A New Family of Probability Distributions and Asymptotics of Classical and LOCC Conversions

Wataru Kumagai\textsuperscript{1}, Masahito Hayashi\textsuperscript{2,3}

\textsuperscript{1}Faculty of Engineering, Kanagawa University, Japan, \textsuperscript{2}Graduate School of Mathematics, Nagoya University, Japan, \textsuperscript{3}Centre for Quantum Technologies, National University of Singapore, Singapore.

Abstract

We consider the optimal approximate conversion between multiple copies of pure entangled states in quantum systems when only local operations and classical communications (LOCC) are allowed. This problem contains a kind of cloning problem with LOCC restriction as a special case. To derive the asymptotic LOCC conversion rate, we consider two kinds of approximate conversions, deterministic conversion and majorization conversion, for probability distributions, and solve their asymptotic conversion rates up to the second order. Then, the asymptotic LOCC conversion rate is obtained via the natural relation between the LOCC conversion and the majorization conversion. To derive these asymptotic rates, we introduce new probability distributions named Rayleigh-normal distributions. The family of Rayleigh-normal distributions includes a Rayleigh distribution and coincides with the standard normal distribution in the limit case, and the optimal conversion rate is represented by Rayleigh-normal distribution in a unified manner.

1 Introduction

The topics of this paper consist of three different areas, quantum information theory, classical information theory, and probability theory although the obtained results in these areas are correlated with each other. Hence, we describe the obtained results separately in this section although the most important result is the LOCC conversion.

1.1 Contribution to Probability Theory

We introduce a new family of probability distributions with one-parameter $v \in [0, +\infty)$ on real numbers called Rayleigh-normal distributions in order to describe the optimal LOCC conversion rate. Rayleigh-normal distributions contain a Rayleigh distribution at $v = 1$ and the standard normal distribution.
with the limits \( v \to 0, +\infty \). Thus, the family connects a Rayleigh distribution and the standard normal distribution, which is the origin of the name of Rayleigh-normal distribution. The Rayleigh-normal distributions is defined as the solution of an optimization problem for continuous probability distributions on real numbers. Although the definition seems very abstract, the cumulative distribution function \( Z_v \) can be explicitly given dependently of the parameter \( v \in [0, \infty) \). Then, we can plot the graphs of the cumulative distribution functions as Fig. 1. The family of Rayleigh-normal distributions has a symmetry for the reflection with respect to the parameter \( v \) and other properties. Based on these useful properties, the Rayleigh-normal distributions characterize the asymptotic second order coefficients of the LOCC conversion as well as the conversion of probability distributions. That is, the Rayleigh-normal distributions work as a kind of information quantities in the second order asymptotics. The notion of the Rayleigh-normal distributions will be used in the later parts in introduction.

1.2 Contribution to Quantum Information Theory

In quantum information theory, various quantum tasks have been proposed and a specific entangled state is often required to implement those tasks. When some distant parties want to implement such quantum tasks, they have to prepare the desired entangled state in advance. In the situation, LOCC is a fundamental method to prepare an entangled state shared between distant places. Based on the problem, we consider LOCC conversion between the multiple copies of general pure entangled states on bipartite systems in this paper. We especially focus on the maximum convertible number of copies from \( \psi \) to \( \phi \) by LOCC under a permissible accuracy \( 0 < \nu < 1 \) defined as follows:

\[
L_n(\psi, \phi|\nu) := \max_{F:LOCC} \{ L \in \mathbb{N} | \Gamma(\psi^\otimes n, \phi^\otimes L) \geq \nu \},
\]

where \( F \) is the fidelity between quantum states. This number represents how many copies of the target entangled state \( \phi \) can be generated from a given entangled state \( \psi^\otimes n \) by LOCC under the accuracy constraint. As a fundamental result of LOCC conversion, Bennett et. al. [4] showed that the optimal LOCC conversion rate from a pure entangled state \( \psi \) to another one \( \phi \) is the ratio \( S_\psi/S_\phi \) of von Neumann entropies of their reduced density matrices. In other words, they proved the first-order asymptotic expansion \( L_n(\psi, \phi|\nu) = S_\psi/S_\phi n + o(n) \). However, the asymptotic expansion does not give good approximation to the maximum convertible number \( L_n(\psi, \phi|\nu) \) because the first order coefficient \( S_\psi/S_\phi \) does not reflect the fidelity \( \nu \). Thus, higher-order asymptotics is required to accurately evaluate its behaviour. To resolve this problem, using the cumulative distribution function \( \Phi \) of the standard normal distribution, Kumagai and Hayashi [24] explicitly derived derived accuracy of entanglement dilution and entanglement concentration, which corresponds to the case when \( \psi \) or \( \phi \) is the maximally entangled state. The result of [24] implies that the second-order optimal LOCC conversion rates in entanglement dilution and entanglement concentration are
represented by \( \Phi \) as
\[
L_n(\psi, \phi | \nu) = \frac{S_\psi}{S_\phi} n + \text{const}_{\psi, \phi} \Phi^{-1}(1 - \nu^2)\sqrt{n} + o(\sqrt{n}),
\]
where \( \text{const}_{\psi, \phi} \) is a constant given in (148) and (149). In fact, besides entanglement dilution and entanglement concentration, the cumulative standard normal distribution function \( \Phi \) commonly appears in the second-order rates for typical quantum information-processing tasks including quantum hypothesis testing [25, 34], classical-quantum channel coding [35], quantum fixed-length source coding [34, 11], data compression with quantum side information [34], randomness extraction against quantum side information [34], and noisy dense coding [11], as was also pointed out in Datta and Leditzky [11].

In this paper, we consider LOCC conversion when \( \phi \) or \( \psi \) are not maximally entangled. This setting is more important when the entangled states are used for measurement-based quantum computation [14] and quantum channel estimation [22] because these tasks require specific entangled states that are not necessarily maximally entangled. Thus, these tasks require us to efficiently generate non-maximally entangled states by LOCC conversion. Surprisingly, it is shown that the second-order optimal LOCC conversion rate between general pure states \( \psi \) and \( \phi \) can not be represented by the cumulative distribution function of the standard normal distribution but by that of the Rayleigh-normal distribution as follows
\[
L_n(\psi, \phi | \nu) = \frac{S_\psi}{S_\phi} n + \frac{Z_{\psi, \phi}^{-1} (1 - \nu^2)}{D_{\psi, \phi}} \sqrt{n} + o(\sqrt{n}),
\]
where \( D_{\psi, \phi} \) and \( C_{\psi, \phi} \) are certain constants. In this point, our result is different from conventional behavior of second-order rates and is quite nontrivial. In particular, when either the initial state \( \psi \) or the target entangled state \( \phi \) is a maximally entangled state, the cumulative Rayleigh-normal distribution function coincides with the cumulative standard normal distribution function \( \Phi \) and the above expansion (3) recovers (2).

Next, as a special situation of LOCC conversion, we focus on the case when the target entangled state is the same with the given entangled state (i.e. \( \phi = \psi \)). In this special case, the formula (3) can be simplified to
\[
L_n(\psi, \psi | \nu) \cong n + \frac{\sqrt{8V_\psi \log \nu^{-1}}}{S_\psi} \sqrt{n},
\]
where \( V_\psi \) is a constant depending on \( \psi \). This special case can be regarded as a special type of asymptotic cloning problem, which has some interest recently. However, our problem is different from conventional settings of cloning in the following points. Although the knowledge of the state to be cloned is not perfect in the conventional setting, our setting assumes the perfect knowledge for the entangled state to be cloned. The essential point of our setting is the restriction of our operations to LOCC operations. No additional entangled resource are
not prepared. This is contrastive with the LOCC cloning by the papers [32, 31] because their setting assumes an imperfect knowledge for the entangled state to be cloned and additional limited entangled resource. To distinguish their setting, we call our setting the LOCC cloning with perfect knowledge, and call their setting the LOCC cloning with imperfect knowledge.

To characterize the performance, Chiribella et al. [10] introduced the replication rate as the rate of the incremental copies. The formula [11] shows that the replication rate of the LOCC cloning with perfect knowledge is 1/2.

1.3 Contribution to Classical Information Theory

We consider two kinds of conversions for probability distributions called deterministic conversion and majorization conversion. The deterministic conversion is more natural setting from the viewpoint of classical information theory because it describe natural random number conversion. However, analysis of the majorization conversion is more important from the viewpoint of quantum information theory because LOCC conversion for pure entangled states is mathematically reduced to the majorization conversion of probability distribution. That is, the above mentioned result in quantum information theory is obtained via the analysis on majorization conversion. Since these two problems are deeply related to each other, solving one problem requires the analysis of the other problem. Hence, we treat both problems in this paper.

In this paper, we consider the both kinds of conversion between independent and identical distributions of two given distribution $P$ and $Q$ in the asymptotic setting. To directly apply the results of majorization conversion to the LOCC conversion, as accuracy of classical conversions, we adopt the fidelity (or Bhattacharyya coefficient) $F$ between probability distributions. Then, we mainly focus on the following values called the maximum convertible number of the target distribution $Q$ with $n$-fold independent and identical distribution $P^n$ of the distribution $P$ for a permissible accuracy $0 < \nu < 1$ by majorization conversions

$$L_n^M(P, Q|\nu) := \max_{W} \{ L \in \mathbb{N}| F(W'(P^n), Q^L) \geq \nu, W' : \text{majorization conversion} \},$$

and that by deterministic conversions

$$L_n^D(P, Q|\nu) := \max_{W} \{ L| F(W(P^n), Q^L) \geq \nu, W : \text{deterministic conversion} \}.$$

Those numbers represent how many copies of the target probability distribution $Q$ can be generated from the initial probability distribution $P^n$ under the accuracy constraint $\nu$. It is known that the first order coefficient of $L_n^D(P, Q|\nu)$ is the ratio of the Shannon entropies $H(P)$ and $H(Q)$ [15] and does not depend on the accuracy $\nu$. Recently, as a more precise asymptotic characterization, the second order asymptotics attracts much attention [21, 33, 20]. When either initial or target probability distribution is uniform, the asymptotic expansions of these numbers are solved up to the second order $\sqrt{n}$ [20, 30]. However, in the non-uniform case (i.e. neither given nor target probability distribution is uniform),

$$L_n^M(P, Q|\nu) := \max_{W} \{ L \in \mathbb{N}| F(W'(P^n), Q^L) \geq \nu, W' : \text{majorization conversion} \},$$

and that by deterministic conversions

$$L_n^D(P, Q|\nu) := \max_{W} \{ L| F(W(P^n), Q^L) \geq \nu, W : \text{deterministic conversion} \}.$$
the corresponding result has not been shown in the context of second-order asymptotics.

In this paper, we show that the asymptotic behaviour of the maximum convertible numbers is described by the inverse function of a cumulative Rayleigh-normal distribution with certain constants $D_{P,Q}$ and $C_{P,Q}$ as follows

$$L_n^D(P,Q|\nu) \cong L_n^M(P,Q|\nu) = \frac{H(P)}{H(Q)}n + \frac{Z_{C_{P,Q}}^{-1}(1-\nu^2)}{D_{P,Q}} \sqrt{n} + o(\sqrt{n}),$$

where $\cong$ shows that the difference between the left and the right side terms is $o(\sqrt{n})$ at most and $H(P)$ and $H(Q)$ are the Shannon entropies. The asymptotic expansion (5) is very similar to the form in (3). In fact, it is shown that (5) is essentially equivalent to (3) in Section 5.1.

1.4 Organization of This Paper

The paper is organized as follows. In Section 2, we introduce a new family of probability distributions on real numbers. It describes the optimal conversion rate under the accuracy constraint in Section 4. In Section 3, as problems in classical information theory, we formulate two kinds of approximate conversion problems between two probability distributions by using the majorization condition and the deterministic transformation, respectively. Then, we define the maximum convertible numbers and describe their properties in non-asymptotic setting. In Section 4, we consider derive the asymptotic expansion of these numbers up to the second order $\sqrt{n}$. In these derivation, we divide our setting into two cases, uniform case and non-uniform case. The non-uniform case itself does not contain the uniform case, however, we show that the results in the uniform case can be regarded as the limit of the results in the general case. In Section 5, we apply the results of Sections 4 to the LOCC conversion. Then, we obtain the optimal LOCC conversion rate between general pure states up to the second-order $\sqrt{n}$. As a special case, we derive the rate of the incremental copies and the optimal coefficient for the LOCC cloning with the perfect knowledge. In Section 6, we give the conclusion.

2 Rayleigh-Normal Distribution

2.1 Introduction of Rayleigh-Normal Distribution

In this section, we introduce a new probability distribution family on $\mathbb{R}$ with one parameter which connects the normal distribution and the Rayleigh distribution. A function $Z$ on $\mathbb{R}$ is generally called a cumulative distribution function if $Z$ is right continuous, monotonically increasing and satisfies $\lim_{x \to -\infty} Z(x) = 0$ and $\lim_{x \to \infty} Z(x) = 1$. Then, there uniquely exists a probability distribution on $\mathbb{R}$ whose cumulative distribution coincides with $Z$. That is, given a cumulative distribution function in the above sense, it determines a probability distribution.
on \( \mathbb{R} \). To define the new probability distribution family, we give its cumulative distribution function.

For \( \mu \in \mathbb{R} \) and \( v \in \mathbb{R}_+ \), let \( \Phi_{\mu,v} \) and \( N_{\mu,v} \) be the cumulative distribution function and the probability density function of the normal distribution with the mean \( \mu \) and the variance \( v \). We denote \( \Phi_{0,1} \) and \( N_{0,1} \) simply by \( \Phi \) and \( N \). Using the continuous fidelity (or the Bhattacharyya coefficient) for continuous probability distributions \( p \) and \( q \) on \( \mathbb{R} \) defined by

\[
\mathcal{F}(p, q) := \int_{\mathbb{R}} \sqrt{p(x)q(x)} \, dx,
\]

we define the following function.

**Definition 1** For \( v > 0 \), a Rayleigh-normal distribution function \( Z_v \) on \( \mathbb{R} \) is defined by

\[
Z_v(\mu) = 1 - \sup_{A} \mathcal{F} \left( \frac{dA}{dx}, N_{\mu,v} \right)^2,
\]

where \( A : \mathbb{R} \to [0, 1] \) runs over continuous differentiable monotone increasing functions satisfying \( \Phi \leq A \leq 1 \) in RHS.

The Rayleigh-normal distribution function is proven to be a cumulative distribution function later, and thus, it determines a probability distribution on \( \mathbb{R} \). The graphs of the Rayleigh-normal distributions can be described as in Fig. 1 by Proposition 4 in subsection 2.2. We note that Rayleigh distribution is included in Weibull distribution and Weibull-normal distribution is already proposed [9], however, our Rayleigh-normal distribution is different from the Weibull-normal distribution in [9] because a Rayleigh distribution with a specific scale parameter is included in the family of Rayleigh-normal distributions and is not in that of the Weibull-normal distributions.

### 2.2 Properties of Rayleigh-Normal Distribution

In the following, we show that it connects the Rayleigh distribution with scale parameter \( \sqrt{2} \) and the standard normal distribution. To give an explicit form of the Rayleigh-normal distribution functions, we prepare two lemmas.

