\
\leq u^s \frac{H_{\mu}(e^{sI_l(X; E|W_E \times P_{mix,x})})}{L^s/2}
\]
(111)
hold for \(0 < s \leq 1\) and \(\mu_s := \min\{4 + \sqrt{v_s}, 4 + \sqrt{\lambda_s}\} e^s\).

Lemma 10: We suppose that the coset \([Y]_{C_1[Z]} \in X/C_1[Z]\) is chosen with the uniform distribution for the given code pair \(C_2[Z] \subset C_1[Z]\) and that the ensemble \(\{C_1[Z]\}\) satisfies Condition \([1]\) as a subset of \(X\). The relation
\[
E_{Z,Y} \epsilon(W_B|\Phi_{[Y]_{C_1[Z]}, C_2[Z]}) \\
\leq 4(ML)^s \epsilon - sI_{-X} (X; B|W_B \times P_{mix,x})
\]
(112)
holds for \(0 < s \leq 1\).

Here, when \(X\) is given as a vector space \(\mathbb{F}_q^k\) of a finite field \(\mathbb{F}_q\), we construct an ensemble of the code pairs \(\{C_2[Z] \subset C_1[Z]\}\) such that the ensemble of submodules \(\{C_2[Z]\}\) and \(\{C_1[Z]\}\) satisfies Condition \([1]\) as subsets of \(X\). In the following construction, we choose \(C_1[Z]\) and \(C_2[Z]\) as \(l_1\)-dimensional and \(l_2\)-dimensional subspaces. First, we fix \((k - l_2) \times (l_2 - l_1)\) matrix \(D\) with the rank \(l_2 - l_1\). Let \(Z\) be the Toeplitz matrix of the size \((k - l_2) \times l_2\), which contains \(k - 1\) random variables taking values in the finite field \(\mathbb{F}_q\). When the codes \(C_1[Z]\) and \(C_2[Z]\) are given by the images of the random matrix \(\begin{pmatrix} I & 0 \\ Z & D \end{pmatrix}\) and \(\begin{pmatrix} I & Z \\ Z' \end{pmatrix}\), the ensemble \(\{C_2[Z]\}\) satisfies Condition \([1]\).

Furthermore, let \(Z'\) be the Toeplitz matrix of the size \((k - l_1) \times (l_1 - l_2)\), which contains \(k - l_2 - 1\) random variables taking values in the finite field \(\mathbb{F}_q\). When \(C_1[Z, Z']\) and \(C_2[Z]\) are given by the images of the random matrix \(\begin{pmatrix} I & 0 \\ Z & Z' \end{pmatrix}\) and \(\begin{pmatrix} I \\ Z \end{pmatrix}\), the ensemble \(\{C_1[Z, Z]\}\) satisfies Condition \([1]\) as subsets of \(X\) as well as \(\{C_2[Z]\}\).

Proof of Lemma 9 Using (50) and (56), we obtain
\[
E_{Z,E}[Y]_{[C_1[Z]} I_{mix}(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]}) \\
\leq E_{Z,E}[Y]_{[C_1[Z]} I(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]} \times P_{mix,[Y]_{C_1[Z]}, C_2[Z}||W_E \circ P_{mix,x}) \\
= E_{Z,E}[Y]_{[C_1[Z]} \frac{1}{|X|} \sum_{[x] \in [Y]_{C_1[Z]}, C_2[Z]} D(W_E \circ P_{mix,[x]}||W_E \circ P_{mix,x}) \\
= E_{Z,E}[Y] D(W_E \circ P_{mix,C_2[Z]}||W_E \circ P_{mix,x}) \\
\leq u^s \frac{H_{\mu}(e^{sI_l(X; E|W_E \times P_{mix,x})})}{L^s/2},
\]
(113)
which implies (109). Similarly, using (58) and the convexity of \(x \mapsto e^x\), we obtain
\[
E_{Z,E}[Y]_{[C_1[Z]} e^{sI(M; E|W_E \circ \Gamma_{[Y]_{C_1[Z]}, C_2[Z]} \times P_{mix,[Y]_{C_1[Z]}, C_2[Z}}||W_E \circ P_{mix,x}) \\
= E_{Z,E}[Y]_{[C_1[Z]} e^{s\sum_{[x] \in [Y]_{C_1[Z]} } \Phi_D(W \circ P_{mix,[x]}||W \circ P_{mix,x})} \\
\leq E_{Z,E}[Y] e^{sD(W \circ P_{mix,C_2[Z]}||W \circ P_{mix,x})} \\
\leq u^s (1 + \frac{e^{sI_l(X; E|W_E \times P_{mix,x})}}{L^s})
\]
(115)
for $0 < s \leq 1$, where (14) and (15) follow from the convexity of $x \mapsto e^x$ and (58), respectively. Relation (30) guarantees that
\[
e^{sI_{\text{mix}}(M; E|W_E) \gamma_{|C_1[Z]} c_2[Z]} \leq e^{sI(M; E|W_E \gamma_{|C_1[Z]} c_2[Z])} P_{\text{mix}}(Y, c_2[Z]|W_E \gamma_{|C_1[Z]} c_2[Z])|W_E \gamma_{|C_1[Z]} c_2[Z]|P_{\text{mix}}(E),
\]
(116)
Thus, combination of (115) and (116) implies (110).
Next, we show (111). For a given value $Z$, (47) implies that
\[
E[|Y|_{C_1[Z]} |Zd_{1, \text{mix}}(M; E|W_E \gamma_{|Y|_{C_1[Z]} C_2[Z]})] = e_{d_{1}'(|X|; |X|_{C_1[Z]} E|W_E \times P_{\text{mix}})} = 2d_{1}'(|X|; E|W_E \times P_{\text{mix}}) = 2d_{1}'(|X|; E|W_E \times P_{\text{mix}}),
\]
(117)
where the final equation follows from the fact that the random variable $|X|_{C_1[Z]}$ is given as a function of $|X|$. Since the ensemble of submodules $\{C_2[Z]\}$ satisfies Condition 1 as a subset of $X$, the ensemble of the hash functions $X \mapsto |X|_{C_2[Z]}$ satisfies Condition 3. Hence, Lemmas 11 and 12 guarantee
\[
E[Zd_{1}'(|X|; E|W_E \times P_{\text{mix}})] = \mu_s(\|X\|/L) \frac{\mu_s G(X; E|W_E \times P_{\text{mix}})}{L^{1/2}} = \mu_s e^{sI_{\text{mix}}(X; E|W_E \times P_{\text{mix}})}
\]
(118)
where $\sigma_E$ is $(\sum_x P_{\text{mix},X}(x) W_{z_{1+s}}) \frac{1}{1+z_{1+s}} / \text{Tr} (\sum_x P_{\text{mix},X}(x) W_{z_{1+s}})$. Combination of (117) and (118) yields (111).

Proof of Lemma 10
Since the ensemble of submodules $\{C_1[Z]\}$ satisfies Condition 1 as a subset of $X$, the proof given in (13) is valid with $P_X = P_{\text{mix},X}$. Thus, the ensemble expectation of the average error probability concerning decoding the input message is bounded by $4(ML)^s e^{-sI_{\text{mix}}(X; B|W_B \times P_{\text{mix},X})}$ for $0 < s \leq 1$. That is,
\[
E_Z E_{Y \in X/C_1[Z]} [1 + e^{-sI_{\text{mix}}(X; B|W_B \times P_{\text{mix},X})}] \leq 4(ML)^s e^{-sI_{\text{mix}}(X; B|W_B \times P_{\text{mix},X})}.
\]
(119)

