Quantum Blind Data Compression and Structure of Quantum Operations Approximately Preserving Quantum States

by

Kohdai Kuroiwa

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2021

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, we investigate a variation of quantum information processing tasks, *blind data compression*, and we analyze an approximation of a structure of a set of quantum states closely related to the task, which is called the *Koashi-Imoto (KI) decomposition*.

While various quantum information processing tasks have been extensively investigated in the framework of quantum Shannon theory, a problem called *blind quantum data compression* is considered as one of the most general forms of quantum data compression. It is known that its optimal compression rate within an asymptotically vanishing error is given by using the KI decomposition. However, it is also argued that allowing even an extremely small approximation causes a significant change in the compression rate. The sensitivity of the compression rate to approximations originates from the sensitivity of the KI decomposition. In this thesis, taking advantage of the sensitivity, we construct a novel protocol for blind quantum data compression that may perform remarkably well under the existence of finite approximations. Furthermore, to acquire insights into the instability of the KI decomposition and to analyze an approximation of the KI decomposition with finite approximations allowed, we investigate a structure of quantum channels that may lead to further understanding of an approximate structure of quantum states that is essential for more sophisticated error analysis of blind compression.

Our results shed light on an instability of the rate of blind quantum data compression against approximations. Our compression protocol makes the data transmission with approximations much more efficient. Furthermore, our results on the approximation of the KI decomposition provides us with insights into an approximate KI decomposition of quantum states that is essential to conduct more rigorous and general analysis of blind data compression, as well as contributes to foundation of quantum mechanics from the perspective of what restrictions are imposed to quantum operations when they cause a small disturbance. We believe that our work paves the way to further investigation of blind quantum data compression with finite approximations, and our results make substantial progress towards the general analysis of approximate KI decomposition, which is essential not only for the study of blind quantum data compression but also for investigation of other quantum phenomena characterized by the KI decomposition.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Debbie Leung, for her enthusiastic guidance and support throughout my master's program. She taught me important skills and honest attitudes required for being a scientific researcher. I would also like to thank her for patiently instructing me during the period of remote work due to the outbreak of COVID-19. I owe a lot to her, and I enjoyed conducting research under her supervision.

I sincerely appreciate my collaborators, Anurag Anshu and Felix Leditzky, for their helpful and fruitful discussions. Some parts of my research would not be possible without their advice.

I also highly appreciate Prof. Crystal Senko, Prof. John Watrous, and Prof. Norbert Lütkenhaus for being parts of my advisory committee and examination committee.

Finally, I would like to thank my girlfriend Reika Fujimura, my family, and my friends for their love and support.

Table of Contents

List of Figures vii

List of Tables ix

1 Introduction 1

2 Backgrounds 4

2.1 Basic Mathematical Notation 4

2.2 Quantum Systems and Operators 5

2.2.1 Quantum Systems 5

2.2.2 Operators on Quantum Systems 7

2.2.3 Decomposition of Linear Operators 12

2.3 Quantum Channels 15

2.4 Norms and Quantum Entropy 18

2.4.1 Norms of Linear Operators 18

2.4.2 Operator Functions and Quantum Entropy 21

2.5 Structure of Quantum Ensembles 24

2.6 Quantum Data Compression 29

3 Blind Quantum Data Compression 34

3.1 Setup of Blind Quantum Data Compression 34

3.2 Previous Results on Blind Data Compression 37
3.2.1 Asymptotic Optimal Rate of Blind Data Compression 37
3.2.2 Blind Compression of Classical Ensembles under Local Error Criterion 39
3.3 Our Results: Data Compression Protocol with Finite Approximation under Local Error Criterion 40
3.3.1 Procedure of Our Protocol 40
3.3.2 Reduction of rates with a finite error 43
3.3.3 Approximation of Classical Ensembles 49
3.4 Summary and Discussions 62

4 Structure of Quantum Operations Approximately Preserving Quantum Ensemble 63
4.1 Brief Overview of Structure of Quantum Operations that Preserve a Quantum Ensemble 64
4.2 Our results: Approximate Structure of Quantum Operation Almost Preserving Quantum Ensemble 66
4.3 Summary and Discussions 78

5 Conclusion 79

References 81

List of Figures

2.1 The schematic diagram showing the basic setup of quantum data compression tasks. The sender encodes a target state (red-colored in the figure) picked from a given ensemble and send it to the receiver via noiseless quantum channels. The receiver decodes the transmitted state to recover the original state (blue-colored in the figure). 30

3.1 The schematic diagram showing the setup of blind quantum data compression. In this task, the sender is given a target state (red-colored in the figure) from the referee without its description. Then, the sender encodes the state and send it to the receiver via noiseless quantum channels. The receiver decodes the transmitted state to recover the original state (blue-colored in the figure). 35

3.2 Graph of the compression rate as a function of finite allowed error. The term "rate" refers to the upper bound shown in Eq. (3.97). The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetic and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively. 57

3.3 Graph of the actual error as a function of finite allowed error. Here, "actual error" means the distance between the initial state and the resulting state in terms of the trace norm. The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetic and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively. 57

3.4 Graph of the compression rate as a function of finite allowed error and the corresponding fitting curves. The blue and red points represent the averages of the results of the arithmetic and geometric mean methods. The indigo curve is the fitting function corresponding to the blue points; the brown curve is the fitting function corresponding to the red points. 58

3.5 Graph of the compression rate as a function of the dimension of system. The term "rate" refers to the upper bound shown in Eq. (3.97). The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetical and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively. 59

3.6 Graph of the actual error as a function of the dimension of system. Here, "actual error" means the distance between the initial state and the resulting state in terms of the trace norm. The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetical and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively. 59

3.7 Graph of (the compression rate by the arithmetic mean method) − (the compression rate by the geometric mean method) as a function of the dimension of system. The horizontal axis is expressed in log scale. 60

3.8 Graph of the compression rate as a function of the dimension of system and the corresponding fitting curves. The blue and red points represent the averages of the results of the arithmetic and geometric mean methods. The indigo curve is the fitting function corresponding to the blue points; the brown curve is the fitting function corresponding to the red points. 61

List of Tables

3.1 The results of the fitting of the graph of the compression rate shown in Figure 3.4. In the table, a and b are defined in Eq. (3.98). 58

3.2 The results of the fitting of the graph of the compression rate shown in Figure 3.8. In the table, a and b are defined in Eq. (3.99). 61

Chapter 1

Introduction

Quantum information science has been trying to utilize fundamental properties of quantum mechanics in order to achieve advantages over our classical information technologies. Among various quantum information processing tasks, a data compression problem called blind quantum data compression has been widely investigated owing to their general and practical setup. In Ref. [1], the authors obtained the optimal compression rate of the quantum blind compression by using a structure of quantum states. However, while the compression rate achieved in Ref. [1] is optimal with asymptotically vanishing error, recently, it has been discussed that the error analysis in the paper is still loose [2, 3]. Indeed, even a diminutive approximation causes a large change in the compression rate due to the fact that the structure of states is extremely unstable against approximations [2]. In this thesis, taking advantage of the sensitivity, we construct a novel protocol for blind quantum data compression that may perform significantly better than the one Ref. [1] proposed when approximations are allowed. Furthermore, aiming to have insights into the instability of the structure to analyze blind compression with errors or approximations more generally, we investigate a structure of quantum channels, which should be closely related to an approximate structure of quantum states demanded for more accurate error analysis of blind compression.

In quantum mechanics, when we attempt to identify a given quantum state, the description of the state may change as a result of measurements. To study this restriction of quantum states, researchers have been studied conditions under which such disturbance does not occur. Various investigations have been previously conducted, and they revealed properties and structures of operations and states so as to achieve the read-out of infor-

mation from quantum states without a disturbance. In particular, Ref. [4] disclosed that a set of quantum states has its own structure, called the Koashi-Imoto (KI) decomposition, with which we can identify its classical parts, non-redundant quantum parts, and redundant parts. The KI decomposition can be used to characterize various quantum mechanical properties such as quantum information processing tasks [1,5,6], quantum Markov chains [7–9], and the recoverability of quantum operations [10, 11].

In particular, it has been revealed that the KI decomposition can be applied to efficient transmission of quantum data [1], which is called *blind quantum data compression*. This problem serves as a part of *quantum Shannon theory*, where we exploit quantum-mechanical phenomena to send quantum or classical information through quantum information channels. These information processing problems were initially motivated by classical Shannon theory. See, for example, Refs. [12–15] for problems considered in classical Shannon theory. Various previous research has extensively investigated its quantum variety, for example, *quantum data compression* [1–3, 16–34], *quantum channel capacity* [35–44], *quantum state merging* [45, 46], *quantum state redistribution* [47–49], and other more involved tasks including those leading to *protocol family* [50–58]. Blind quantum data compression, which is a data compression problem in which the sender does not know the description of quantum states aimed to transmit, has been getting attraction [1–3], and this problem is considered as the most general form of the quantum data compression. In Ref. [1], its optimal compression rate was given by utilizing KI decomposition; however, in Refs. [2,3], it was argued that the error analysis conducted in Ref. [1] is loose because KI decomposition is quite sensitive to approximations. In particular, Ref. [2] introduced a compression protocol for blind compression of classical data whose rate is considerably small with finite approximations allowed. Nevertheless, the compression of general quantum states with approximations has not been formally discussed before.

In this thesis, we investigate blind quantum data compression protocol with finite approximations. Focusing on the instability of the optimal compression rate against approximations, we propose a novel data compression protocol for this problem. Our protocol exhibits a substantial reduction of the compression rate compared with the rate obtained in Ref. [1], which is optimal under asymptotically vanishing errors. Moreover, using our protocol, we investigated blind compression of classical states. We conducted numerical experiments and showed the performance of the protocol for classical ensembles. To further investigate blind compression protocol with approximations, an approximated version of the KI decomposition is essential. To investigate an approximation of KI decomposition, we consider a structure of quantum channels. As noted in the previous research [2], analysis of an approximation of KI decomposition is intractable because the decomposition is sensitive to approximations. Indeed, even an extremely small approximation to a single state leads

to completely different KI decomposition. In Ref. [4], along with the KI decomposition of quantum states, a structure of quantum channels that preserves a set of quantum states is also considered. Extending discussions in Ref. [4], we prove that a quantum channel has an approximate block structure when it approximately preserves a set of quantum states. The extent of approximation depends on the extent of preservability of the given quantum channel, and when the channel perfectly preserves the set of states, the channel has an exact block structure. Our analysis provides an insight into how KI decomposition can be approximated, which may lead to strict and general error analysis of blind quantum data compression.

The rest of this thesis is organized as follows. In Chapter 2, we review backgrounds of our research and overview basic concepts required to understand our results. In Chapter 3, we analyze a variation of quantum data compression tasks, *blind quantum data compression*, with finite approximations. In Chapter 4, to obtain an insight into an approximate structure of quantum ensembles, we investigate a structure of quantum channels that almost preserves a quantum ensemble. Finally, in Chapter 5, we summarize and discuss our results and show further directions and interests of our research.

Chapter 2

Backgrounds

In this chapter, we briefly review backgrounds and preliminaries for the thesis. In Sec. 2.1, we show our notation for basic mathematical concepts. In Sec. 2.2, we review mathematical concepts of quantum systems and operators defined on a quantum system. We also introduce several kinds of decomposition of operators. In Sec. 2.3, we review the definition of quantum channels and useful representations of quantum channels. In Sec. 2.4, we review norms on a space of operators and entropic quantities. In Sec. 2.5, we briefly overview a structure of quantum ensemble, which is known as Koashi-Imoto decomposition. Finally, in Sec. 2.6, we show the basic setup of quantum data compression and classify several classes of quantum data compression problems.

2.1 Basic Mathematical Notation

In this section, we briefly introduce our notation for basic concepts.

We let $\mathbb{R}$ denote the set of real numbers and $\mathbb{C}$ denote the set of complex numbers. For a finite nonempty set Σ, $\mathbb{C}^{\Sigma}$ is a vector space over $\mathbb{C}$ with dimension $|\Sigma|$ that is formed by the set of functions from Σ to $\mathbb{C}$.

For a given complex number $x \in \mathbb{C}$, $\mathrm{Re}(x)$ denotes the real part of x, and $\mathrm{Im}(x)$ denotes the imaginary part of x. The complex conjugate of a complex number x is denoted by $\overline{x}$.

For real numbers $x, y \in \mathbb{R}$, $\max\{x, y\}$ represents the larger value of x and y; $\min\{x, y\}$ represents the smaller value of x and y.

2.2 Quantum Systems and Operators

In this section, we review the basic concepts of quantum systems and operators on a quantum system.

2.2.1 Quantum Systems

Throughout this thesis, we consider quantum systems, which are governed by the law of quantum mechanics. A quantum system is represented by a complex Hilbert space, that is, a complex complete Euclidean inner-product space. For a quantum system A, we let $\mathcal{H}_A$ denote the corresponding complex Hilbert space. For convenience, we also say that $\mathcal{H}_A$ is a quantum system. In this thesis, we only consider a finite-dimensional quantum system; that is, we only need to deal with a finite-dimensional complex Euclidean inner-product space. Hereafter, we call a finite-dimensional quantum system just a quantum system for brevity. Then, for a quantum system $\mathcal{H}$, we let $D_{\mathcal{H}}$ denote the dimension of the system. An arbitrary quantum system $\mathcal{H}$ can be written as

$$\mathcal{H} = \mathbb{C}^{\Sigma} \tag{2.1}$$

for an alphabet, i.e., a finite nonempty set, Σ with $|\Sigma| = D_{\mathcal{H}}$. When $\Sigma = \{1, 2, \ldots, n\}$ for some positive integer n, we also write

$$\mathcal{H} = \mathbb{C}^{n}. \tag{2.2}$$

In this thesis, we adopt the Dirac notation; that is, we let $|x\rangle$ ("*ket* x") denote a vector in a quantum system $\mathcal{H}$. The dual of $|x\rangle \in \mathcal{H}$ is denoted by $\langle x|$ ("*bra* x"). When a given quantum system $\mathcal{H}$ can be written as $\mathcal{H} = \mathbb{C}^{\Sigma}$ for some alphabet Σ, a vector $|x\rangle \in \mathcal{H}$ can be expressed as a tuple labeled by Σ; that is, $|x\rangle = (x_a)_{a\in\Sigma}$ with complex numbers $x_a \in \mathbb{C}$ for all $a \in \Sigma$.

Now, we define the *inner-product* on quantum systems.

Definition 2.1 (Inner-Product)**.** Let $\mathcal{H} = \mathcal{C}^{\Sigma}$ be a quantum system. Then, we define the inner-product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ on the system $\mathcal{H}$ as

$$\langle |x\rangle, |y\rangle \rangle \coloneqq \sum_{a\in\Sigma} \overline{x_a} y_a \tag{2.3}$$

for all vectors $|x\rangle, |y\rangle \in \mathcal{H}$. In this thesis, we write

$$\langle x|y\rangle \coloneqq \langle |x\rangle, |y\rangle \rangle \tag{2.4}$$

to represent the inner-product.

The inner-product $\langle \cdot, \cdot \rangle$ defined in Definition 2.1 satisfies the following properties, which qualify a function of vectors to be a inner-product of the vector space over $\mathbb{C}$.

1. For all vectors $|x\rangle$, $|y\rangle$, and $|z\rangle$, and for all complex numbers α and β, we have
$$\langle |x\rangle , \alpha |y\rangle + \beta |z\rangle \rangle = \alpha \langle x|y\rangle + \beta \langle x|z\rangle . \tag{2.5}$$
2. For all vectors $|x\rangle$ and $|y\rangle$, we have
$$\langle x|y\rangle = \overline{\langle y|x\rangle}. \tag{2.6}$$
3. For all vectors $|x\rangle$, we have
$$\langle x|x\rangle \geqq 0. \tag{2.7}$$
Furthermore, $\langle x|x\rangle = 0$ if and only if $|x\rangle = 0$.

Now, we define a concept of *basis* of a quantum system $\mathcal{H}$.

Definition 2.2 (Basis)**.** Let $\mathcal{H}$ be a quantum system. Then, a set of vectors $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ is said to be a basis of $\mathcal{H}$ if the following conditions are satisfied.

1. The vectors in $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ are linearly independent; that is, if it holds that
$$\sum_{i=1}^{D_{\mathcal{H}}} c_i |i\rangle = 0 \tag{2.8}$$
for some set of complex numbers $\{c_i \in \mathbb{C} : 1 \leqq i \leqq D_{\mathcal{H}}\}$, we have
$$c_1 = c_2 = \cdots = c_{D_{\mathcal{H}}} = 0. \tag{2.9}$$
2. The set $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ spans the space; that is, for any vector $|x\rangle \in \mathcal{H}$, there exists a set of complex numbers $\{c_i \in \mathbb{C} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ such that
$$|x\rangle = \sum_{i=1}^{D_{\mathcal{H}}} c_i |i\rangle . \tag{2.10}$$

In this thesis, we usually consider an *orthonormal basis*, which is a basis with orthonormality. When a basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ is orthonormal; that is, if the vectors in the set satisfies

$$\langle i|j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \tag{2.11}$$

the basis is called an orthonormal basis. Hereafter, the term "basis" will be referred to as "orthonormal basis".

To represent a composition of multiple quantum systems, we use the tensor product of the corresponding Hilbert spaces, which is called a *composite system*.

Definition 2.3 (Composite Systems)**.** Let A and B be quantum systems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. The composite system AB of A and B is the quantum system corresponding to the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

2.2.2 Operators on Quantum Systems

For quantum systems $\mathcal{H}_A$ and $\mathcal{H}_B$, let $\mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ denote the set of linear operators from $\mathcal{H}_A$ to $\mathcal{H}_B$. Taking alphabets Σ_A and Σ_B such that $\mathcal{H}_A = \mathbb{C}^{\Sigma_A}$ and $\mathcal{H}_B = \mathbb{C}^{\Sigma_B}$, for an arbitrary linear operator $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, we can consider a matrix representation of X:

$$X = \sum_{a \in \Sigma_A, b \in \Sigma_B} X_{ba} \, |b\rangle \langle a| \tag{2.12}$$

with

$$X_{ba} \coloneqq \langle b|X|a\rangle \coloneqq \langle |b\rangle, X\,|a\rangle \rangle, \tag{2.13}$$

where $\{|a\rangle \in \mathcal{H}_A : a \in \Sigma_A\}$ and $\{|b\rangle \in \mathcal{H}_B : b \in \Sigma_B\}$ forms orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. We write $\mathrm{L}(\mathcal{H}) \coloneqq \mathrm{L}(\mathcal{H}, \mathcal{H})$ when the input space and the output space are both $\mathcal{H}$. Operators in a space $\mathrm{L}(\mathcal{H})$ is called *square* operators.

A linear operator can be regarded as a linear function of vectors. As an example of linear operators, we give the *identity operator*, which works as an identity function.

Definition 2.4 (Identity Operator)**.** Let $\mathcal{H}_A$ be a quantum system. The identity operator $I_{\mathcal{H}_A} \in \mathrm{L}(\mathcal{H}_A)$ is defined as the unique operator such that

$$I_{\mathcal{H}_A} |u\rangle = |u\rangle \tag{2.14}$$

for all $|u\rangle \in \mathcal{H}_A$. In this thesis, we may write I_A instead of $I_{\mathcal{H}_A}$.

Here, we define the kernel and the image of linear operators.

Definition 2.5 (Kernel and Image)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ be a linear operator. Then, the kernel of X is a subspace of $\mathcal{H}_A$ defined as

$$\ker(X) \coloneqq \{|u\rangle \in \mathcal{H}_A : X|u\rangle = 0\}. \tag{2.15}$$

In addition, the image of X is a subspace of $\mathcal{H}_B$ defined as

$$\mathrm{im}(X) \coloneqq \{X|u\rangle : |u\rangle \in \mathcal{H}_A\}. \tag{2.16}$$

When a linear operator is given, we can consider a conjugated linear operator to the given operator, which is called the *adjoint operator*.