**Lemma 2** When \( v < 1 \), the equation with respect to \( x \)

\[
\frac{N(x)}{N_{\mu,v}(x)} = \frac{1 - \Phi(x)}{1 - \Phi_{\mu,v}(x)}
\]

has the unique solution \( \beta_{\mu,v} < \frac{\mu}{1-v} \).

**Proof:** The existence of the unique solution is equivalent to the existence of the unique zero point of the function

\[
f(x) := (1 - \Phi_{\mu,v}(x)) - (1 - \Phi(x)) \frac{N_{\mu,v}(x)}{N(x)}.
\]
Figure 1: The graphs of the Rayleigh-normal distributions. The black, purple, green, blue and red lines represent the Rayleigh-normal distributions with parameter \( v = 0, 1/10, 1/6, 1/3 \) and 1. The Rayleigh-normal distribution with parameter \( v > 1 \) can be transformed to that with parameter \( v < 1 \) by Lemma 5.

Since

\[
\frac{df}{dx} = -\frac{d}{dx} \left( \frac{N_{\mu,v}(x)}{N} \right) (1 - \Phi),
\]

the function \( f \) is strictly monotonically decreasing when \( x < \arg\min(\frac{N}{N_{\mu,v}}) = \frac{\mu}{1 - v} \) and is strictly monotonically increasing when \( x > \frac{\mu}{1 - v} \). Since

\[
\lim_{x \to -\infty} f(x) = 1, \quad \lim_{x \to \infty} f(x) = 0,
\]

the function \( f \) has the unique zero point \( \beta_{\mu,v} < \frac{\mu}{1 - v} \) due to the intermediate value theorem.

**Lemma 3** When \( v > 1 \), the equation with respect to \( x \)

\[
\frac{N(x)}{N_{\mu,v}(x)} = \frac{\Phi(x)}{\Phi_{\mu,v}(x)}
\]

has the unique solution \( \alpha_{\mu,v} > \frac{\mu}{1 - v} \).

**Proof:** The existence of the unique solution is equivalent to the existence of the unique zero point of the function \( f(x) = (N(x)/N_{\mu,v}(x))\Phi_{\mu,v}(x) - \Phi(x) \). Since

\[
\frac{df}{dx} = \frac{d}{dx} \left( \frac{N}{N_{\mu,v}} \right) \Phi_{\mu,v},
\]

7
the function \( f \) is strictly monotonically increasing when \( x < \text{argmax}(N/N_{\mu,v}) = \frac{\mu}{1-v} \) and is strictly monotonically decreasing when \( x > \frac{\mu}{1-v} \). Since
\[
\lim_{x \to -\infty} f(x) = 0, \quad \lim_{x \to \infty} f(x) = -1,
\] (14)
the function \( f \) has the unique zero point \( \alpha_{\mu,v} > \frac{\mu}{1-v} \) due to the intermediate value theorem.

We denote the cumulative distribution function of the Rayleigh distribution with scale parameter \( \sigma > 0 \) by
\[
R_{\sigma}(x) = \begin{cases} 
1 - e^{-\frac{x^2}{2\sigma^2}} & \text{if } x > 0 \\
0 & \text{if } x \leq 0
\end{cases}
\]
Then, a family of Rayleigh-normal distribution functions is represented as follows. In particular, it includes the Rayleigh distribution with scale parameter \( \sigma = \sqrt{2} \) when \( v = 1 \).

**Theorem 4** For \( v > 0 \), the following holds
\[
Z_{\nu}(\mu) = \begin{cases} 
1 - ((1 - \Phi(\beta_{\mu,v}) \sqrt{1 - \Phi_{\mu,v}(\beta_{\mu,v})} + I_{\mu,v}(\beta_{\mu,v}))^2 & \text{if } v < 1 \\
R_{\sqrt{2}(\mu)} & \text{if } v = 1 \\
1 - ((\Phi(\alpha_{\mu,v}) \Phi_{\mu,v}(\alpha_{\mu,v}) + I_{\mu,v}(\infty) - I_{\mu,v}(\alpha_{\mu,v}))^2 & \text{if } v > 1,
\] (15)
where
\[
I_{\mu,v}(x) := \sqrt{2v} \frac{e^{-\frac{x^2}{2v}}}{\Phi \left( \sqrt{\frac{1+v}{v}} \right)} (x),
\] (16)
\[
I_{\mu,v}(\infty) := \lim_{x \to \infty} I_{\mu,v}(x) = \sqrt{2} \frac{e^{-\frac{\mu^2}{2(1+v)}}}{\Phi \left( \sqrt{\frac{1+v}{v}} \right)}.
\] (17)
Theorem 4 is proven in Appendix A. From Theorem 4, we can show a kind of symmetry of the family of the Rayleigh-normal distribution functions about the inversion of \( v \) as follows.

**Lemma 5**
\[
Z_{\nu}(\mu) = Z_{\nu^{-1}} \left( \frac{\mu}{\sqrt{v}} \right).
\] (18)

**Proof:** We first show
\[
\alpha_{\mu,v} = -\sqrt{\frac{v}{1-v}} + \mu.
\] (19)
for \( v > 1 \). We define a function \( g \) on \( \mathbb{R} \) as \( g(x) := \sqrt{v}x - \mu \). Since the solution \( \alpha_{\mu,v} \) of the equation (12) is unique, we only have to show that \(-g(\beta_{\mu,v})\) is a solution of the equation (12). Note that the function \( g \) is represented as
\[
g(x) = \Phi_{\sqrt{v}/\mu,v}^{-1} \circ \Phi(x) = \Phi_{\sqrt{v}/\mu,v}^{-1} \circ \Phi(x).
\] (20)
Then, the following holds

$$\frac{\Phi(-g(\beta_{\mu,v}))}{\Phi_{\mu,v}(-g(\beta_{\mu,v}))} = \frac{1-\Phi(g(\beta_{\mu,v}))}{1-\Phi_{\mu,v}(g(\beta_{\mu,v}))} = \frac{1-\Phi_{\mu,v}(\beta_{\mu,v})}{1-\Phi(\beta_{\mu,v})}.$$

$$= \frac{N(\beta_{\mu,v})}{N(\beta_{\mu,v})} = \frac{\Phi(g(\beta_{\mu,v}))}{\Phi_{\mu,v}(g(\beta_{\mu,v}))}.$$

Thus, we have (19). From (20), we have

$$\Phi(\alpha_{\mu,v}) = 1 - \Phi_{\mu,v}(g(\beta_{\mu,v})) \quad (21)$$

$$\Phi_{\mu,v}(\alpha_{\mu,v}) = 1 - \Phi_{\mu,v}(\beta_{\mu,v}). \quad (22)$$

From direct calculation,

$$I_{\mu,v}(\infty) - I_{\mu,v}(\alpha_{\mu,v}) = I_{\mu,v}(\beta_{\mu,v}). \quad (23)$$

holds, and thus, we obtain (18).

By Proposition 18, the behaviour of the Rayleigh-normal distribution function $Z_v$ for $v > 1$ can be represented by that for $v < 1$.

Next we show that the family of Rayleigh-normal distribution function includes the standard normal distribution function as its extreme distribution.

**Proposition 6**

$$\lim_{v \to 0} Z_v(\mu) = \lim_{v \to \infty} Z_v(\sqrt{v}\mu) = \Phi(\mu). \quad (24)$$

**Proof:** Since the first equation obviously holds because of Lemma 5, we show only $\lim_{v \to 0} Z_v(\mu) = \Phi(\mu)$ in the following.

First, the definition in (16) implies

$$0 \leq I_{\mu,v}(\beta_{\mu,v}) \leq I_{\mu,v}(\infty) \xrightarrow{v \to 0} 0. \quad (25)$$

Moreover, as shown below,

$$\lim_{v \to 0} \Phi(\beta_{\mu,v}) = \Phi(\mu) \quad (26)$$

$$\lim_{v \to 0} \Phi_{\mu,v}(\beta_{\mu,v}) = 0. \quad (27)$$

The relations (25), (26) and (27) yield the equation $\lim_{v \to 0} Z_v(\mu) = \Phi(\mu)$.

Thus, all we have to do is to show (26) and (27). First, we will show (26). In order to show (26), it is enough to prove that $\lim_{v \to 0} \beta_{\mu,v} = \mu$. First, we have $\beta_{\mu,v} < \frac{\mu}{1 - v}$ from Lemma 2. Since $\lim_{v \to 0} \frac{\mu}{1 - v} = \mu$, we obtain $\lim_{v \to 0} \beta_{\mu,v} \leq \mu$. Therefore, (26) follows. The rest of the proof is similar to that for (27).
\( \mu \). Next, we set the function \( f_{\mu,v}(x) \) as \( (1 - \Phi_{\mu,v}(x)) - (1 - \Phi(x)) \frac{N_{\mu,v}(x)}{N(x)} \) and take an arbitrary \( x \in \mathbb{R} \) such that \( x < \mu \). Since \( \lim_{v \to 0} N_{\mu,v}(x) = 0 \) and \( \lim_{v \to 0} \Phi_{\mu,v}(x) > 0 \), in other words, \( x \) is not a zero point of \( f_{\mu,v} \) when \( v \) is close to 0. Thus, we obtain \( \lim_{v \to 0} \beta_{\mu,v} \geq \mu \) since \( \beta_{\mu,v} \) is the unique zero point of \( f_{\mu,v} \). Therefore, \( \lim_{v \to 0} \beta_{\mu,v} = \mu \) holds.

Next, we will show (27). In order to show (27), it is enough to prove that \( \lim_{v \to 0} \beta_{\mu,v} - \mu \sqrt{v} = -\infty \) by the definition of \( \Phi_{\mu,v} \). Since \( \beta_{\mu,v} < \frac{\mu}{1-v} \) and \( \lim_{v \to 0} \frac{1}{1-v} = \mu \), \( \beta_{\mu,v} \) is bounded above by some constant \( \gamma \) as \( \beta_{\mu,v} < \gamma \) when \( v \) is close to 0, and then, we have the following inequality:

\[
\frac{N(\beta_{\mu,v})}{N_{\mu,v}(\beta_{\mu,v})} = \frac{1 - \Phi(\beta_{\mu,v})}{1 - \Phi_{\mu,v}(\beta_{\mu,v})} \geq \Phi(-\gamma). \tag{28}
\]

Thus, the following holds

\[
2 \log \Phi(-\gamma) \leq -\log \frac{N(\beta_{\mu,v})}{N_{\mu,v}(\beta_{\mu,v})} = (1 - v) \left( \frac{\beta_{\mu,v} - \mu(1-v)^{-1}}{\sqrt{v}} \right)^2 + \log v - \left( \frac{\mu}{1-v} \right)^2. \tag{29}
\]

Therefore, we have

\[
\lim_{v \to 0} \left( \frac{\beta_{\mu,v} - \mu(1-v)^{-1}}{\sqrt{v}} \right)^2 = \infty. \tag{30}
\]

Since Lemma guarantees that \( \beta_{\mu,v} - \mu \frac{1}{1-v} < 0 \), we obtain

\[
\lim_{v \to 0} \frac{\beta_{\mu,v} - \mu}{\sqrt{v}} = \lim_{v \to 0} \frac{\beta_{\mu,v} - \mu(1-v)^{-1}}{\sqrt{v}} = -\infty. \tag{31}
\]

Thus, the family of Rayleigh-normal distribution functions connects the cumulative distribution functions of the Rayleigh distribution and the standard normal distribution. In the following, we define as \( Z_0 := \Phi \).

Finally, we give the following most basic property of the Rayleigh-normal distribution function.

**Proposition 7** The Rayleigh-normal distribution function \( Z_v \) is a cumulative distribution function for each \( v \geq 0 \).

**Proof:** From Lemma, it is enough to treat the case when \( 0 \leq v < 1 \). This proposition is obvious for \( v = 0 \) by the definition \( Z_0 := \Phi \). In the following, we fix \( 0 < v < 1 \). First, we show that \( Z_v(\mu) \) is continuous. From Lemma, \( \beta_{\mu,v} \) is differentiable, especially continuous, with respect to \( \mu \). Thus, \( Z_v(\mu) \) is continuous from Theorem. Next we show \( \lim_{\mu \to -\infty} Z_v(\mu) = 0 \). From Lemma, the inequality \( \beta_{\mu,v} < \frac{\mu}{1-v} \) holds and thus \( \lim_{\mu \to -\infty} \beta_{\mu,v} = -\infty \). Similarly, from Lemma
The inequality $\alpha \frac{\mu}{v} > \sqrt{v(1-v^{-1})}$ holds and thus $\lim_{\mu \to -\infty} \alpha \frac{\mu}{v} = \infty$. Therefore, we obtain $\lim_{\mu \to -\infty} \Phi(\beta_{\mu,v}) = 0$, $\lim_{\mu \to -\infty} \Phi_{\mu,v}(\beta_{\mu,v}) = \lim_{\mu \to -\infty} \Phi(\alpha \frac{\mu}{v}) = 0$. Since $\lim_{\mu \to -\infty} I_{\mu,v}(\beta_{\mu,v}) \leq \lim_{\mu \to -\infty} I_{\mu,v}(\infty) = 0$, we have $\lim_{\mu \to -\infty} Z_v(\mu) = 0$ from Theorem 4. Next we show $\lim_{\mu \to -\infty} \Phi(\beta_{\mu,v}) = 0$, $\lim_{\mu \to -\infty} \Phi_{\mu,v}(\beta_{\mu,v}) = \lim_{\mu \to -\infty} \Phi(-\alpha \frac{\mu}{v}) = 0$. Since $\lim_{\mu \to -\infty} I_{\mu,v}(\beta_{\mu,v}) \leq \lim_{\mu \to -\infty} I_{\mu,v}(\infty) = 0$, we have $\lim_{\mu \to -\infty} Z_v(\mu) = 0$ from Theorem 4. Finally, we show that $Z_v(\mu)$ is monotonically increasing. We define a shift operator $S_\mu$ for a map $A : \mathbb{R} \to \mathbb{R}$ by $(S_\mu A)(x) := A(x - \mu)$. Then we have $\mathcal{F}(S_\mu p, S_\mu q) = \mathcal{F}(p, q)$. Thus when we set as

$$\mathcal{A}(\mu) := \left\{ A : \mathbb{R} \to [0, 1] \mid A \text{ is a continuous differentiable monotone increasing function such that } \Phi_{\mu,1} \leq A \leq 1 \right\},$$

we obtain the following form of the Rayleigh-normal distribution function

$$Z_v(\mu) := 1 - \sup_{A \in \mathcal{A}(0)} \mathcal{F} \left( \frac{dA}{dx}, N_{\mu,v} \right)^2 = 1 - \sup_{A \in \mathcal{A}(0)} \mathcal{F} \left( S_{-\mu} \frac{dA}{dx}, S_{-\mu} N_{\mu,v} \right)^2 = 1 - \sup_{A \in \mathcal{A}(0)} \mathcal{F} \left( \frac{d(S_{-\mu} A)}{dx}, N_{0,v} \right)^2 = 1 - \sup_{A \in \mathcal{A}(0)} \mathcal{F} \left( \frac{dA}{dx}, N_{0,v} \right)^2.$$

For $\mu < \tau$, $\mathcal{A}(\tau) \supset \mathcal{A}(\tau)$ holds, and thus we obtain $Z_v(\mu) \leq Z_v(\tau)$. By Proposition 7 the set of the functions $Z_v$ determines a family of probability distributions on $\mathbb{R}$ with one parameter $v \geq 0$. We call the probability distribution determined by $Z_v$ the Rayleigh-normal distribution. As shown in Theorem 10 the family of probability distribution functions can represent the optimal conversion rate in the second order asymptotics.