2) Additive case: Next, we consider the case of additive channels. Assume that the channel $W_E$ is called additive, i.e., the set $X$ has a structure of module and there exist a state $\rho$ and a projective representation $U$ of $X$ such that $W_E = U_x \rho U_x^\dagger$. In this case, the relations
\[
I_{\text{mix}}(M; E|W_E \gamma_{|C_1[C_2]} |C_2[C_2]) = I_{\text{mix}}(M; E|W_E \gamma_{|Y|_{C_1[Z]} C_2[Z]}),
\]
(120)
hold for $y \in X$. Since $W_E \gamma_{|C_1[C_2]}$ is an additive channel, (52) guarantees that
\[
I_{\text{mix}}(M; E|W_E \gamma_{|C_1[C_2]} |C_2[C_2]) = I_{\text{mix}}(M; E|W_E \gamma_{|C_1[C_2]}).
\]
(121)
Further, any additive channel $W_B$ satisfies that
\[
\epsilon(W_B | \Phi_{|Y|_{C_1[Z]} C_2[Z]} | C_2[Z]) = \epsilon(W_B | \Phi_{C_1[Z]} C_2[Z]).
\]
(122)

Hence, (109), (110), (111), and (112) are simplified to
\[
E[Z e^{sI_{\text{mix}}(M; E|W_E \gamma_{|C_1[Z]} C_2[Z])}] \leq \frac{e^{sI_{\text{mix}}(X; E|W_E \times P_{\text{mix},X})}}{L^{1/2} S},
\]
(123)
\[
E[Z e^{sI_{\text{mix}}(M; E|W_E \gamma_{|C_1[Z]} C_2[Z])}] \leq \frac{e^{sI_{\text{mix}}(X; E|W_E \times P_{\text{mix},X})}}{L^{1/2} S},
\]
(124)
\[
E[Z \epsilon(W_B | \Phi_{C_1[Z]} C_2[Z]) \leq 4(ML)^s e^{-sI_{\text{mix}}(X; B|W_B \times P_{\text{mix},X})}.
\]
(125)
for $0 < s \leq 1$. Further, $I_{\text{mix}}(X; E|W_E \times P_{\text{mix},X})$ and $I_{\text{mix}}^s(X; E|W_E \times P_{\text{mix},X})$ can be calculated as
\[
I_{\text{mix}}(X; E|W_E \times P_{\text{mix},X}) = H(\rho) - H(\rho),
\]
\[
I_{\text{mix}}(X; E|W_E \times P_{\text{mix},X}) = H(\rho) - H(\rho),
\]
(126)
\[
I_{\text{mix}}^s(X; E|W_E \times P_{\text{mix},X}) = H(\rho) - H(\rho),
\]
(127)
for $0 < s \leq 1$. Further, $I_{\text{mix}}(X; E|W_E \times P_{\text{mix},X})$ and $I_{\text{mix}}^s(X; E|W_E \times P_{\text{mix},X})$ can be calculated as
\[
I_{\text{mix}}(X; E|W_E \times P_{\text{mix},X}) = H(\rho) - H(\rho),
\]
(128)
\[
I_{\text{mix}}^s(X; E|W_E \times P_{\text{mix},X}) = H(\rho) - H(\rho),
\]
(129)
where \( \mathbf{p} = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} [u_x \rho u_x^\dagger] \). Especially, when \( \rho \) is pure,

\[
s I_s(X; E|W_E \times P_{\text{mix}, \mathcal{X}}) = \log \text{Tr} \mathbf{p}^{1-s}
\]

(126)

\[
I(X; E|W_E \times P_{\text{mix}, \mathcal{X}}) = H(\mathbf{p})
\]

(127)

\[
\frac{s}{1 + s} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{A}}) = \log \text{Tr} \mathbf{p}^{1/(1+s)}.
\]

(128)

Now, we consider the case when the code \( C_1 \) is fixed and only \( C_2 \) is randomly chosen.

**Lemma 11:** Assume that the ensemble \( \{C_2[Z]\} \) satisfies Condition [1] as a subset of \( C_1 \). When the channels \( W_B \) and \( W_E \) are additive, the relations

\[
E_Z I_{\max}(M; E|W_E \circ \Gamma_{C_1, C_2}[Z]) \leq \eta \left( \frac{H_s}{s^2} e^{-\frac{s}{2}} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{X}}), \log d_E \right),
\]

(129)

\[
E_Z d_{1, \text{mix}}(M; E|W_E \circ \Gamma_{C_1, C_2}[Z]) \leq \frac{H_s}{s^2} e^{-\frac{s}{2}} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{X}})
\]

(130)

hold for \( 0 < s \leq 1 \).

Therefore, even if we fixed our error correcting code to \( C_1 \), we can find a subcode \( C_2 \subset C_1 \) satisfying that

\[
I_{\max}(M; E|W_E \circ \Gamma_{C_1, C_2}) \leq 2 \eta \left( \frac{H_s}{s^2} e^{-\frac{s}{2}} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{X}}), \log d_E \right),
\]

(131)

\[
d_{1, \text{mix}}(M; E|W_E \circ \Gamma_{C_1, C_2}) \leq 2 \frac{H_s}{s^2} e^{-\frac{s}{2}} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{X}})
\]

for \( 0 < s \leq 1 \).

**Proof:** In this case, we have \( d_{1, \text{mix}}(M; E|W_E \circ \Gamma_{C_1, C_2}) = d_1^f([X]; E|W_E \times P_{\text{mix}, C_1}) \). Hence, (118) implies that

\[
E_Z d_{1, \text{mix}}(M; E|W_E \circ \Gamma_{C_1, C_2}) \leq \frac{H_s e^{-\frac{s}{2}} I_s^G([X]; E|W_E \times P_{\text{mix}, C_1})}{L^\gamma}.
\]

(131)

Since \( W_E \) is additive, \( I_s^G([X]; E|W_E \times P_{\text{mix}, C_1}) = I_s^G([X]; E|W_E \times P_{\text{mix}, C_1} + x) \). Hence, Lemma [3] yields

\[
e^{-\frac{s}{1 + s}} I_s^G([X]; E|W_E \times P_{\text{mix}, C_1}) = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} e^{-\frac{s}{1 + s}} I_s^G([X]; E|W_E \times P_{\text{mix}, C_1} + x)
\]

\[
\leq e^{-\frac{s}{1 + s}} I_s^G([X]; E|W_E \times P_{\text{mix}, \mathcal{X}}),
\]

which implies that

\[
I_s^G([X]; E|W_E \times P_{\text{mix}, C_1}) \leq I_s^G([X]; E|W_E \times P_{\text{mix}, \mathcal{X}}).
\]

(132)

The combination of (131) and (132) yields (130). Further, attaching [51] to (130), we obtain

\[
E_Z I_{\text{mix}}(M; E|W_E \circ \Gamma_{C_1, C_2}[Z]) \leq \eta \left( \frac{H_s}{s^2} e^{-\frac{s}{2}} I_s^G(X; E|W_E \times P_{\text{mix}, \mathcal{X}}), \log d_E \right),
\]

(133)

for \( 0 < s \leq 1 \) due to the concavity of the map \( x \mapsto \eta(x, \log d_E) \). Finally, (120) and (121) imply (129).

Indeed, when the fixed submodule \( C_1 \) is isomorphic to a vector space of a finite field \( \mathbb{F}_q \), we can construct an ensemble of submodules \( \{C_2[Z]\} \) of \( C_1 \) satisfying Condition [1] by the same method as that given in Section [11].

**B. The case when uniform distribution is unavailable**

Now, we consider the case when the uniform distribution is not available for encoding the message \( [x] \in C_1/C_2 \). In this case, we assume that a module \( \mathcal{A} \) with the cardinality \( |\mathcal{A}| \) and a distribution \( P_A \) on the module \( \mathcal{A} \) is available for this purpose.

We employ a submodule \( C_1 \) of \( \mathcal{A} \) with the cardinality \( |\mathcal{A}| \) and an injective homomorphism \( f \) from \( \mathcal{A} \) to \( C_1 \). We fix a set of representatives \( \{x_1, \ldots, x_M\} \) of all elements of \( C_1/f(A) \). Then, we can define the affine map \( f_0(a) := f(a) + x \).