Definition 2.6 (Adjoint Operator)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ be a linear operator. The adjoint operator $X^\dagger \in \mathrm{L}(\mathcal{H}_B, \mathcal{H}_A)$ of X is a linear operator that satisfies

$$\langle |v\rangle, X|u\rangle \rangle = \langle X^\dagger |v\rangle, |u\rangle \rangle \tag{2.17}$$

for all $|u\rangle \in \mathcal{H}_A$ and $|v\rangle \in \mathcal{H}_B$.

Remark 2.7. Since quantum systems are finite here, for a given linear operator $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, its adjoint operator $X^\dagger$ is uniquely determined. Indeed, taking alphabets Σ_A and Σ_B such that $\mathcal{H}_A = \mathbb{C}^{\Sigma_A}$ and $\mathcal{H}_B = \mathbb{C}^{\Sigma_B}$, for fixed bases $\{|a\rangle \in \mathcal{H}_A : a \in \Sigma_A\}$ and $\{|b\rangle \in \mathcal{H}_B : \Sigma_B\}$, the adjoint operator $X^\dagger$ of

$$X = \sum_{a,b} x_{ba} |b\rangle \langle a| \tag{2.18}$$

is given by

$$X^\dagger = \sum_{a,b} \overline{x}_{ba} |a\rangle \langle b|, \tag{2.19}$$

where $\overline{x}$ represents the complex conjugate of a complex number $x \in \mathbb{C}$.

Here, we introduce the *trace* function, which is a useful function to a linear operator to a real number.

Definition 2.8 (Trace)**.** Let $\mathcal{H}$ be a quantum system. The trace function $\mathrm{Tr} : \mathrm{L}(\mathcal{H}) \to \mathbb{C}$ is the unique linear function satisfying

$$\mathrm{Tr}(|v\rangle \langle u|) = \langle v|u\rangle \tag{2.20}$$

for all $|u\rangle, |v\rangle \in \mathcal{H}$.

Remark 2.9. Suppose that $\mathcal{H} = \mathbb{C}^{\Sigma}$ for some alphabet Σ. When we fix an orthonormal basis $\{|i\rangle \in \mathcal{H} : i \in \Sigma\}$, the trace of a linear operator $X \in \mathrm{L}(\mathcal{H})$ can be written as

$$\mathrm{Tr}(X) = \sum_{x \in \Sigma} X_{ii} = \sum_{x \in \Sigma} \langle i|X|i\rangle \,. \tag{2.21}$$

Note that the trace function does not depend on the choice of basis.

Using the trace function, we can introduce an inner product on $\mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ as

$$\langle X, Y\rangle \coloneqq \mathrm{Tr}(X^{\dagger}Y) \tag{2.22}$$

for all $X, Y \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$.

When two square operators are given, we can consider the *commutator* of two operators.

Definition 2.10 (Commutator)**.** Let $\mathcal{H}$ be a quantum system, and let $X, Y \in \mathrm{L}(\mathcal{H})$ be linear operators. The commutator of X and Y is defined as

$$[X, Y] \coloneqq XY - YX. \tag{2.23}$$

When $[X, Y] = 0$, X and Y are said to be commuting.

Remark 2.11. Since we consider a finite quantum system $\mathcal{H}$ here, two linear operators $X, Y \in \mathrm{L}(\mathcal{H})$ are commuting if and only if these two operators are simultaneously diagonalizable; i.e., X and Y are commuting if and only if there exists an orthonormal basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ of $\mathcal{H}$ and sets of complex numbers $\{x_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ and $\{y_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ such that

$$X = \sum_{i=1}^{D_{\mathcal{H}}} x_i \,|i\rangle \langle i| \,, \tag{2.24}$$

$$Y = \sum_{i=1}^{D_{\mathcal{H}}} y_i \,|i\rangle \langle i| \,. \tag{2.25}$$

In the following, we introduce several classes of linear operators. The first one is *normal operators*. A normal operator is defined as a square operator that is commuting with its adjoint operator.

Definition 2.12 (Normal Operators)**.** Let $\mathcal{H}$ be a quantum system. Then, a linear operator $N \in \mathrm{L}(\mathcal{H})$ is said to be a normal operator if

$$[N, N^{\dagger}] = NN^{\dagger} - N^{\dagger}N = 0. \tag{2.26}$$

As a special class of normal operators, we can consider *Hermitian operators*. A Hermitian operator is defined as a square operator whose adjoint operator is equal to itself.

Definition 2.13 (Hermitian Operators)**.** Let $\mathcal{H}$ be a quantum system. Then, an linear operator $O \in \mathrm{L}(\mathcal{H})$ is said to be a Hermitian operator if

$$O^\dagger = O. \tag{2.27}$$

By definition, it naturally holds that a Hermitian operator is a normal operator. In fact, for a Hermitian operator $O \in \mathrm{L}(\mathcal{H})$ it holds that $\langle u|O|u\rangle$ is a real number for all $|u\rangle \in \mathcal{H}$. Taking this fact into account, we move to a next class of linear operators, *positive operators*.

Definition 2.14 (Positive Operators)**.** Let $\mathcal{H}$ be a quantum system. Then, an linear operator $P \in \mathrm{L}(\mathcal{H})$ is said to be a positive (semi-definite) operator if

$$\langle u|P|u\rangle \geqq 0. \tag{2.28}$$

for all $|u\rangle \in \mathcal{H}$. In particular, if P satisfies that

$$\langle u|P|u\rangle > 0. \tag{2.29}$$

for all $|u\rangle \in \mathcal{H}$, P is said to be a positive definite operator.

By definition, we can see that a positive operator is trivially a Hermitian operator. Indeed, the meaning of the term *positive* is clear, compared with the property of Hermitian operators.

Now, we focus on special positive operators. The first one is *projection operators*. A projection operator works as a projection onto a subspace of a given quantum system.

Definition 2.15 (Projection Operators)**.** Let $\mathcal{H}$ be a quantum system. A positive operator $P \in \mathrm{L}(\mathcal{H})$ is said to be the projection operator if it holds that $P^2 = P$.

Let $\mathcal{G}$ is a subsystem of $\mathcal{H}$; that is, $\mathcal{G}$ is a Hilbert space included in $\mathcal{H}$. Then, the projection operator $P_\mathcal{G}$ corresponding to $\mathcal{G}$ is uniquely determined as the projection operator $P_\mathcal{G} \in \mathrm{L}(\mathcal{H})$ such that $\mathrm{im}(P_\mathcal{G}) = \mathcal{G}$.

Then, the next subclass of positive operators is *density operators*, which stipulate a *state* of a given quantum system.

Definition 2.16 (Density Operators)**.** Let $\mathcal{H}$ be a quantum system. Then, a positive operator on the system is said to be a density operator if it holds that

$$\mathrm{Tr}(\rho) = 1. \tag{2.30}$$

The set of density operators on the system $\mathcal{H}$ is denoted by $\mathrm{D}(\mathcal{H})$.

In this thesis, we call a density operator belonging to $\mathrm{D}(\mathcal{H})$ a *quantum state* on a system $\mathcal{H}$. When a quantum state $\rho \in \mathrm{D}(\mathcal{H})$ can be written as $\rho = |\psi\rangle\langle\psi|$ with some $|\psi\rangle$, we call ρ a *pure state*. In this thesis, we also refer to a normalized vector $|\psi\rangle \in \mathcal{H}$ as a pure state. In addition, when a basis of a given quantum system is fixed, a state with a diagonal matrix representation is called a *classical state* in the basis. In fact, when a system $\mathcal{H} = \mathbb{C}^{\Sigma}$ and an orthonormal basis of the system is given, we can construct a classical state from a probability distribution over Σ; for an orthonormal basis $\{|a\rangle : a \in \Sigma\}$ and a probability distribution $\{p_a : a \in \Sigma\}$, a quantum state

$$\rho \coloneqq \sum_{a\in\Sigma} p_a |a\rangle\langle a| \tag{2.31}$$

is classical. For example, considering an operator

$$\rho \coloneqq \frac{I_{\mathcal{H}}}{D_{\mathcal{H}}} \tag{2.32}$$

on a system $\mathcal{H}$, we can see that this is a quantum state; in particular, this is a classical state. Note that this state corresponds to the uniform distribution over $\{i : 1 \leqq i \leqq D_{\mathcal{H}}\}$. The state defined as Eq. (2.32) is called a *flat state*.

Here, let us move to another special class of linear operators that are not normal. *Isometry operators*, or more simply, *isometries*, are introduced as follows.

Definition 2.17 (Isometries)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. A linear operator $U \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ is said to be an isometry if it holds that

$$U^\dagger U = I_A. \tag{2.33}$$

The set of isometries from $\mathcal{H}_A$ to $\mathcal{H}_B$ is denoted by $\mathrm{U}(\mathcal{H}_A, \mathcal{H}_B)$. When $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}$, an isometry $U \in \mathrm{U}(\mathcal{H}_A, \mathcal{H}_B)$ is called a unitary operator, and we write $\mathrm{U}(\mathcal{H}_A, \mathcal{H}_B) = \mathrm{U}(\mathcal{H})$.

Isometry/unitary operators are considered as reversible operations for quantum states.

2.2.3 Decomposition of Linear Operators

Here, we overview several useful forms of linear operators. First, we review the *spectral decomposition* of normal operators.

Theorem 2.18 (Spectral Decomposition)**.** *Let $\mathcal{H}$ be a quantum system, and let $X \in \mathrm{L}(\mathcal{H})$ be a normal operator. Then, there exists a set of complex numbers $\{\lambda_i(X) \in \mathbb{C} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ and an orthonormal basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ of $\mathcal{H}$ such that*

$$X = \sum_{i=1}^{D_{\mathcal{H}}} \lambda_i(X) \, |i\rangle \langle i| \, . \tag{2.34}$$

The decomposition (2.34) is called a *spectral decomposition* of X. For each i, $\lambda_i(X)$ is called an eigenvalue of X, and $|i\rangle$ is called the eigenvector with the eigenvalue $\lambda_i(X)$.

Using the spectral decomposition, we give the definition of the support of a normal operator.

Definition 2.19 (Support)**.** Let $\mathcal{H}$ be a quantum system, and let $X \in \mathrm{L}(\mathcal{H})$ be a normal operator. The support of X, denoted by $\mathrm{supp}(X)$, is defined as a subspace of $\mathcal{H}$ spanned by the eivenvectors of X with nonzero eigenvalues. For a normal operator X with a spectral decomposition

$$X = \sum_{i=1}^{D_{\mathcal{H}}} \lambda_i(X) \, |i\rangle \langle i| \, , \tag{2.35}$$

the support of X is

$$\mathrm{supp}(X) = \mathrm{span}(\{|i\rangle : \lambda_i(X) \neq 0\}), \tag{2.36}$$

where $\mathrm{span}(\cdot)$ denotes the space spanned by a given set of vectors.

When a given operator is Hermitian or positive, we can say something more about the eigenvalues of the operator.

Lemma 2.20. *Let $\mathcal{H}$ be a quantum system, and let $O \in \mathrm{L}(\mathcal{H})$ be a Hermitian operator. Consider a spectral decomposition*

$$O = \sum_{i=1}^{D_{\mathcal{H}}} \lambda_i(O) \, |i\rangle \langle i| \, . \tag{2.37}$$

Then, all the eigenvalues $\{\lambda_i(O) \in \mathbb{C} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ are real numbers. Moreover, if O is a positive semi-definite operator, the eigenvalues are all non-negative; if O is a positive definite operator, the eigenvalues are all positive.

According to the Lemma 2.20, when $O \in \mathrm{L}(\mathcal{H})$ is a Hermitian operator, we can consider the following decomposition

$$O \coloneqq \sum_{i:\lambda_i(O)>0} \lambda_i(O) \left|i\right\rangle \left\langle i\right| - \sum_{i:\lambda_i(O)\leqq 0} \left|\lambda_i(O)\right| \left|i\right\rangle \left\langle i\right| \tag{2.38}$$

since the eigenvalues $\lambda_i(O)$ are real. Then, both of the first term and the second term are positive operators on $\mathcal{H}$, and they are orthogonal. Letting O_+ denote the first term and O_- denote the second term, we have

$$O = O_+ - O_-. \tag{2.39}$$

Theorem 2.21 (Jordan-Haan Decomposition)**.** *Let $\mathcal{H}$ be a quantum system, and let $O \in \mathrm{L}(\mathcal{H})$ be a Hermitian operator. Then, the Jordan-Haan decomposition of O is defined as the unique decomposition of O into two positive operators:*

$$O = O_+ - O_-, \tag{2.40}$$

where

$$O_+ \coloneqq \sum_{i:\lambda_i(O)>0} \lambda_i(O) \left|i\right\rangle \left\langle i\right| \tag{2.41}$$

$$O_- \coloneqq \sum_{i:\lambda_i(O)\leqq 0} \left|\lambda_i(O)\right| \left|i\right\rangle \left\langle i\right|. \tag{2.42}$$

The expressions considered above are only for normal operators or Hermitian operators. Here, we give a useful decomposition for general linear operators, which is called the *singular value decomposition*.

Theorem 2.22 (Singular Value Decomposition)**.** *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be quantum systems. Let $X \in \mathrm{L}(\mathcal{H}_1, \mathcal{H}_2)$ be a linear operator. Then, there exists an orthonormal basis $\{\left|u_i\right\rangle \in \mathcal{H}_1 : 1 \leqq i \leqq D_{\mathcal{H}_1}\}$ of $\mathcal{H}_1$, an orthonormal basis $\{\left|v_i\right\rangle \in \mathcal{H}_2 : 1 \leqq i \leqq D_{\mathcal{H}_2}\}$ of $\mathcal{H}_2$, a positive integer $1 \leqq r \leqq \min\{D_{\mathcal{H}_1}, D_{\mathcal{H}_2}\}$, and a set of positive real numbers $\{s_i(X) : 1 \leqq i \leqq r\}$ such that*

$$X = \sum_{i=1}^{r} s_i(X) \left|v_i\right\rangle \left\langle u_i\right|. \tag{2.43}$$

The decomposition (2.43) is called a singular value decomposition of X. The positive real numbers $s_i(X)$ are called the singular values of X. In this thesis, we take a singular value decomposition such that $s_1(X) \geqq s_2(X) \geqq \cdots \geqq s_r(X)$.

Remark 2.23 (Matrix Representation of Singular Value Decomposition). Let $\mathcal{H}_A = \mathbb{C}^n$ and $\mathcal{H}_B = \mathbb{C}^m$ quantum systems for some positive integers n and m. Suppose that a singular value decomposition of X is given as

$$X = \sum_{i=1}^{r} s_i(X) \left|v_i\right\rangle \left\langle u_i\right| \tag{2.44}$$

for some orthogonal bases $\{|u_i\rangle \in \mathcal{H}_A : 1 \leqq i \leqq n\}$ and $\{|v_i\rangle \in \mathcal{H}_B : 1 \leqq i \leqq n\}$. Let us fix an orthonormal basis $\{|i\rangle \in \mathbb{C}^r : 1 \leqq i \leqq r\}$ of $\mathbb{C}^r$. Then, the singular value decomposition in this matrix representation is given by

$$X = WSV^\dagger, \tag{2.45}$$

where

$$W \coloneqq \sum_{i=1}^{r} \left|v_i\right\rangle \left\langle i\right|, \tag{2.46}$$

$$V \coloneqq \sum_{i=1}^{r} \left|u_i\right\rangle \left\langle i\right|, \tag{2.47}$$

$$S \coloneqq \sum_{i=1}^{r} s_i(X) \left|i\right\rangle \left\langle i\right|. \tag{2.48}$$

Here, $W \in \mathrm{U}(\mathbb{C}^r, \mathcal{H}_B)$ and $V \in \mathrm{U}(\mathbb{C}^r, \mathcal{H}_A)$ are isometries, and $S \in \mathrm{L}(\mathbb{C}^r)$ is a positive definite operator that is diagonal in the given basis.

Finally, we introduce a polar decomposition of a square operator.

Theorem 2.24 (Polar Decomposition)**.** *Let $\mathcal{H}$ be a quantum system, and let $X \in \mathrm{L}(\mathcal{H})$ be a square operator. Then, there exist a unitary operator $U \in \mathrm{U}(\mathcal{H})$ and a positive operator $P \in \mathrm{L}(\mathcal{H})$ such that*

$$X = UP. \tag{2.49}$$

Considering the discussion in Remark 2.23, when we have a singular value decomposition $X = WSV^\dagger$, we can construct a polar decomposition by taking

$$U = WV^\dagger, \tag{2.50}$$

$$P = VSV^\dagger. \tag{2.51}$$

2.3 Quantum Channels

In this section, we briefly overview the definition and basic properties of quantum channels. Intuitively, a quantum channel is an operation that converts a given quantum state into some quantum state. More formally, a quantum channel is a linear function of linear operators that maps a density operator to a density operator.

First, we review the definition of linear maps.

Definition 2.25 (Linear Maps)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. Then, a map $\mathcal{N} : \mathrm{L}(\mathcal{H}_A) \to \mathrm{L}(\mathcal{H}_B)$ is said to be a linear map if it holds that

$$\mathcal{N}(\alpha X + \beta Y) = \alpha \mathcal{N}(X) + \beta \mathcal{N}(Y) \tag{2.52}$$

for all $X, Y \in \mathrm{L}(\mathcal{H}_A)$ and $\alpha, \beta \in \mathbb{C}$. The set of linear maps from $\mathrm{L}(\mathcal{H}_A)$ to $\mathrm{L}(\mathcal{H}_B)$ is denoted by $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$. When $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}$, we write $\mathcal{L}(\mathcal{H}) \coloneqq \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ for brevity.

As an special example of linear maps, we can consider the *identity map*, which does nothing to an input.

Definition 2.26 (Identity Map)**.** Let $\mathcal{H}_A$ be a quantum system. The identity map $\mathbb{1}_{\mathcal{H}_A} \in \mathcal{L}(\mathcal{H}_A)$ is defined as the unique linear map satisfying

$$\mathbb{1}_{\mathcal{H}_A}(X) = X \tag{2.53}$$

for all $X \in \mathrm{L}(\mathcal{H})$. We also write $\mathbb{1}_A \coloneqq \mathbb{1}_{\mathcal{H}_A}$ for convenience.

When we are given a state on some composite system, we can consider a *composite map* on the system.

Definition 2.27 (Composite Maps)**.** Let $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}'_1$, and $\mathcal{H}'_2$ be quantum systems. Let $\mathcal{N}_1 \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}'_1)$ and $\mathcal{N}_2 \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}'_2)$ be linear maps. Then, the composite map $\mathcal{N}_1 \otimes \mathcal{N}_2 \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{H}'_1 \otimes \mathcal{H}'_2)$ is defined as the unique linear map satisfying

$$(\mathcal{N}_1 \otimes \mathcal{N}_2)(X_1 \otimes X_2) = \mathcal{N}_1(X_1) \otimes \mathcal{N}_2(X_2) \tag{2.54}$$

for all $X_1 \in \mathrm{L}(\mathcal{H}_1)$ and $X_2 \in \mathrm{L}(\mathcal{H}_2)$.

Extending the trace function (Definition 2.8), we give the definition of the *partial trace*, which is defined as a composite map of the trace and the identity map. The partial trace takes the trace over a subsystem of a given composite system.

Definition 2.28 (Partial Trace)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_E$ be quantum systems. The partial trace $\mathrm{Tr}_E : \mathrm{L}(\mathcal{H}_A \otimes \mathcal{H}_E) \to \mathrm{L}(\mathcal{H}_A)$ is defined as the composite map

$$\mathrm{Tr}_E \coloneqq \mathbb{1}_A \otimes \mathrm{Tr}, \tag{2.55}$$

where Tr is the trace on the system $\mathcal{H}_E$.

Here, We review two important subclasses of linear maps, which are essential to consider quantum channels. The first one is *completely positive maps*.

Definition 2.29 (Completely Positive Maps)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $\mathcal{N} \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ be a linear map. Then, $\mathcal{N}$ is said to be completely positive if $(\mathcal{N} \otimes \mathbb{1}_E)(P)$ is positive for all quantum systems $\mathcal{H}_E$ and all positive operators P on $\mathcal{H}_A \otimes \mathcal{H}_E$.