3 Conversions for Probability Distributions: Non-Asymptotic Setting

In Sections 3 and 4, we focus on mathematical aspects for two kinds of conversions called deterministic conversion and majorization conversion for probability distributions, and their roles in quantum information theory will be explained in Section 5.1. Through this paper, the entropy of a probability distribution plays an essential role and is required not to be 0 in several situations. Thus we assume that the entropy of a probability distribution is not 0 in this paper, or equivalently, the support size of a probability distribution is not 1.

3.1 Deterministic Conversion

In order to describe the accuracy of approximate conversion, we introduce a function $F$ called the fidelity or the Bhattacharyya coefficient between proba-
probability distributions over the same discrete set $\mathcal{Y}$ as
\[ F(Q, Q') := \sum_{y \in \mathcal{Y}} \sqrt{Q(y)} \sqrt{Q'(y)}. \] (32)

The fidelity $F$ represents how close two probability distributions are and relates to the Hellinger distance $d_H$ as $d_H(\cdot, \cdot) = \sqrt{1 - F(\cdot, \cdot)}$ [30].

In order to give a deterministic conversion, for a probability distribution $P$ on a finite set $\mathcal{X}$ and a map $W : \mathcal{X} \to \mathcal{Y}$, we define the probability distribution $W(P)$ on $\mathcal{Y}$ by $W(P)(y) := \sum_{x \in W^{-1}(y')} P(x)$. Then, we call the conversion from $P$ to $W(P)$ the deterministic conversion by $W$. When a permissible accuracy $0 < \nu < 1$ is fixed, we define the maximum convertible number $L$ of $Q^L$ which can be approximated from $P$ by deterministic conversions as
\[ L^D(P, Q|\nu) := \max_W \{L | F(W(P), Q^L) \geq \nu, W : \mathcal{X} \to \mathcal{Y}^L \}. \] (33)

One of the main topics of the paper is to analyze the above maximum convertible number by deterministic conversions. In order to rewrite $L^D(P, Q|\nu)$, we define the maximum fidelity $F^D$ from $P$ on $\mathcal{X}$ to $Q$ on $\mathcal{Y}$ among deterministic conversions by
\[ F^D(P \to Q) := \max_W \{F(W(P), Q)|W : \mathcal{X} \to \mathcal{Y} \}. \] (34)

Then $L^D$ is rewritten as
\[ L^D(P, Q|\nu) = \max_{L \in \mathbb{N}} \{L | F^D(P \to Q^L) \geq \nu \}. \] (35)

We denote the maximum convertible number from $n$-i.i.d. $P^n$ to i.i.d. of $Q$ with a permissible accuracy $0 < \nu < 1$ by deterministic conversions as
\[ L^D_n(P, Q|\nu) := L^D(P^n, Q|\nu). \]

### 3.2 Majorization Conversion

In order to relax the condition for conversion, we introduce the concept of majorization. For a probability distribution $P$ on a finite set $\mathcal{X}$, let $P_i$ be a sequence $\{P_i\}_{i=1}^{|\mathcal{X}|}$ where $|\mathcal{X}|$ represents the cardinality of the set $\mathcal{X}$ and $P_i$ is the $i$-th element of $\{P(x)\}_{x \in \mathcal{X}}$ sorted in decreasing order for $1 \leq i \leq |\mathcal{X}|$. When probability distributions $P$ and $Q$ satisfy $\sum_{i=1}^l P_i \leq \sum_{i=1}^l Q_i$ for any $l$, it is said that $P$ is majorized by $Q$ and written as $P \prec Q$. Here, we note that the sets where $P$ and $Q$ are defined do not necessarily coincide with each other, and the majorization relation is a partial order on a set of probability distributions on finite sets [26] [2]. When $P \prec P'$, we call the conversion from $P$ to $P'$ a majorization conversion. Majorization conversions have an operational meaning in quantum settings as we will see in Section 5. We give two important remarks on the majorization. The first one is that the majorization relation
$P \prec W(P)$ holds for a probability distribution $P$ on a finite set $\mathcal{X}$ and a map $W : \mathcal{X} \to \mathcal{Y}$. This fact means that a deterministic conversion is a kind of majorization conversions and derives the following relations

$$F^M(P \to Q) \geq F^D(P \to Q), \quad (36)$$

$$L^M(P,Q|\nu) \geq L^D(P,Q|\nu), \quad (37)$$

$$L_n^M(P,Q|\nu) \geq L_n^D(P,Q|\nu). \quad (38)$$

These inequalities play an quite essential role in the asymptotics of the maximum convertible numbers by those conversions. The second one is that, when the support size of a probability distribution $P$ is less than or equal to $L$ and $U_L$ is the uniform distribution with support size $L$, we have $U_L \prec P$. This fact is necessary in the analysis for the quantum operation called entanglement concentration which is treated in Section 5.

Here, we define the maximum convertible number $L$ of $Q^L$ which can be approximated from $P$ under a permissible accuracy $0 < \nu < 1$ among majorization conversions as

$$L^M(P,Q|\nu) := \max_{P'} \{ L \in \mathbb{N} | F(P',Q^L) \geq \nu, P \prec P' \}. \quad (39)$$

To analyze the above maximum convertible number by majorization conversion is also one of the main topics in the paper beside to treat that by deterministic conversions. In order to rewrite $L^M(P,Q|\nu)$, we introduce the maximum fidelity among the majorization conversions as

$$F^M(P \to Q) := \max_{P'} \{ F(P',Q)|P \prec P' \} \quad (40)$$

where $P$ and $Q$ are probability distributions on $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then $L^M$ is rewritten as

$$L^M(P,Q|\nu) = \max_{L \in \mathbb{N}} \{ L | F^M(P \to Q^L) \geq \nu \}. \quad (41)$$

We also denote the maximum convertible number from $n$-i.i.d. of $P$ to i.i.d. of $Q$ under a permissible accuracy $0 < \nu < 1$ by majorization conversions as

$$L_n^M(P,Q|\nu) := L^M(P^n,Q|\nu).$$

This quantity plays an important role in quantum information theory as is shown in Section 5.

Next, we prepare two lemmas for discussions in latter parts. First, the following is easily proven by using the Schwartz inequality.

**Lemma 8** [38] For a probability distribution $Q$ and a natural number $L$, let $D_L(Q)$ be defined as follows:

$$D_L(Q)(j) := \begin{cases} \frac{Q^L(j)}{\sum_{i=1}^{L} Q^L(i)} & \text{if } 1 \leq j \leq L \\ 0 & \text{if } L + 1 \leq j. \end{cases} \quad (42)$$
Then, for the uniform distribution $U_L$ whose support size is $L$,

$$F^M(U_L \to Q) = F(\mathcal{D}_L(Q), Q^\downarrow) = \sqrt{\sum_{i=1}^{L} Q^\downarrow(i)} \quad (43)$$

holds.

Similarly, the following lemma holds.

**Lemma 9** For a probability distribution $P$ and a natural number $L$, let $\mathcal{C}_L(P)$ be defined as follows:

$$\mathcal{C}_L(P)(j) := \begin{cases} \frac{P^\downarrow(j)}{\sum_{i=J_{P,L}}^{L} P^\downarrow(i)} & \text{if } 1 \leq j < J_{P,L} \\ \frac{\sum_{i=J_{P,L}}^{L} P^\downarrow(i)}{L+1-J_{P,L}} & \text{if } J_{P,L} \leq j \leq L \end{cases} \quad (44)$$

where

$$J_{P,L} := \max\{1\} \cup \left\{ 2 \leq j \leq L \left| \frac{\sum_{i=J_{P,L}}^{L} P^\downarrow(i)}{L+1-j} < P^\downarrow(j-1) \right. \right\}. \quad (48)$$

Then, the following equation holds

$$F^M(P \rightarrow U_L) = F(\mathcal{C}_L(P), U_L)$$

$$= \sqrt{\frac{1}{L} \sum_{j=1}^{J_{P,L}-1} \sqrt{P^\downarrow(j)} + \sqrt{(L+1-J_{P,L}) \sum_{i=J_{P,L}}^{L} P^\downarrow(i)}} \quad (45)$$

**Proof:** It is enough to show that probability distribution in (44) satisfies the following equation:

$$F^M(P \rightarrow U_L) = F(\mathcal{C}_L(P), U_L). \quad (46)$$

We take $Q = (Q(1), \ldots, Q(L))$ as a probability distribution which satisfies $P \prec Q$, $Q(i) \geq Q(i+1)$ and

$$F^M(P \rightarrow U_L) = F(Q, U_L). \quad (47)$$

We assume that $Q \neq \mathcal{C}_L(P)$ and derive contradiction in the following. Since $Q \neq \mathcal{C}_L(P)$, the set $\{1 \leq k \leq L | Q(k) > \mathcal{C}_L(P)(k)\}$ is not empty. Let

$$k_0 := \min\{1 \leq k \leq L | Q(k) > \mathcal{C}_L(P)(k)\}. \quad (48)$$

When $k_0 > J_{P,L}$, $Q(k) = \mathcal{C}_L(P)(k)$ holds for $k = 1, \ldots, J_{P,L}$. Since $\mathcal{C}_L(P)$ is uniform on $\{J_{P,L} + 1, \ldots, L\}$, the following holds by the Schwartz inequality:

$$F(\mathcal{C}_L(P), U_L) > F(Q, U_L). \quad (49)$$
On the other hand, since $P \prec C_L(P)$, the following holds:

$$F(C_L(P), U_L) \leq F^M(P \rightarrow U_L) = F(Q, U_L). \quad (50)$$

The inequalities (49) and (50) are contradictory to each other. When $k_0 \leq J_{P, L}$, let

$$l_0 := \max\{1 \leq k \leq L | Q(k_0) = Q(k)\}. \quad (51)$$

Then, $Q(l_0) > Q(l_0 + 1)$ and $Q(l_0) > C_L(P)(l_0)$ hold since $Q(l_0) - C_L(P)(l_0) \geq Q(k_0) - C_L(P)(k_0) > 0$. For $\epsilon := \min\{\frac{1}{2}(Q(l_0) - Q(l_0 + 1)), Q(l_0) - C_L(P)(l_0)\} > 0$, we define a probability distribution as

$$Q'(k) = \begin{cases} 
Q(l_0) - \epsilon & \text{if } k = l_0 \\
Q(l_0 + 1) + \epsilon & \text{if } k = l_0 + 1 \\
Q(k) & \text{otherwise.}
\end{cases} \quad (52)$$

Then, note that $Q'(k) = C_L(P)(k) = P(k)$ for $k < k_0$ by the definition of $k_0$ and $C_L(P)$. Thus, $Q'$ satisfies $P \prec Q' \prec Q$, $Q'(i) \geq Q'(i + 1)$ and

$$F(Q', U_L) > F(Q, U_L) = F^M(P \rightarrow U_L) \quad (53)$$

by the Schwartz inequality. On the other hand, since $P \prec Q'$, the following holds:

$$F(Q', U_L) \leq F^M(P \rightarrow U_L). \quad (54)$$

The inequalities (53) and (54) are contradictory to each other.

4 Conversions for Probability Distributions: Asymptotic Setting

We will derive the asymptotic expansion formulas for $L_n^D(P, Q|\nu)$ and $L_n^M(P, Q|\nu)$. Those asymptotic expansions are also called the second-order asymptotic expansions in information theory because not only the first-order term $n$ but also the second largest order term $\sqrt{n}$ is treated. Then, the coefficients of $n$ and $\sqrt{n}$ are called the first-order rate and the second-order rate, respectively.

To begin with, we note that the first-order asymptotics of $L_n^D(P, Q|\nu)$ and $L_n^M(P, Q|\nu)$ are potentially done as follows:

$$\lim_{n \to \infty} \frac{L_n^D(P, Q|\nu)}{n} = \lim_{n \to \infty} \frac{L_n^M(P, Q|\nu)}{n} = \frac{H(P)}{H(Q)}. \quad (55)$$

where $H(P)$ is the entropy of $P$, i.e. $H(P) := -\sum P(x) \log_2 P(x)$. The equation between LHS and RHS is obtained from the results about the intrinsic randomness and the resolvability in classical information theory [15]. Similarly, the equation between the middle side and the right side is obtained from the results about entanglement concentration and dilution in quantum information theory [4].
4.1 Asymptotic Expansion Formula

When we introduce constants as
\[ V(P) := \sum_x P(x)(-\log_2 P(x) - H(P))^2, \]
\[ D_{P,Q} := \frac{H(Q)}{\sqrt{V(P)}}, \]
\[ C_{P,Q} := \frac{H(P)}{V(P)} \left( \frac{H(Q)}{V(Q)} \right)^{-1}, \]
the following theorem gives asymptotic expansion of two maximum convertible number with accuracy constraint. Then, although the direct calculation of \( L_n(P,Q|\nu) \) from its definition is very hard when \( n \) is large, our asymptotic expansion formula enables us to highly accurately evaluate it.

**Theorem 10** Let \( P \) and \( Q \) be arbitrary probability distributions on finite sets. Then, the following asymptotic expansion holds for any \( \nu \in (0,1) \)
\[ L_n^D(P,Q|\nu) \approx L_n^{M}(P,Q|\nu) \approx \frac{H(P)}{H(Q)} n + \frac{Z_{C_{P,Q}}^{-1}(1-\nu^2)}{D_{P,Q}} \sqrt{n}, \]
where \( \approx \) shows that the difference between the left and the right side terms is \( o(\sqrt{n}) \) at most.

The graphs of the second-order rates of \( L_n^D(P,Q|\nu) \) are described as in Fig. 2 with respect to the value of \( C_{P,Q} \).

Here, for non-uniform distributions \( P \) and \( Q \), the second-order rate in (59) admits another form
\[ \frac{Z_{C_{P,Q}}^{-1}(1-\nu^2)}{D_{P,Q}} = \left( \frac{H(P)}{H(Q)} \right)^{\frac{1}{2}} \frac{Z_{C_{Q,P}}^{-1}(1-\nu^2)}{D_{Q,P}} \]
from symmetry of Rayleigh-normal distributions represented by Lemma 5 and the equation \( C_{P,Q}^{-1} = C_{Q,P} \).