Based on the above structure, we construct a code for a quantum wiretap channel \( \Phi_{C_1, f, \{x_1, \ldots, x_M\}} = (M, \{f_0 \circ P_A\}_{[x_m] \in C_1/f(A)}, \{D_{[x]}\}_{[x] \in C_1/f(A)} \) as follows. The encoding distribution corresponding to the message \( [x_m] \in C_1/f(A) \) is given as the distribution \( f_0 \circ P_A \) on the coset \( [x_m] = x_m + f(A) \), and the decoding \( D_{[x]} \) is given as the subset \( \sum_{x \in x + f(A)} D_{x} \). As a generalization, for \( x \in \mathcal{X} \), we consider a code \( \Phi_{C_1, f, \{x_1, \ldots, x_M\}} = (M, \{f_0 \circ P_A\}_{[x_m] \in C_1/f(A)}, \{D_{[x]}\}_{[x] \in C_1/f(A)} \) as follows. The encoding distribution corresponding to the message \( [x + x_m] \in C_1/f(A) \) is given as the distribution \( f_0 \circ P_A \) on the coset \( [x + x_m] = x + x_m + f(A) \), and the decoding \( D_{[x]} \) is given as the subset \( \sum_{x \in x + f(A)} D_{x} \). When the hash function \( f_2 \) is randomly chosen as a homomorphism from \( \mathcal{A} \) to \( \mathcal{X} \) according to a random variable \( X \), the code \( C_1[Z] \) is also chosen satisfying that \( f_2(A) \subset C_1[Z] \) and \( |C_1[Z]|/f_2(A) = M \).

Since the distribution \( f_0 \circ P_A \) on \( \mathcal{X} \) depends on the random variable \( M \) on \( \mathcal{M} \), it can be regarded as the transition matrix \( m \mapsto f_0 \circ P_A \). In order to clarify this point, this transition matrix is denoted by \( P_{X|M}[f_0 \circ P_A] \). Hence, the transition matrix \( (\tilde{x}, m) \mapsto f_0 \circ P_A \) is denoted by \( P_{X,M}[f_0 \circ P_A] \).
Lemma 12: When the random variable $\bar{X} \in \bar{X} = \mathcal{X}$ is subject to the uniform distribution and $\{f_\mathcal{Z}\}$ satisfies Condition 2, we obtain the following.

\[
E_{\bar{X}} I_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A]]) \leq \frac{u^s e^{-sH_{1+}(P_A)} e^{sI_{\mathcal{X}}(X;E[W_E \times P_{\text{mix},\mathcal{X}}])}}{s} \quad (134)
\]

\[
E_{\bar{X}} e^{sI_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A])} \leq e^s (1 + e^{-sH_{1+}(P_A)} e^{sI_{\mathcal{X}}(X;E[W_E \times P_{\text{mix},\mathcal{X}}])}) \quad (135)
\]

Here, we construct an ensemble $\{\mathcal{Z}_i; f_\mathcal{Z}, \{x_1, x_2, \ldots, x_M, \mathcal{Z}\}\}$ such that $\{f_\mathcal{Z}\}$ satisfies Condition 2 when $\mathcal{X}$ is given as a vector space $\mathbb{F}^m_q$ of a finite field $\mathbb{F}_q$. In the following construction, we choose $C_1[\mathcal{Z}]$ as an $l_1$-dimensional subspace and $\mathcal{A}$ as the $l_2$-dimensional space $\mathbb{F}^l_q$. First, we fix $(k-l_2) \times (l_1 - l_2)$ matrix $D$ with the rank $l_2 - l_1$. Let $\mathcal{Z}$ be the Toeplitz matrix of the size $(k - l_2) \times l_2$, which contains $k - 1$ random variables taking values in the finite field $\mathbb{F}_q$, and $\mathcal{Z}'$ be the random matrix taking values in the set of invertible matrices of the size $l_2 \times l_2$ with the uniform distribution. We choose $f_{\mathcal{Z}}, \mathcal{Z}$ to be the multiplication of the random matrix $(\mathcal{Z}', \mathcal{Z})^T$ with two independent random variables $\mathcal{Z}'$ and $\mathcal{Z}$, and $C_1[\mathcal{Z}, \mathcal{Z}']$ to be the image of the random matrix $\begin{pmatrix} \mathcal{Z}' & 0 \\ \mathcal{Z} \\ \mathcal{Z} \end{pmatrix}$. The coset representatives are chosen to be the image of the matrix $\begin{pmatrix} 0 & \mathcal{Z} \end{pmatrix}$. Then, $\{f_{\mathcal{Z}}, \mathcal{Z}\}$ satisfies Condition 2.

Proof of Lemma 12:

\[
E_{\bar{X}} I_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A]]) \leq E_{\bar{X}} I_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A]) \times P_{\text{mix},\mathcal{X}} || W_E \circ P_{\text{mix},\mathcal{X}})
\]

\[
= E_{\bar{X}} \sum_{m=1}^{M} D(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}})
\]

\[
= E_{\bar{X}} D(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}}) \quad (137)
\]

Similarly, we obtain

\[
E_{\bar{X}} e^{sI_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A])} \leq E_{\bar{X}} e^{sI_{\text{mix}}^\mathcal{Z}(M; E[W_E \circ P_{\mathcal{X}|M}[f_{\mathcal{Z}} \circ P_A]) \times P_{\text{mix},\mathcal{X}} || W_E \circ P_{\text{mix},\mathcal{X}})
\]

\[
= E_{\bar{X}} e^{sD(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}})}
\]

\[
= E_{\bar{X}} e^{sD(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}})} \quad (138)
\]

\[
= E_{\bar{X}} e^{sD(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}})}
\]

\[
= E_{\bar{X}} e^{sD(W_E \circ f_{\mathcal{Z}} \circ P_A || W_E \circ P_{\text{mix},\mathcal{X}})} \quad (139)
\]

where (138) follows from the convexity of $x \mapsto e^x$. These upper bounds do not depend on the choice of the set of representatives $\{x_1, \ldots, x_M\}$.

Then, (59) and (61) imply that

\[
E_{\bar{X}} D(W_E \circ f_{\mathcal{Z}} \circ P_A || W \circ P_{\text{mix},\mathcal{X}}) \leq \frac{u^s e^{-sH_{1+}(P_A)} e^{sI_{\mathcal{X}}(X;E[W_E \times P_{\text{mix},\mathcal{X}}])}}{s} \quad (140)
\]

\[
E_{\bar{X}} e^{sD(W_E \circ f_{\mathcal{Z}} \circ P_A || W \circ P_{\text{mix},\mathcal{X}})} \leq e^s (1 + e^{-sH_{1+}(P_A)} e^{sI_{\mathcal{X}}(X;E[W_E \times P_{\text{mix},\mathcal{X}}])}) \quad (141)
\]

Hence, combination of (137) and (140) yields (134), and combination of (139) and (141) yields (135).
Relation (47) yields that
\[
E_\tilde{x} d_1 \text{mix}(M; E|W_E \circ P_X|M[fz_\tilde{x} + x_M \circ P_A])
\]
\[
= d'_1(M; E, \tilde{x}|W_E \circ P_X|M[fz_\tilde{x} + x_M \circ P_A] \times P_{\text{mix},X|\tilde{x}})
\]
\[
\leq 2d'_1(M, \tilde{x}; E|W_E \circ P_X|M[fz_\tilde{x} + x_M \circ P_A] \times P_{\text{mix},X|\tilde{x}})
\]
\[
= 2 \sum_{\tilde{x} \in \tilde{x}, m \in M} \frac{1}{|X| \cdot |M|} ||W_E \circ fz_{\tilde{x} + x_M} \circ P_A - W_E \circ P_{\text{mix},\tilde{x}}||_1
\]
\[
= 2 \sum_{\tilde{x} \in \tilde{x}, m \in M} \frac{1}{|X|} ||W_E \circ fz_{\tilde{x}} \circ P_A - W_E \circ P_{\text{mix},\tilde{x}}||_1
\]
\[
= 2d'_1(\tilde{x}; E|W_E \circ P_X|M[fz_{\tilde{x}} \circ P_A] \times P_{\text{mix},\tilde{x}}).
\]
(142)

Since the ensemble of the hash functions \((x, a) \mapsto fz(a) + x \in X\) satisfies Condition [3] Lemma [14] yields that
\[
E_fd'_1(\tilde{x}; E|W_E \circ P_X|M[fz_{\tilde{x}} \circ P_A] \times P_{\text{mix},\tilde{x}})
\]
\[
\leq \mu_\tilde{x}|X| \mathbb{E}_{H_1 + s(X,E)}[|W_E \times P_{\text{mix},X,K}\circ P_A|] \leq \mu_\tilde{x}|X| \mathbb{E}_{H_1 + s(X,E)}[|W_E \times P_{\text{mix},X,K}\circ P_A|] \leq \mu_\tilde{x} e^{s(I(X,E|W_E \times P_{\text{mix},X,K}\circ P_A) - H_{1 + s}(P_A))}
\]
\[
= \mu_\tilde{x} e^{s(I(X,E|W_E \times P_{\text{mix},X,K}\circ P_A) - H_{1 + s}(P_A))},
\]
(143)

where \(\sigma_E = (\sum_x P_{\text{mix},X}(x)W_{E|x}^{1+s})^{1/s} / \text{Tr}(\sum_x P_{\text{mix},X}(x)W_{E|x}^{1+s})^{1/s}\). Hence, combination of (122) and (143) implies (136).