The second one is *trace-preserving maps*.

Definition 2.30 (Trace-Preserving Maps)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $\mathcal{N} \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ be a linear map. Then, $\mathcal{N}$ is said to be trace-preserving if it holds that

$$\mathrm{Tr}(\mathcal{N}(X)) = \mathrm{Tr}(X) \tag{2.56}$$

for all $X \in \mathrm{L}(\mathcal{H}_A)$.

Now, we give the formal definition of quantum channels.

Definition 2.31 (Quantum Channels)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. A linear map $\mathcal{N} \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ is said to be a quantum channel if it satisfies the following conditions:

1. The map $\mathcal{N}$ is linear;
2. The map $\mathcal{N}$ is completely positive;
3. The map $\mathcal{N}$ is trace-preserving.

In other words, a quantum channel is a completely positive and trace-preserving (CPTP) linear map. The set of quantum channels from $\mathrm{L}(\mathcal{H}_A)$ to $\mathrm{L}(\mathcal{H}_B)$ is denoted by $\mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$. When $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}$, we write $\mathrm{C}(\mathcal{H}) \coloneqq \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$ for brevity.

There are several ways to explicitly characterize a quantum channel. Here, we review two representations of quantum channels, which are known as *Stinespring representation* and *Kraus representation*.

Proposition 2.32 (Stinespring Representations)**.** *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be finite-dimensional quantum systems. Let $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$ be a quantum channel. Then, there exists a quantum system $\mathcal{H}_E$, some reference state $\omega_E \coloneqq |\omega\rangle\langle\omega|_E \in \mathrm{D}(\mathcal{H}_E)$, and an isometry $U \in \mathrm{U}(\mathcal{H}_A \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_B \otimes \mathcal{H}_E)$ such that*

$$\mathcal{N}(X) = \mathrm{Tr}_E(U(X \otimes \omega_E)U^\dagger) \tag{2.57}$$

for all $X \in \mathrm{L}(\mathcal{H}_A)$, where $\mathcal{H}_E^{(\omega)}$ is the subspace of $\mathcal{H}_E$ spanned by $\{|\omega\rangle_E\}$. In this thesis, we call the isometry U a Stinespring isometry of the channel $\mathcal{N}$.

Proposition 2.33 (Kraus Representations)**.** *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. Let $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$ be a quantum channel. Then, there exists an alphabet Σ and a collection of linear operators $\{E_a \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B) : a \in \Sigma\}$ such that*

$$\mathcal{N}(X) = \sum_{a\in\Sigma} E_a X E_a^\dagger, \tag{2.58}$$

$$\sum_{a\in\Sigma} E_a^\dagger E_a = I_A. \tag{2.59}$$

In this thesis, we call each operator E_a a Kraus operator for the channel $\mathcal{N}$.

When we are given a Kraus representation of a quantum channel $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$, we can construct a Stinespring representation and vice versa. Suppose that we have a Kraus representation

$$\{E_a \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B) : a \in \Sigma\} \tag{2.60}$$

of a quantum channel $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$. Then, for a fixed basis $\{|a\rangle \in \mathbb{C}^\Sigma : a \in \Sigma\}$, we can construct a Stinespring representation

$$U = \sum_{a\in\Sigma} E_a \otimes |a\rangle\langle\omega|_E \in \mathrm{U}(\mathcal{H}_A \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_B \otimes \mathcal{H}_E). \tag{2.61}$$

On the other hand, suppose that we have a Stinespring representation $U \in \mathrm{U}(\mathcal{H}_A \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_B \otimes \mathcal{H}_E)$ with some reference state ω_E. Then, for a fixed basis $\{|a\rangle : a \in \Sigma\}$ of $\mathcal{H}_E$, we have a decomposition of U,

$$U = \sum_{a\in\Sigma} E_a \otimes |a\rangle\langle\omega|_E, \tag{2.62}$$

such that $E_a \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ for all $a \in \Sigma$. Then,

$$\{E_a \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B) : a \in \Sigma\} \tag{2.63}$$

forms a Kraus representation of $\mathcal{N}$.

Finally, we show an example of quantum channels.

Example 2.1 (Dephasing Channels)**.** Let $\mathcal{H}$ be a quantum system. Fix an orthonormal basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ of the system $\mathcal{H}$. Then, the dephasing channel $\Delta \in \mathrm{C}(\mathcal{H})$ with respect to this basis is defined as

$$\Delta(X) \coloneqq \sum_{i=1}^{D_{\mathcal{H}}} X_{ii} \, |i\rangle \langle i| \tag{2.64}$$

for all $X \in \mathrm{L}(\mathcal{H})$ with matrix representation

$$X = \sum_{i,j=1}^{D_{\mathcal{H}}} X_{ij} \, |i\rangle \langle j| \, . \tag{2.65}$$

A dephasing channel is considered as an operation that erases all the off diagonal elements of an input with respect to a given basis. The dephasing channel defined in Eq. (2.64) is indeed a quantum channel; its Kraus representation can be given as $\{|i\rangle \langle i| : 1 \leqq i \leqq D_{\mathcal{H}}\}$.

2.4 Norms and Quantum Entropy

In this section, we introduce norms of linear operators and quantum entropic quantities of quantum states.

2.4.1 Norms of Linear Operators

First, to introduce concepts of *size* and *distance* to linear operators, we consider *norm* of linear operators.

Definition 2.34 (Norm)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. A *norm* on a space $\mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ is a function $\| \cdot \| : \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B) \to \mathbb{R}$ satisfying the following three conditions.

1. For all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that $\|X\| \geqq 0$. We have $\|X\| = 0$ if and only if $X = 0$.

2. For all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ and $\alpha \in \mathbb{C}$, it holds that $\|\alpha X\| = |\alpha| \|X\|$.

3. For all $X, Y \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that $\|X + Y\| \leqq \|X\| + \|Y\|$.

In this thesis, in particular, we use the *Schatten p-norms*.

Definition 2.35 (Schatten p-norms)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ are quantum systems, and let $p \geqq 1$ be a positive real number. Then, the Schatten p-norm $\|\cdot\|_p$ on a space $\mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ is defined as

$$\|X\|_p \coloneqq \left(\mathrm{Tr}\left((X^\dagger X)^{\frac{p}{2}}\right)\right)^{\frac{1}{p}} \tag{2.66}$$

for all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$.

In fact, using a singular value decomposition of $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$

$$X = \sum_{i=1}^{r} s_i(X) \left|v_i\right\rangle \left\langle u_i\right|, \tag{2.67}$$

it holds that

$$\|X\|_p = \left(\sum_{i=1}^{r} s_i(X)^p\right)^{\frac{1}{p}}. \tag{2.68}$$

By taking a limit, we can also consider the Schatten ∞-norm.

Definition 2.36 (Schatten ∞-norm)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ are quantum systems. Then, the Schatten ∞-norm $\|\cdot\|_p$ on a space $\mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ is defined as a limit of p-norm:

$$\|X\|_\infty \coloneqq \lim_{p \to \infty} \|X\|_p = s_1(X) \tag{2.69}$$

for all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$.

We mainly use the Schatten 1-, 2-, and ∞-norms. The Schatten 1-norm is also called the *trace norm*, and we have

$$\|X\|_1 = \sum_{i=1}^{r} s_i(X). \tag{2.70}$$

The trace norm has a property called *monotonicity*, which is also referred to as *data processing inequality* of the trace distance.

Lemma 2.37. *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ are quantum systems, let $X \in \mathrm{L}(\mathcal{H}_A)$ be a linear operator, and let $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_B)$ be a quantum channel. Then, it holds that*

$$\|\mathcal{N}(X)\|_1 \leqq \|X\|_1. \tag{2.71}$$

This Lemma ensures that the application of a quantum channel to a linear operator does not increase the trace norm of the operator.

The Schatten 2-norm is also called the *Frobenius norm*, and we have

$$\|X\|_2 = \sqrt{\langle X, X \rangle} = \sqrt{\sum_{i=1}^{r} s_i(X)^2}. \tag{2.72}$$

Indeed, for a matrix representation

$$X = \sum_{\substack{a \in \Sigma_A \\ b \in \Sigma_B}} X_{ba} \, |b\rangle \langle a| \tag{2.73}$$

of $X \in \mathrm{L}(\mathbb{C}^{\Sigma_A}, \mathbb{C}^{\Sigma_B})$, we have

$$\|X\|_2 = \sqrt{\sum_{\substack{a \in \Sigma_A \\ b \in \Sigma_B}} |X_{ba}|^2}. \tag{2.74}$$

Here, we give the basic properties of the Schatten norms. For more detailes, see textbooks such as Refs. [59–61].

Proposition 2.38. *Let $\mathcal{H}_A$, $\mathcal{H}_B$, $\mathcal{H}_C$, and $\mathcal{H}_D$ be quantum systems. Let $p, q \in [1, \infty]$.*

1. *For all $1 \leqq p \leqq q \leqq \infty$ and for all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that*

$$\|X\|_p \geqq \|X\|_q. \tag{2.75}$$

2. *For all $X \in \mathrm{L}(\mathcal{H}_B, \mathcal{H}_C)$, $U \in \mathrm{U}(\mathcal{H}_C, \mathcal{H}_D)$, and $V \in \mathrm{U}(\mathcal{A}_A, \mathcal{H}_B)$, it holds that*

$$\|UXV\|_p = \|X\|_p. \tag{2.76}$$

3. *Suppose that it holds that*

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{2.77}$$

Then, for all $X, Y \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that

$$|\langle X, Y \rangle| \leqq \|X\|_p \|Y\|_q. \tag{2.78}$$

4. For all $X \in \mathrm{L}(\mathcal{H}_C, \mathcal{H}_D)$, $Y \in \mathrm{L}(\mathcal{H}_B, \mathcal{H}_C)$, and $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that

$$\|XYZ\|_p \leqq \|X\|_\infty \|Y\|_p \|Z\|_\infty. \tag{2.79}$$

5. For all $X \in \mathrm{L}(\mathcal{H}_B, \mathcal{H}_C)$ and $Y \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that

$$\|XY\|_p \leqq \|X\|_p \|Y\|_p. \tag{2.80}$$

6. For all $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$, it holds that

$$\|X^\dagger\|_p = \|X\|_p. \tag{2.81}$$

2.4.2 Operator Functions and Quantum Entropy

Here, we review the quantum entropy. To define the quantum entropy, an extension of a complex-valued function called *operator function* is needed.

Definition 2.39 (Operator Functions)**.** Let $f : \mathbb{C} \to \mathbb{C}$ be a function. Then, the extension of the function f to a space of normal operators is given by

$$f(X) \coloneqq \sum_{a \in \Sigma} f(\lambda_a(X)) \, |a\rangle \langle a| \tag{2.82}$$

for all quantum systems $\mathcal{H} = \mathbb{C}^\Sigma$ and all normal operator $X \in \mathrm{L}(\mathcal{H})$ with the spectral decomposition

$$X \coloneqq \sum_{a \in \Sigma} \lambda_a(X) \, |a\rangle \langle a| \, . \tag{2.83}$$

Using the extension of $\log_2$, we can define the quantum entropy, which is a quantum variant of the Shannon entropy.

Definition 2.40 (Quantum Entropy)**.** Let $\mathcal{H}_A$ be a quantum system, and let $\rho \in \mathrm{D}(\mathcal{H}_A)$ be a state on the system. Then, the entropy of the state ρ is defined as

$$\mathrm{H}(A)_\rho \coloneqq \mathrm{H}(\rho) \coloneqq -\mathrm{Tr}[\rho \log_2 \rho]. \tag{2.84}$$

Indeed, letting

$$\rho = \sum_{i=1}^{D_A} \rho_i \, |i\rangle \langle i| \tag{2.85}$$

be a spectral decomposition of ρ, the quantum entropy is given as

$$H(\rho) = -\sum_{i=1}^{D_A} \rho_i \log_2(\rho_i). \tag{2.86}$$

When a given quantum system is a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, we can consider the quantum entropy with respect to each subsystem. For example, if the state of the system is $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the quantum entropy with respect to the system A is

$$\mathrm{H}(A)_\rho \coloneqq -\mathrm{Tr}(\rho_A \log_2 \rho_A), \tag{2.87}$$

where $\rho_A \coloneqq \mathrm{Tr}_A(\rho)$.

Here, we introduce the binary entropy function, which can be interpreted as a special example of the quantum entropy.

Definition 2.41 (Binary Entropy Function)**.** The binary entropy function $h_2 : [0,1] \to \mathbb{R}$ is a real-valued function defined as

$$h_2(x) \coloneqq -x \log_2 x - (1-x) \log_2(1-x) \tag{2.88}$$

for $x \in [0,1]$. Here, we regard $0 \log_2 0 = 0$.

Considering the definition of the quantum entropy, for a given $x \in [0,1]$, the binary entropy $h_2(x)$ can be regarded as the entropy of a classical state

$$\rho \coloneqq x \left|0\right\rangle \left\langle 0\right| + (1-x) \left|1\right\rangle \left\langle 1\right|. \tag{2.89}$$

Using the definition of the quantum entropy, we can define other entropic quantities. First, we review the quantum conditional entropy, which is a quantum variant of the conditional entropy.

Definition 2.42 (Quantum Conditional Entropy)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a state on the composite system. Then, the quantum conditional entropy $\mathrm{H}(A|B)_\rho$ conditioned on B is defined as

$$\mathrm{H}(A|B)_\rho \coloneqq \mathrm{H}(AB)_\rho - \mathrm{H}(B)_\rho. \tag{2.90}$$

Furthermore, we can consider *quantum mutual information* in a similar manner.

Definition 2.43 (Quantum Mutual Information)**.** Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems, and let $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a state on the composite system. Then, the quantum mutual information is defined as

$$I(A:B) \coloneqq I(\mathcal{H}_A : \mathcal{H}_B) \coloneqq \mathrm{H}(A)_\rho + \mathrm{H}(B)_\rho - \mathrm{H}(AB)_\rho. \tag{2.91}$$

By definition, we can also write the quantum mutual information $I(A:B)_\rho$ as

$$\begin{aligned} I(A:B)_\rho &= \mathrm{H}(A)_\rho - \mathrm{H}(A|B)_\rho \\ &= \mathrm{H}(B)_\rho - \mathrm{H}(B|A)_\rho. \end{aligned} \tag{2.92}$$

Finally, we introduce *quantum conditional mutual information* defined as follows.

Definition 2.44 (Quantum Conditional Mutual Information)**.** Let $\mathcal{H}_A$, $\mathcal{H}_{A'}$, $\mathcal{H}_B$, $\mathcal{H}_{B'}$, and $\mathcal{H}_C$ be quantum systems, and let $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be a state on the composite system. Then, the quantum conditional mutual information $I(A:B|C)_\rho$ conditioned on C is defined as

$$I(A:B|C) \coloneqq I(\mathcal{H}_A : \mathcal{H}_B|\mathcal{H}_C) \coloneqq I(AC:B) - I(C:B). \tag{2.93}$$

Now, we state several useful properties of quantum entropic quantities [59–61].

Proposition 2.45. *Let $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_C$ be quantum systems. Then, the following statements hold.*

1. *For all states $\rho \in \mathrm{D}(\mathcal{H}_A)$, the quantum entropy is bounded as*

$$0 \leqq \mathrm{H}(A)_\rho \leqq \log_2 D_A. \tag{2.94}$$

2. *For all states $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the quantum mutual information is bounded as*

$$0 \leqq I(A:B)_\rho \leqq 2\log_2[\min\{D_A, D_B\}]. \tag{2.95}$$

In particular, if system A is classical; that is, if ρ can be written as

$$\rho = \sum_{i=1}^{D_A} p_i \left|i\right\rangle \left\langle i\right|_A \otimes \rho_B^{(i)}, \tag{2.96}$$

it holds that

$$0 \leqq I(A:B)_\rho \leqq \log_2[\min\{D_A, D_B\}] \tag{2.97}$$

3. *For all states* $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$*, the quantum mutual information is bounded as*

$$0 \leqq I(A:B|C)_\rho \leqq 2\log_2[\min\{D_A, D_B\}]. \tag{2.98}$$

In particular, if a system A *is classical; that is, if* ρ *can be written as*

$$\rho = \sum_{i=1}^{D_A} p_i \left|i\right\rangle \left\langle i\right|_A \otimes \rho_{BC}^{(i)}, \tag{2.99}$$

it holds that

$$0 \leqq I(A:B)_\rho \leqq \log_2 D_A. \tag{2.100}$$

4. *For all states* $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ *and for all quantum channels* $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_{A'})$ *and* $\mathcal{M} \in \mathrm{C}(\mathcal{H}_B, \mathcal{H}_{B'})$*, it holds that*

$$I(A':B')_{\mathcal{N}\otimes\mathcal{M}(\rho)} \leqq I(A:B)_\rho. \tag{2.101}$$

5. *For all states* $\rho \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ *and for all quantum channels* $\mathcal{N} \in \mathrm{C}(\mathcal{H}_A, \mathcal{H}_{A'})$ *and* $\mathcal{M} \in \mathrm{C}(\mathcal{H}_B, \mathcal{H}_{B'})$*, it holds that*

$$I(A':B'|C)_{\mathcal{N}\otimes\mathcal{M}(\rho)} \leqq I(A:B|C)_\rho. \tag{2.102}$$

These entropic quantities have appeared as optimal rates or costs of various quantum information processing tasks. The optimality shown in previous research can be considered as operational interpretations of these quantities. For example, the quantum entropy is operationally interpreted as the optimal rate of the quantum data compression of pure-state quantum ensembles [16]. In addition, the conditional quantum entropy is regarded as the optimal rate of the state merging [45], and the conditional quantum mutual information is related to the optimal rate of quantum state redistribution [47].

2.5 Structure of Quantum Ensembles

Here, we review the definition of quantum ensembles and overview a structure of quantum ensembles.

Definition 2.46 (Quantum Ensembles)**.** Let $\mathcal{H}$ be a quantum system, and let Σ be an alphabet. A quantum ensemble Φ is a set of pairs of a positive real number and a quantum state

$$\Phi \coloneqq \{(p_a, \rho_a) \in \mathbb{R} \times \mathrm{D}(\mathcal{H}) : a \in \Sigma\}, \tag{2.103}$$

where $\{p_a : a \in \Sigma\}$ forms a probability distribution; that is,

$$0 \leqq p_a \leqq 1 \tag{2.104}$$

for all $a \in \Sigma$ and

$$\sum_{a \in \Sigma} p_a = 1. \tag{2.105}$$

We also write

$$\Phi = \{p_a, \rho_a\}_{a \in \Sigma}. \tag{2.106}$$

In addition, the average state ρ_Φ of a quantum ensemble $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$ is defined as

$$\rho_\Phi \coloneqq \sum_{a \in \Sigma} p_a \rho_a. \tag{2.107}$$

Here, we define a special class of quantum ensembles, which is called *classical ensembles*.

Definition 2.47 (Classical Ensembles)**.** Let $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$ be a quantum ensemble. The ensemble Φ is said to be *classical ensemble* if ρ_a are commuting to each other for all $a \in \Sigma$.

Remark 2.48. When $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$ is a classical ensemble, ρ_a are simultaneously diagonalizable for all $a \in \Sigma$; that is, there exists an orthonormal basis of $\mathcal{H}$ such that for all $a \in \Sigma$, the basis gives a spectral decomposition of ρ_a. Therefore, a classical ensemble is a quantum ensemble composed of classical states for some orthonormal basis.

For a quantum ensemble $\{p_a, \rho_a\}_{a \in \Sigma}$, if ρ_a is pure for all $a \in \Sigma$, the ensemble is called a *pure-state ensemble*. When a pure-state ensemble $\{p_a, \rho_a\}_{a \in \Sigma}$ is given, taking vectors $|\psi_a\rangle$ such that $\rho_a = |\psi_a\rangle \langle\psi_a|$, we also write $\{p_a, |\psi_a\rangle\}_{a \in \Sigma}$.