We give concrete forms of (59) for specific cases. Here, let \( U_m \) be the uniform distribution on the set \( \{1,\ldots,d\} \). When \( P \) is the uniform distribution \( U_m \), the right hand side of (60) can be defined for \( P = U_m \) although \( V(P) = 0 \) and thus \( D_{U_m,P} \) and \( C_{U_m,P} \) are not defined. In this sense, Theorem 10 holds for \( P = U_m \) and we have
\[ L_n^D(U_m,Q|\nu) \approx L_n^{M}(U_m,Q|\nu) \approx \frac{\log m}{H(Q)} n + \frac{\sqrt{V(Q)\log m}}{H(Q)^{\frac{1}{2}}} \Phi^{-1}(1-\nu^2)\sqrt{n}. \]

When \( Q \) is the uniform distribution \( U_m \), \( D_{P,U_m} = \frac{\log m}{\sqrt{V(P)}} \) and \( C_{P,U_m} = 0 \). Since we defined \( Z_0 \) as the cumulative distribution function \( \Phi \) of the standard normal distribution in Section 2, we have
\[ L_n^D(P,U_m|\nu) \approx L_n^{M}(P,U_m|\nu) \approx \frac{H(P)}{\log m} n + \frac{\sqrt{V(P)\log m}}{\log m} \Phi^{-1}(1-\nu^2)\sqrt{n}. \]
Figure 2: The second-order rates in (60) of \( L^D_n(P, Q|\nu) \) behaves as above with respect to accuracy \( \nu \in (0, 1) \). The black, purple, green, blue and red lines correspond to the cases when \( C_{P, Q} = 0, 1/10, 1/6, 1/3 \) and 1 under \( D_{P, Q} = 1 \). The case \( C_{P, Q} > 1 \) can be transformed to that with parameter \( C_{P, Q} < 1 \) by Lemma 5. Only when \( C_{P, Q} = 1 \), the second-order rate is always non-negative and goes to 0 when \( \nu \) tends to 1. On the other hand, when \( C_{P, Q} \neq 1 \), the second-order rate goes to \(-\infty\) when \( \nu \) tends to 1.

When \( Q \) is equal to \( P \), since \( Z_1 \) is the cumulative distribution function \( R_{\sqrt{\tau}} \) of the Rayleigh distribution from (65) and \( R_{\sqrt{\tau}}^{-1}(1 - \nu^2) = \sqrt{8 \log \nu^{-1}} \), we have

\[
L^D_n(P, P|\nu) \approx L^M_n(P, P|\nu) \approx n + \frac{\sqrt{8V(P) \log \nu^{-1}}}{H(P)} \sqrt{n}. \tag{63}
\]

In the remaining of this section, we give the proof of Theorem 10. To do so, we define as

\[
F^i_{P, Q}(b) := \lim_{n \to \infty} F^i \left( P^n \to Q^\frac{H(P)}{n} \nu^{-1} + b \sqrt{n} \right) \tag{64}
\]

for \( i = D \) or \( M \) and show the following lemma.

**Lemma 11** Let \( P \) and \( Q \) be arbitrary probability distributions on finite sets. Then, the following holds for any \( b \in \mathbb{R} \)

\[
F^D_{P, Q}(b) = F^M_{P, Q}(b) = \sqrt{1 - Z_{C_{P, Q}}(bD_{P, Q})}. \tag{65}
\]
For non-uniform distributions $P$ and $Q$, we obtain

$$Z_{C,P,Q}(bD_{P,Q}) = Z_{C,Q,P}(bD_{Q,P})$$

from Lemma. Then the right hand side of (65) is defined for $P = U_m$.

Theorem is easily obtained from Lemma as follows. For arbitrary $\nu \in (0, 1)$ and $\epsilon > 0$, we have

$$F_{i}^{-1}(\nu) = \frac{Z_{C,P,Q}^{-1}(1 - \nu^2)}{D_{P,Q}}$$

from Lemma. Since

$$F^i \left( P^n \to Q^{\frac{H(P)}{\alpha(n)}} + F_{i}^{-1}(\nu - \epsilon) \sqrt{n} \right) > \nu$$

holds for $i = D$ and $M$, $L_i^e(P, Q|\nu)$ is greater than or equal to the right side in (69). Similarly,

$$F^i \left( P^n \to Q^{\frac{H(P)}{\alpha(n)} + (F_{i}^{-1}(\nu) + \epsilon) \sqrt{n}} \right) < \nu$$

holds. Thus, $L_i^e(P, Q|\nu)$ for $i = D$ and $M$ is less than or equal to RHS in (69). Therefore, Theorem is obtained.

From the above discussion, all we have to do is to show Lemma. Since (36) implies that

$$F_{P,Q}^D(b) \leq F_{P,Q}^M(b),$$

it is enough to show the following two inequalities to obtain (65):

$$F_{P,Q}^D(b) \geq \Phi \left( -\sqrt{\frac{H(Q)^3}{V(Q)\log m}} \right),$$

$$F_{P,Q}^M(b) \leq \Phi \left( -\sqrt{\frac{H(Q)^3}{V(Q)\log m}} \right).$$

We separately prove the inequalities and for the uniform case (i.e. $P$ or $Q$ is uniform) and the non-uniform case (i.e. both $P$ and $Q$ are non-uniform).

### 4.2 Uniform Distribution Case

In this subsection, we prove and for the uniform case.

#### 4.2.1 Source Distribution $P$ is Uniform

Let $P$ and $Q$ be the uniform distribution $U_m$ and a non-uniform probability distribution on a finite set, respectively. We note that

$$\sqrt{1 - Z_{C,U_m,Q}(bD_{U_m,Q})} = \Phi \left( -\sqrt{\frac{H(Q)^3}{V(Q)\log m}} b \right).$$
For the above value, we prove (71) and (72). Let \( S_n^P(x) := \{1, 2, ..., \lceil 2^{H(P)n + x\sqrt{\pi}} \rceil \} \) and \( S_n^P(x, x') := S_n^P(x') \setminus S_n^P(x) \). Then we prepare the following lemma.

**Lemma 12** For an arbitrary distribution \( P \) and non-uniform distribution \( Q \) on finite sets,

\[
\lim_{n \to \infty} Q^{n+1}(S_n^Q(x)) = \Phi \left( \frac{x}{\sqrt{\max(Q)}} \right), \quad (74)
\]

\[
\lim_{n \to \infty} Q^{H(P)n+b\sqrt{\pi}+1}(S_n^P(x)) = \Phi \left( \frac{\sqrt{H(Q)}}{\sqrt{H(P)\max(Q)}}(x-bH(Q)) \right), \quad (75)
\]

In particular, when both \( P \) and \( Q \) are non-uniform distributions,

\[
\lim_{n \to \infty} Q^{H(P)n+b\sqrt{\pi}}(S_n^P(x)) = \Phi_{P,Q,b} \left( \frac{x}{\sqrt{\min(P)}} \right), \quad (76)
\]

where \( \Phi_{P,Q,b} := \Phi_{bD,P,Q,C_P,Q} \).

**Proof:** First, we show (74). Let \( \tilde{S}_n^P(x) := \{i \in \mathbb{N} | P^{n+1}(i) \geq 2^{-H(P)n - x\sqrt{\pi}} \} \) and \( \tilde{S}_n^P(x, x') := \tilde{S}_n(x') \setminus \tilde{S}_n(x) \). Then, the followings are obtained by the central limit theorem:

\[
\lim_{n \to \infty} P^{n+1}(\tilde{S}_n^P(x)) = \Phi \left( \frac{x}{\sqrt{\min(P)}} \right). \quad (77)
\]

Next, we will show that the following holds for an arbitrary \( \delta > 0 \):

\[
\tilde{S}_n^P(x) \subset S_n^P(x) \subset S_n^P(x + \delta) \quad (78)
\]

when \( n \in \mathbb{N} \) is large enough. Since \( \tilde{S}_n^P(x) \subset S_n^P(x) \) is obviously holds for any \( n \in \mathbb{N} \), it is enough to show \( S_n^P(x) \subset S_n^P(x + \delta) \).

We note that \( P^{n+1}(i) - 2^{-H(P)n-(x+\frac{\delta}{2})\sqrt{\pi}} \geq 0 \) holds if and only if \( i \in \tilde{S}_n(x + \frac{\delta}{2}) \). Thus, we have the following inequality for an arbitrary subset \( S \subset \mathbb{N} \):

\[
P^{n+1} \left( \tilde{S}_n \left( x + \frac{\delta}{2} \right) \right) - 2^{-H(P)n-(x+\frac{\delta}{2})\sqrt{\pi}} \tilde{S}_n \left( x + \frac{\delta}{2} \right) \geq P^{n+1}(S) - 2^{-H(P)n-(x+\frac{\delta}{2})\sqrt{\pi}} |S|. \quad (79)
\]

In particular, when \( S = \tilde{S}_n^P(x + \delta) \), we obtain

\[
|\tilde{S}_n^P \left( x + \frac{\delta}{2}, x + \delta \right) | \geq 2^{H(P)n+(x+\frac{\delta}{2})\sqrt{\pi}} P^{n+1} \left( \tilde{S}_n^P \left( x + \frac{\delta}{2}, x + \delta \right) \right) \quad (80)
\]

from (79). Since

\[
\lim_{n \to \infty} P^{n+1} \left( S_n^P \left( x + \frac{\delta}{2}, x + \delta \right) \right) = \Phi \left( \frac{x + \delta}{\sqrt{\min(P)}} \right) - \Phi \left( \frac{x + \frac{\delta}{2}}{\sqrt{\min(P)}} \right) > 0 \quad (81)
\]
from (77), the following holds for large enough \( n \in \mathbb{N} \):
\[
2^{H(P)n+(x+\frac{\delta}{2})\sqrt{\pi}n^\frac{3}{2}}} \left( S_n^P \left( x + \frac{\delta}{2}, x + \delta \right) \right) \geq 2^{H(P)n+x\sqrt{\pi}}. \tag{82}
\]
Therefore, the inequality
\[
|\hat{S}_n^P(x+\delta)| \geq |S_n^P \left( x + \frac{\delta}{2}, x + \delta \right) | \geq 2^{H(P)n+x\sqrt{\pi}} = |S_n^P(x)| \tag{83}
\]
holds from (80) and (82) for large enough \( n \in \mathbb{N} \), and implies (78). Thus (77) and (78) imply (74).

Then, we show (75). For an arbitrarily small \( \epsilon > 0 \), the following holds for large enough \( n \)
\[
S_n^Q \left( \frac{x - bH(Q)}{\frac{H(P)}{H(Q)}}, \frac{b}{\sqrt{H(Q)}} + \epsilon \right) \subset S_n^P(\epsilon) \subset S_n^Q \left( \frac{x - bH(Q)}{\frac{H(P)}{H(Q)}}, \frac{b}{\sqrt{H(P)}} + \epsilon \right). \tag{84}
\]
Thus, we obtain (75) as follows
\[
\lim_{n \to \infty} Q_{\frac{H(P)}{H(Q)} n + b\sqrt{n}} \left( S_n^P(x) \right) = \lim_{n \to \infty} Q_{\frac{H(P)}{H(Q)} n + b\sqrt{n}} \left( S_n^Q \left( \frac{x - bH(Q)}{\frac{H(P)}{H(Q)}}, \frac{b}{\sqrt{H(P)}} + \epsilon \right) \right) = \Phi \left( \frac{x - bH(Q)}{\frac{H(P)}{H(Q)}}, \frac{b}{\sqrt{H(P)}} \right) \tag{85}
\]
\[
= \Phi_{P,Q,b} \left( \frac{x}{\sqrt{H(P)}} \right), \tag{86}
\]
where the second equation is obtained from (74).

First, we prove (74). For an arbitrary \( \gamma > 0 \), it is easily verified that there exists a map \( W_n : \mathbb{N} \to \mathbb{N} \) such that
\[
Q_{\frac{H(P)}{H(Q)} n + b\sqrt{n}}(j) \leq W_n(U_m^n)(j) + 2^{-(\log m)n} \tag{87}
\]
for \( j \in S_n^{U_m}(-\gamma) \). Then

\[
F_D(U_n \to Q^{\log_m n + b\sqrt{n}}) \geq F_D(W_n(U_n), Q^{\log_m n + b\sqrt{n}})
\]

\[
= \sum_{j \in S_n^{U_m}(-\gamma)} \sqrt{W_n(U_n)(j)} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)}
\]

\[
\geq \sum_{j \in S_n^{U_m}(-\gamma)} \sqrt{\max\{Q^{\log_m n + b\sqrt{n}}(j) - 2 - (\log m)n, 0\}} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)}
\]

\[
\geq \sum_{j \in S_n^{U_m}(-\gamma)} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)} - \sum_{j \in S_n^{U_m}(-\gamma)} \sqrt{2 - (\log m)n} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)}
\]

\[
\geq Q^{\log_m n + b\sqrt{n}}(S_n^{U_m}(-\gamma)) - \sqrt{2 - (\log m)n} \sqrt{|S_n^{U_m}(-\gamma)|} \sqrt{Q^{\log_m n + b\sqrt{n}}(j)}
\]

and the second term goes to 0 as \( n \) tends to \( \infty \). Thus,

\[
\lim_{n \to \infty} \inf F_D(U_n \to Q^{\log_m n + b\sqrt{n}}) \geq \lim_{\gamma \to 0} \lim_{n \to \infty} \inf Q^{\log_m n + b\sqrt{n}}(S_n^{U_m}(-\gamma)) = \Phi\left(-\sqrt{\frac{H(Q)^3}{V(Q)\log m}}\right),
\]

where we used (75) in the above final equation.

Next, we prove (72).

\[
\lim_{n \to \infty} F_M(U_n \to Q^{\log_m n + b\sqrt{n}})^2 = \lim_{n \to \infty} F_M(U_2^{\log_m n} \to Q^{\log_m n + b\sqrt{n}})^2
\]

\[
= \lim_{k \to \infty} F_M(U_2^{H(Q)k - \sqrt{H(Q)\log m} b\sqrt{k}} \to Q^k)^2
\]

\[
= \lim_{k \to \infty} Q^{k\downarrow} \left(S_n^{Q} \left(-\sqrt{\frac{H(Q)^3}{\log m} b}\right)\right) \quad (88)
\]

\[
= \Phi\left(-\sqrt{\frac{H(Q)^3}{V(Q)\log m}}\right), \quad (89)
\]

where we used (74) in the above final equation.
4.2.2 Target Distribution $Q$ is Uniform

Let $P$ and $Q$ be a non-uniform probability distribution on a finite set and the uniform distribution $U_m$, respectively. We note that

$$\sqrt{1 - Z_{C,P,U_m}(bD_{P,U_m})} = \sqrt{\Phi\left(-\frac{b \log m}{\sqrt{V(P)}}\right)}.$$  \hfill (90)

For the above value, we prove (71) and (72). To do so, we prepare the following lemma.