When the channel \(W_E\) is additive, the relations
\[
I_{\text{mix}}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A]) = I_{\text{mix}}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A])
\]
\[
d_1 \text{mix}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A]) = d_1 \text{mix}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A])
\]
hold for any \(x \in X\). Hence, Lemma [12] can be simplified to
\[
E_fd'_1(\tilde{x}; E|W_E \circ P_X|M[fz_{\tilde{x}} \circ P_A] \times P_{\text{mix},\tilde{x}})
\]
\[
\leq \mathbb{E}_{H_1 + s(X,E)}[|W_E \times P_{\text{mix},X,K}\circ P_A|] \leq \mu_\tilde{x} e^{s(I(X,E|W_E \times P_{\text{mix},X,K}\circ P_A) - H_{1 + s}(P_A))},
\]
(144)

Now, similar to Lemma [11], we consider the case when the code \(C_1\) is fixed and only \(C_2\) is randomly chosen.

Lemma 13: Assume that \(\{fz\}\) is an ensemble of functions from \(A\) to \(C_1\) and satisfies Condition [1] When the channels \(W_B\) and \(W_E\) are additive, the relations
\[
E_f I_{\text{mix}}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A]) \leq \eta(2\mu_e e^{s(I(X,E|W_E \times P_{\text{mix},X,K}\circ P_A) - H_{1 + s}(P_A))}, \log d_E)
\]
(147)
\[
E_f d_1 \text{mix}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A]) \leq 2\mu_e e^{s(I(X,E|W_E \times P_{\text{mix},X,K}\circ P_A) - H_{1 + s}(P_A))}
\]
(148)
hold for \(0 < s \leq 1\).

Proof: Relation (148) can be shown as follows. Relation (124) with \(X = C_1\) implies that
\[
E_f d_1 \text{mix}(M; E|W_E \circ P_X|M[fz_{x_M} \circ P_A]) \leq 2\mu_e e^{s(I(X,E|W_E \times P_{\text{mix},C_1} - H_{1 + s}(P_A))}
\]
Hence, using (132), we obtain (148). Relation (147) can be obtained in the same way as (129) in Lemma [11].

Therefore, even though the uniform distribution is not available, if a distribution close to the uniform distribution is available, there exists a code with a performance similar to Lemma [11].

VII. APPLICATION TO PAULI CHANNEL

As a simple example, we treat a Pauli channel in the \(d\)-dimensional system \(H\). First, we define the discrete Weyl-Heisenberg representation \(W\) for \(\mathbb{Z}_d^2\):
\[
X := \sum_{j=1}^d |j + 1\rangle \langle j|, \quad Z := \sum_{j=1}^d \omega^j |j\rangle \langle j|
\]
(149)
\[
W(x, z) := X^x Z^z,
\]
(150)
Then, in the following, we assume that the eavesdropper can access all of the environment of the channel $\Lambda_{P_{X,Z}}$. When the state $|j\rangle$ is input to the channel $\Lambda_{P_{X,Z}}$, the environment system is spanned by the basis $\{|x,z\}\rangle$ and the state $W_{E|j}$ of the environment is given as

$$W_{E|j} = \sum_{z} P_z(z) |j, z : P_{X,Z} \rangle \langle j, z : P_{X,Z}|,$$

(152)

where

$$|j, z : P_{X,Z}\rangle := \sum_{x=1}^{d} \omega^{jx} \sqrt{P_{X|Z}(x|z)} |x, z\rangle.$$

(153)

Then, the average state is

$$\overline{\rho}_E = \sum_{x,z} P_{X,Z}(x,z) |x,z\rangle \langle x,z| = \sum_{z} P_z(z) \overline{\rho}_{E|z},$$

(154)

where

$$\overline{\rho}_{E|z} := \sum_{x} P_{X|Z}(x|z) |x,z\rangle \langle x,z|.$$

(155)

Then, $I_s(X;E|W_E \times P_{mix,\chi})$ and $I(X;E|W_E \times P_{mix,\chi})$ are calculated by using the conditional entropy $H(X|Z|P_{X,Z})$ and the conditional Rényi entropy $H_{1-s}(X|Z|P_{X,Z})$ as

$$I_s(X;E|W_E \times P_{mix,\chi}) = \frac{1}{s} \sum_j D(W_{E|j} || \overline{\rho}_E) = \frac{1}{s} \sum_j \sum_z P_z(z) D(|j, z : P_{X,Z}\rangle \langle j, z : P_{X,Z}| || \overline{\rho}_{E|z}) = \frac{1}{s} \sum_j \sum_z P_z(z) H(X|P_{X,Z} = z) = H(X|Z|P_{X,Z}),$$

$$I(X;E|W_E \times P_{mix,\chi}) = \frac{1}{1-s} \sum_j \sum_z P_z(z) H(X|P_{X,Z} = z) = H(X|Z|P_{X,Z}),$$

$$e^{sl_s(X;E|W_E \times P_{mix,\chi})} = \frac{1}{d} \sum_j \sum_z P_z(z) H(X|P_{X,Z} = z) = H(X|Z|P_{X,Z}),$$

$$e^{sl_s(X;E|W_E \times P_{mix,\chi})} = \frac{1}{d} \sum_j \sum_z P_z(z) H(X|P_{X,Z} = z) = H(X|Z|P_{X,Z}),$$

which implies

$$I_s(X;E|W_E \times P_{mix,\chi}) = H_{1-s}(X|Z|P_{X,Z}).$$

Similarly, we obtain the simplification:

$$e^{sl_s(X;E|W_E \times P_{mix,\chi})} = e^{sl_s(X;E|W_E \times P_{mix,\chi})} = H_{1-s}(X|Z|P_{X,Z}),$$

which implies

$$I_s(X;E|W_E \times P_{mix,\chi}) = H_{1-s}(X|Z|P_{X,Z}).$$

When we employ the same ensemble of codes as in Subsection VI-A, Inequalities (122), (123), and (124) imply that

$$\mathbb{E}_z I_{max}(M;E|W_E \circ \Gamma_{C_1[Z],C_2[Z]} \leq \frac{e^{sl_s(X|Z|P_{X,Z})}}{L_s^s},$$

(156)

$$\mathbb{E}_z e^{sl_s(M;E|W_E \circ \Gamma_{C_1[Z],C_2[Z]} \leq e^{sl_s(X|Z|P_{X,Z})}}(1 + \frac{e^{sl_s(X|Z|P_{X,Z})}}{L_s^s}),$$

(157)

$$\mathbb{E}_z d_{1,max}(M;E|W_E \circ \Gamma_{C_1[Z],C_2[Z]} \leq 2\mu_s \frac{e^{sl_s(X|Z|P_{X,Z})}}{L_s^{s/2}}.$$
As is mentioned in Subsection II-B, the function $h$ where $e$ state is using the relation between the eavesdropper's Holevo information and the virtual phase error probability [16], we obtain

$$e_R(R|W_E, P_{mix,x}) = \max_{0 \leq s \leq 1} sR - sH_{1-s}(X|P_{X,Z})$$

(159)

$$e_G(R|W_E, P_{mix,x}) = \max_{0 \leq s \leq 1} \left(\frac{s}{2}R - \frac{s}{2}H_{1/(1+s)}(X|P_{X,Z})\right).$$