Suppose that we have a quantum ensemble $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$. Reference [4] gave a structure of a given quantum ensemble, which is called the *Koashi-Imoto (KI) decomposition* or the *KI structure*. Intuitively, when we have a quantum ensemble, we can decompose each state in the ensemble into the following three parts: *classical parts*, *non-redundant quantum parts*, and *redundant parts*. Here, we show the formal statement as a theorem.

Theorem 2.49 (KI Decomposition [4])**.** *Let $\mathcal{H}$ be a quantum system. Let $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$ be an ensemble on the system; that is, for all $a \in \Sigma$, $\rho_a \in \mathrm{D}(\mathcal{H})$. Then, there exists a unique decomposition of the quantum system*

$$\mathcal{H} \coloneqq \bigoplus_{l \in \Xi} \mathcal{H}_Q^{(l)} \otimes \mathcal{H}_R^{(l)} \tag{2.108}$$

and the corresponding isometry

$$\Gamma_\Phi \in \mathrm{U}\left(\mathcal{H}, \bigoplus_l \mathcal{H}_Q^{(l)} \otimes \mathcal{H}_R^{(l)}\right) \tag{2.109}$$

satisfying the following conditions:

1. *For all* $a \in \Sigma$,

$$\Gamma_\Phi \rho_a \Gamma_\phi^\dagger = \bigoplus_{l \in \Xi} q^{(a,l)} \rho_Q^{(a,l)} \otimes \rho_R^{(l)}. \tag{2.110}$$

   Here $q^{(a,l)} : l \in \Xi$ *forms a probability distribution over labels* $l \in \Xi$ *for all* $a \in \Sigma$, $\rho_Q^{(a,l)} \in \mathrm{D}(\mathcal{H}_Q^l)$ *is a density operator on system* $\mathcal{H}_Q^l$, *depending on both* $a \in \Sigma$ *and* $l \in \Xi$, *and* $\rho_R^{(l)} \in \mathrm{D}(\mathcal{H}_R^l)$ *is a density operator on system* $\mathcal{H}_R^l$, *which is independent of* $a \in \Sigma$.

2. *For all* $l \in \Xi$, *if a projection operator* $P : \mathcal{H}_Q^{(l)} \to \mathcal{H}_Q^{(l)}$ *satisfies*

$$P q^{(a,l)} \rho_Q^{(a,l)} = q^{(a,l)} \rho_Q^{(a,l)} P \tag{2.111}$$

   for all $a \in \Sigma$, *then* $P = I_{\mathcal{H}_Q^{(l)}}$ *or* $P = 0$.

3. *For all* $l, l' \in \Xi$ *such that* $l \neq l'$, *there exists no isometry* $V \in \mathrm{U}(\mathcal{H}_Q^{(l)}, \mathcal{H}_Q^{(l')})$ *such that*

$$V q^{(a,l)} \rho_Q^{(a,l)} = \alpha q^{(a,l')} \rho_Q^{(a,l')} V \tag{2.112}$$

   with some positive real number α *for all* $a \in \Sigma$.

The first statement in Theorem 2.49 shows the form of the KI decomposition. The second and third statements ensure that the decomposition (2.110) is maximal; that is, we cannot further refine the structure. Indeed, the second one states that we cannot further decompose each block of the KI decomposition; the third one means that we cannot relate a block of the decomposition to another block. In the decomposition shown in Eq. (2.110), we can observe that all the quantum states in the given ensemble Φ can be decomposed in a block-diagonal structure. Note that $\rho_R^{(l)} \in \mathrm{D}(\mathcal{H}_R^{(l)})$ does not depend on $a \in \Sigma$; thus it is called redundant because it does not contain the information of the label $a \in \Sigma$. Hereafter, when we specify the KI decomposition of a given ensemble $\Phi = \{p_a, \rho_a\}_{a \in \Sigma}$, we may omit Γ_Φ and write

$$\rho_a = \bigoplus_{l \in \Xi} q^{(a,l)} \rho_Q^{(a,l)} \otimes \rho_R^{(l)} \tag{2.113}$$

for brevity.

Once the KI decomposition of a given ensemble is obtained, we can define two quantum channels that correspond to the operation taking off the ensemble's redundant parts and the operation putting on the redundant parts respectively.

Theorem 2.50. *Let $\mathcal{H}$ be a quantum system. Let $\Phi = \{p_a, \rho_a\}_{a\in\Sigma}$ be an ensemble on the system; that is, for all $a \in \Sigma$, $\rho_a \in \mathrm{D}(\mathcal{H})$. Consider the KI decomposition of Φ:*

$$\mathcal{H} \coloneqq \bigoplus_{l\in\Xi} \mathcal{H}_Q^{(l)} \otimes \mathcal{H}_R^{(l)} \tag{2.114}$$

such that

$$\rho_a = \bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)} \otimes \rho_R^{(l)}. \tag{2.115}$$

Then, there exist quantum channels $\mathcal{K}_{\mathrm{off}}$ and $\mathcal{K}_{\mathrm{on}}$ such that

$$\mathcal{K}_{\mathrm{off}}(\rho_a) = \bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)}, \tag{2.116}$$

$$\mathcal{K}_{\mathrm{on}}\left(\bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)}\right) = \rho_a \tag{2.117}$$

for all $a \in \Sigma$.

Proof. We can give these two quantum channels explicitly by using Kraus representations. First, $\mathcal{K}_{\mathrm{off}}$ is given by Kraus operators

$$A_{j_l}^{(l)} \coloneqq I_{\mathcal{H}_Q^{(l)}} \otimes \langle j_l| \,, \tag{2.118}$$

where $\{|j_l\rangle : j_l\}$ is a orthonormal basis of $\mathcal{H}_Q^{(l)}$ for all $l \in \Xi$. Kraus operators $A_{j_l}^{(l)}$ apply to the lth block, and $\{A_{j_l}^{(l)} : j_l\}$ forms the partial trace over $\mathcal{H}_R^{(l)}$. Indeed, we have that

$$\begin{aligned}
\sum_{l\in\Xi}\sum_{j_l} A_{j_l}^{(l)} \rho_a (A_{j_l}^{(l)})^\dagger &= \bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)} \otimes \left(\sum_{j_l} \langle j_l|\rho_R^{(l)}|j_l\rangle\right) \\
&= \bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)} \otimes \mathrm{Tr}(\rho_R^{(l)}) \\
&= \bigoplus_{l\in\Xi} q^{(a,l)} \rho_Q^{(a,l)},
\end{aligned} \tag{2.119}$$

for all $a \in \Sigma$. In addition, $\{A_{j_l}^{(l)} : l \in \Xi, j_l\}$ forms a valid Kraus representation because it holds that

$$\begin{aligned}\sum_{l\in\Xi}\sum_{j_l}(A_{j_l}^{(l)})^\dagger A_{j_l}^{(l)} &= \sum_{l\in\Xi}\sum_{j_l} I_{\mathcal{H}_Q^{(l)}} \otimes |j_l\rangle\langle j_l| \\ &= \sum_{l\in\Xi} I_{\mathcal{H}_Q^{(l)}} \otimes \left(\sum_{j_l} |j_l\rangle\langle j_l|\right) \\ &= \sum_{l\in\Xi} I_{\mathcal{H}_Q^{(l)}} \otimes I_{\mathcal{H}_R^{(l)}} \\ &= I_{\mathcal{H}}.\end{aligned} \tag{2.120}$$

Therefore, we can construct $\mathcal{K}_{\text{off}}$ by a Kraus representation

$$\{A_{j_l}^{(l)} : l \in \Xi, j_l\}. \tag{2.121}$$

Next, $\mathcal{K}_{\text{on}}$ is given by Kraus operators

$$A_{k_l}^{(l)} \coloneqq I_{\mathcal{H}_Q^{(l)}} \otimes \sqrt{r_{k_l}}\,|k_l\rangle\,, \tag{2.122}$$

where $\{|k_l\rangle : k_l\}$ is an orthonormal basis of $\mathcal{H}_Q^{(l)}$ for all $l \in \Xi$ corresponding to a spectral decomposition

$$\rho_R^{(l)} \coloneqq \sum_{k_l} r_{k_l}\,|k_l\rangle\langle k_l| \tag{2.123}$$

with eigenvalues $r_{k_l} \geqq 0$. Kraus operators $A_{k_l}^{(l)}$ apply to the lth block, and $\{A_{k_l}^{(l)} : k_l\}$ forms the construction of $\rho_R^{(l)}$ on system $\mathcal{H}_R^{(l)}$. Indeed, we have that

$$\begin{aligned}\sum_{l\in\Xi}\sum_{k_l} A_{k_l}^{(l)}\left(\bigoplus_{l\in\Xi} q^{(a,l)}\rho_Q^{(a,l)}\right)(A_{k_l}^{(l)})^\dagger &= \bigoplus_{l\in\Xi} q^{(a,l)}\rho_Q^{(a,l)} \otimes \left(\sum_{k_l} r_{k_l}\,|k_l\rangle\langle k_l|\right) \\ &= \bigoplus_{l\in\Xi} q^{(a,l)}\rho_Q^{(a,l)} \otimes \rho_R^{(l)} \\ &= \rho_a\end{aligned} \tag{2.124}$$

for all $a \in \Sigma$. In addition, $\{A_{k_l}^{(l)} : l \in \Xi, k_l\}$ forms a valid Kraus representation because it

holds that

$$\begin{aligned}\sum_{l\in\Xi}\sum_{k_l}(A_{k_l}^{(l)})^\dagger A_{k_l}^{(l)} &= \sum_{l\in\Xi}\sum_{k_l} I_{\mathcal{H}_Q^{(l)}} \otimes r_{k_l} \langle k_l|k_l\rangle \\ &= \sum_{l\in\Xi} I_{\mathcal{H}_Q^{(l)}} \otimes \left(\sum_{k_l} r_{k_l}\right) \\ &= \sum_{l\in\Xi} I_{\mathcal{H}_Q^{(l)}} \otimes 1 \\ &= I_{\mathcal{H}_Q^{(l)}}.\end{aligned} \tag{2.125}$$

Therefore, we can construct $\mathcal{K}_{\rm on}$ by a Kraus representation

$$\{A_{k_l}^{(l)} : l \in \Xi, k_l\}. \tag{2.126}$$

□

Thus, the operation $\mathcal{K}_{\rm off}$ can be considered as an operation taking off the redundant parts of a given quantum ensemble by tracing out subsystems corresponding to the redundant parts. The operation $\mathcal{K}_{\rm on}$ is regarded as an operation attaching back the redundant parts of a given quantum ensemble by constructing quantum states on subsystems corresponding to the redundant parts. Using $\mathcal{K}_{\rm off}$ and $\mathcal{K}_{\rm on}$, we can reversibly remove the redundant parts from the ensemble. In this thesis, we refer to $\mathcal{K}_{\rm off}$ and $\mathcal{K}_{\rm on}$ as the *KI operations*.

2.6 Quantum Data Compression

In this section, we explain basic concepts of quantum data compression. First, we show a general setup for quantum data compression. The aim of the task of quantum data compression is to send a given state as efficiently as possible using noiseless quantum channels, i.e., the identity maps.

The sender receives a quantum state from the referee, i.e., the source of quantum data. Then, the sender compresses the given state and transmits the compressed state to the receiver through noiseless quantum channels. After receiving the state, the receiver decompresses the state to recover the original state. Figure 2.1 shows a schematic diagram of the quantum data compression task.

More formally, let $\mathcal{H}$ be a quantum system and let $\{p_a, \rho_a\}_{a\in\Sigma}$ be a quantum ensemble on this system. The referee independently draws a state from the ensemble according to the

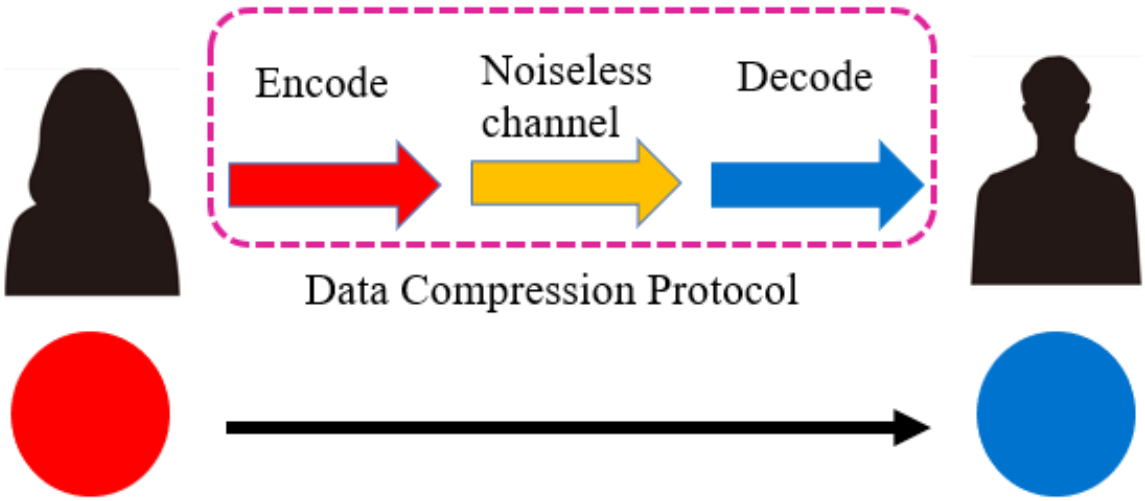

Figure 2.1: The schematic diagram showing the basic setup of quantum data compression tasks. The sender encodes a target state (red-colored in the figure) picked from a given ensemble and send it to the receiver via noiseless quantum channels. The receiver decodes the transmitted state to recover the original state (blue-colored in the figure).

probability distribution n times. Thus, states drawn at different rounds are independent and identically distributed (i.i.d.). The resulting state is denoted by $\rho_{a^n} \coloneqq \rho_{a_1} \otimes \rho_{a_2} \otimes \cdots \rho_{a_n}$ where a_i is the label of the state drawn in the ith round. The referee gives the state ρ_{a^n} to the sender. The sender applies an encoding channel to compress the state and send it to the receiver via the identity maps. The receiver decompresses the state by applying a decoding channel. Here, we require that the receiver's recovered state should be close to the original state from the referee. The ratio between the size of the compressed state and the number n of drawn states is called the *rate* of the data compression, or simply, the *compression rate*. We aim to minimize the compression rate by constructing a good protocol.

Historically, there are multiple setups for quantum data compression. Mainly, the following points have been extensively discussed:

1. Whether the given ensemble consists of pure states or mixed states;
2. Whether the referee tells the sender the label of the drawn state;
3. What is the definition of "closeness" between the recovered state and the original state.

As for the second point, the former, where the sender knows the label, is called *visible compression*, and the latter, where the sender does not know the label, is called *blind compression*. By definition, a compression protocol for the blind compression can naturally be exploited as a protocol for the visible compression.

For the last point, i.e., the definition of the closeness, two types of error criteria have been considered. The first one is *global error criterion*, for which the whole states should be close. The other one is called *local error criterion*, for which the two states should be close letter-wisely. Here, when the global error criterion is satisfied, the local error criterion is automatically implied. On the other hand, the global error criterion does not generally hold even if the local error criterion is assumed.

We briefly review major previous results on quantum data compression. While data compression problems had been broadly studied classically, quantum data compression was firstly discussed in Refs. [16–18]. In these references, as the first step for understandings of quantum data compression problems, the authors adopted *blind* compression of *pure-state ensembles* under *global error criterion*. It was shown that the optimal compression rate is the quantum entropy of the average state of the given pure-state ensemble; that is, letting

$\{p_a, |\psi_a\rangle\}_{a\in\Sigma}$ denote a given ensemble, the optimal rate R^* is given as

$$R^* \coloneqq \mathrm{H}\left(\sum_{a\in\Sigma} p_a \,|\psi_a\rangle\,\langle\psi_a|\right). \tag{2.127}$$

In particular, a compression protocol proposed in Ref. [16] is currently called *Schumacher compression*. More details about Schumacher compression are reviewed in various textbooks such as Refs. [59–61].

After blind compression of pure states, the visible compression began to be considered. In Refs. [24,25], a lower bound of the compression rate of the visible and blind compression under the global error criterion was given. Indeed, for a given quantum ensemble $\Phi = \{p_a, \rho_a\}_{a\in\Sigma}$, the optimal rate is lower-bounded by the Holevo information

$$I(\Phi) \coloneqq \mathrm{H}\left(\sum_{a\in\Sigma} p_a\rho_a\right) - \sum_{a\in\Sigma} p_a\,\mathrm{H}(\rho_a). \tag{2.128}$$

Observe that for a pure-state ensemble, the Holevo information is equal to the entropy of the average state since the entropy of a pure state is zero. Therefore, it is deduced that the optimal rate of the visible compression is also the entropy of the average state. Surprisingly, for compression of a pure-state ensemble under the global error criterion, whether the compression is visible or blind does not make any difference.

Then, what is interesting next is the compression of general mixed-state ensemble. Historically, the optimal rate of the visible compression is called *effective information*; that of blind compression is called *passive information*. Letting $I_e(\Phi)$ denote the effective information of a quantum ensemble Φ and $I_p(\Phi)$ denote the passive information of a quantum ensemble Φ, we have the following relation:

$$I(\Phi) \leqq I_e(\Phi) \leqq I_p(\Phi) \tag{2.129}$$

since the visible compression can be regarded as a subclass of blind compression. The quantity $I_d \coloneqq I_p - I_e$ is called *information defect*, and it characterizes a difference between the visible and blind compression tasks. In Ref. [24,25], a lower bound of the information defect was given, and Refs. [26,27] showed examples for which the information defect is strictly positive. Moreover, the optimal rate of the visible compression was derived in Ref. [28], and the optimal rate is given by the entropy of an extension of a given state. In Ref. [29], the authors gave another representation of the optimal rate of the visible compression. On the other hand, in Ref. [1], the optimal rate of blind compression was

studied. The main difficulty of blind compression is that the sender does not know the label of a given state. The authors proposed a protocol in which the sender only sends the essential parts of a given ensemble, and the compression rate achieved by the protocol is given by the KI decomposition [4]. In fact, the rate of blind compression derived in Ref. [1] is optimal under both the global and local error criteria.

Here, let us note that we mainly discuss quantum data compression without any assistance; however, data compression tasks with various assistance, e.g., entanglement and shared randomness, have also been widely investigated [30–34].

In this thesis, we investigate blind compression of mixed-state ensembles under local error criterion. It has been observed that the optimal rate of blind compression is quite sensitive to approximations even though an allowed approximation is diminutive [2]. Considering the instability of the optimal rate against approximations, we analyze blind compression problem with finite local approximations.

Chapter 3

Blind Quantum Data Compression

In this chapter, we discuss a quantum information processing task, namely, blind quantum data compression. We first review the basic setup of blind quantum data compression in Sec. 3.1. Then, we review previous important results in Sec. 3.2. We show and discuss our results on blind quantum data compression with finite local approximations allowed in Sec. 3.3. We construct a novel protocol that performs well compared with the previous one that works with asymptotically vanishing errors. We also show numerical experiments for blind compression of 2-*state* classical ensembles. We summarize and discuss our results in Sec. 3.4.

3.1 Setup of Blind Quantum Data Compression

In this section, we provide the basic setup of blind quantum data compression.

Blind data compression is a quantum information processing task between two parties, the sender and the receiver, in which fundamental quantum properties emerge. In the procedure, the sender aims to asymptotically send quantum data without knowing its actual description to the receiver as efficiently as possible.