**Lemma 13** Let $P$ be a non-uniform distribution and $A$ be a continuous differentiable monotone increasing function satisfying $\Phi \leq A \leq 1$. Then, we set a function $y(x) : \mathbb{R} \rightarrow \mathbb{R}$ and $\alpha_n(x)$ as

$$y(x) := \sqrt{V(P)}\Phi^{-1}\left(A\left(\frac{x}{\sqrt{V(P)}}\right)\right),$$  \hfill (91)

$$\alpha_n(x) := \min_{k \in S_n^p(y(x))} P^n\downarrow(k) = P^n\downarrow(\lceil 2^{H(P)n + y(x)} \sqrt{n} \rceil),$$  \hfill (92)

we have the following for $p, \epsilon > 0$

$$\alpha_n(x + \epsilon)|S_n^P(x - p, x)| \leq 2^{-\epsilon \sqrt{n}}.$$  \hfill (93)

**Proof:** The definition of $\alpha_n$ implies that

$$\alpha_n(x + \epsilon) \leq 2^{-(nH(P) + \sqrt{n}y(x+\epsilon))}.$$  \hfill (94)

Since $|S_n^P(x - p, x)| \leq 2^{nH(P) + \sqrt{n}x}$, we obtain

$$\alpha_n(x + \epsilon)|S_n^P(x - p, x)| \leq 2^{-\epsilon \sqrt{n}(y(x+\epsilon)-x)}.$$  \hfill (95)

Since $A \geq \Phi$, we have $x \leq y(x)$. Hence, we have $2^{-\epsilon \sqrt{n}(y(x+\epsilon)-x)} \leq 2^{-\epsilon \sqrt{n}(x+\epsilon)-x} = 2^{-\epsilon \sqrt{n}}$, which implies (93).
First, we prove (71). When $A = \Phi$, we have $y(x) = x$ in (91). Then

$$F_D(P^n \to U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}})$$

$$\geq \sum_{j \in S_p^n(x-p,x)} \sqrt{W_n,\ell(P^n,\ell)(j)} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

$$\geq \sum_{j \in S_p^n(x-p,x)} \sqrt{\max \{P_n^*(x,p,\varepsilon)(j) - \alpha_n(x+\varepsilon,0) \}} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

$$\geq \sum_{j \in S_p^n(x-p,x)} \sqrt{P_n^*(x,p,\varepsilon)(j)} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

$$- \sum_{j \in S_p^n(x-p,x)} \sqrt{\alpha_n(x+\varepsilon)} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

$$= \sqrt{P^n(S_p^n(x+\varepsilon,x+\varepsilon+p)) \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}(S_p^n(x-p,x))}$$

$$- \sum_{j \in S_p^n(x-p,x)} \sqrt{\alpha_n(x+\varepsilon)} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

Using the Schwartz inequality, the second term of (97) can be evaluated as

$$\sum_{j \in S_p^n(x-p,x)} \sqrt{\alpha_n(x+\varepsilon)} \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}}$$

$$\leq \sqrt{\alpha_n(x+\varepsilon)} \sqrt{|S_p^n(x-p,x)| \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}(S_p^n(x-p,x))}}$$

$$\leq 2^{-\varepsilon \sqrt{n}} \to 0.$$  

(98)

When we set $x$ to $b \log m$, the following holds:

$$\liminf_{n \to \infty} F_D(P^n \to U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}})$$

$$\geq \liminf_{n \to \infty} \sqrt{P^n(S_p^n(b+\varepsilon,b+\varepsilon+p)) \sqrt{U_{\frac{H(P)}{m} \sqrt{n + b \sqrt{n}}}(S_p^n(b-p,b))}}$$

$$= \sqrt{\Phi \left( \frac{b \log m + \varepsilon + p}{\sqrt{V(P)}} \right) - \Phi \left( \frac{b \log m + \varepsilon}{\sqrt{V(P)}} \right)}$$

$$\to 0$$

(100)

$$p \to \infty$$

$$\varepsilon \to 0$$

$$1 - \Phi \left( \frac{b \log m}{\sqrt{V(P)}} \right)$$

(101)
where (100) follows from Lemma 12.

Next, we prove (72). By Lemmas 4 and 5 in [19], for an arbitrary positive integers $L_n \geq L'_n$, the following inequality holds:

$$F^M(P^n \to U^L_n)^2 \leq \frac{1}{2L_n} \left( \sqrt{\left| \{ P^n \geq 1/2L'_n \} \right|} \cdot \sqrt{P^n \cdot \left( P^n \geq 1/2L'_n \right)} \right) + \sqrt{2L_n - \left| \{ P^n \geq 1/2L'_n \} \right|} \cdot \sqrt{1 - P^n \cdot \left( P^n \geq 1/2L'_n \right)}^2.$$  

When $L_n = an + b\sqrt{n}$ and $L'_n = an + (b - \lambda)\sqrt{n}$, LHS in (102) goes to $F^M_{P,U_m}(b)$.

Moreover, RHS in (102) goes to $\sqrt{\Phi \left( \frac{-b\log m}{\sqrt{V(P)}} \right)}$ taking $\lim_{n \to \infty}$ and $\lim_{\lambda \to +0}$.

Thus we obtain (72).

**Remark 14** For probability distributions $P$ and $Q$ on finite sets $X$ and $Y$, the approximate conversion problem from i.i.d. of $P$ to that of $Q$ has been treated as the random number generation. In particular, when $P$ or $Q$ is a uniform distribution, the problems have been well-known as the resolvability problem and the intrinsic randomness respectively [15], and analyzed not only in the context of the first-order asymptotic theory but also that in the second asymptotic theory [30, 20]. Hayashi [20] treated the intrinsic randomness and Nomura and Han [30] treated the resolvability besides the intrinsic randomness in the framework of the second-order asymptotics. Since they do not adopted the Hellinger distance which has a one-to-one correspondence with the fidelity but the total variation distance as the error function, the formulation in this paper and theirs does not completely coincide with each other. On the other hand, Tomamichel and Hayashi [34] consider randomness extraction against quantum side information in the second-order asymptotics and adopted the fidelity to measure accuracy of the operations. Since the intrinsic randomness in this paper is regarded as randomness extraction without quantum side information in [34, 62] can be directly obtained from Lemma 16 in [34].

### 4.3 Non-Uniform Distribution Case

In this subsection, we prove (71) and (72) for the non-uniform case. We define the normal distribution determined by distributions $P$, $Q$ and a real number $b$ as $N_{P,Q,b} := N_{bD_{P,Q},C_{P,Q}}$. Then, note that

$$\sqrt{1 - Z_{C_{P,Q}}(bD_{P,Q})} = \sup_A \mathcal{F} \left( \frac{dA}{dx}, N_{P,Q,b} \right)$$  

by the definition, where supremum is taken over the functions satisfying the conditions in Definition 4.
4.3.1 Direct Part

In this subsection, we prove (71) for the non-uniform case. To prove (71), it is enough to show

\[ F_{P,Q}^{D}(b) \geq F \left( \frac{dA}{dx}, N_{P,Q,b} \right) - \epsilon. \]  

(104)

for an arbitrary continuous differentiable monotone increasing function \( A \) satisfying \( \Phi \leq A \leq 1 \) and an arbitrary \( \epsilon > 0 \).

First, for an arbitrary \( \epsilon > 0 \), we choose \( \lambda > 0 \) which satisfies

\[ \int_{(-\infty,-\lambda)\cup(\lambda,\infty)} \sqrt{\frac{dA}{dx}(x)} \sqrt{N_{P,Q,b}(x)} \, dx \leq \epsilon. \]  

(105)

For \( I \in \mathbb{N}, 0 \leq i \leq I \), and \( \lambda > -\lambda \), we set sequences as

\[ x_{I}^{i} := \sqrt{V(P)} \left(-\lambda + \frac{2\lambda}{I} i\right), \quad y_{I}^{i} := y(x_{I}^{i}). \]  

(106)

Here we introduce a probability distribution \( P'_{n,I} \). For any \( j \in \bigcup_{i=1}^{I} S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i}) \), we note that there uniquely exists \( i \) such that \( j \in S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i}) \). Then we define \( P'_{n,I} \) as follows

\[ P'_{n,I}(j) = \frac{P_{n}^{\lambda}(S_{n}^{P}(y_{I+1}^{j}, y_{I+2}^{j}))}{Q^{H_{n}(P)}_{n} + b\sqrt{n}_{i}} \frac{Q^{H_{n}(P)}_{n} + b\sqrt{n}_{i}}{Q^{H_{n}(P)}_{n} + b\sqrt{n}_{i}}(j) \]  

(107)

for \( j \in S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i}) \), and \( P'_{n,I}(j) \) is arbitrary for \( j \in \mathbb{N} \setminus \bigcup_{i=1}^{I} S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i}) \) as long as \( P'_{n,I} \) is a probability distribution.

When we set \( \alpha_{n}^{i} \) as \( \alpha_{n}(x_{I+i+1}) \) defined in (92), because

\[ \sum_{j \in S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i})} P'_{n,I}(j) = \sum_{k \in S_{n}^{P}(y_{I+1}^{i}, y_{I+2}^{i})} P_{n}^{\lambda}(k), \]  

(108)

there exists a deterministic map \( W_{n,I} : \mathbb{N} \to \mathbb{N} \) for large enough \( n \in \mathbb{N} \) such that

\[ W_{n}(S_{n}^{P}(y_{I+1}^{j}, y_{I+2}^{j})) = S_{n}^{P}(x_{I-1}^{i}, x_{I}^{i}), \]  

\[ P'_{n,I}(j) \leq W_{n}(P_{n}^{\lambda})(j) + \alpha_{n}^{i}. \]  

(109)

(110)
for any $i = 1, \ldots, I - 2$ and $j \in S_n^P(x_{i-1}, x_i')$. Then, the following holds:

$$ F^P(P^i \rightarrow Q^{H(P)n + b\sqrt{n}}) \geq F(W_n(P^i), Q^{H(P)n + b\sqrt{n}I}) $$

$$ \geq \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{W_n(P^i(j))} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ \geq \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{\max(P^i_{n,j}(j) - a^I_{n,i}, 0)} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ \geq \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{P^i_{n,j}(j)} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ - \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{a^I_{n,i}} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)}, \quad (111) $$

where the final inequality follows from $\sqrt{x - y} \geq \sqrt{x} - \sqrt{y}$ for any $x \geq y \geq 0$. Using the definition (107) of $P^i_{n,j}(j)$, the first term of (111) can be calculated as

$$ \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{P^i_{n,j}(j)} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ = \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{\frac{P^i_{n,j}(S_n^P(y_{i+1}, y_{i+2}))}{Q^{H(P)n + b\sqrt{n}I}(S_n^P(x_{i-1}, x_i'))}} Q^{H(P)n + b\sqrt{n}I}(j) $$

$$ = \sum_{i=1}^{I-2} \sqrt{\frac{P^i_{n,j}(S_n^P(y_{i+1}, y_{i+2}))}{Q^{H(P)n + b\sqrt{n}I}(S_n^P(x_{i-1}, x_i'))}} Q^{H(P)n + b\sqrt{n}I}(j) $$

$$ = \sum_{i=1}^{I-2} \sqrt{P^i_{n,j}(S_n^P(y_{i+1}, y_{i+2}))} \sqrt{Q^{H(P)n + b\sqrt{n}I}(S_n^P(x_{i-1}, x_i'))} \quad (112) $$

Using the Schwartz inequality, the second term of (111) can be evaluated as

$$ \sum_{i=1}^{I-2} \sum_{j \in S_n^P(x_{i-1}, x_i')} \sqrt{a^I_{n,i}} \sqrt{Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ \leq \sum_{i=1}^{I-2} \sqrt{a^I_{n,i}} \sqrt{\sum_{j \in S_n^P(x_{i-1}, x_i')} Q^{H(P)n + b\sqrt{n}I}(j)} $$

$$ \leq \sum_{i=1}^{I-2} 2^{-\sqrt{(2\lambda)V(P)n/2I}} = (I - 2)2^{-\sqrt{(2\lambda)V(P)n/2I}} \xrightarrow{n \to \infty} 0. \quad (113) $$
Then, the following holds:

\[ F_{P,Q}(b) = \liminf_{n \to \infty} F_D(P^n \to Q^H_n b) \] (114)

\[ \geq \liminf_{n \to \infty} \sum_{i=1}^{I-2} \sqrt{P^n_i(S^n_{P_i}(y_{i+1}^I, y_{i+2}^I))} \sqrt{Q^H_n b} \sqrt{N_{P,Q,b}(x_{i-1}^I, x_i^I)} \] (115)

\[ = \sum_{i=1}^{I-2} \left[ \Phi \left( \frac{y_{i+2}^I}{\sqrt{V(P)}} \right) - \Phi \left( \frac{y_{i+1}^I}{\sqrt{V(P)}} \right) \right] \Phi_{P,Q,b} \left( \frac{x_i^I}{\sqrt{V(P)}} \right) - \Phi_{P,Q,b} \left( \frac{x_{i-1}^I}{\sqrt{V(P)}} \right) \] (116)

\[ = \sum_{i=1}^{I-2} \left[ \Phi \left( \frac{y_{i+2}^I}{\sqrt{V(P)}} \right) - \Phi \left( \frac{y_{i+1}^I}{\sqrt{V(P)}} \right) \right] \Phi_{P,Q,b} \left( \frac{x_i^I}{\sqrt{V(P)}} \right) - \Phi_{P,Q,b} \left( \frac{x_{i-1}^I}{\sqrt{V(P)}} \right) \]

\[ \geq \liminf_{n \to \infty} \sum_{i=1}^{I-2} \sqrt{P^n_i(S^n_{P_i}(y_{i+1}^I, y_{i+2}^I))} \sqrt{Q^H_n b} \sqrt{N_{P,Q,b}(x_{i-1}^I, x_i^I)} \] (117)

\[ = \int_{-\lambda + \frac{2\lambda}{T}(I-1)} \sqrt{N_{P,Q,b}(x, x - \frac{2\lambda}{T})} \] (118)

where (115) follows from (111), (112) and (113), (116) follows from Lemma 12, and (118) follows from (105).

4.3.2 Converse Part

In this subsection, we prove (72) for the non-uniform case. We first show the following lemma.
Lemma 15 Let \( a = \{a_i\}_{i=0}^I \) and \( b = \{b_i\}_{i=0}^I \) be probability distributions and satisfy \( \frac{a_{i+1}}{b_{i+1}} > \frac{a_i}{b_i} \). When \( c = \{c_i\}_{i=0}^I \) is a probability distribution and satisfies
\[
\sum_{i=0}^k a_i \leq \sum_{i=0}^k c_i \quad (k = 0, 1, \ldots, I) \tag{119}
\]
the following holds:
\[
\sum_{i=0}^I \sqrt{a_i \sqrt{b_i}} \geq \sum_{i=0}^I \sqrt{c_i \sqrt{b_i}}. \tag{120}
\]
Moreover, the equation holds for \( c \) if and only if \( c = a \).

Proof: First, let \( D(a) \) be the set of probability distributions on \( \{0, 1, \ldots, I\} \) whose element \( c \) satisfies (119). Then, since the function \( f_b(c) := F(b, c) \) on \( D(a) \) defined by the fidelity \( F \) is continuous, there exists \( c_0 \in D(a) \) which attains the maximum of \( f_b \). For an arbitrary \( c \neq a \) in \( D(a) \), we will show that \( c \) does not attain the maximum of \( f_b \) in the following. Then, it implies that \( a \) is the unique point in \( D(a) \) which attains the maximum of \( f_b \).

Note that there exists \( J \in \{0, 1, \ldots, I\} \) such that \( a_J > c_J \) and \( a_{J+1} = c_{J+1}, \ldots, a_J = c_J \) since \( a \) and \( c \) are different probability distributions and satisfy (119). Therefore, there exist two numbers \( i < j \) in \( \{0, 1, \ldots, I\} \) such that \( a_i < c_i, \ a_{i+1} = c_{i+1}, \ldots, a_{j-1} = c_{j-1} \) and \( a_j > c_j \). Then we have \( \frac{a_i}{b_i} > \frac{a_j}{b_j} \) since \( \frac{a_{i+1}}{b_{i+1}} > \frac{a_i}{b_i} \).