(160)

When the random variables $X$ and $Z$ are independent under the distribution $P_{X,Z}$, i.e., $P_{X,Z}(x, z) = P_X(x)P_Z(z)$, the state of the environment is written in the form

$$W_{E|j} = |j : P_X \rangle \langle j : P_X| \otimes \sum_z P_Z(z)|z\rangle\langle z|$$

(161)

$$|j : P_X\rangle := \sum_{x=1}^{d} \omega^{jx} \sqrt{P_X(x)|x\rangle}.$$  

(162)

Since $\sum_z P_X(x)|x\rangle\langle x|$ does not depend on $j$, the state $\rho_{E|j}$ can be essentially regarded as $|j : P_X\rangle \langle j : P_X|$. Then, the average state is

$$\mathcal{D} = \sum_x P_X(x)|x\rangle\langle x|.$$  

(163)

In this case, we obtain the lower bound of the exponential decreasing rate of the eavesdropper’s Holevo information:

$$e_R(R|W_E, P_{mix,x}) = \max_{0 \leq s \leq 1} sR - sH_{1-s}(P_X),$$

(164)

and

$$e_G(R|W_E, P_{mix,x}) = \max_{0 \leq s \leq 1} \left(\frac{s}{2}R - \frac{s}{2}H_{1/(1+s)}(P_X)\right).$$

(165)

As is shown in Lemma [7], $e_R(R|W_E, P_{mix,x})$ is better than $e_G(R|W_E, P_{mix,x})$. A similar analysis has been done by the stabilizer formalism by Tsurumaru et al [12]. Their approach evaluates the virtual phase error probability, which is less than

$$\epsilon(n, R) := \min_{0 \leq s \leq 1} e^{n(-sR+sH_{1/(1+s)}(P_X))}.$$  

(166)

Using the relation between the eavesdropper’s Holevo information and the virtual phase error probability [16], we obtain

$$E_Z I_{mix}(M; E|W_E \circ \Gamma_{C_1[z], C_2[z]}) \leq h(\epsilon(n, R)) + n\epsilon(n, R) \log d,$$

(167)

where $h(x) := -x \log x - (1-x) \log(1-x)$. Using their approach, we obtain the lower bound of the exponential decreasing rate of the eavesdropper’s Holevo information:

$$\max_{0 \leq s \leq 1} sR - sH_{1/(1+s)}(P_X) = 2e_G(R|W_E, P_{mix,x}).$$

(168)

As is mentioned in Subsection II-B, the function $s \mapsto H_{1+s}(P_X)$ is monotone decreasing, i.e., $sH_{1-s}(P_X) \geq sH_{1/(1+s)}(P_X)$. Hence, the exponent $2e_G(R|W_E, P_{mix,x})$ by Tsurumaru et al [12] is greater than our exponents $e_R(R|W_E, P_{mix,x})$ and $e_G(R|W_E, P_{mix,x})$. Its numerical verification is given in Fig [1].

Next, we focus on $L_1$ distinguishability. As is mentioned in Subsection II-B, we have two lower bounds of exponents $\frac{1}{2}e_R(R|W_E, P_{mix,x})$ and $e_G(R|W_E, P_{mix,x})$ under this criterion, and the inequality $\frac{1}{2}e_R(R|W_E, P_{mix,x}) \leq e_G(R|W_E, P_{mix,x})$ holds. Using the relation between the universal composability and the virtual phase error probability [17], we obtain

$$E_Z d_{1, mix}(M; E|W_E \circ \Gamma_{C_1[z], C_2[z]}) \leq 2\sqrt{2}\sqrt{\epsilon(n, R)}.$$  

(169)

Using their approach, we obtain the lower bound of the exponential decreasing rate of $L_1$ distinguishability:

$$\frac{1}{2} \max_{0 \leq s \leq 1} sR - sH_{1/(1+s)}(P_X) = e_G(R|W_E, P_{mix,x}).$$

(170)
That is, the exponent by Tsurumaru et al.\cite{12} is the same as our better exponent $e_G(R|W; P_{\text{mix}}, X)$. Hence, we can conclude that our method is not better than the stabilizer formalism when the phase error occurs independently of the bit error in the Pauli channel.

However, when the phase error depends on the bit error in the Pauli channel, the stabilizer formalism cannot provide the error exponent so clearly. In this case, in the evaluation of phase error probability, we have to account into error probability in the error correction concerning the bit error. So, it is not easy to derive a simple bound for the virtual phase error probability as (168). Our bounds (156) and (158) derive lower bounds (159) and (160) of the exponential decreasing rates of the eavesdropper’s information and $L_1$ distinguishability. Therefore, our method has an advantage over the stabilizer formalism when the phase error depends on the bit error. Further, by using (157), our method provides the equivocation rate while the stabilizer formalism cannot derive the equivocation rate even in the independent case.

VIII. Conclusion

We have given a protocol for quantum wiretap channel with an auxiliary random variable subject to a non-uniform distribution. Then, when the distribution of the auxiliary random variable is not the uniform distribution but is close to the uniform distribution, we have derived an upper bound for exponential decreasing rate of leaked information in the quantum mutual information criterion and $L_1$ distinguishability. Further, we have derived the equivocation rate for a given quantum wiretap channel. For a practical construction, we have proposed a code for quantum wiretap channel that requires only a pair of a linear code and an auxiliary random variable subject to a non-uniform distribution. We have also derived an upper bound for the leaked information for this protocol. The organization of the case with a linear code is different from the case with a non-linear code with respect to the construction of the code ensemble and the evaluation of average performance. We can apply the same discussion from the evaluation of average performance to the existence of codes and the asymptotic analysis. Hence, we have omitted the latter part in Section VII. The average performances of the code ensemble based on linear codes have been evaluated only with the single-shot form. These results are summarized in Table I. Finally, we have treated Pauli channel as a typical example.

Further, in the evaluations (67) and (81), the factor $\lambda_s$ increases only linearly under a general setting. Hence, a part of our results derived from them are expected to be applied to the non-stationary and non-memoryless case, e.g., the Markovian case. Unfortunately, we could not prove Lemma 2 for the non-commutative case. The extension of this lemma to the non-commutative case is a future study.

| TABLE I |

| Summary of obtained bounds |

| Error correcting code | general random code | randomized linear code | fixed linear code |
|------------------------|---------------------|------------------------|------------------|
| Auxiliary random number | non-uniform | uniform | additive | non-uniform | additive | uniform |
| Channel | $I_{\text{mix}}(W_E \circ \Gamma)$ | $d_{1, \text{mix}}(W_E \circ \Gamma)$ | $e^{sI_{\text{mix}}(W_E \circ \Gamma)}$ | $I_{\text{mix}}(W_E \circ \Gamma)$ | $d_{1, \text{mix}}(W_E \circ \Gamma)$ | $e^{sI_{\text{mix}}(W_E \circ \Gamma)}$ |
| Bound of average | (65) | (109) | (129) | (134) | (144) | (129) | (144) |
| Key inequality | (66) | (67) | (111) | (124) | (130) | (140) | (130) | (140) |
| | (66) | (67) | (111) | (124) | (130) | (140) | (130) | (140) |

Key inequality expresses the key inequality for deriving the evaluation. – means that there is no specialized discussion.
ACKNOWLEDGMENTS

The author is grateful to the reviewer for giving helpful comments for the first version. He is partially supported by a MEXT Grant-in-Aid for Young Scientists (A) No. 20686026 and Grant-in-Aid for Scientific Research (A) No. 23246071. He is partially supported by 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.