We provide a more formal description of the setup below. Suppose that $\{p_x, \rho_x\}_{x\in\Sigma}$ is a quantum ensemble corresponding to quantum data the sender want to send. For each $x \in \Sigma$, the state $\rho_x \in \mathrm{D}(\mathcal{H}_A)$ is a density operator defined on a quantum system A with a Hilbert space $\mathcal{H}_A$, and $\{p_x : x \in \Sigma\}$ is the probability distribution. The referee, or the source of the

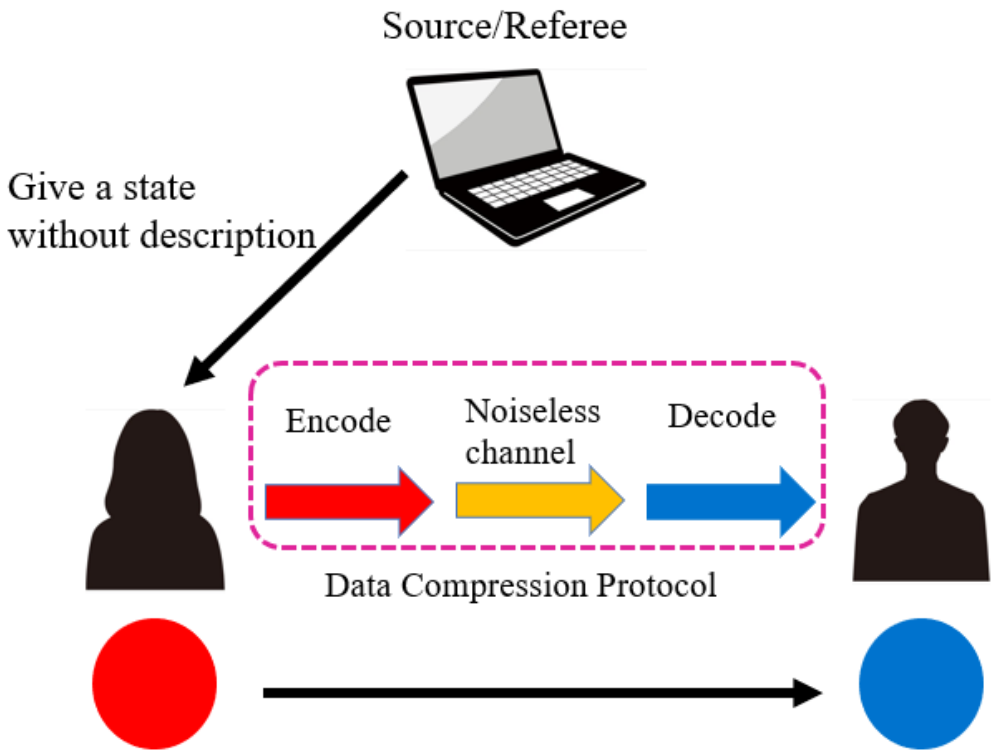

Figure 3.1: The schematic diagram showing the setup of blind quantum data compression. In this task, the sender is given a target state (red-colored in the figure) from the referee without its description. Then, the sender encodes the state and send it to the receiver via noiseless quantum channels. The receiver decodes the transmitted state to recover the original state (blue-colored in the figure).

data, draws a state from the ensemble n times independently, so that the resulting n-fold state is independent and identically distributed (i.i.d.). Let ρ_{x_k} denote a quantum state drawn in the kth round, and write the n-fold state as $\rho_{x^n} := \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n}$. Then, the referee gives the state ρ_{x^n} to the sender. After receiving the state ρ_{x^n} from the referee, the sender encodes the state by an encoding channel $\mathcal{E}_n \in \mathrm{C}(\mathcal{H}_A^{\otimes n}, \mathcal{H}_{M_n})$ to compress the state into $\mathcal{E}_n(\rho_{x_n})$ on a quantum system M_n. When the assistance of entanglement is allowed, the sender encodes the given state and a half of a given entangled state together. Here, we note that the sender does not know the label x_n of the given state while the sender knows the description of the ensemble. Therefore, the compression scheme cannot be state-specific and must only depend on the ensemble. The sender transmits the compressed state $\mathcal{E}_n$ to the receiver through a noiseless quantum channel. The receiver decompresses the state by a decoding channel $\mathcal{D}_n \in \mathrm{C}(\mathcal{H}_{M_n}, \mathcal{H}_B^{\otimes n})$ so that the resulting state $\mathcal{D}_n \circ \mathcal{E}_n(\rho_{x_n})$ on a system $B^n := B_1 \cdots B_n$ is close to the initial state ρ_{x_n}. The sequential operation $\mathcal{D}_n \circ \mathcal{E}_n$ given by the encoding and decoding channels is called a *protocol*.

The rate of the communication is given by the number of qubits needed for the sender to send the state $\mathcal{E}_n(\rho_n)$ to the receiver; that is, letting R_n denote the rate, we have

$$R_n \coloneqq \frac{\log_2 |M_n|}{n}. \tag{3.1}$$

There are two kinds of scenarios for what types of error we allow: the global error criterion and the local error criterion. Under the global error criterion, the error $\epsilon > 0$ of the protocol must satisfy that

$$\|\mathcal{D}_n \circ \mathcal{E}_n(\rho_{x^n}) - \rho_{x^n}\|_1 \leqq \epsilon; \tag{3.2}$$

on the other hand, under the local error criterion, the error ϵ must satisfy that

$$\left\|\mathrm{Tr}_{B_1 \cdots B_{k-1}, B_{k+1} \cdots B_n}(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{x^n})) - \rho_{x_k}\right\|_1 < \epsilon \tag{3.3}$$

for all integers $1 \leqq k \leqq n$. Thus, the global error criterion ensures that the whole of resulting state is close to the original state; the local error criterion guarantees that the resulting state is letter-wisely close to the originals state. Note that the global error criterion trivially implies that the local error criterion.

We call a pair $(\mathcal{E}_n, \mathcal{D}_n)$ with an encoding channel $\mathcal{E}_n$ and a decoding channel $\mathcal{D}_n$ an (n, R_n, ϵ) code if the protocol $\mathcal{D}_n \circ \mathcal{E}_n$ yields the rate R_n within error ϵ.

Then, we say that the rate R is achievable if for any $\epsilon > 0$ and $\delta > 0$, there exists sufficiently large n such that we have an $(n, R + \delta, \epsilon)$ code. Similarly, we say that the rate R is achievable within error ϵ if for any $\delta > 0$, there exists sufficiently large n such that we can construct an $(n, R + \delta, \epsilon)$ code.

3.2 Previous Results on Blind Data Compression

In this section, we briefly review the previous results on blind data compression.

3.2.1 Asymptotic Optimal Rate of Blind Data Compression

In Ref. [19], an achievable rate of blind state compression was investigated. Indeed, for a given ensemble $\{p_x, \rho_x\}_{x\in\Sigma}$, it was proved that the quantum entropy of the average state

$$\mathrm{H}\left(\sum_{x\in\Sigma} p_x\rho_x\right) \tag{3.4}$$

is an achievable rate. It had been open whether this rate is optimal or not before Ref. [1] conducted a more detailed analysis.

In Ref. [1], the authors exploited the Koashi-Imoto(KI) decomposition [4] of a given ensemble $\{p_x, \rho_x\}_x$:

$$\rho_x = \bigoplus_{l\in\Xi} q^{(x,l)}\rho_Q^{(x,l)} \otimes \rho_R^{(l)}, \tag{3.5}$$

which is introduced in Theorem 2.49. Then, by Theorem 2.50, we can define quantum channels $\mathcal{K}_{\mathrm{off}}$ and $\mathcal{K}_{\mathrm{on}}$ satisfying the following:

$$\mathcal{K}_{\mathrm{off}}(\rho_x) = \bigoplus_{l\in\Xi} q^{(x,l)}\rho_Q^{(x,l)}, \tag{3.6}$$

$$\mathcal{K}_{\mathrm{on}}\left(\bigoplus_{l\in\Xi} q^{(x,l)}\rho_Q^{(x,l)}\right) = \rho_x \tag{3.7}$$

for all $x \in \Sigma$. An operation $\mathcal{K}_{\mathrm{off}}$ represents a quantum operation taking the redundant parts of a given ensemble; $\mathcal{K}_{\mathrm{on}}$ represents a quantum operations putting the redundant parts back to the reduced states. With these quantum channels, we can observe that the sender does not necessarily send the redundant parts of a target ensemble if the sender and the receiver agree on the ensemble in the scenario of blind data compression. If they both knows the description of the ensemble, they also agree on $\mathcal{K}_{\mathrm{off}}$ and $\mathcal{K}_{\mathrm{on}}$; that is, they can freely take off and put on the redundant parts. In addition, Ref. [1] also proved that Eq. (3.4) is optimal even under local error criterion. To summarize, we have the following theorem.

Theorem 3.1. *Let $\{p_x, \rho_x\}$ be a quantum ensemble with KI decomposition*

$$\rho_x = \bigoplus_{l\in\Xi} q^{(x,l)} \rho_Q^{(x,l)} \otimes \rho_R^{(l)}. \tag{3.8}$$

Then, the optimal rate R^* of blind data compression within asymptotically vanishing errors for this ensemble is

$$R^* = \mathrm{H}\left(\sum_{x\in\Sigma} p_x \bigoplus_{l\in\Xi} q^{(x,l)} \rho_Q^{(x,l)}\right). \tag{3.9}$$

To show the optimality of the rate, the authors of Ref. [1] introduces two error functions of compression protocol Λ_n:

$$f(\Lambda_n) \coloneqq 1 - \sum_{x^n\in\Sigma^n} \mathrm{F}(p_{x^n}, \rho_{x^n}), \tag{3.10}$$

$$g(\Lambda_n) \coloneqq h_2(\Delta) + \Delta \log_2(D_A - 1), \tag{3.11}$$

with

$$\Delta \coloneqq 1 - \sum_{i=1}^{D_A} \lambda_i(\rho) \langle i|\Lambda_n(|i\rangle\langle i|)|i\rangle, \tag{3.12}$$

where ρ is the average state of the given ensemble with the spectral decomposition

$$\rho = \sum_{i=1}^{D_A} \lambda_i(\rho) |i\rangle\langle i|. \tag{3.13}$$

Indeed, it was shown that when the protocol Λ_n is given, the compression rate R is

$$R \geqq R^* - g(\Lambda_n). \tag{3.14}$$

The authors proved the optimality by showing that when $f(\Lambda_n)$ goes to zero, $g(\Lambda_n)$ also converges to zero; that is, for an infinitesimal error, R^* is the smallest compression rate that can be achievable.

Recently, Ref. [3] further investigated the optimality of the compression rate under the global error criterion. The authors introduced an error function that represents how an error allowed in the protocol affects the compression rate. Even though the sensitivity of the compression rate against error has not been revealed completely, properties of the error function shown in Ref. [3] might lead to insightful discussions for this problem.

3.2.2 Blind Compression of Classical Ensembles under Local Error Criterion

In this subsection, we review blind compression of classical ensembles with local errors. In Ref. [2], the authors formalized blind compression of classical ensembles with the assistance of entanglement. Taking advantage of the classicality and using properties of quantum entropic quantities (see Proposition 2.45 in Chapter 2), they proved a new lower bound of the compression rate under the local error criterion.

Theorem 3.2 (Lower Bound of Rate of Blind Compression with Entanglement Assistance)**.** *Let $\mathcal{H}$ be a quantum system, and let $\{p_x, \rho_x\}_{x\in\Sigma}$ be a classical ensemble on the system $\mathcal{H}$. Letting R be a rate of blind compression with entanglement assistance under the local error criterion, it holds that*

$$R \geqq \min_{\mathcal{F}:\mathcal{H}\to\mathcal{H}\otimes\mathcal{H}'} \left(I(\mathcal{H}:\mathcal{H}'|\mathcal{H}_X)_{\mathcal{F}(\rho)}\right) + I(\mathcal{H}_X:\mathcal{H})_\rho - \epsilon\log_2|\Sigma| - 1, \tag{3.15}$$

where $\mathcal{H}'$ is a copy of the system $\mathcal{H}$, $\mathcal{H}_X$ is a quantum system corresponding to the alphabet Σ; that is, $\mathcal{H}_X \coloneqq \mathbb{C}^\Sigma$, and a map $\mathcal{F}$ satisfies

$$\frac{1}{2}\|\mathrm{Tr}_{\mathcal{H}}(\mathcal{F}(\rho_x)) - \rho_x\|_1 \leqq \epsilon \tag{3.16}$$

for all $x \in \Sigma$.

Using this bound, the authors found an example of classical ensemble that cannot be well compressed; that is, the lower bound of the rate is close to $\log_2 D_{\mathcal{H}}$, which is the tight upper bound of the rate.

Theorem 3.3. *Let $\mathcal{H}$ be a quantum system, and let $\{p_x, \rho_x\}_{x=1,2}$ be a classical ensemble on the system $\mathcal{H}$ where $p_1 = p_2 = 1/2$ and*

$$\rho_1 \coloneqq \sum_{i=1}^{D_{\mathcal{H}}} \frac{1}{D_{\mathcal{H}}} |i\rangle\langle i|, \tag{3.17}$$

$$\rho_2 \coloneqq \sum_{i=1}^{D_{\mathcal{H}}} \frac{2i}{D_{\mathcal{H}}(D_{\mathcal{H}}+1)} |i\rangle\langle i|, \tag{3.18}$$

where $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ forms an orthonormal basis of $\mathcal{H}$. Then, the compression rate for this ensemble is at least $(\log_2 D_{\mathcal{H}}) - 7$.

On the other hand, a compression protocol for blind compression of classical ensembles without any assistance that may lead to a small compression rate was also proposed in Ref. [2]. The protocol works for all two-state classical ensembles.

Theorem 3.4. *Let $\mathcal{H}$ be a quantum system, and let $0 < \epsilon < 1/2$ and $0 < \delta < 1$ be positive real numbers. Then, for a positive number n, there exists an $(n, R, \epsilon + \delta)$-protocol for two-state classical ensembles such that*

$$R \leqq 2 \log_2 \log_2 \frac{D_{\mathcal{H}}}{\delta} + 2 \log_2 \frac{1}{\epsilon} + 3. \tag{3.19}$$

3.3 Our Results: Data Compression Protocol with Finite Approximation under Local Error Criterion

In this section, we show our results on the quantum blind compression with finite local approximations; in particular, we show our novel protocol and present some examples for which the protocol leads to large reduction of the compression rate compared to the previous results with asymptotically vanishing errors. Moreover, as the first step of general understandings of blind compression with finite approximations, we focus on quantum ensembles consisting of two classical states. Through numerical experiments, we reveal the performance of our protocol for classical ensembles.

3.3.1 Procedure of Our Protocol

Here, we consider blind data compression under the local error criterion; that is, we only require the resulting state to be close to the original state letter-wisely. In addition, we allow finite approximations to the resulting state under this error criterion.

First, notice that the compression scheme shown in Sec. 3.2 can be summarized as follows.

1. After receiving a state ρ_{x^n} from the referee, the sender applies $\mathcal{K}_{\text{off}}$ to remove the redundant parts of the state and obtain $\mathcal{K}_{\text{off}}^{\otimes n}(\rho_{x^n})$.

2. The sender encodes the state $\mathcal{K}_{\text{off}}^{\otimes n}(\rho_{x^n})$ using an encoding channel $\mathcal{E}_n$ to obtain $\mathcal{E}_n \circ \mathcal{K}_{\text{off}}^{\otimes n}(\rho_{x^n})$.

3. The sender sends the state $\mathcal{E}_n \circ \mathcal{K}_{\rm off}^{\otimes n}(\rho_{x^n})$ via a noiseless quantum channel to the receiver.

4. The receiver decodes the transmitted state using a decoding channel $\mathcal{D}_n$ and obtains $\mathcal{D}_n \circ \mathcal{E}_n \circ \mathcal{K}_{\rm off}^{\otimes n}(\rho_{x^n})$.

5. The receiver put the redundant parts back using the quantum channel $\mathcal{K}_{\rm on}$; the resulting state is $\mathcal{K}_{\rm on}^{\otimes n} \circ \mathcal{D}_n \circ \mathcal{E}_n \circ \mathcal{K}_{\rm off}^{\otimes n}(\rho_{x^n})$.

Our main idea here is that when we allow a finite approximation to the resulting state, we may use quantum channels $\Lambda_s^{(n)}$ and $\Lambda_r^{(n)}$, potentially acting jointly on the entire block, for smoothing of a given state instead of $\mathcal{K}_{\rm off}$ and $\mathcal{K}_{\rm on}$. Then, if the resulting state $\Lambda_s^{(n)} \circ \mathcal{D}_n \circ \mathcal{E}_n \circ \Lambda_r^{(n)}(\rho_{x^n})$ is close to the given state ρ_{x^n} within a given finite error, we may achieve a smaller rate with the smoothing scheme $\Lambda_s^{(n)}$ and $\Lambda_r^{(n)}$ than the optimal rate with a vanishing error attained by $\mathcal{K}_{\rm off}$ and $\mathcal{K}_{\rm on}$. Even for letter-wise smoothing, one may wonder if we can use the KI operations from a different quantum ensemble for smoothing instead of using the KI operaations $\mathcal{K}_{\rm off}$ and $\mathcal{K}_{\rm on}$ for the given ensemble. In the following, we prove that this scheme indeed performs well.

More formally, suppose that we have a quantum ensemble $\{p_x, \rho_x\}_{x\in\Sigma}$ and that we allow the resulting state to be letter-wise different from the original state up to $\epsilon > 0$. We consider all possible approximate ensembles $\{p_x, \tilde{\rho}_x\}_{x\in\Sigma}$, whose states should be close to the states in the original ensemble. Each such approximate ensemble defines KI operations $\tilde{\mathcal{K}}_{\rm off}$ and $\tilde{\mathcal{K}}_{\rm on}$. Now the question is, if we apply the KI operations of an approximate ensemble to the original ensemble, how much error do we incur and what rate can be achieved? Hence, in the following, we consider using the KI operations $\tilde{\mathcal{K}}_{\rm off}$ and $\tilde{\mathcal{K}}_{\rm on}$ of $\{p_x, \tilde{\rho}_x\}_{x\in\Sigma}$, which is obtained as in Theorem 2.50, as a smoothing scheme $\Lambda_s^{(n)}$ and $\Lambda_r^{(n)}$ for $\{p_x, \rho_x\}_{x\in\Sigma}$.

We now present our protocol for blind data compression more formally.

1. The sender and receiver first agree on an approximate ensemble $\{p_x, \tilde{\rho}_x\}_{x\in\Sigma}$ of the original given ensemble $\{p_x, \rho_x\}_{x\in\Sigma}$ such that for all $x \in \Sigma$,

$$\|\tilde{\mathcal{K}}_{\rm on} \circ \tilde{\mathcal{K}}_{\rm off}(\rho_x) - \rho_x\|_1 \leqq \epsilon, \tag{3.20}$$

   and

$$\mathrm{H}\left(\sum_x p_x \tilde{\mathcal{K}}_{\rm off}(\rho_x)\right) \leqq \mathrm{H}\left(\sum_x p_x \mathcal{K}_{\rm off}(\rho_x)\right). \tag{3.21}$$

2. The referee gives a state ρ_{x^n} generated by n independent and identically distributed (i.i.d.) draws from $\{p_x, \rho_x\}_{x\in\Sigma}$ to the sender.

3. The sender applies $\tilde{\mathcal{K}}_{\mathrm{off}}$ to the given en state, which yields $\tilde{\mathcal{K}}_{\mathrm{off}}^{\otimes n}(\rho_{x^n})$.

4. The sender applies an encoding operation $\mathcal{E}_n$ to obtain $\mathcal{E}_n \circ \tilde{\mathcal{K}}_{\mathrm{off}}^{\otimes n}(\rho_{x^n})$.

5. The sender transmits the encoded state $\mathcal{E}_n \circ \tilde{\mathcal{K}}_{\mathrm{off}}^{\otimes n}(\rho_{x^n})$ to the receiver via a noiseless quantum channel.

6. Upon receiving the state, the receiver applies the decoding channel $\mathcal{D}_n$ corresponding to $\mathcal{E}_n$, which yields $\mathcal{D}_n \circ \mathcal{E}_n \circ \tilde{\mathcal{K}}_{\mathrm{off}}^{\otimes n}(\rho_{x^n})$.

7. Finally, the receiver applies $\tilde{\mathcal{K}}_{\mathrm{on}}$ to obtain $\tilde{\mathcal{K}}_{\mathrm{on}}^{\otimes n} \circ \mathcal{D}_n \circ \mathcal{E}_n \circ \tilde{\mathcal{K}}_{\mathrm{off}}^{\otimes n}(\rho_{x^n})$, which must be close to the original state within the local error ϵ.