Hence, for a small constant \( \epsilon > 0 \), the following holds:
\[
\sqrt{c_i \sqrt{b_i}} + \sqrt{c_j \sqrt{b_j}} < \sqrt{c_i - \epsilon \sqrt{b_i}} + \sqrt{c_j + \epsilon \sqrt{b_j}}. \tag{121}
\]

When we set \( c' \) as \( c'_i := c_i - \epsilon, c'_j := c_j + \epsilon \) and \( c'_k = c_k \) for \( k \neq i, j \), \( c' \) is in \( D(a) \) and \( f_b(c') > f_b(c) \) from (121). Therefore, \( c \) does not attain the maximum of \( f_b \), and only \( a \) attain the maximum.

Then we obtain the following statement.

Lemma 16 Assume that real numbers \( t \leq t' \) satisfy the following condition (\( * \)):

\( * \) There exist \( s \) and \( s' \) which satisfy the following three conditions:

(I) \( s \leq t \leq t' \leq s' \),

(II) \( \frac{\Phi(t)}{\Phi_{\mu,v}(t)} = \frac{N(s)}{N_{\mu,v}(s)}, \frac{1 - \Phi(t')}{1 - \Phi_{\mu,v}(t')} = \frac{N(s')}{N_{\mu,v}(s')}, \)

(III) \( \frac{N(x)}{N_{\mu,v}(x)} \) is strictly monotonically decreasing on the interval \( (s, s') \).

Then the following inequality holds
\[
F_{P,Q}^N(b) \leq \sqrt{\Phi(t)} \sqrt{\Phi_{P,Q,b}(t)} + \int_t^{t'} \sqrt{N(x)} \sqrt{N_{P,Q,b}(x)} dx + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{P,Q,b}(t')} \tag{122}
\]
Proof: It is enough to show that
\[
\limsup_{n \to \infty} F(P_{n}^{t}, Q_{n}^{H(P)_{n}+b_{n}/\sqrt{n}}) \\
\leq \sqrt{\Phi(t)} \sqrt{\Phi_{P,Q,b}(t)} + \int_{t}^{t'} \sqrt{N(x)} \sqrt{N_{P,Q,b}(x)} dx \\
+ \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{P,Q,b}(t')}.
\]
(123)

for an arbitrary sequence \(\{P_{n}^{t}\}_{n=1}^{\infty}\) of probability distributions such that \(P_{n}^{t} \to P_{n}\). For simplicity, we denote \(\sqrt{V(P_{n}^{t})x} \to \bar{x}\) for an arbitrary real number \(x\).

When we set a sequence \(\{x_{i}^{t}\}_{i=0}^{l}\) for \(I \in \mathbb{N}\) as \(x_{i}^{t} := t + \frac{t-i}{l}i\), we have the following by the monotonicity of the fidelity [29]:
\[
\begin{align*}
F(P_{n}^{t}, Q_{n}^{H(P)_{n}+b_{n}/\sqrt{n}}) \\
\leq \sqrt{P_{n}^{t}(S_{n}^{P_{n}}(\bar{x}_{0}^{t}))} \sqrt{Q_{n}^{H(P)_{n}+b_{n}/\sqrt{n}}(S_{n}^{P_{n}}(\bar{x}_{0}^{t}))} \\
+ \sum_{i=1}^{l} \sqrt{P_{n}^{t}(S_{n}^{P_{n}}(\bar{x}_{i-1}^{t}, \bar{x}_{i}^{t}))} \sqrt{Q_{n}^{H(P)_{n}+b_{n}/\sqrt{n}}(S_{n}^{P_{n}}(\bar{x}_{i-1}^{t}, \bar{x}_{i}^{t}))} \\
+ \sqrt{1 - P_{n}^{t}(S_{n}^{P_{n}}(\bar{x}_{l}^{t}))} \sqrt{1 - Q_{n}^{H(P)_{n}+b_{n}/\sqrt{n}}(S_{n}^{P_{n}}(\bar{x}_{l}^{t}))}.
\end{align*}
\]
(124)

Here, we denote the right hand side of (124) by \(R_{t}(n)\). Then, we can choose a subsequence \(\{n_{i}\}_{i=1}^{n} \subset \{n\}\) such that \(\lim_{n \to \infty} R_{t}(n_{i}) = \limsup R_{t}(n)\) and the limits \(c_{0}^{t} := \lim_{l \to \infty} P_{n_{i}}^{t}(S_{n_{i}}^{P}(\bar{x}_{0}^{t}))\), \(c_{i}^{t} := \lim_{l \to \infty} P_{n_{i}}^{t}(S_{n_{i}}^{P}(\bar{x}_{i-1}^{t}, \bar{x}_{i}^{t}))\) and \(c_{l+1}^{t} := 1 - \lim_{l \to \infty} P_{n_{i}}^{t}(S_{n_{i}}^{P}(\bar{x}_{l}^{t}))\) exist for \(i = 1, \ldots, I\). Hence, we obtain
\[
\limsup_{n \to \infty} F(P_{n}^{t}, Q_{n}^{t}) \leq \limsup_{n \to \infty} R_{t}(n) = \lim_{l \to \infty} R_{t}(n_{i}) \\
= \sqrt{c_{0}^{t} \Phi_{P,Q,b}(x_{0})} \\
+ \sum_{i=1}^{l} \sqrt{c_{i}^{t} \Phi_{P,Q,b}(x_{i}^{t}) - \Phi_{P,Q,b}(x_{i-1}^{t})} \\
+ \sqrt{c_{l+1}^{t} \Phi_{P,Q,b}(x_{l}^{t})},
\]
(125)

where we used Lemma 12 in the last equality.

When we set as \(a_{0} := \Phi(x_{0}^{t}), a_{i} := \Phi(x_{i}^{t}) - \Phi(x_{i-1}^{t}), a_{i+1} = 1 - \Phi(x_{i}^{t}), b_{0} = \Phi_{P,Q,b}(x_{0}), b_{i} = \Phi_{P,Q,b}(x_{i}^{t}) - \Phi_{P,Q,b}(x_{i-1}^{t})\) and \(b_{l+1} = 1 - \Phi_{P,Q,b}(x_{l}^{t})\), those satisfy the assumptions of Lemma 13 as follows. First, \(a_{0}/b_{0} = N(s)/N_{P,Q,b}(s)\) and \(a_{i+1}/b_{i+1} = N(s)/N_{P,Q,b}(s)\) hold by the assumption (II). Moreover, there exist \(z_{i} \in [x_{i-1}^{t}, x_{i}^{t}]\) for \(i = 1, \ldots, I\) such that \(a_{i}/b_{i} = N(z_{i})/N_{P,Q,b}(z_{i})\) for \(i = 1, \ldots, I\) due to the mean value theorem. Then \(z_{i} \in (s, s')\) holds because of the relation \(t = x_{0}^{t} \leq x_{i-1}^{t} \leq z_{i} \leq x_{i}^{t} \leq x_{l}^{t} = t'\) and the assumption (I). Since
\( N(x)/N_{P,Q,b}(x) \) is monotonically decreasing on \((s, s')\) by the assumption (III), we have \( a_{i-1}/b_{i-1} \geq a_i/b_i \) for \( i = 1, ..., I + 1 \). Moreover,

\[
\sum_{i=0}^{k} a_{i} = \Phi(x_i^k) = \lim_{i \to \infty} P_{n_{i}}^{m_{i}}(S_{n_{i}}^{m_{i}}(x_i^k)) \leq \lim_{i \to \infty} P_{n_{i}}^{m_{i}}(S_{n_{i}}^{m_{i}}(x_i^k)) = \sum_{i=0}^{k} c_{i}^{k} \]

holds for \( k = 0, 1, ..., I \) since \( P_{n_{i}} \prec P_{n_{i}}^{m_{i}} \), and \( \sum_{i=0}^{I+1} a_{i} = 1 = \sum_{i=0}^{I+1} c_{i}^{k} \) holds.

From the above discussion, we can use Lemma 15. Therefore, the following hold:

\[
\limsup_{n \to \infty} F(P_{n}^{m_{i}}, Q_{n}^{m_{i}}) \leq \sqrt{c_{0}^{k} \Phi(x_{0}^{k})} + \sum_{i=1}^{l} \sqrt{c_{i}^{k} \Phi(x_{i}^{k}) - \Phi(x_{i}^{k-1})} \]

\[
+ \sqrt{c_{0}^{k} \Phi(x_{0}^{k})} \sqrt{1 - \Phi(x_{0}^{k})} \leq \sqrt{\Phi(t)} \sqrt{\Phi_{P,Q,b}(t)} \]

\[
+ \sum_{i=1}^{l} \sqrt{\Phi(x_{i}^{k}) - \Phi(x_{i-1}^{k})} \sqrt{\Phi_{P,Q,b}(x_{i}^{k})} - \Phi_{P,Q,b}(x_{i-1}^{k}) \]

\[
+ \sqrt{1 - \Phi(t') \sqrt{1 - \Phi_{P,Q,b}(t')}}. \tag{126} \]

where we used \( x_{0}^{k} = t \) and \( x_{i}^{k} = t' \). By taking limit \( I \to \infty \),

\[
\limsup_{n \to \infty} F(P_{n}^{m_{i}}, Q_{n}^{m_{i}}) \leq \sqrt{\Phi(x_{0}^{k}) \Phi_{P,Q,b}(x_{0}^{k})} \]

\[
+ \lim_{i \to \infty} \sum_{i=1}^{l} \sqrt{\Phi(x_{i}^{k}) - \Phi(x_{i-1}^{k})} \sqrt{\Phi_{P,Q,b}(x_{i}^{k})} - \Phi_{P,Q,b}(x_{i-1}^{k}) \]

\[
+ \sqrt{1 - \Phi(x_{0}^{k}) \sqrt{1 - \Phi_{P,Q,b}(x_{0}^{k})}} \]

\[
= \sqrt{\Phi(t)} \sqrt{\Phi_{P,Q,b}(t)} + \int_{t}^{t'} \sqrt{N(x)} \sqrt{N_{P,Q,b}(x)} dx \]

\[
+ \sqrt{1 - \Phi(t') \sqrt{1 - \Phi_{P,Q,b}(t')}}. \]

Here we introduce a function \( A_{\mu,v} : \mathbb{R} \to [0,1] \) with parameters \( \mu \in \mathbb{R} \) and \( v > 0 \) which is separately defined with respect to the value of \( v \) as follows. When \( v = 1 \),

\[
A_{\mu,1}(x) = A_{\mu,1}(x) := \begin{cases} 
\Phi_{\mu,1}(x) & \text{if } \mu < 0 \\
\Phi(x) & \text{if } \mu \geq 0.
\end{cases} \tag{127} \]
Figure 3: Let $v > 1$. The dashed, the normal and the thick lines show $\Phi$, $\Phi_{\mu,v}$ and $A_{\mu,v}$, respectively. Then, $A = A_{\mu,v}$ attains the supremum in (7).

When $v > 1$,

$$A_{\mu,v}(x) := \left\{ \begin{array}{ll}
\frac{\Phi(\alpha_{\mu,v})}{\Phi_{\mu,v}(\alpha_{\mu,v})}\Phi_{\mu,v}(x) & \text{if } x \leq \alpha_{\mu,v} \\
\Phi(x) & \text{if } \alpha_{\mu,v} \leq x.
\end{array} \right. \tag{128}$$

When $v < 1$,

$$A_{\mu,v}(x) := \left\{ \begin{array}{ll}
\Phi(x) & \text{if } x \leq \beta_{\mu,v} \\
1 - \frac{1 - \Phi(\beta_{\mu,v})}{1 - \Phi_{\mu,v}(\beta_{\mu,v})}(1 - \Phi_{\mu,v}(x)) & \text{if } \beta_{\mu,v} \leq x.
\end{array} \right. \tag{129}$$

The function $A_{\mu,v}$ is represented in Figs. 3 and 4. Then the following lemma is essential to prove (H2).

**Lemma 17** For an arbitrary $\epsilon > 0$, there exist real numbers $t \leq t'$ which satisfy the condition $(\ast)$ in Lemma [7] and the following inequality

$$\Phi(t) \sqrt{\Phi_{\mu,v}(t)} + \int_t^{t'} \sqrt{N(x)} \sqrt{N_{\mu,v}(x)} dx + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{\mu,v}(t')} \leq \mathcal{F} \left( \frac{dA_{\mu,v}}{dx}, N_{\mu,v} \right) + \epsilon \tag{130}$$

**Proof:** First, we treat the case when $v = 1$. By the Schwartz inequality, the left hand side of (130) is less than or equal to 1. When $\mu \geq 0$, since $A_{\mu,1} = \Phi_{\mu,1}$ satisfies $\mathcal{F} \left( \frac{dA_{\mu,1}}{dx}, N_{\mu,1} \right) = 1$, (130) holds. In the following, we treat the case when $\mu < 0$. For an arbitrary $\epsilon > 0$, we take a real number $\lambda > 0$ which satisfies

$$\sqrt{\Phi(-\lambda)} \sqrt{\Phi_{\mu,v}(-\lambda)} + \sqrt{1 - \Phi(\lambda)} \sqrt{1 - \Phi_{\mu,v}(\lambda)} < \epsilon. \tag{131}$$
Figure 4: Let $v < 1$. The dashed, the normal and the thick lines show $\Phi$, $\Phi_{\mu,v}$ and $A_{\mu,v}$, respectively. Then, $A = A_{\mu,v}$ attains the supremum in [7].

Here we set a sequence $\{x^i_I\}_{i=0}$ for $I \in \mathbb{N}$ as $x^i_I := -\lambda + \frac{2\lambda}{i}$. Then, we verify that $-\lambda$ and $\lambda$ satisfy the condition $(*)$ of Lemma 23. First, (III) in $(*)$ of Lemma 23 holds from Lemma 22. Since

$$\lim_{x \to -\infty} N_{\mu,v}(x) = \infty, \quad \lim_{x \to \infty} \frac{N(x)}{N_{\mu,v}(x)} = 0$$

(132)

and $\frac{N(x)}{N_{\mu,v}(x)}$ is strictly monotonically decreasing on $\mathbb{R}$ from Lemma 22, there uniquely exist $s$ and $s'$ which satisfy $s < s'$,

$$\frac{\Phi(-\lambda)}{\Phi_{\mu,v}(-\lambda)} = \frac{N(s)}{N_{\mu,v}(s)}, \quad \frac{1-\Phi(\lambda)}{1-\Phi_{\mu,v}(\lambda)} = \frac{N(s')}{N_{\mu,v}(s')}.$$  

(133)

Thus, we obtained (II) in $(*)$ of Lemma 23. Moreover, the inequalities $s \leq -\lambda < \lambda \leq s'$ hold as follows. To show $s < -\lambda$, it is enough to prove that

$$\frac{N(s)}{N_{\mu,v}(s)} \geq \frac{N(-\lambda)}{N_{\mu,v}(-\lambda)}$$

because $\frac{N(x)}{N_{\mu,v}(x)}$ is monotonically decreasing. We have

$$\frac{N(s)}{N_{\mu,v}(s)} = \frac{\Phi(-\lambda)}{\Phi_{\mu,v}(-\lambda)} = \frac{\Phi(-\lambda) - \Phi(w)}{\Phi_{\mu,v}(-\lambda) - \Phi_{\mu,v}(w)} = \frac{N(s(w))}{N_{\mu,v}(s(w))} \geq \frac{N(-\lambda)}{N_{\mu,v}(-\lambda)}$$

(134)

where the existence of $s(w) \in (w, -\lambda)$ in (134) is guaranteed by the mean value theorem and the last inequality follows from the monotone decrease of $\frac{N(x)}{N_{\mu,v}(x)}$.  