REFERENCES

[1] M. Hayashi, Quantum Information: An Introduction, Springer (2006).
[2] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” IEEE Trans. Inform. Theory, vol. 24, No. 3, 339–348, 1978.
[3] L. Carter and M. Wegman, “Universal classes of hash functions,” J. Comput. Sys. Sci., vol. 18, No. 2, 143–154, 1979.
[4] H. Krawczyk, LFSR-based hashing and authentication. Advances in Cryptology 94. Lecture Notes in Computer Science, vol. 839, Springer-Verlag, pp 129–139, 1994.
[5] M. Hayashi, “General non-asymptotic and asymptotic formulas in channel resolvability and identification capacity and its application to wire-tap channel,” IEEE Trans. Inform. Theory, vol. 52, No. 4, 1562–1575, 2006.
[6] A. D. Wyner, “The wire-tap channel,” Bell. Sys. Tech. Jou., vol. 54, 1355–1387, 1975.
[7] I. Devetak, “The private classical information capacity and quantum information capacity of a quantum channel,” IEEE Trans. Inform. Theory, vol. 51, No. 1, 44–55, 2005.
[8] A. Winter, A. C. A. Nascimento, and H. Imai, “Commitment Capacity of Discrete Memoryless Channels,” IEEE Trans. Inform. Theory, vol. 18, No. 7, 4700-4717, (2013).
[9] R. G. Gallager, Information Theory and Reliable Communication, John Wiley & Sons, 1968.
[10] R. Renner, “Security of Quantum Key Distribution,” PhD thesis, Dipl. Phys. ETH, Switzerland, 2005. arXiv:quantph/0512258.
[11] I. Csiszár, “Almost Independence and Secrecy Capacity,” Problems of Information Transmission, vol. 32, no.1, pp.40-47, 1996.
[12] T. Tsurumaru, M. Hayashi, “Dual universality of hash functions and its applications to quantum cryptography,” IEEE Trans. Inform. Theory, vol. 59, No. 7, 4700-4717, (2013).
[13] M. Hayashi, “Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding,” Physical Review A, vol. 76, 062301 (2007).
[14] I. Devetak, P. W. Shor, “The capacity of a quantum channel for simultaneous transmission of classical and quantum information,” Comm. Math. Phys., 256, 287 (2005).
[15] M. Hayashi, “Practical Evaluation of Security for Quantum Key Distribution,” Physical Review A, vol. 74, 022307 (2006).
[16] M. Hayashi, “Upper bounds of eavesdropper’s performances in finite-length code with the decoy method,” Physical Review A, vol. 76, 012329 (2007); Physical Review A, vol. 79, 019901(E) (2009).
[17] M. Hayashi, “Large deviation analysis for quantum security via smoothing of Rényi entropy of order 2,” arXiv:1202.0422 (2012).
[18] M. Hayashi, “Tight exponential analysis of universally composable privacy amplification and its applications,” arXiv:1010.1358 (2010). Accepted in IEEE Trans. Inform. Theory.
[19] M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” IEEE Trans. Inform. Theory, vol. 57, No. 6, 3989–4001, (2011).
[20] T. S. Han and S. Verdu, “Approximation theory of output statistics,” IEEE Trans. Inform. Theory, vol. 39, no. 3, pp. 752–772, (1993).
[21] M. Hayashi, R. Matsumoto, “Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages,” arXiv:1202.1332 (2012).
[22] M. Hayashi, “Precise evaluation of leaked information with universal2 privacy amplification in the presence of quantum attacker,” Proceedings of the IEEE International Symposium on Information Theory (ISIT 2012), Cambridge, MA, July 2012. arXiv:1202.0601.
[23] M. Hayashi, “Information Spectrum Approach to Second-Order Coding Rate in Channel Coding,” IEEE Trans. Inform. Theory, vol. 55, No.11, 4947 - 4966 (2009).
[24] Y. Polyanskiy, H.V. Poor, and S. Verdu, “Channel coding rate in the finite blocklength regime,” IEEE Trans. Inform. Theory, vol. 56, 2307 - 2359 (2010).
[25] S. Watanabe “Private and Quantum Capacities of More Capable and Less Noisy Quantum Channels,” Phys. Rev. A 85, 012326 (2012)
[26] J. Körner and A. Sgarro, “Universally attainable error exponents for broadcast channels with degraded message sets,” IEEE Trans. Inform. Theory, vol. 26, no. 6, pp. 670–679, Nov. 1980.
[27] T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing,” IEEE Trans. Inform. Theory, 46, 2428–2433 (2000).

APPENDIX A

PROOF OF LEMMA [1]

When $s \geq 0$, in order to show (34), it is enough to show that

$$\min_{E} \sum_{x} P_{X}(x) \operatorname{Tr} W_{x}^{1+s} \sigma_{E}^{-s} = (\operatorname{Tr} \sum_{x} P_{X}(x) W_{x}^{1+s}) \frac{1}{1+s}.$$ 

The reverse operator Holder inequality

$$\operatorname{Tr} X Y \geq (\operatorname{Tr} X^{1/(1+s)})^{1+s} (\operatorname{Tr} Y^{-1/s})^{-s}$$

holds for two non-negative matrices $X$ and $Y$. Then,

$$\operatorname{Tr} \left( \sum_{x} P_{X}(x) W_{x}^{1+s} \right) \sigma_{E}^{-s} \geq \left( \operatorname{Tr} \left( \sum_{x} P_{X}(x) W_{x}^{1+s} \right) \frac{1}{1+s} \right)^{1+s} \left( \operatorname{Tr} \sigma_{E}^{-s-1/s} \right)^{-s} = \left( \operatorname{Tr} \left( \sum_{x} P_{X}(x) W_{x}^{1+s} \right) \frac{1}{1+s} \right)^{1+s}.$$ 

The equality holds when $\sigma_{E} = \left( \sum_{x} P_{X}(x) W_{x}^{1+s} \right)^{1+s} \frac{1}{1+s} / \left( \sum_{x} P_{X}(x) W_{x}^{1+s} \right)^{1+s}$. 

When $s < 0$, in order to show (34), it is enough to show that

$$\max_{E} \sum_{x} P_{X}(x) \operatorname{Tr} W_{x}^{1+s} \sigma_{E}^{-s} = (\operatorname{Tr} \sum_{x} P_{X}(x) W_{x}^{1+s}) \frac{1}{1+s}.$$
The operator Holder inequality

\[ \text{Tr } XY \geq (\text{Tr } X^{1/(1+s)})^{1+s} (\text{Tr } Y^{-1/s})^{-s} \]

holds for two non-negative matrixes \( X \) and \( Y \). Similarly, we have

\[ \text{Tr} \left( \sum_x P_X(x) W_x^{1+s} \right) \sigma_E^{-s} \leq \left( \text{Tr} \left( \sum_x P_X(x) W_x^{1+s} \right) \right)^{1/s}. \]

The equality holds when \( \sigma_E = (\sum_x P_X(x) W_x^{1+s})^{1/s} / \text{Tr} (\sum_x P_X(x) W_x^{1+s})^{1/s}. \)

**APPENDIX B**

**PROOF OF (47)**

Define \( W_y := \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W_{x,y} \). Then,

\[ \| W_y - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \|_1 \leq \left\| \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W_{x,y} - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \right\|_1 \]

\[ \leq \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} \| W_{x,y} - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \|_1. \]

Thus,

\[ d'_1(\mathcal{X}, \mathcal{Y}, E|\rho_{\mathcal{X}, \mathcal{Y}, E}) = \sum_x P_{\text{mix}, \mathcal{X} \times \mathcal{Y}}(x) \sum_y P_{\text{mix}, \mathcal{Y}}(y) \| W_{x,y} - W_y \|_1 \]

\[ \leq \sum_x P_{\text{mix}, \mathcal{X} \times \mathcal{Y}}(x) \sum_y P_{\text{mix}, \mathcal{Y}}(y) \| W_{x,y} - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \|_1 \]

\[ = \sum_x P_{\text{mix}, \mathcal{X} \times \mathcal{Y}}(x) \sum_y P_{\text{mix}, \mathcal{Y}}(y) \| W_{x,y} - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \|_1 \]

\[ \leq 2 \sum_x P_{\text{mix}, \mathcal{X} \times \mathcal{Y}}(x) \sum_y P_{\text{mix}, \mathcal{Y}}(y) \| W_{x,y} - W \circ P_{\text{mix}, \mathcal{X} \times \mathcal{Y}} \|_1 \]

\[ = 2d'_1(\mathcal{X}, \mathcal{Y}, E|\rho_{\mathcal{X}, \mathcal{Y}, E}). \]