At the fourth step of the protocol, we virtually consider blind compression of ensemble $\{p_x, \tilde{\mathcal{K}}_{\mathrm{off}}(\rho_x)\}_{x\in\Sigma}$. Hence, the compression rate achieved by this protocol is

$$R = \mathrm{H}\left(\sum_{x\in\Sigma} p_x \tilde{\mathcal{K}}_{\mathrm{off}}(\rho_x)\right) \tag{3.22}$$

within local error ϵ.

Then, we are interested in when the condition (3.20) is satisfied. Here, we give a sufficient condition for Eq. (3.20); we prove that $\tilde{\mathcal{K}}_{\mathrm{off}}$ and $\tilde{\mathcal{K}}_{\mathrm{on}}$ satisfy the local error criterion for $\{p_x, \rho_x\}_{x\in\Sigma}$ if $\{p_x, \tilde{\rho}_x\}_{x\in\Sigma}$ is close enough to the original ensemble.

Proposition 3.5. *Let $\{p_x, \rho_x\}_{x\in\Sigma}$ be a quantum ensemble. Let $\{p_x, \tilde{\rho}_x\}_{x\in\Sigma}$ be a quantum ensemble such that*

$$\|\rho_x - \tilde{\rho}_x\|_1 \leqq \frac{\epsilon}{2} \tag{3.23}$$

for all x. Then, it holds that

$$\|\tilde{\mathcal{K}}_{\mathrm{on}} \circ \tilde{\mathcal{K}}_{\mathrm{off}}(\rho_x) - \rho_x\|_1 \leqq \epsilon. \tag{3.24}$$

Proof. It holds that

$$\begin{aligned} \|\tilde{\mathcal{K}}_{\mathrm{on}} \circ \tilde{\mathcal{K}}_{\mathrm{off}}(\rho_x) - \rho_x\|_1 &\leqq \|\tilde{\mathcal{K}}_{\mathrm{on}} \circ \tilde{\mathcal{K}}_{\mathrm{off}}(\rho_x) - \tilde{\mathcal{K}}_{\mathrm{on}} \circ \tilde{\mathcal{K}}_{\mathrm{off}}(\tilde{\rho}_x)\|_1 + \|\tilde{\rho}_x - \rho_x\|_1 && (3.25)\\ &\leqq 2\|\tilde{\rho}_x - \rho_x\|_1 && (3.26)\\ &\leqq \epsilon. && (3.27) \end{aligned}$$

□

Here, we would like to remark that once we have the protocol presented above, we also have a possibly better protocol. In fact, letting $\Lambda_{\rm off}$ and $\Lambda_{\rm on}$ denote the KI operations for $\{p_x, \tilde{\mathcal{K}}_{\rm off}(\rho_x)\}_{x\in\Sigma}$, by applying $\Lambda_{\rm off}^{\otimes n}$ to $\tilde{\mathcal{K}}_{\rm off}^{\otimes n}(\rho_{x^n})$, we can remove the redundant parts of $\{p_x, \tilde{\mathcal{K}}_{\rm off}(\rho_x)\}_{x\in\Sigma}$, leading to a slightly better rate

$$R = \mathrm{H}\left(\sum_{x\in\Sigma} p_x \Lambda_{\rm off}\circ\tilde{\mathcal{K}}_{\rm off}(\rho_x)\right). \tag{3.28}$$

3.3.2 Reduction of rates with a finite error

Here, we present two main examples for our protocol. The first one helps us understand the procedure and performance of our protocol in an intuitive way; the second one shows a large reduction of the compression rate compared to the case where finite approximations are not allowed.

First, we consider a four-dimensional two-state ensemble to intuitively understand the protocol.

Example 3.1. Let $\epsilon > 0$ be a fixed positive number sufficiently smaller than $1/2$. Consider the following two density operators.

$$\rho_1 := \frac{1}{4}\begin{pmatrix} 2 & 1-2\epsilon & 0 & 0 \\ 1-2\epsilon & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.29}$$

$$\rho_2 := \frac{1}{4}\begin{pmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & 2-\epsilon & 1 \\ 0 & 0 & 1 & 2-\epsilon \end{pmatrix}. \tag{3.30}$$

Then, we consider an ensemble $\{p_x, \rho_x\}_{x=1,2}$ where

$$p_1 = p_2 = \frac{1}{2}. \tag{3.31}$$

Here, we construct an approximate ensemble $\{p_x, \tilde{\rho}_x\}_{x=1,2}$ as

$$\tilde{\rho}_1 := \frac{1}{4}\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.32}$$

$$\tilde{\rho}_2 := \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}. \tag{3.33}$$

Observe that $\tilde{\rho}_1$ and $\tilde{\rho}_2$ can be written as

$$\tilde{\rho}_1 := |0\rangle\langle 0| \otimes \frac{1}{4}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \tag{3.34}$$

$$\tilde{\rho}_2 := |1\rangle\langle 1| \otimes \frac{1}{4}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{3.35}$$

Therefore,

$$\omega := \frac{1}{4}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{3.36}$$

can be regarded as a redundant part of the approximate ensemble. Then, we can define the KI operations corresponding to this redundant part as follows.

$$\tilde{\mathcal{K}}_{\mathrm{off}}(\cdot) := \mathrm{Tr}_B(\cdot), \tag{3.37}$$

$$\tilde{\mathcal{K}}_{\mathrm{on}}(\cdot) := ((|0\rangle\langle 0| \cdot |0\rangle\langle 0|) + (|1\rangle\langle 1| \cdot |1\rangle\langle 1|)) \otimes \omega. \tag{3.38}$$

Then, with these operations, we have

$$\tilde{\mathcal{K}}_{\mathrm{off}}(\rho_1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |0\rangle\langle 0|, \tag{3.39}$$

$$\tilde{\mathcal{K}}_{\mathrm{off}}(\rho_2) = \frac{1}{2}\begin{pmatrix} \epsilon & 0 \\ 0 & 2-\epsilon \end{pmatrix} = \frac{\epsilon}{2}|0\rangle\langle 0| + \left(1 - \frac{\epsilon}{2}\right)|1\rangle\langle 1|. \tag{3.40}$$

In addition, it holds that

$$\tilde{\mathcal{K}}_{\rm on} \circ \tilde{\mathcal{K}}_{\rm off}(\rho_1) = \frac{1}{4}\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = |0\rangle\langle 0| \otimes \omega = \tilde{\rho}_1, \tag{3.41}$$

$$\tilde{\mathcal{K}}_{\rm on} \circ \tilde{\mathcal{K}}_{\rm off}(\rho_2) = \frac{1}{8}\begin{pmatrix} 2\epsilon & \epsilon & 0 & 0 \\ \epsilon & 2\epsilon & 0 & 0 \\ 0 & 0 & 4-2\epsilon & 2-\epsilon \\ 0 & 0 & 2-\epsilon & 4-2\epsilon \end{pmatrix} = \left(\frac{\epsilon}{2}|0\rangle\langle 0| + \left(1-\frac{\epsilon}{2}\right)|1\rangle\langle 1|\right) \otimes \omega. \tag{3.42}$$

Then, we have that

$$\|\tilde{\mathcal{K}}_{\rm on} \circ \tilde{\mathcal{K}}_{\rm off}(\rho_1) - \rho_1\|_1 = \left\| \frac{1}{2}\begin{pmatrix} 0 & \epsilon & 0 & 0 \\ \epsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\|_1 = \epsilon, \tag{3.43}$$

$$\|\tilde{\mathcal{K}}_{\rm on} \circ \tilde{\mathcal{K}}_{\rm off}(\rho_2) - \rho_2\|_1 = \left\| \frac{1}{8}\begin{pmatrix} 0 & \epsilon & 0 & 0 \\ \epsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & -\epsilon \\ 0 & 0 & -\epsilon & 0 \end{pmatrix} \right\|_1 = \frac{\epsilon}{2} \leqq \epsilon. \tag{3.44}$$

Therefore, the KI operations $\tilde{\mathcal{K}}_{\rm off}$ and $\tilde{\mathcal{K}}_{\rm on}$ satisfy the local error criterion a finite error ϵ.

We now see that this example exhibits a reduction of the rate compared with the compression rate under asymptotically vanishing errors. Let R_0 denote the optimal rate for $\{p_x, \rho_x\}_{x\in\Sigma}$; let R denote the rate for $\{p_x, \rho_x\}_{x\in\Sigma}$ obtained by $\tilde{\mathcal{K}}_{\rm off}$ and $\tilde{\mathcal{K}}_{\rm on}$. Then, we have

$$R_0 = {\rm H}\left(\frac{1}{2}\rho_1 + \frac{1}{2}\rho_2\right) \approx \log_2 4 = 2, \tag{3.45}$$

$$R = {\rm H}\left(\frac{1}{2}\tilde{\mathcal{K}}_{\rm off}(\rho_1) + \frac{1}{2}\tilde{\mathcal{K}}_{\rm off}(\rho_2)\right) \approx \log_2 2 = 1. \tag{3.46}$$

Then, after this approximation, the compression rate becomes almost the half of the original rate.

Next, we show an example for which a finite approximation dramatically changes a KI structure of a given ensemble. In fact, for this example, we can see a large reduction of

the compression rate compared with the compression rate under asymptotically vanishing errors.

Example 3.2. Let $\omega_a \in \mathcal{D}(\mathcal{H}_a)$ and $\omega_b \in \mathcal{D}(\mathcal{H}_b)$ be density operators. Define $2N$-dimensional density operators

$$\sigma_1 \coloneqq \frac{1}{4N}\begin{pmatrix} 2 & \epsilon & 0 & \cdots & \epsilon \\ \epsilon & 2 & \epsilon & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \epsilon & 2 & \epsilon \\ \epsilon & 0 & \cdots & \epsilon & 2 \end{pmatrix} \tag{3.47}$$

$$\sigma_2 \coloneqq \frac{1}{4N}\begin{pmatrix} 1+2\epsilon & & & & & & \\ & \ddots & & & & & \\ & & 1+2\epsilon & & & & \\ & & & 3-2\epsilon & & & \\ & & & & \ddots & & \\ & & & & & 3-2\epsilon \end{pmatrix}. \tag{3.48}$$

Here, the off diagonal elements of σ_2 are all zero, and we omit these elements in the matrix form (3.48). With these density operators, let us define a quantum ensemble $\{p_x, \rho_x\}_{x=1,2}$ with

$$p_1 = p_2 = \frac{1}{2}, \tag{3.49}$$

$$\rho_1 \coloneqq \frac{1}{3}\omega_a \oplus \frac{1}{3}\sigma_1 \oplus \frac{1}{3}\omega_b, \tag{3.50}$$

$$\rho_2 \coloneqq \frac{1}{6}\omega_a \oplus \frac{1}{3}\sigma_2 \oplus \frac{1}{2}\omega_b, \tag{3.51}$$

where ω_a and ω_b are some density operators. Here, we can notice that ρ_1 and ρ_2 already have redundant parts ω_a and ω_b. Now, we introduce an approximate ensemble of $\{p_x, \tilde{\rho}_x\}_{x=1,2}$ with

$$\tilde{\rho}_1 \coloneqq \frac{1}{3}\omega_a \oplus \frac{1}{3}\tilde{\sigma}_1 \oplus \frac{1}{3}\omega_b, \tag{3.52}$$

$$\tilde{\rho}_2 \coloneqq \frac{1}{6}\omega_a \oplus \frac{1}{3}\tilde{\sigma}_2 \oplus \frac{1}{2}\omega_b, \tag{3.53}$$

where

$$\tilde{\sigma}_1 \coloneqq \frac{1}{2N}\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{3.54}$$

$$\tilde{\sigma}_2 \coloneqq \frac{1}{4N}\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 3 & & \\ & & & & \ddots & \\ & & & & & 3 \end{pmatrix}. \tag{3.55}$$

We note that $\tilde{\rho}_1$ and $\tilde{\rho}_2$ can be written as

$$\tilde{\rho}_1 = \frac{1}{2}\tilde{\omega}_a \oplus \frac{1}{2}\tilde{\omega}_b, \tag{3.56}$$

$$\tilde{\rho}_2 = \frac{1}{4}\tilde{\omega}_a \oplus \frac{3}{4}\tilde{\omega}_b, \tag{3.57}$$

where

$$\tilde{\omega}_a \coloneqq \frac{2}{3}\left(\omega_a \oplus \frac{1}{2N}I\right), \tag{3.58}$$

$$\tilde{\omega}_b \coloneqq \frac{2}{3}\left(\frac{1}{2N}I \oplus \omega_b\right). \tag{3.59}$$

Then, considering the KI operations $\tilde{\mathcal{K}}_{\text{off}}$ and $\tilde{\mathcal{K}}_{\text{on}}$ corresponding to this structure, we have

$$\tilde{\mathcal{K}}_{\text{off}}(\rho_1) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.60}$$

$$\tilde{\mathcal{K}}_{\text{off}}(\rho_2) = \frac{1}{12}\begin{pmatrix} 3+2\epsilon & 0 \\ 0 & 9-2\epsilon \end{pmatrix}, \tag{3.61}$$

and

$$\tilde{\mathcal{K}}_{\text{on}} \circ \tilde{\mathcal{K}}_{\text{off}}(\rho_1) = \frac{1}{2}\tilde{\omega}_a \oplus \frac{1}{2}\tilde{\omega}_b = \tilde{\rho}_1 \tag{3.62}$$

$$\tilde{\mathcal{K}}_{\text{on}} \circ \tilde{\mathcal{K}}_{\text{off}}(\rho_2) = \left(\frac{1}{4}+\frac{\epsilon}{6}\right)\tilde{\omega}_a \oplus \left(\frac{3}{4}-\frac{\epsilon}{6}\right)\tilde{\omega}_b. \tag{3.63}$$

Therefore, it holds that

$$\left\|\tilde{\mathcal{K}}_{\rm on}\circ\tilde{\mathcal{K}}_{\rm off}(\rho_1)-\rho_1\right\|_1 \leqq \epsilon, \tag{3.64}$$

$$\left\|\tilde{\mathcal{K}}_{\rm on}\circ\tilde{\mathcal{K}}_{\rm off}(\rho_2)-\rho_2\right\|_1 = \frac{4\epsilon}{9} \leqq \epsilon, \tag{3.65}$$

$$\tag{3.66}$$

implying that this pair of KI operations yields a compression protocol within a finite error ϵ.

Letting R_0 denote the optimal rate for $\{p_x,\rho_x\}_x$ without any approximations and R denote the rate for $\{p_x,\rho_x\}_x$ obtained by $\tilde{\mathcal{K}}_{\rm off}$ and $\tilde{\mathcal{K}}_{\rm on}$, we have

$$R_0 = {\rm H}\left(\frac{1}{2}\mathcal{K}_{\rm off}(\rho_1)+\frac{1}{2}\mathcal{K}_{\rm off}(\rho_2)\right) \gtrsim \log_2 N, \tag{3.67}$$

$$R = {\rm H}\left(\frac{1}{2}\tilde{\mathcal{K}}_{\rm off}(\rho_1)+\frac{1}{2}\tilde{\mathcal{K}}_{\rm off}(\rho_2)\right) \leqq \log_2 2 = 1, \tag{3.68}$$

which shows the reduction of the compression rate to a constant rate, which is independent of the size of the system.

Thus, as these examples show, our compression protocol performs well compared to the case where finite approximations are not allowed when we have an approximate ensemble of a given quantum ensemble, which has large redundant parts in terms of the KI decomposition. In the first example, we can generate a redundant part ω defined in Eq. (3.36), and in the secomd example, we can generate large redundant parts $\tilde{\omega}_a$ and $\tilde{\omega}_b$ shown in Eqs. (3.58) and (3.59). Then, by not sending these redundant parts, we can achieve small compression rates. Remarkably, in Example 3.2, we can see that even a small approximation leads to a constant compression rate independent of the dimension of the system.

Despite these examples showing large reductions in rates of blind compression, we would like to remark that we do not necessarily find a good approximation. For example, if an allowed error is much smaller than ϵ in Examples 3.1 and 3.2, we cannot apply the same approximation anymore. To further advance the study of blind quantum compression with finite approximations, a general investigation of conditions under which we can successfully find a good approximate ensemble of a given ensemble is needed. It is also interesting to study a more general compression protocol for blind quantum compression with finite approximations.

3.3.3 Approximation of Classical Ensembles

In this section, we discuss an approximate blind compression of classical ensembles, aiming to obtain an insight into blind compression of general quantum ensembles with finite approximations. Classical ensembles are a special subclass of quantum ensembles; hence, analyses of classical ensembles may lead us to a discovery also applicable to the general case. Here, to see properties of blind compression of classical ensembles concisely, we consider two-state classical ensembles, which consist of two classical states.

To consider a classical ensemble, we need to fix a basis of a given quantum system. We show that when we approximate a classical ensemble with respect to the same given basis, we only have to consider an approximation of diagonal elements.

Proposition 3.6. *Let $\mathcal{H}$ be a quantum system. Let $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ be quantum states on the system. Suppose that ρ and σ forms a classical ensemble; that is, we can write*

$$\rho = \sum_{i=1}^{D_{\mathcal{H}}} \rho_i |i\rangle \langle i| , \tag{3.69}$$

$$\sigma = \sum_{i=1}^{D_{\mathcal{H}}} \sigma_i |i\rangle \langle i| \tag{3.70}$$

for some orthonormal basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ of $\mathcal{H}$ and some probability distributions $\{\rho_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ and $\{\sigma_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$. Suppose that ρ and σ can be approximated as

$$\left\| \rho - \bigoplus_{l \in \Xi} q^{(\rho,l)} \omega_Q^{(\rho,l)} \otimes \omega_R^{(l)} \right\|_1 \leqq \epsilon, \tag{3.71}$$

$$\left\| \sigma - \bigoplus_{l \in \Xi} q^{(\sigma,l)} \omega_Q^{(\sigma,l)} \otimes \omega_R^{(l)} \right\|_1 \leqq \epsilon \tag{3.72}$$

in the same basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$. Then, we can find an approximation such that all $\omega_Q^{(\rho,l)}$, $\omega_Q^{(\sigma,l)}$, and $\omega_R^{(l)}$ are diagonal.

Proof. Let us take some $l \in \Xi$, and consider $q^{(\rho,l)} \omega_Q^{(\rho,l)} \otimes \omega_R^{(l)}$ and $\bigoplus_{l \in \Xi} q^{(\sigma,l)} \omega_Q^{(\sigma,l)} \otimes \omega_R^{(l)}$. Let $\rho^{(l)}$ and $\sigma^{(l)}$ be the corresponding block of ρ and σ respectively. Taking the dephasing

channel Δ with respect to the basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$, we have

$$\begin{aligned} \left\| \rho^{(l)} - \bigoplus_{l \in \Xi} q^{(\rho,l)} \Delta(\omega_Q^{(\rho,l)}) \otimes \Delta(\omega_R^{(l)}) \right\|_1 &= \left\| \Delta(\rho^{(l)}) - \Delta\left(\bigoplus_{l \in \Xi} q^{(\rho,l)} \omega_Q^{(\rho,l)} \otimes \omega_R^{(l)} \right) \right\|_1 \\ &\leqq \left\| \rho^{(l)} - \bigoplus_{l \in \Xi} q^{(\rho,l)} \omega_Q^{(\rho,l)} \otimes \omega_R^{(l)} \right\|_1 . \end{aligned} \tag{3.73}$$

Similarly, it holds that

$$\left\| \sigma^{(l)} - \bigoplus_{l \in \Xi} q^{(\sigma,l)} \Delta(\omega_Q^{(\sigma,l)}) \otimes \Delta(\omega_R^{(l)}) \right\|_1 \leqq \left\| \sigma^{(l)} - \bigoplus_{l \in \Xi} q^{(\sigma,l)} \omega_Q^{(\sigma,l)} \otimes \omega_R^{(l)} \right\|_1 . \tag{3.74}$$

Therefore, diagonal states $\bigoplus_{l \in \Xi} q^{(\rho,l)} \Delta(\omega_Q^{(\rho,l)}) \otimes \Delta(\omega_R^{(l)})$ and $\bigoplus_{l \in \Xi} q^{(\sigma,l)} \Delta(\omega_Q^{(\sigma,l)}) \otimes \Delta(\omega_R^{(l)})$ are also approximations of ρ and σ. □

Therefore, when a basis is fixed, an approximation of a given classical state is the same with an approximation of a probability distribution corresponding to the classical state. To consider an approximation of a probability distribution, *binning* is a useful method, where we divide probabilities in the distribution into several groups (bins) and replace each probability with the average value of the bin the probability belongs to.