32
Thus, $s < -\lambda$ holds. Similarly, $\lambda < s'$ can be obtained. Therefore, we obtained (I) in (⋆) of Lemma 23. Then, we have the following inequality
\[
\sqrt{\Phi(-\lambda)}\sqrt{\Phi_{\mu,v}(-\lambda)} + \int_{-\lambda}^{\lambda} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \sqrt{1-\Phi(\lambda)}\sqrt{1-\Phi_{\mu,v}(\lambda)} = \int_{\mathbb{R}} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \epsilon
\] (136)
\[
\mathcal{F}\left(\frac{dA_{\mu,1}}{dx}, N_{\mu,v}\right) + \epsilon.
\] (137)
Thus, the proof is completed for the case of $v = 1$.

Next, we treat the case when $v > 1$. We take a constant $\lambda \in \mathbb{R}$ which satisfies $\alpha_{\mu,v} < \lambda$ and $\sqrt{1-\Phi(\lambda)}\sqrt{1-\Phi_{\mu,v}(\lambda)} < \epsilon$. We verify that $\alpha_{\mu,v}$ and $\lambda$ satisfy the condition (⋆) of Lemma 23 in the following. Then there exists $s'$ such that
\[
\frac{1-\Phi(\lambda)}{1-\Phi_{\mu,v}(\lambda)} = \frac{N(s')}{N_{\mu,v}(s')},
\]
and $\lambda \leq s'$ by the mean value theorem. Moreover, since $\alpha_{\mu,v}$ satisfies (12), $\alpha_{\mu,v}$ can be taken as $s = t$ in Lemma 23. Thus, the conditions (I) and (II) in (⋆) hold. Next, since $\frac{N(s)}{N_{\mu,v}(s)}$ is monotonically decreasing on $(\alpha_{\mu,v}, s')$ from Lemma 22 and Lemma 2 the condition (III) in (⋆) holds. Therefore, $\alpha_{\mu,v}$ and $\lambda$ satisfy (⋆). Then the following holds
\[
\sqrt{\Phi(\alpha_{\mu,v})}\sqrt{\Phi_{\mu,v}(\alpha_{\mu,v})} + \int_{\alpha_{\mu,v}}^{\lambda} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \sqrt{1-\Phi(\lambda)}\sqrt{1-\Phi_{\mu,v}(\lambda)} \leq \sqrt{\Phi(\alpha_{\mu,v})}\sqrt{\Phi_{\mu,v}(\alpha_{\mu,v})} + \int_{\alpha_{\mu,v}}^{\infty} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \epsilon,
\]
\[
= \mathcal{F}\left(\frac{dA_{\mu,v}}{dx}, N_{\mu,v}\right) + \epsilon.
\] (138)
Thus, the proof is completed for the case of $v > 1$.

Finally, we treat the case when $v < 1$. We take a constant $\lambda \in \mathbb{R}$ which satisfies $\lambda < \beta_{\mu,v}$ and $\sqrt{\Phi(\lambda)}\sqrt{\Phi_{\mu,v}(\lambda)} < \epsilon$. To use Lemma 23, we verify that $\lambda$ and $\beta_{\mu,v}$ satisfy the condition (⋆) of Lemma 23 in the following. First, there exists a real number $s$ such that
\[
\frac{\Phi(\lambda)}{\Phi_{\mu,v}(\lambda)} = \frac{N(s)}{N_{\mu,v}(s)},
\]
and $s \leq \lambda$ by the mean value theorem. Moreover, since $\beta_{\mu,v}$ satisfies (3), $\beta_{\mu,v}$ can be taken as $t' = s'$ in Lemma 23. Thus, the conditions (I) and (II) in (⋆) hold. Next, since $\frac{N(s)}{N_{\mu,v}(s)}$ is monotonically decreasing on $(s, \beta_{\mu,v})$ from Lemma 22 and Lemma 2 the condition (III) in (⋆) holds. Therefore, $\lambda$ and $\beta_{\mu,v}$ satisfy the condition (⋆). Then the following holds
\[
\sqrt{\Phi(\lambda)}\sqrt{\Phi_{\mu,v}(\lambda)} + \int_{\lambda}^{\beta_{\mu,v}} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \sqrt{1-\Phi(\lambda)}\sqrt{1-\Phi_{\mu,v}(\lambda)} = \int_{-\infty}^{\beta_{\mu,v}} \sqrt{N(x)}\sqrt{N_{\mu,v}(x)}dx + \sqrt{1-\Phi(\lambda)}\sqrt{1-\Phi_{\mu,v}(\lambda)} + \epsilon,
\]
\[
= \mathcal{F}\left(\frac{dA_{\mu,v}}{dx}, N_{\mu,v}\right) + \epsilon.
\] (139)
Thus, the proof is completed for the case of $v < 1$.

From Lemmas 16 and 17, we obtain

$$F_{P,Q}^{M}(b) \leq F \left( \frac{dA_{bD_{P,Q}}}{dx}, N_{P,Q,b} \right) \leq \sup_{A} F \left( \frac{dA}{dx}, N_{P,Q,b} \right). \quad (140)$$

Thus the proof of (72) is completed.

5 Application to Quantum Information Theory

In this section, we apply the second-order asymptotics to the approximate conversion between two bipartite entangled pure states by LOCC and the cloning for a known entangled pure state by LOCC.

5.1 LOCC Conversion

Entanglement is used in several quantum informational operations [23, 31, 8, 5, 12, 3, 11], and the conversion of entangled states by LOCC has been studied in both the non-asymptotic case [1, 37, 28, 38] and the asymptotic case [4, 7, 17]. In this section, as an application to quantum information theory, we treat problems of the approximate conversion between pure entangled states by LOCC. Entanglement of a pure state $\psi$ is characterized by the von Neumann entropy $S_{\psi}$ of its partial density matrix. For example, the pure state $\psi$ is entangled if and only if $S_{\psi} \neq 0$. Since several values cannot be defined for the case $S_{\psi} = 0$, we assume that pure states are entangled in this section.

As a typical conversion under the LOCC condition, we focus on entanglement concentration, in which, an i.i.d. pure state of $\psi$ is converted to multiple copies of the EPR state. It is known that the optimal conversion rate is the von Neumann entropy $S_{\psi}$ of its partial density matrix [4]. That is, it is possible to generate multi copies $\psi_{EPR}^{n+o(n)}$ of the EPR state from $n$-copies $\psi^{\otimes n}$ under the condition that the fidelity between the generated state and the target state asymptotically goes to 1. However, the converse does not hold, that is, even when the number of EPR states to be generated has the asymptotic expansion of the form of $S_{\psi} n + o(n)$, it is necessarily not possible to generate them under the condition that the fidelity between the generated state and the target state asymptotically goes to 1. In order to treat the error of LOCC conversion more precisely, we need to deal with the second order asymptotics. That is, the asymptotically possible fidelity $\nu$ between the generated state and the target state depends on the coefficient of the order $\sqrt{n}$. A similar problem occurs in entanglement dilution, in which, the multiple copies of the EPR state are converted to the multiple copies of a target pure entangled state. That is, in entanglement dilution, the asymptotically possible fidelity $\nu$ between the generated state and the target state also depends on the coefficient of the order $\sqrt{n}$. Such relations in entanglement concentration and dilution were studied in [16, 18, 24]. However, the existing studies dealt with the relation between the asymptotic fidelity and the coefficient of the order $\sqrt{n}$ only when the initial or
the target state is the EPR state, and thus they did not investigate this relation when both of the initial and the target states are non-EPR states. In the following, we treat more general LOCC conversions including entanglement concentration and entanglement dilution under the fidelity constraint, and clarify the relation between the second-order rate of the conversion and the fidelity $\nu$.

Before going to the asymptotics of LOCC conversion, we give some notations and remarks. In the following, we employ the fidelity $F$ to describe the accuracy of LOCC conversions. The following value represents the maximum fidelity of LOCC conversion for states $\psi$ and $\phi$:

$$F(\psi \to \phi) := \max_{\Gamma: \text{LOCC}} F(\Gamma(\psi), \phi).$$

(141)

Let $P_\psi$ and $P_\phi$ be the probability distributions which consist of the squared Schmidt coefficients for pure entangled states $\psi$ and $\phi$, respectively. Then, we have the relation between the fidelity $F(\psi, \phi)$ between pure entangled states and that $F(P_\psi, P_\phi)$ between probability distributions as

$$F(P_\psi^L, P_\phi^L) = \max_{U_A, U_B: \text{unitary}} F((U_A \otimes U_B)\psi, \phi).$$

(142)

Since $\psi$ is transformed to $\phi$ by LOCC if and only if $P_\psi \prec P_\phi$ where $\prec$ is the majorization relation given in Subsection 2.2, (142) implies the following relation for pure states $\psi$ and $\phi$:

$$F(\psi \to \phi) = F^M(P_\psi \to P_\phi).$$

(143)

We define the maximum convertible number for $\phi$ from $n$-copies of $\psi$ by LOCC under a permissible accuracy $0 < \nu < 1$ as follows:

$$L_n(\psi, \phi|\nu) := \max_{\Gamma: \text{LOCC}} \{ L \in \mathbb{N} | F(\Gamma(\psi \otimes n), \phi \otimes L) \geq \nu \}$$

(144)

$$= \max \{ L \in \mathbb{N} | F(\psi \otimes n \to \phi \otimes L) \geq \nu \}. \tag{145}$$

Since $P_{\psi \otimes n} = P_n^\psi$, the following holds:

$$L_n(\psi, \phi|\nu) = L_n^M(P_\psi, P_\phi|\nu). \tag{146}$$

Let $D_{\psi, \phi} := D_{P_\psi, P_\phi}$ and $C_{\psi, \phi} := C_{P_\psi, P_\phi}$. Since $H(P_\psi) = S_\psi$ and $H(P_\phi) = S_\phi$, the asymptotic expansion of the maximum convertible number $L_n(\psi, \phi|\nu)$ is obtained from Theorem 10 as follows.

**Theorem 18**

$$L_n(\psi, \phi|\nu) \approx \frac{S_\psi}{S_\phi} n + \frac{Z_{C_{\psi, \phi}}^{-1}(1 - \nu^2)}{D_{\psi, \phi}} \sqrt{n}. \tag{147}$$

In particular, when the initial state is the maximally entangled state $\psi_m^{\text{max}}$ on $\mathbb{C}^m \otimes \mathbb{C}^m$, Theorem 18 derives

$$L_n(\psi_m^{\text{max}}, \phi|\nu) \approx \frac{\log m}{S_\phi} n + \sqrt{\frac{V_\phi \log m}{S_\phi^3}} \Phi^{-1}(1 - \nu^2) \sqrt{n}. \tag{148}$$
where $V_{\psi} := V(P_{\psi})$. Similarly, when the target state is the maximally entangled state $\psi_{m}^{\text{max}}$, we have

$$L_n(\psi, \psi_{m}^{\text{max}}|\nu) \approx \frac{S_\psi}{\log m} n + \frac{\sqrt{V_{\psi}}}{\log m} \Phi^{-1}(1 - \nu^2) \sqrt{n}. \tag{149}$$

Remark 19 Bennett et.al \cite{4} gave the first-order rate of $L_n(\psi, \phi|\nu)$. Moreover, when $\psi$ or $\phi$ is the EPR state (i.e. the cases of entanglement dilution or entanglement concentration), Hayden and Winter \cite{18} and Harrow and Lo \cite{16} pointed out that the second-order of $L_n(\psi, \phi|\nu)$ is $\sqrt{n}$ and its second-order rate depends on the permissible accuracy for those operations. However, the explicit form of the second-order rate has not been obtained even for entanglement dilution and concentration. On the other hand, when $\psi$ or $\phi$ is the EPR state, Theorem 18 gives the explicit second-order rate of $L_n(\psi, \phi|\nu)$ in \cite{59}, which coincides with the result in \cite{24}, and hence, our results provide a refinement of the existing studies. Moreover, we also derived the second-order rates of $L_n(\psi, \phi|\nu)$ when both $\psi$ and $\phi$ are non-EPR pure states. Therefore, we obtain the second-order expansion of $L_n(\psi, \phi|\nu)$ in all cases as long as both $\psi$ and $\phi$ are entangled pure state.

Remark 20 We mention the relation between conversion problems in classical and quantum information theory. Due to the results of Nielsen \cite{28}, the approximate conversion problem between pure states on bipartite systems is induced into that of probability distribution under the majorization condition by considering the squared Schmidt coefficients of the pure states. Since the squared Schmidt coefficients of a maximally entangled state form a uniform distribution, in particular, those of the EPR state form the uniform distribution over $\{0, 1\}$, it is thought that entanglement dilution and concentration in quantum information theory correspond to the resolvability and the intrinsic randomness in classical information theory.

5.2 LOCC Cloning with Perfect Knowledge

Due to the no-cloning theorem, we can not generate a complete copy of an unknown quantum state. Then, in studies of the cloning of an unknown quantum state, an approximate cloning method and the evaluation of its accuracy have been mainly treated \cite{39, 6, 13}. On the other hand, even when the state to be copied is known, it is impossible to perfectly copy the state when the state is entangled and our operations are limited to LOCC. In the following, we treat such a case. Thus, we assume that we know entangled state to be copied, but, we can use only LOCC for copying. We note that existing studies \cite{1, 32} discussed similar cloning problem\footnote{The papers \cite{1, 32} discussed local copying, and a limited amount of the EPR states are prepared as a resource for copying, only LOCC is allowed for our operation, and we only know that the state to be copied belongs to the set of candidate of the states. It is required to copy the unknown state perfectly by using the same amount of the EPR states as the number of required clones.}, however, the setting is different from ours because their
setting assumes an imperfect knowledge for the entangled state to be cloned and additional limited entangled resource. To distinguish their setting, we call our setting the LOCC cloning with perfect knowledge, and call their setting the LOCC cloning with imperfect knowledge. In this paper, we investigate LOCC cloning with perfect knowledge when the initial entangled state is \( n \)-copies of \( \psi \) and the target state is \( L_n \)-copies of \( \psi \) with \( L_n \geq n \). That is, we analyze how large number \( L_n \) of copies we can generate under the condition that the fidelity between the transformed state from the initial state by LOCC and the target entangled state is greater than a permissible accuracy \( \nu \). The maximal number \( L_n(\psi|\nu) \) of \( L_n \) given above is formulated by

\[
L_n(\psi|\nu) := \max_{\Gamma: \text{LOCC}} \{ L | F(\Gamma(\psi^\otimes n), \psi^\otimes L) \geq \nu \},
\]

which equals \( L_n(\psi, \psi|\nu) \) by the definition in (144). Then we obtain the following asymptotic expansion from (143).