**APPENDIX C**

**PROOF OF (52)**

For an element \( x' \in \mathcal{X} \), we denote the addition map \( x \mapsto x + x' \) by \( \text{Add}_{x'} \). For a distribution \( P_X \) on \( \mathcal{X} \), we consider the distribution \( \text{Add}_{x'} \circ P_X \) on \( \mathcal{X} \). The symmetry of the channel \( W \), we have

\[ I(\mathcal{X}; E|W \times P_X) = I(\mathcal{X}; E|W \times (\text{Add}_{x'} \circ P_X)) \]

for any \( x' \in \mathcal{X} \). Using (30), we also have

\[ \sum_{x' \in \mathcal{X}} \frac{1}{|\mathcal{X}|} I(\mathcal{X}; E|W \times (\text{Add}_{x'} \circ P_X)) \leq \sum_{x' \in \mathcal{X}} \frac{1}{|\mathcal{X}|} I(\mathcal{X}; E|W \times (\text{Add}_{x'} \circ P_X)||W \circ P_{\text{mix}, \mathcal{X}}) \]

\[ = \sum_{x' \in \mathcal{X}} \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} P_X(x - x') D(W_x || W \circ P_{\text{mix}, \mathcal{X}}) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} D(W_x || W \circ P_{\text{mix}, \mathcal{X}}) = I_{\text{mix}}(W). \]

Combining (173) and (174), we obtain \( I(\mathcal{X}; E|W \times P_X) \leq I_{\text{mix}}(\mathcal{X}; E|W) \), which implies (53).
APPENDIX D
SECRET KEY GENERATION WITH UNIVERSAL COMPOSABILITY

We assume that Alice and Bob share a common classical random number \(a \in {\mathcal A}\), and Eve has a quantum state \(\rho_a \in {\mathcal H}_E\), which is correlated to the random number \(a\). The task is to extract a common random number \(f(a)\) from the random number \(a \in {\mathcal A}\), 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 \(a \in {\mathcal A}\). Now, we focus on an ensemble of the functions \(f_Z\) from \(A\) to \(\{1, \ldots, M\}\), where \(Z\) denotes a random variable describing the stochastic behavior of the function \(f\). An ensemble of the functions \(f_Z\) is called universal\(_2\) when it satisfies the following condition\(3\):

**Condition 3**: \(\forall a_1 \neq \forall a_2 \in {\mathcal A}\), the probability that \(f_Z(a_1) = f_Z(a_2)\) is at most \(\frac{1}{M}\).

For example, when \(M\) is an arbitrary integer and the cardinality \(|{\mathcal A}|\) is an arbitrary multiple of \(M\), an ensemble \(\{f_Z\}\) satisfying the above condition is given in the following way. First, we fix a function \(f\) from \(A\) to \(\{1, \ldots, M\}\) such that the cardinality \(|f^{-1}\{i\}|\) is \(\frac{1}{M}\). We randomly choose a permutation \(g \in S_A\) on \(A\) with the uniform distribution, where \(S_A\) denotes the set of permutation on \(A\). So, we can make a random function \(\{f \circ g\}_{g \in S_A}\). This ensemble satisfies Condition 3.

**Lemma 14 (\[18\] Lemma 33)**: When an ensemble of the functions \(f_Z\) from \(A\) to \(\{1, \ldots, M\}\) satisfies Condition 3 any density matrix \(\sigma_E\) on the system \({\mathcal H}_E\) satisfies

\[
E_Z d_1'(f_Z(A); E|\rho_{E,A}) \\
\leq (4 + \sqrt{v})M^{s/2}e^{-\frac{1}{2}H_{1+s}(A|E|\rho_{E,A},\sigma_E)} (175)
\]

\[
E_Z d_1'(f_Z(A); E|\rho_{E,A}) \\
\leq (4 + \sqrt{|A|})M^{s/2}e^{-\frac{1}{2}H_{1+s}(A|E|\rho_{E,A},\sigma_E)} + \frac{1}{8}, (176)
\]

where \(v\) is the number of eigenvalues of \(\sigma_E\) and \(\lambda\) is defined as the real number \(\log a_1 - \log a_0\) by using the maximum eigenvalue \(a_1\) and the minimum eigenvalue \(a_0\) of \(\sigma\).

APPENDIX E
PROOF OF LEMMAS 4 AND 6

First, we show Lemma 4.

\[
E_{\Phi} D(W \circ \Phi \circ P_A || W \circ P_X) = E_{\Phi} \text{Tr} \left( \sum_a P_A(a)W_{\Phi(a)} \right) (\log(\sum_a P_A(a)W_{\Phi(a)}) - \log(W \circ P_X))
\]

\[
= E_{\Phi} \text{Tr} \left( \sum_a P_A(a)W_{\Phi(a)} \right) (\log(\sum_a P_A(a)W_{\Phi(a)}) - \log(W \circ P_X))
\]

\[
= \text{Tr} \sum_a E_{\Phi(a)} P_A(a)W_{\Phi(a)} (\log(P_A(a)W_{\Phi(a)}) + \sum_{a' \neq a} P_A(a')W_{\Phi(a')} - \log(W \circ P_X))
\]

\[
\leq \text{Tr} \sum_a E_{\Phi(a)} P_A(a)W_{\Phi(a)} (\log(P_A(a)W_{\Phi(a)}) + E_{\Phi(a)} \sum_{a' \neq a} P_A(a')W_{\Phi(a')}) - \log(W \circ P_X))
\]

\[
= \text{Tr} \sum_a E_{\Phi(a)} P_A(a)W_{\Phi(a)} (\log(P_A(a)W_{\Phi(a)}) + \sum_{a' \neq a} P_A(a')W \circ P_X) - \log(W \circ P_X))
\]

\[
\leq \text{Tr} \sum_a E_{\Phi(a)} P_A(a)W_{\Phi(a)} (\log(P_A(a)W_{\Phi(a)}) + W \circ P_X) - \log(W \circ P_X))
\]

\[
\leq \frac{1}{s} \text{Tr} \sum_a E_{\Phi(a)} P_A(a)W_{\Phi(a)} (\log(\sum_a P_A(a)) \log(I + v P_A(a)E_{\Phi(a)}(W_{\Phi(a)})(W \circ P_X)^{-1}))
\]

\[
\leq \frac{v}{s} \sum_a E_{\Phi(a)} P_A(a) W_{\Phi(a)} (E_{\Phi(a)}(W_{\Phi(a)}))^{-1} (W \circ P_X)^{-s}
\]

\[
= \frac{v}{s} \sum_a E_{\Phi(a)} P_A(a) (1+s) \text{Tr} (E_{\Phi(a)}(W_{\Phi(a)}))^{-1} (W \circ P_X)^{-s}
\]

\[
= \frac{v}{s} \sum_a P_A(a) (1+s) \text{Tr} P_X(x) (E_{\Phi(a)}(W_{\Phi(a)}))^{-1} (W \circ P_X)^{-s} = \frac{v}{s} e^{-sH_{1+s}(X;E|E_{\Phi(a)}(W_{\Phi(a)})(W \circ P_X))}
\]

\[
\leq \frac{v}{s} e^{-sH_{1+s}(X;E|E\times P_X)},
\]

which implies (53). In the above derivation, (177) follows from the concavity of \(x \rightarrow \log x\), (178) follows from \(\sum_{a' \neq a} P_A(a') \leq 1\), (179) follows from (13), (180) follows from the following inequality, and (181) follows from (43). The inequality \((x+y)^s \leq \frac{v}{s} e^{-sH_{1+s}(X;E|E\times P_X)}\)
Next, we show (54). Since $s \mapsto D^s_s(W \circ \Phi \circ P_A||W \circ P_X)$ is monotone increasing, we obtain $s \mapsto D^s_s(W \circ \Phi \circ P_A||W \circ P_X) \leq s \mapsto D^s_s(A; E|W \circ \Phi \circ P_A||W \circ P_X)$. Thus,