Definition 3.7 (Binning)**.** Let $\{p_a : a \in \Sigma\}$ be a probability distribution. Consider a division of an alphabet Σ

$$\Sigma = \bigcup_{k=1}^{m} \Sigma_k \tag{3.75}$$

where Σ_k is an alphabet, that is, a nonempty and finite set, for all k, and it holds that

$$\Sigma_k \cap \Sigma_l = \emptyset \tag{3.76}$$

when $k \neq l$. Therefore, the union in Eq. (3.75) is indeed a disjoint union. Then, for each k, Σ_k is called a bin, and an integer m is the number of bins. For each bin Σ_k, we can consider a set of probabilities $\{p_a : a \in \Sigma_k\}$. Let $p^{(k)}$ is the average value of the set $\{p_a : a \in \Sigma_k\}$. Then, we can construct a new probability distribution $\{p'_a : a \in \Sigma\}$ by replacing p_a with $p'_a = p^{(k)}$ when $a \in \Sigma_k$. This method generating a probability distribution $\{p'_a : a \in \Sigma\}$ from a given distribution $\{p_a : a \in \Sigma_k\}$ is called *binning*.

When probabilities in a bin $\{p_a : a \in \Sigma_k\}$ are all close to each other, the binning serves as an approximation method of probability distributions; that is, it is also regarded as a good approximation method of classical states.

Now, we consider applying the binning method to the compression of classical ensembles. As we saw in Sec. 3.2, in Ref. [2], a protocol to approximately compress a two-state classical ensemble was proposed. We observe that this method exploits binning to approximate the diagonal elements of given states. Here, instead of thinking general two-state classical ensembles, we discuss a special two-state classical ensemble where one of the two states is the flat state, motivated by the incompressibility shown in Theorem 3.3. We propose two methods to approximate a given ensemble, and numerically compare the performance of the two methods.

Suppose that $\mathcal{H}$ be a quantum system, and consider two classical states $\{\rho, \sigma\}$ where ρ is the flat state. Take an orthonormal basis $\{|i\rangle \in \mathcal{H} : 1 \leqq i \leqq D_{\mathcal{H}}\}$ of $\mathcal{H}$ such that we can write

$$\rho = \sum_{i=1}^{D_{\mathcal{H}}} \frac{1}{D_{\mathcal{H}}} |i\rangle \langle i| , \tag{3.77}$$

$$\sigma = \sum_{i=1}^{D_{\mathcal{H}}} p_i |i\rangle \langle i| , \tag{3.78}$$

where $\{p_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ forms a probability distribution with $p_1 \geqq p_2 \geqq \cdots \geqq p_{D_{\mathcal{H}}}$. We consider an approximation of $\{p_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ within an allowed error $\epsilon > 0$, that is, a probability distribution $\{p'_i : 1 \leqq i \leqq D_{\mathcal{H}}\}$ such that

$$\sum_{i=1}^{D_{\mathcal{H}}} |p_i - p'_i| \leqq \epsilon. \tag{3.79}$$

To construct an approximation leading to large redundant parts, we propose the following two method, namely, the *arithmetic mean method* and the *geometric mean method*.

- Arithmetic Mean Method

  1. First, for given $\epsilon > 0$, find the largest positive integer k_1 such that

     $$|p_1 - p_{k_1}| \leqq \frac{\epsilon}{D_{\mathcal{H}}}. \tag{3.80}$$

     Define a set of positive integers $I_1 \coloneqq \{1, \ldots, k_1\}$.

2. For $i > 1$, find the largest positive integer k_i such that

$$\left| p_{k_{i-1}+1} - p_{k_i} \right| \leqq \frac{\epsilon}{D_{\mathcal{H}}}. \tag{3.81}$$

Define a set of positive integers $I_i \coloneqq \{k_{i-1}+1, \ldots, k_i\}$.

3. Repeat Step 2 until we find a positive integer L such that $k_L = D_{\mathcal{H}}$.
4. For $j \in I_i$ with some $1 \leqq i \leqq L$, replace p_j as

$$p_j \mapsto p_j^{(A)} \coloneqq \frac{1}{|I_i|} \sum_{m \in I_i} p_m, \tag{3.82}$$

and define

$$\sigma^{(A)} \coloneqq \sum_{i=1}^{D_{\mathcal{H}}} p_i^{(A)} |i\rangle \langle i| \tag{3.83}$$

- Geometric Mean Method

1. First, for given $\epsilon > 0$, find the largest positive integer k_1 such that

$$\frac{p_{k_1}}{p_1} \geqq \frac{1}{1+\epsilon}. \tag{3.84}$$

Define a set of positive integers $I_1 \coloneqq \{1, \ldots, k_1\}$.

2. For $i > 1$, find the largest positive integer k_i such that

$$\frac{p_{k_i}}{p_{k_{i-1}+1}} \geqq \frac{1}{1+\epsilon}. \tag{3.85}$$

Define a set of positive integers $I_i \coloneqq \{k_{i-1}+1, \ldots, k_i\}$.

3. Repeat Step 2 until we find a positive integer L such that $k_L = D_{\mathcal{H}}$.
4. For $j \in I_i$ with some $1 \leqq i \leqq l$, replace p_j as

$$p_j \mapsto p_j^{(G)} \coloneqq \frac{1}{|I_i|} \sum_{m \in I_i} p_m, \tag{3.86}$$

and define

$$\sigma^{(G)} \coloneqq \sum_{i=1}^{D_{\mathcal{H}}} p_i^{(G)} |i\rangle \langle i| . \tag{3.87}$$

First, we show that these approximation methods are valid; that is, the resulting states $\sigma^{(A)}$ and $\sigma^{(G)}$ are close enough to the original state σ.

Proposition 3.8. *For a positive real number $\epsilon > 0$, let $\sigma^{(A)}$ and $\sigma^{(G)}$ be the approximated states of a given diagonal state σ by Arithmetic Mean Method and Geometric Mean Method respectively. Then, it holds that*

$$\|\sigma^{(A)} - \sigma\|_1 \leqq \epsilon, \tag{3.88}$$
$$\|\sigma^{(G)} - \sigma\|_1 \leqq \epsilon. \tag{3.89}$$

Proof. It holds that

$$\begin{aligned}
\|\sigma - \sigma^{(A)}\|_1 &= \sum_{i=1}^{D_{\mathcal{H}}} |p_i - p_i^{(A)}| \\
&= \sum_{i=1}^{L} \sum_{j \in I_i} |p_j - p_j^{(A)}| \\
&\leqq \sum_{i=1}^{L} \sum_{j \in I_i} |p_{k_{i-1}+1} - p_{k_i}| \\
&\leqq \sum_{i=1}^{L} \sum_{j \in I_i} \frac{\epsilon}{D_{\mathcal{H}}} \\
&= \epsilon.
\end{aligned} \tag{3.90}$$

The first inequality follows because the difference $|p_j - p_j^{(A)}|$ is upper-bounded by the difference between the largest value and the smallest value of the set $\{p_i : i \in I_j\}$, which is given by $|p_{k_{i-1}+1} - p_{k_i}|$. The second inequality follows from Eq. (3.81).

On the other hand, it holds that

$$\begin{aligned}
\|\sigma - \sigma^{(G)}\|_1 &= \sum_{i=1}^{D_{\mathcal{H}}} |p_i - p_i^{(G)}| \\
&= \sum_{i=1}^{L} \sum_{j \in I_i} |p_j - p_j^{(G)}| \\
&= \sum_{i=1}^{L} \sum_{j \in I_i} p_j^{(G)} \left| \frac{p_j}{p_j^{(G)}} - 1 \right|.
\end{aligned} \tag{3.91}$$

Here, by Eq. (3.85), we have

$$\frac{1}{1+\epsilon} \leqq \frac{p_j}{p_j^{(G)}} \leqq 1+\epsilon, \tag{3.92}$$

which implies that

$$-\epsilon \leqq \frac{-\epsilon}{1+\epsilon} \leqq \frac{p_j}{p_j^{(G)}} - 1 \leqq \epsilon, \tag{3.93}$$

Combining Eqs. (3.91) and (3.93), we have that

$$\begin{aligned} \|\sigma - \sigma^{(G)}\|_1 &= \sum_{i=1}^{L} \sum_{j \in I_i} p_j^{(G)} \left| \frac{p_j}{p_j^{(G)}} - 1 \right| \\ &\leqq \sum_{i=1}^{L} \sum_{j \in I_i} p_j^{(G)} \epsilon \\ &= \epsilon. \end{aligned} \tag{3.94}$$

□

Therefore, $\{\rho, \sigma'\}$ with $\sigma' = \sigma^{(A)}$ or $\sigma^{(G)}$ is considered as an approximation of $\{\rho, \sigma\}$. Consider a quantum ensemble formed by $\{\rho, \sigma'\}$ and quantum channels $\mathcal{K}_{\text{off}}$ and $\mathcal{K}_{\text{on}}$ with respect to this ensemble. Since

$$\mathcal{K}_{\text{off}}(\rho) = \sum_{i=1}^{L} \frac{|I_i|}{D_{\mathcal{H}}} |i\rangle \langle i|, \tag{3.95}$$

$$\mathcal{K}_{\text{off}}(\sigma) = \sum_{i=1}^{L} \left(\sum_{m \in I_i} p_m \right) |i\rangle \langle i|, \tag{3.96}$$

the rate R of our protocol for this ensemble is

$$R = \mathrm{H}\left(p_\rho \mathcal{K}_{\text{off}}(\rho) + p_\sigma \mathcal{K}_{\text{off}}(\sigma)\right) \leqq \log_2 L. \tag{3.97}$$

Now, we compare the performance of the two approximation methods. It is of interest to determine which of the two binning methods performs better. In fact, the protocol shown in Ref. [2] exploits geometric mean, and one may wonder whether arithmetic mean yields better results. While analytical discussion is complicated, we conduct numerical experiments to see tendency of performance of these approximation methods.

1. First, we evaluate the performance of the two approximation methods with various values of allowed error ϵ. In the calculation, we fix the space to $\mathcal{H} = \mathbb{C}^{1024}$; that is, $D_{\mathcal{H}} = 1024 = 2^{10}$. We generate 1000 random diagonal states σ on this space. In more detail, we randomly generate a real number in the range $[0, 1]$ according to the uniform distribution 1024 times to obtain a vector on $\mathbb{R}^{1024}$. Then, we normalize the obtained vector to obtain a probability distribution, and we regard this probability distribution as a classical state. For each state, we create the approximate states $\sigma^{(A)}$ and $\sigma^{(G)}$ by using the arithmetic mean method and the geometric mean method.

   Then, considering $\sigma^{(A)}$ and $\sigma^{(G)}$, we compare the rates of the two methods. We also compare differences between original states and resulting states. As shown in Figures 3.2 and 3.3, the arithmetic mean method performs better than the geometric mean method on average. In particular, the results show the tendency that the difference between the two methods becomes large as the allowed error increases.

   Furthermore, we estimate the curves in the graph. Observing the graph shown in Figure 3.2, we adopt the following function

$$f(x) = \log_2 D_{\mathcal{H}} - a\left(1 - \mathrm{e}^{-b\sqrt{x}}\right) \tag{3.98}$$

   with parameters a and b as a fitting function. This function explains the tendency that the rate approaches $\log_2 D_{\mathcal{H}}$ as the error becomes small, and it also expresses the sudden decrease around the error $1/D_{\mathcal{H}}$. However, we do not believe that this function explains the full dependency of the rate on the error, because of the sensitive behaviour of the rate against errors when errors are roughly larger than $1/\sqrt{D_{\mathcal{H}}}$. For more general and deep understandings of the dependency, we need further investigations of the compression protocols.

   The results shown in Figure 3.4 and Table 3.1 indicate that the fitting function defined in Eq. (3.98) is the correct function characterizing the compression rate as a function of the error while we have not theoretically and analytically demonstrated it. Since the compression rate is strictly upper-bounded by $\log_2 D_{\mathcal{H}}$ when we do not allow any approximation, the second term $a\left(1 - \mathrm{e}^{-b\sqrt{x}}\right)$ of Eq. (3.98) is considered to represent the degree of reduction caused by an approximation. Although we need further investigations to fully characterize the performance of these approximation methods, our results roughly disclosed the tendency.

2. Next, we investigate the dependence of the performance of the two approximation methods on the dimension of the space. Here, we adopt the error $\epsilon = 1/\sqrt{D_{\mathcal{H}}}$ considering the discussion in Ref. [2] stating that this size of error $f(\Lambda_n) \approx 1/\sqrt{D_{\mathcal{H}}}$

(see Eq. (3.10)) leads to instability of the error function $g(\Lambda_n)$ (see Eq. (3.11)). We generate 1000 random diagonal states σ on this space in the same way as in the first experiment. For each state, we create the approximated states $\sigma^{(A)}$ and $\sigma^{(G)}$ by using the arithmetic mean method and the geometric mean method.

Then, considering $\sigma^{(A)}$ and $\sigma^{(G)}$, we compare the rates of the two methods. We also compare differences between original states and resulting states. As shown in Figures 3.5 and 3.6, the arithmetic mean method performs better than the geometric mean method on average.

To investigate the dependence of the difference between the two methods on the dimension, we also plotted the difference in Figure 3.7. As seen in the graph, the difference becomes large as the dimension gets large, and it linearly depends on the logarithm of dimension.

Moreover, we estimate the curves in the graph. Observing the graph shown in Figure 3.2, we adopt the following function

$$f(x) = a \log_2 x + b \tag{3.99}$$

with parameters a and b as a fitting function. The results shown in Figure 3.8 and Table 3.2 imply that the fitting function defined in Eq. (3.99) is appropriate for characterizing the compression rate as a function of the dimension. Since the compression rate is strictly upper-bounded by $\log_2 D_{\mathcal{H}}$ when we do not allow any approximation, the coefficient a represents the degree of reduction caused by an approximation.

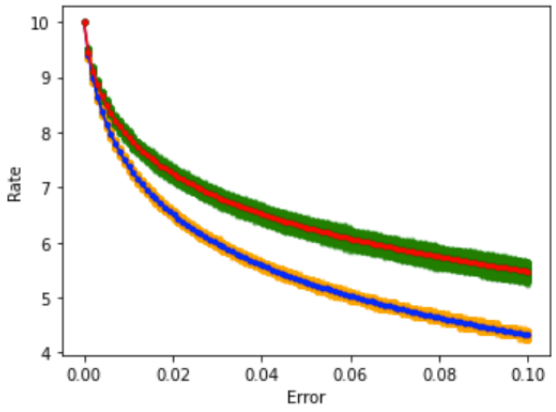

Figure 3.2: Graph of the compression rate as a function of finite allowed error. The term "rate" refers to the upper bound shown in Eq. (3.97). The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetic and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively.

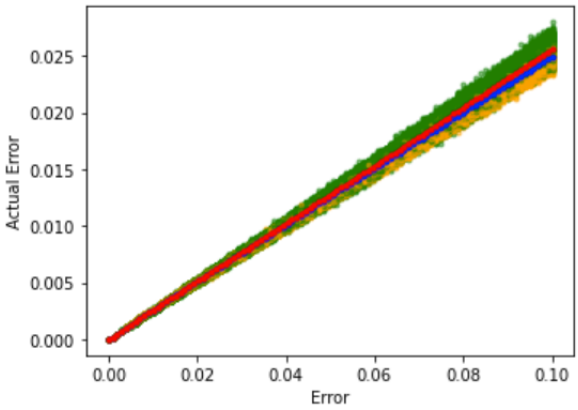

Figure 3.3: Graph of the actual error as a function of finite allowed error. Here, "actual error" means the distance between the initial state and the resulting state in terms of the trace norm. The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetic and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively.

	Arithmetical	Geometrical
a	7.856	6.264
b	4.079	4.036

Table 3.1: The results of the fitting of the graph of the compression rate shown in Figure 3.4. In the table, a and b are defined in Eq. (3.98).

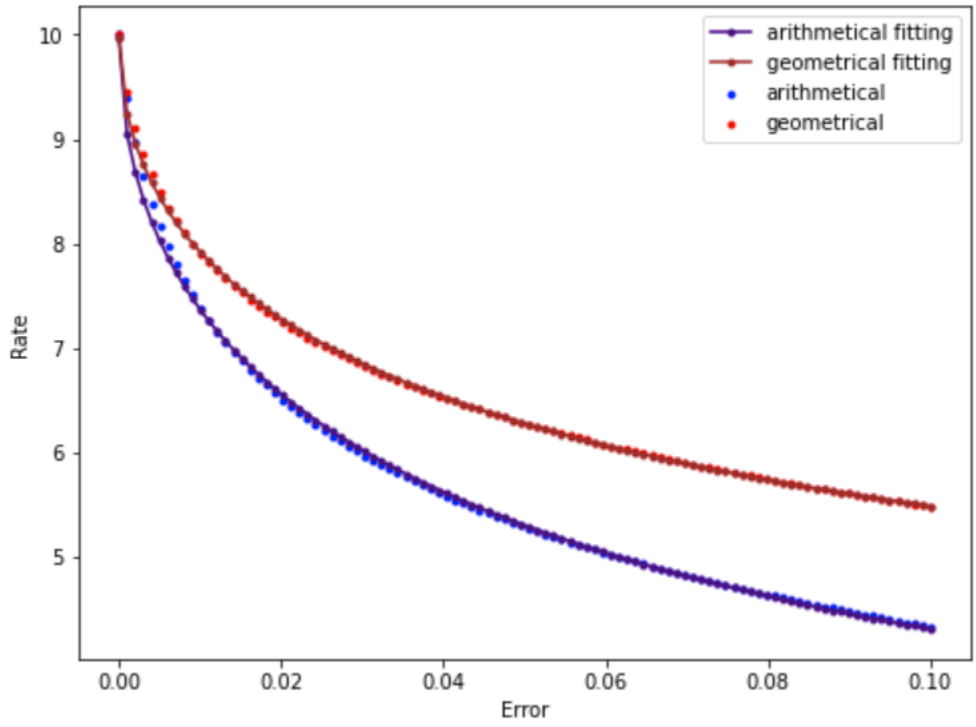

Figure 3.4: Graph of the compression rate as a function of finite allowed error and the corresponding fitting curves. The blue and red points represent the averages of the results of the arithmetic and geometric mean methods. The indigo curve is the fitting function corresponding to the blue points; the brown curve is the fitting function corresponding to the red points.

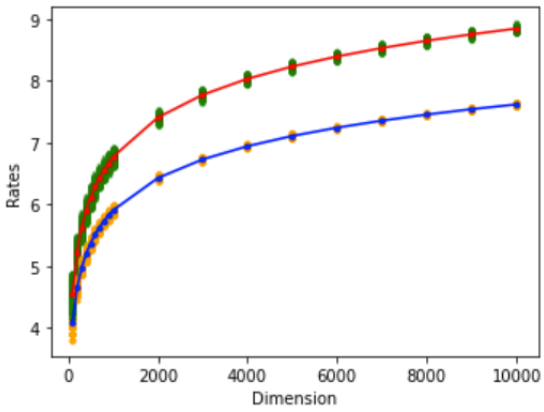

Figure 3.5: Graph of the compression rate as a function of the dimension of system. The term "rate" refers to the upper bound shown in Eq. (3.97). The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetical and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively.