**Theorem 21**

\[
L_n(\psi|\nu) \approx n + \frac{\sqrt{8V_{\psi}\log\nu^{-1}}}{S_{\psi}} \sqrt{n}.
\]

Thus, when the initial state is the i.i.d. entangled state \( \psi^\otimes n \) of a non-maximally entangled state \( \psi \), the incremental number \( L_n(\psi|\nu) - n \) of copies by LOCC cloning with perfect knowledge has the order of \( \sqrt{n} \). On the other hand, when \( \psi \) is a maximally entangled state, \( V_{\psi} = 0 \) and thus the incremental number of copies by LOCC cloning with perfect knowledge does not have the order of \( \sqrt{n} \). Indeed, since

\[
\max_{\Gamma: \text{LOCC}} F(\Gamma(\psi_{EPR}^\otimes n), \psi_{EPR}^\otimes L) = \sqrt{2n - L}.
\]

holds by Lemma 8, we obtain

\[
L_n(\psi_{EPR}|\nu) = \lfloor n + 2 \log \nu^{-1} \rfloor,
\]

where \( \lfloor \cdot \rfloor \) represents the floor function, and the incremental number \( L_n(\psi|\nu) - n \) is bounded by a constant \( 2 \log \nu^{-1} \) for any \( n \) unlike a non-maximally state.

According to Chiribella et al. [10], we define the replication rate as the limit

\[
r(\psi, \nu) := \lim_{n \to \infty} \log \frac{L_n(\psi|\nu)}{n}.
\]

Then, the rate \( r(\psi, \nu) \) can be characterized as follows

\[
r(\psi, \nu) = \begin{cases} \frac{1}{2} & \text{when } V_{\psi} \neq 0 \\ 0 & \text{when } V_{\psi} = 0. \end{cases}
\]

**6 Conclusion**

We have addressed approximation conversion problems of probability distributions by deterministic and majorization conversions under a permissible accuracy. We have found that two conversion methods are related as in [36, 38].
and have derived the asymptotic expansion of the maximum convertible number up to the order $\sqrt{n}$ for the both kinds of conversion problems between two i.i.d. probability distributions. To derive the computable form of the second-order rate of the asymptotic expansion, the problem has been divided into the uniform case and the non-uniform case. However, we note that the maximum convertible numbers for two kinds of conversions are equivalent to each other in all cases up to the order $\sqrt{n}$ as stated in Theorem 10. A key to derive the asymptotic expansion is to introduce Rayleigh-normal distribution and to investigate its properties. In particular, the optimal conversion rate is described by the Rayleigh-normal distribution function (or the maximum continuous fidelity) for the non-uniform case. Thereafter, as applications to quantum information theory, we have addressed LOCC conversion problem between bipartite entangled pure states including entanglement concentration and entanglement dilution. Then, we have derived the asymptotic expansion of the maximum convertible number using the results for majorization conversion of probability distributions. In particular, we have clarified the relation between the second-order rate and the accuracy of LOCC conversion. As a special case, we have introduced the notion of LOCC cloning with the perfect knowledge. Using the results for LOCC conversion, we have derived the rate of the incremental copies and the optimal coefficient in this setting.

The following problems can be considered as future problems. First, this paper assumes the independent and identical distributed condition for the sequences of distributions to be converted. However, the actual sequences of distributions or the pure entangled states might have correlation in practice. Hence, it is an interesting open problem to extend the obtained result to the case of correlated sequences of distributions or entangled pure states [27], e.g., the Markovian case. Next, only pure states have been treated in quantum information setting although mixed entangled states may appear in practice. So, an extension to the case of mixed states is required as a future study. Finally, we point out the significance of analysis in a finite-length setting. We have analyzed the asymptotic performance of approximate conversions in this paper. On the other hand, we can operate an input state only with a finite length. Therefore, it is needed to analyze the approximate conversion problems in a finite-length setting. For entanglement dilution and the concentration, the recent paper [24] dealt with an analysis in a finite-length setting and derived its numerical results. However, no result investigates the finite-length setting of the case when the initial state and the target state are non-EPR states.

Acknowledgments

We would like to thank Mr. Hiroyasu Tajima for his helpful comments. WK was partially supported by Grant-in-Aid for JSPS Fellows No. 233283. MH is partially supported by a MEXT Grant-in-Aid for Scientific Research (A) No. 23246071 and the National Institute of Information and Communication Technology (NICT), Japan. The Centre for Quantum Technologies is funded
A Derivation of An Explicit Form of The Rayleigh-Normal Distribution

We prove Theorem 4 which gives an explicit form of the Rayleigh-normal distribution. To do so, we first prepare two lemmas.

Lemma 22 The ratio \( N(x) \) is strictly monotonically decreasing only on the interval \( \mathcal{A}_{\mu, v} \) defined by

\[
\mathcal{A}_{\mu, v} = \begin{cases} \mathbb{R} & (v = 1 \text{ and } \mu > 0) \\ \emptyset & (v = 1 \text{ and } \mu \leq 0) \\ (\mu(v - 1)^{-1}, \infty) & (v > 1) \\ (-\infty, \mu(v - 1)^{-1}) & (v < 1), \end{cases}
\]

where \( \emptyset \) is the empty set.

Proof: The area \( \mathcal{A}_{P, Q, b} \) is easily derived from

\[
\frac{N(x)}{N_{\mu, v}(x)} = \begin{cases} e^{\frac{\mu^2}{2}} e^{-\mu x} & (v = 1) \\ \frac{1}{\sqrt{2\pi}e^{v(\mu - 1)^{-1}} e^{\frac{1}{v-1}(x-\mu)^2}} & (v \neq 1). \end{cases}
\]

Lemma 23 Assume that real numbers \( t \leq t' \) satisfy the condition (*) in Lemma 16. Then the following inequality holds

\[
\sup_{A} \mathcal{F} \left( \frac{dA}{dx}, N_{\mu, v} \right) \leq \sqrt{\Phi(t)} \sqrt{\Phi_{\mu, v}(t)} + \int_{t}^{t'} \sqrt{N(x)} \sqrt{N_{\mu, v}(x)} dx + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{\mu, v}(t')}.
\]

Proof: When we set a sequence \( \{x_i\}_{i=0}^{\infty} \) for \( I \in \mathbb{N} \) as \( x_i := t + \frac{t'}{I} + i \), we have the following by the Schwartz inequality for an arbitrary \( A \) satisfying the conditions
in definition 1

\[ F\left( \frac{dA}{dx} N_{\mu,v} \right) = \int_{t}^{t'} \sqrt{dA(x)} \sqrt{N_{\mu,v}(x)} dx + \int_{t'}^{\infty} \sqrt{dA(x)} \sqrt{N_{\mu,v}(x)} dx \]

(158)

\[ \sum_{i=1}^{t} \int_{x_{i-1}^{t}}^{x_{i}^{t}} \sqrt{dA(x)} \sqrt{N_{\mu,v}(x)} dx \]

(159)

\[ \leq \sqrt{A(t)} \sqrt{\Phi_{\mu,v}(t)} + \sqrt{1 - A(t')} \sqrt{1 - \Phi_{\mu,v}(\lambda)} \]

(160)

\[ + \sum_{i=1}^{t} \sqrt{A(x_{i}^{t}) - A(x_{i-1}^{t})} \sqrt{\Phi_{\mu,v}(x_{i}^{t}) - \Phi_{\mu,v}(x_{i-1}^{t})} \]

\[ \leq \sqrt{\Phi(t)} \sqrt{\Phi_{\mu,v}(t)} + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{\mu,v}(t')} \]

(161)

\[ + \sum_{i=1}^{t} \sqrt{\Phi(x_{i}^{t}) - \Phi(x_{i-1}^{t})} \sqrt{\Phi_{\mu,v}(x_{i}^{t}) - \Phi_{\mu,v}(x_{i-1}^{t})} \]

\[ = \sqrt{\Phi(t)} \sqrt{\Phi_{\mu,v}(t)} + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{\mu,v}(t')} \]

(162)

\[ + \sum_{i=1}^{t} \sqrt{\Phi(x_{i}^{t}) - \Phi(x_{i-1}^{t})} \sqrt{\Phi_{\mu,v}(x_{i}^{t}) - \Phi_{\mu,v}(x_{i-1}^{t})} (x_{i}^{t} - x_{i-1}^{t}) \]

\[ \xrightarrow{t \to \infty} \sqrt{\Phi(t)} \sqrt{\Phi_{\mu,v}(t)} + \sqrt{1 - \Phi(t')} \sqrt{1 - \Phi_{\mu,v}(t')} \]

(163)

where the inequality (160) is obtained from the Schwartz inequality and the inequality (161) is obtained from Lemmas 15 and 22.

Proof of Theorem 4 The function \( A_{\mu,v} \) is a continuous differentiable monotone increasing function satisfying \( \Phi \leq A \leq 1 \). Thus, we obtain

\[ \sup_{A} F\left( \frac{dA}{dx}, N_{\mu,v} \right) = F\left( \frac{dA_{\mu,v}}{dx}, N_{\mu,v} \right) \]

(164)

from Lemmas 23 and 17. From the direct calculation, we can verify that

\[ F\left( \frac{dA_{\mu,v}}{dx}, N_{\mu,v} \right) = \begin{cases} \sqrt{\Phi(\beta_{\mu,v})} \sqrt{\Phi_{\mu,v}(\beta_{\mu,v})} + I_{\mu,v}(\beta_{\mu,v}) & \text{if } v < 1 \\
 e^{-\frac{x_{\mu,v}^{2}}{2}} & \text{if } v = 1 \\
 \sqrt{\Phi(\alpha_{\mu,v})} \Phi_{\mu,v}(\alpha_{\mu,v}) + I_{\mu,v}(\infty) - I_{\mu,v}(\alpha_{\mu,v}) & \text{if } v > 1. \end{cases} \]

Therefore, the proof is completed.
References

[1] F. Anselmi, A. Chefles, and M. B. Plenio, “Local copying of orthogonal entangled quantum states,” New J. Phys., 6(1), 164, (2004).

[2] B. C. Arnold, Majorization and the Lorenz Order: A Brief Introduction, Springer-Verlag, (1986).

[3] I. Bengtsson and K. Zyczkowski, Geometry of Quantum States, Cambridge University Press, (2006).

[4] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, “Concentrating partial entanglement by local operations,” Phys. Rev. A, 53(4), 2046, (1996).

[5] C. H. Bennett, D. P. DiVincenzo, J. Smolin and W. K. Wootters, “Mixed-state entanglement and quantum error correction,” Phys. Rev. A, 54(5), 3824, (1996).

[6] V. Bužek, and M. Hillery, “Universal optimal cloning of arbitrary quantum states: From qubits to quantum registers,” Phys. Rev. Lett., 81(22), 5003, (1998).

[7] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin and A. V. Thapliyal, “Exact and asymptotic measures of multipartite pure-state entanglement,” Phys. Rev. A, 63(1), 012307 (2000).

[8] C. H. Bennett and P. W. Shor, “Quantum information theory,” IEEE Trans. Inform. Theory, 44(6), 2724, (1998).

[9] M. Bourguignon, R. B. Silva and G. M. Cordeiro, “The Weibull-G Family of Probability Distributions,” J. Data Sci., 12 (2014).

[10] G. Chiribella, Y. Yang and A. C. C. Yao, “Quantum replication at the Heisenberg limit,” Nature comm., 4 (2013).

[11] N. Datta and F. Leditzky, “Second-order asymptotics for source coding, dense coding and pure-state entanglement conversions,” arXiv:1403.2543 (2014).

[12] M. J. Donald, M. Horodecki and O. Rudolph, “The uniqueness theorem for entanglement measures,” J. Math. Phys., 43(9), 4252, (2002).

[13] R. Filip, “Quantum partial teleportation as optimal cloning at a distance,” Phys. Rev. A, 69(5), 052301, (2004).

[14] D. Gross, S. T. Flammia and J. Eisert, “Most quantum states are too entangled to be useful as computational resource,” Phys. Rev. Lett., 102(19), 190501, (2009).
[15] T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer, New York, (2003).

[16] A. Harrow and H. K. Lo, “A tight lower bound on the classical communication cost of entanglement dilution,” *IEEE Trans. Inform. Theory*, **50**(2), 319, (2004).

[17] P. Hayden, M. Horodecki and B. M. Terhal, “The asymptotic entanglement cost of preparing a quantum state,” *J. Phys. A*, **34**(35), 6891, (2001).

[18] P. Hayden and A. Winter, “Communication cost of entanglement transformations,” *Phys. Rev. A*, **67**(1), 012326, (2003).

[19] M. Hayashi, “General formulas for fixed-length quantum entanglement concentration,” *IEEE Trans. Inform. Theory*, **52**(5), 1904, (2006).

[20] M. Hayashi, “Second-order asymptotics in fixed-length source coding and intrinsic randomness,” *IEEE Trans. Inform. Theory*, **54**(10), 4619, (2008).

[21] M. Hayashi, “Information-Spectrum Approach to Second-Order Coding Rate in Channel Coding,” *IEEE Trans. Inform. Theory*, **55**(11), 4947, 2009.

[22] M. Hayashi, “Comparison between the Cramer-Rao and the mini-max approaches in quantum channel estimation,” *Comm. Math. Phys.*, **304**(3), 689, (2011).

[23] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, “Quantum Entanglement,” *Rev. Mod. Phys.*, **81**(2), 865, (2009).

[24] W. Kumagai, M. Hayashi, “Entanglement Concentration is Irreversible,” *Phys. Rev. Lett.*, **111**(13), 130407, (2013).

[25] K. Li, “Second order asymptotics for quantum hypothesis testing,” [arXiv:1208.1400] (2012).

[26] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, (1979).

[27] K. Modi, T. Paterek, W. Son, V. Vedral, and M. Williamson, “Unified view of quantum and classical correlations,” *Phys. Rev. Lett.*, **104**(8), 080501, (2010).

[28] M. A. Nielsen, “Conditions for a class of entanglement transformations,” *Phys. Rev. Lett.*, **83**(2), 436, (1999).

[29] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, (2000).

[30] R. Nomura and T. S. Han, “Second-Order Resolvability, Intrinsic Randomness, and Fixed-Length Source Coding for Mixed Sources,” *IEEE Trans. Inform. Theory*, **59**(1), 1, (2013).
[31] M. Ohya and D. Petz, *Quantum entropy and its use*, Springer-Verlag, Heidelberg, (1993).

[32] M. Owari and M. Hayashi, “Local copying and local discrimination as a study for nonlocality of a set of states,” *Phys. Rev. A*, **74**(3), 032108, (2006).

[33] Y. Polyanskiy, H.V. Poor and S. Verdu, “Channel Coding Rate in the Finite Blocklength Regime”, *IEEE Trans. Inform. Theory*, **56**(5), 2307, 2010.

[34] M. Tomamichel and M. Hayashi, “A hierarchy of information quantities for finite block length analysis of quantum tasks,” [arXiv:1208.1478](http://arxiv.org/abs/1208.1478) (2012).

[35] M. Tomamichel and V. Y. Tan, “Second-Order Asymptotics of Classical-Quantum Channels,” [arXiv:1308.6503](http://arxiv.org/abs/1308.6503) (2013).

[36] A. W. Van der Vaart. *Asymptotic Statistics*, Cambridge University Press, (1998).

[37] G. Vidal, “Entanglement of pure states for a single copy,” *Phys. Rev. Lett.*, **83**(5), 1046, (1999).

[38] G. Vidal, D. Jonathan and M. A. Nielsen, “Approximate transformations and robust manipulation of bipartite pure-state entanglement,” *Phys. Rev. A*, **62**(1), 012304, (2000).

[39] R. F. Werner, “Optimal cloning of pure states,” *Phys. Rev. A*, **58**(3), 1827, (1998).