$$E_{\Phi}\exp\left[D^s_s(W \circ \Phi \circ P_A||W \circ P_X)\right] = E_{\Phi}\exp\left(\sum_a P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}\sum_a P_A(a)W_{\Phi(a)})(W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(\sum_a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + \sum_{a' \neq a} P_A(a')W_{\Phi(a')})(W \circ P_X)^{-1/2}s\right)$$

$$\leq \exp\left(\sum_a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + E_{\Phi(a)}\sum_{a' \neq a} P_A(a')W_{\Phi(a')})(W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(\sum a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + \sum_{a' \neq a} P_A(a')W \circ P_X)(W \circ P_X)^{-1/2}s\right)$$

$$\leq \exp\left(\sum a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + W \circ P_X)(W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(\sum a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2} + 1)s\right)$$

$$\leq \exp\left(\sum a E_{\Phi(a)}P_A(a)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2} + 1)s\right)$$

$$= \exp\left(1 + \sum_a E_{\Phi(a)}P_A(a)^{1+s}W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(1 + \sum_a P_A(a)(1+s)W_{\Phi(a)}((W \circ P_X)^{-1/2}(P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(1 + \sum_a P_A(a)(1+s)W_{\Phi(a)}(W \circ P_X)^{-1/2}P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2}s\right)$$

$$= \exp\left(1 + \sum_a P_A(a)(1+s)W_{\Phi(a)}(W \circ P_X)^{-1/2}P_A(a)W_{\Phi(a)} + W \circ P_X)^{-1/2}s\right)$$

In the above derivation, (183) follows from the concavity of $x \mapsto x^s$, (184) follows from $\sum_{a' \neq a} P_A(a') \leq 1$, (185) follows from the inequality $(x+y) \leq x^s + y^s$. Then, we obtain (54).

Next, we show (59) of Lemma 5 by modifying the proof of (53). We introduce the random variable $Z := f_{Z|\tilde{X}}(a) = f_{Z}(a) + \tilde{X}$. The random variable $Z$ is independent of the choice of $f_{Z}$. Since $E_{Z|X}f_{Z|X}(a) = W \circ P_{\text{mix},X}$ for $a \in \mathcal{A}$, the proof of (53) can be applied to the proof of (59) by replacing $\Phi(a)$, $\Phi|\Phi(a)$, and $P_X$ by $Z$, $Z|Z$ and $P_{\text{mix},X}$. The proof of (54) can be applied to the proof of (60) with the same replacement.

**APPENDIX F**

**PROOF OF LEMMA 5**

Next, we show (56) of Lemma 5 by modifying the proof of (59). In this case, for any element $x \in \mathcal{X}$,

$$\frac{P_{Z,\tilde{X}}(x \in C[Z] + \tilde{X}, C[Z] = C)}{P_{Z,\tilde{X}}(C[Z] = C)} = \frac{1}{|\mathcal{X}|} = P_{Z,\tilde{X}}(x \in C[Z] + \tilde{X}).$$

Thus,

$$P_{Z,\tilde{X}}(C[Z] = C|x \in C[Z] + \tilde{X}) = \frac{P_{Z,\tilde{X}}(x \in C[Z] + \tilde{X}, C[Z] = C)}{P_{Z,\tilde{X}}(x \in C[Z] + \tilde{X})} = P_{Z,\tilde{X}}(C[Z] = C).$$

Further, for $x \neq x' \in \mathcal{X}$, when $x \in C[Z] + \tilde{X}$, $x' \in C[Z] + \tilde{X}$ if and only if $x' - x \in C[Z]$. Thus,

$$P_{Z,\tilde{X}}(x' \in C[Z] + \tilde{X}|x \in C[Z] + \tilde{X}) = P_{Z,\tilde{X}}(x' - x \in C[Z]|x \in C[Z] + \tilde{X}) = P_{Z,\tilde{X}}(x' - x \in C[Z]).$$
Therefore,
\[ E_{Z, \hat{X}} D(W_{\text{mix}, C[Z] + \hat{X}} \| W \circ P_{\text{mix}, X}) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} P_{Z, \hat{X}}(x \in C[Z] + \hat{X}) W_x E_{Z|x \in C[Z] + \hat{X}} \left( \log \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in C[Z] + \hat{X}} W_{x'} \right) - \log(W \circ P_{\text{mix}, X}) \right) \]
\[ \leq \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} P_{Z, \hat{X}}(x \in C[Z] + \hat{X}) W_x \left( \log \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in C[Z] + \hat{X}} W_{x'} \right) - \log(W \circ P_{\text{mix}, X}) \right) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( \log \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in \mathcal{X}} W_{x'} \right) - \log(W \circ P_{\text{mix}, X}) \right) \]
\[ \leq \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( \log \left( \frac{1}{|\mathcal{X}|} W_x + W \circ P_{\text{mix}, X} \right) \right) - \log(W \circ P_{\text{mix}, X}) \]  
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \log(I + \frac{1}{|\mathcal{X}|} \mathcal{E}_{W \circ P_{\text{mix}, X}}(W_x)(W \circ P_{\text{mix}, X})^{-1}) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \log(I + \frac{1}{|\mathcal{X}|} \mathcal{E}_{W \circ P_{\text{mix}, X}}(W_x)(W \circ P_{\text{mix}, X})^{-1}) \]
\[ \leq \sum_x \frac{1}{s|\mathcal{X}|} W_x \left( W \circ P_{\text{mix}, X} \right)^{-s} \sum_x \frac{1}{|\mathcal{X}|} s L^s \mathcal{E}_{W \circ P_{\text{mix}, X}}(W_x)(W \circ P_{\text{mix}, X})^{-s} \]
\[ = \sum_x \frac{1}{s|\mathcal{X}|} W_x \left( W \circ P_{\text{mix}, X} \right)^{-s} \sum_x \frac{1}{s|\mathcal{X}|} L^s \mathcal{E}_{W \circ P_{\text{mix}, X}}(W_x)(W \circ P_{\text{mix}, X})^{-s} \]
In the above derivation, (186) follows from the concavity of $x \mapsto \log x$, (187) follows from (13), (188) follows from (182), and (189) follows from (42). Then, we obtain (56).

Next, we show (57).
\[ E_{Z, \hat{X}} D(W_{\text{mix}, C[Z] + \hat{X}} \| W \circ P_{\text{mix}, X}) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} P_{Z, \hat{X}}(x \in C[Z] + \hat{X}) W_x E_{Z|x \in C[Z] + \hat{X}} \left( \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in C[Z] + \hat{X}} W_{x'} \right) \right) \]
\[ \leq \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} P_{Z, \hat{X}}(x \in C[Z] + \hat{X}) W_x \left( \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in \mathcal{X}} W_{x'} \right) \right) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in \mathcal{X}} W_{x'} \right) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( I + \frac{1}{|\mathcal{X}|} \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \right) \left( \frac{1}{|\mathcal{X}|} W_x + \frac{1}{|\mathcal{X}|} \sum_{x' \neq x \in \mathcal{X}} W_{x'} \right) \]
\[ \leq \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( I + \frac{1}{|\mathcal{X}|} \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \right) \left( \frac{1}{|\mathcal{X}|} W_x + W \circ P_{\text{mix}, X} \right) \]
\[ = \text{Tr} \sum_x \frac{1}{|\mathcal{X}|} W_x \left( I + \frac{1}{|\mathcal{X}|} \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \right) \left( \frac{1}{|\mathcal{X}|} W_x + W \circ P_{\text{mix}, X} \right) \]
\[ = 1 + \frac{1}{|\mathcal{X}|} \sum_x \frac{1}{|\mathcal{X}|} \text{Tr} W_x \left( W \circ P_{\text{mix}, X} \right)^{-1/2} \left( \frac{1}{|\mathcal{X}|} W_x + W \circ P_{\text{mix}, X} \right) \]
In the above derivation, (190) follows from the concavity of $x \mapsto x^s$, (191) follows from $(x + y) \leq x^s + y^s$. Thus, similar to (54), we obtain (57).