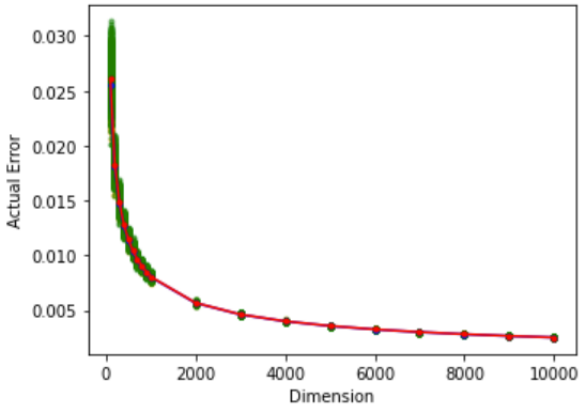

Figure 3.6: Graph of the actual error as a function of the dimension of system. Here, "actual error" means the distance between the initial state and the resulting state in terms of the trace norm. The yellow and green points represent the compression rate of each randomly generated probability distribution obtained by the arithmetical and geometric mean methods respectively. The blue and red points represent the averages of the yellow and green points respectively.

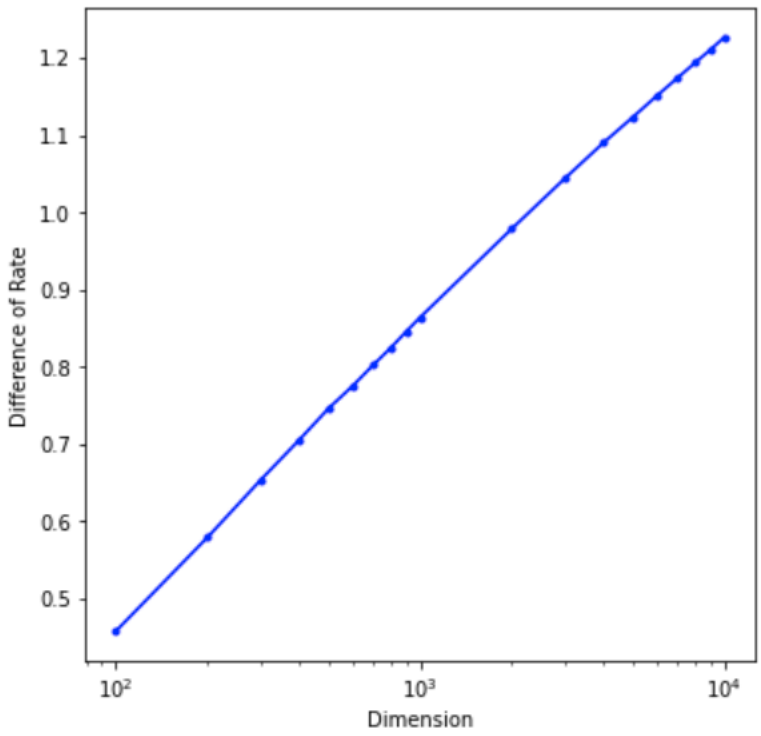

Figure 3.7: Graph of (the compression rate by the arithmetic mean method) − (the compression rate by the geometric mean method) as a function of the dimension of system. The horizontal axis is expressed in log scale.

	Arithmetical	Geometrical
a	0.5258	0.6406
b	0.6486	0.3608

Table 3.2: The results of the fitting of the graph of the compression rate shown in Figure 3.8. In the table, a and b are defined in Eq. (3.99).

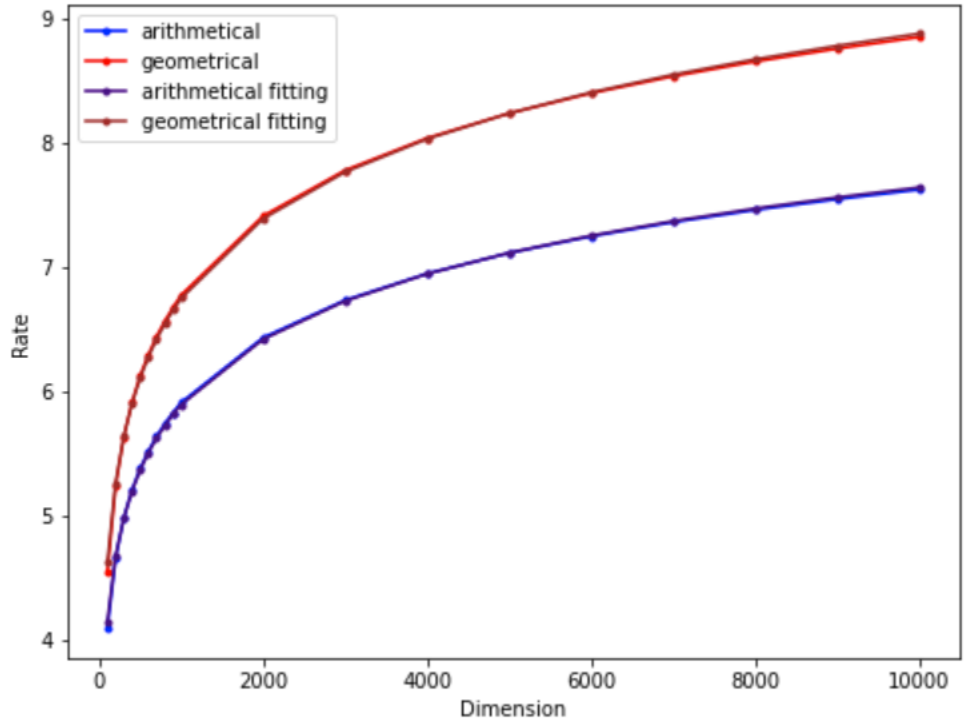

Figure 3.8: Graph of the compression rate as a function of the dimension of system and the corresponding fitting curves. The blue and red points represent the averages of the results of the arithmetic and geometric mean methods. The indigo curve is the fitting function corresponding to the blue points; the brown curve is the fitting function corresponding to the red points.

3.4 Summary and Discussions

In this chapter, we investigated blind compression of quantum ensembles under finite local approximations. In previous research, the optimal rate of blind compression was obtained through the KI decomposition. The main idea was that the sender does not have to send the redundant parts of a given quantum ensemble to effectively transmit a state drawn from the ensemble, even if the sender does not know the label of the given state.

Reviewing the previous research, we focused on the instability of the KI decomposition against approximations. Taking advantage of the sensitivity of KI decomposition, we constructed a compression protocol that shows a substantial reduction of the compression rate with a finite approximation allowed. We explicitly showed a reduction of the compression rate by several examples. Moreover, we investigated blind compression of classical ensembles to analyze general properties of our compression protocol. Slightly modifying our protocol, we proposed two compression methods for two-state classical ensembles including the flat state, namely, the arithmetic mean method and the geometric mean method. We numerically investigated these compression methods of two-state classical ensembles and discovered that the arithmetic mean method performs better than the geometric mean method.

As a future direction, it would be interesting to analyze the protocol more generally. As noted in the previous sections, our protocol does not necessarily show a large reduction of the compression rate for any input. For example, if an allowed approximation is sufficiently small, our protocol might become equal to the optimal compression rate with a vanishing error, obtained in Ref. [1]. The difficulty of the analysis lies in the absence of an approximation of the KI decomposition. While finite approximations can significantly lower the compression rate as we showed in this chapter because of the instability of the KI decomposition, the instability also makes the general analysis intractable. As the first step of the general analysis, it would be interesting to investigate compression protocols on general two-state quantum ensembles. We numerically studied compression protocols for two-state classical ensembles in this chapter, which may give hints for further research of general approximate protocols. For a two-state ensemble, we may consider that one of the two states is diagonal by choosing an appropriate basis. Fixing the basis would be helpful to study a new approximate compression protocol and the dependency of the compression protocol on allowed errors.

Thus, we shed light on practical quantum data compression with the allowance of approximations through our research. Then, it would be interesting to study protocols with finite approximations that work more generally for further understandings of blind quantum data compression.

Chapter 4

Structure of Quantum Operations Approximately Preserving Quantum Ensemble

In this chapter, we discuss an approximate structure of quantum ensembles. In Ref. [4], the authors not only investigated the structure of a given quantum ensemble, called the KI decomposition, they also analyzed a structure of quantum operations that preserves the quantum ensemble, that is, operations with fixed points including the states belonging to the ensemble. As we saw in Chapter. 3, the KI structure plays a fundamental role in blind data compression, and an approximation of the KI structure would advance the analysis of blind data compression. It has been argued, however, that the approximation of the KI structure is generally intractable to analyze because of the algorithmic construction of the KI decomposition [4]. In this chapter, instead of analyzing the approximate KI structure, we investigate an approximate structure of quantum operations. We prove that a quantum channel can be approximated by a block diagonal structure that is determined by a given quantum ensemble if it approximately preserves the ensemble. We first briefly overview previous results in Sec. 4.1. We then show our results in Sec. 4.2. We summarize and discuss our results in Sec. 4.3.

4.1 Brief Overview of Structure of Quantum Operations that Preserve a Quantum Ensemble

In this section, we briefly review the approximate structure of a quantum ensemble. First, recalling Theorem 2.49 in Chapter 2, for a given ensemble $\{p_x, \rho_x\}_{x\in\Sigma}$, we have the unique KI decomposition

$$\rho_x \coloneqq \bigoplus_{l\in\Xi} q^{(x,l)} \rho_Q^{(x.l)} \otimes \rho_R^{(l)}. \tag{4.1}$$

In Ref. [4], they also gives a structure of quantum operations that preserve the quantum states in the ensemble. Suppose that $\mathcal{T}_U \in \mathrm{C}(\mathcal{H}_A)$ is a quantum channel with a Stinespring representation $U \in \mathrm{U}(\mathcal{H}_A \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_A \otimes \mathcal{H}_E)$; that is,

$$\mathcal{T}_U(\cdot) = \mathrm{Tr}_E(U(\cdot \otimes \omega_E)U^\dagger) \tag{4.2}$$

where E is an environment system, $\omega_E \coloneqq |\omega\rangle\langle\omega|_E \in \mathrm{D}(\mathcal{H}_E)$ is a reference state on E, and $\mathcal{H}_E^{(\omega)}$ is the subspace of $\mathcal{H}_E$ spanned by $\{|\omega\rangle_E\}$. Then, suppose that this operation preserves the ensemble; that is, for all $x \in \Sigma$, it holds that

$$\mathcal{T}_U(\rho_x) = \rho_x. \tag{4.3}$$

Then, the problem here is what structure the channel $\mathcal{T}_U$ has. Intuitively, considering the KI structure of the quantum ensemble, one may suppose that U does nothing to the ensemble's non-redundant classical/quantum parts. Indeed, the following theorem captures this intuition.

Theorem 4.1. *Let $\{p_x, \rho_x\}_{x\in\Sigma}$ be a quantum ensemble with the KI decomposition*

$$\rho_x \coloneqq \bigoplus_{l\in\Xi} q^{(x,l)} \rho_Q^{(x.l)} \otimes \rho_R^{(l)} \tag{4.4}$$

for all $x \in \Sigma$. Let $\mathcal{T}_U \in \mathrm{C}(\mathcal{H}_A)$ be a quantum channel satisfying

$$\mathcal{T}_U(\rho_x) = \rho_x \tag{4.5}$$

for all $x \in \Sigma$. Then, the isometry $U \in \mathrm{U}(\mathcal{H}_A \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_A \otimes \mathcal{H}_E)$ corresponding to $\mathcal{T}_U$ can be decomposed into the following form:

$$U = \bigoplus_{l\in\Sigma} I_Q^{(l)} \otimes U^{(l)} \tag{4.6}$$

with some isometries $U^{(l)} \in \mathrm{U}(\mathcal{H}_R^{(l)} \otimes \mathcal{H}_E^{(\omega)}, \mathcal{H}_R^{(l)} \otimes \mathcal{H}_E)$.

Thus, when a quantum channel preserves a given ensemble, an isometry corresponding to its Stinespring representation should have a block structure corresponding to the KI structure of the ensemble. The theorem is based on properties of quantum channels stated in the following theorems, on which we mainly focus in this thesis. The first theorem states that a quantum channel can be decomposed into two parts if it preserves two given states.

Theorem 4.2. *Let $\mathcal{H}$ be a quantum system. Let $\rho \in \mathrm{D}(\mathcal{H})$ and $\rho' \in \mathrm{D}(\mathcal{H})$ be density operators in this quantum system. Consider the decomposition of quantum system*

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2, \tag{4.7}$$

where $\mathcal{H}_1$ is the Hilbert space spanned by the eigenvectors corresponding to the positive eigenvalues of $\rho - \rho'$. Suppose that we have a quantum channel $\mathcal{T}_U$ with Stinespring representation U satisfying

$$\mathcal{T}_U(\rho) = \rho, \tag{4.8}$$
$$\mathcal{T}_U(\rho') = \rho'. \tag{4.9}$$

Then, U can be decomposed into a block diagonal form; that is,

$$U = P_1 U P_1 \oplus P_2 U P_2, \tag{4.10}$$

where P_i is the projection onto the Hilbert space $\mathcal{H}_i$ for $i = 1, 2$.

The second theorem asserts that when we have some block structure of a quantum channel, we can further decompose the channel if a state preserved by the channel has off diagonal elements with respect to the given block structure.

Theorem 4.3. *Let $\mathcal{H}$ be a Hilbert space with decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Let $\rho \in \mathrm{D}(\mathcal{H})$ be a density operator on $\mathcal{H}$ with $P_2 \rho P_1 \neq 0$, where P_i is the projection onto $\mathcal{H}_i$ for $i = 1, 2$. Let $\mathcal{G}_1$ be the support of $P_2 \rho P_1$ and $\mathcal{G}_2$ be the image of $P_2 \rho P_1$, and define $\mathcal{G}_i^{\perp} \coloneqq \mathcal{H}_i - \mathcal{G}_i$ for $i = 1, 2$. Suppose that a pair of unitary operators U_1 and U_2 satisfies*

$$\mathcal{T}_{U_1 \oplus U_2}(\rho) = \rho. \tag{4.11}$$

Then, there exists V_i and $V_i^{\perp}$ such that

$$U_i = V_i \oplus V_i^{\perp} \tag{4.12}$$

for $i = 1, 2$.

This theorem is used to refine a block diagonal structure obtained in Theorem 4.2. While Theorem 4.2 gives a block diagonal structure of the Stinespring representation of a quantum channel preserving a quantum ensemble, it is not guaranteed that the structure of the channel completely corresponds to the KI decomposition of the ensemble in the sense of Theorem 4.1. Theorem 4.3 ensures that when a resulting structure of the quantum channel does not correspond to the KI decomposition of the ensemble; that is, when some states in the ensemble off block diagonal elements with respect to the structure of the channel, we can refine the structure of the channel so that we can resolve the off block diagonal elements. Thus, as we can see in these theorems, when a set of quantum states are given, a quantum channel preserving the set will have a block diagonal structure.

4.2 Our results: Approximate Structure of Quantum Operation Almost Preserving Quantum Ensemble

Here, we show our main results. We consider a variation of Theorem 4.2 and Theorem 4.3, that is, an approximate structure of an isometry corresponding to a channel approximately preserving a given set of states. Before moving to our main theorems corresponding to Theorem 4.2 and Theorem 4.3, we show several useful lemmas.

The first lemma gives an upper bound of the norm of a linear operator restricted to some subspace. When we have the norm of the product of the operator and a linear operator on the subsystem, we can upper-bound the target norm by using the norm we have.

Lemma 4.4. *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be quantum systems. Let $X \in \mathrm{L}(\mathcal{H}_A, \mathcal{H}_B)$ be a linear operator and $Y \in \mathrm{L}(\mathcal{H}_B)$ be a normal operator. Let P_Y be the projection onto the subspace with non-zero eigenvalues of Y. Then, it holds that*

$$\|XP_Y\|_2 \leqq \frac{\|XY\|_2}{s_{\min}(Y)}, \tag{4.13}$$

where $s_{\min}(Y)$ is the smallest singular value of Y.

Proof. Take the basis with which we can write

$$Y = \begin{pmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_k & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} \tag{4.14}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the nonzero eigenvalues of Y. We omit the off diagonal elements, all of which are zero. Hereafter, when we explicitly write a matrix, we may omit elements equal to zero. In this basis, write X as

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \tag{4.15}$$

where $m \coloneqq D_{\mathcal{H}_A}$ and $n \coloneqq D_{\mathcal{H}_B}$. Then, it holds that

$$XY = \begin{pmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} & \cdots & \lambda_k x_{1k} & 0 & \cdots & 0 \\ \lambda_1 x_{21} & \lambda_2 x_{22} & \cdots & \lambda_k x_{2k} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 x_{n1} & \lambda_2 x_{n2} & \cdots & \lambda_k x_{nk} & 0 & \cdots & 0 \end{pmatrix} \tag{4.16}$$

Now, we have that

$$\|XY\|_2^2 = \sum_{\substack{i=1,\ldots,n \\ j=1,\ldots,k}} |x_{ij}|^2 |\lambda_j|^2 \geqq s_{\min}(Y)^2 \sum_{\substack{i=1,\ldots,n \\ j=1,\ldots,k}} |x_{ij}|^2 = s_{\min}(Y)^2 \|XP_Y\|_2^2, \tag{4.17}$$

which implies that

$$\|XP_Y\|_2 \leqq \frac{\|XY\|_2}{s_{\min}(Y)}. \tag{4.18}$$

□

Remark 4.5. We cannot eliminate the $1/s_{\min}(Y)$ term in Eq. (4.18). Let P be a projection operator and suppose that $Y \coloneqq \alpha P$ for some nonzero complex number $\alpha \in \mathbb{C}$. Then, it

holds that

$$\begin{aligned}\|XP_Y\|_2 &= \left\|X\frac{Y}{\alpha}\right\|_2 \\ &= \frac{1}{|\alpha|}\|XY\|_2 \\ &= \frac{1}{s_{\min}(Y)}\|XY\|_2,\end{aligned} \tag{4.19}$$

which implies that the equality in Eq. (4.13) is achieved.

The next lemma gives a fundamental property of a quantum channel that approximately preserves a Hermitian operator. When a quantum channel approximately preserves a Hermitian operator, a Stinespring representation of the channel can be approximately decomposed into two subsystems corresponding to the Jordan-Haan decomposition of the Hermitian operator.

Lemma 4.6. *Let $\mathcal{H}$ be a quantum system. Let $O \in \mathrm{L}(\mathcal{H})$ be a Hermitian operator with at least one positive eigenvalue. Consider the following decomposition of O into two positive operators:*

$$O = O_+ - O_- \tag{4.20}$$

where O_+ is the positive operator corresponding to the positive eigenvalues of O. (That is, O_- is the positive operator corresponding to the non-positive eigenvalues of O.) Let P_+ be the projector onto the space spanned by the eigenstates of O_+ and P_- be the projector onto the space spanned by the eigenstates of O_-. Suppose that a quantum channel $\mathcal{T}_U$ with the Stinespring representation U almost preserves O; that is, for a fixed positive number ϵ, it holds that

$$\|\mathcal{T}_U(O) - O\|_1 \leqq \epsilon. \tag{4.21}$$

Then, it holds that

$$\|[P_+ \otimes \mathbb{1}_E, U](P_+ \otimes \omega_E)\|_2 \leqq \sqrt{\frac{\epsilon}{s_{\min}(O_+)}}. \tag{4.22}$$

Proof. First, by the linearity of $\mathcal{T}_U$, it holds that

$$\begin{aligned}\mathrm{Tr}[P_+\mathcal{T}_U(O)] &= \mathrm{Tr}[P_+\mathcal{T}_U(O_+)] - \mathrm{Tr}[P_+\mathcal{T}_U(O_-)] \\ &= \mathrm{Tr}[\mathcal{T}_U(O_+)] - \mathrm{Tr}[P_-\mathcal{T}_U(O_+)] - \mathrm{Tr}[P_+\mathcal{T}_U(O_-)] \\ &= \mathrm{Tr}[O_+] - \mathrm{Tr}[P_-\mathcal{T}_U(O_+)] - \mathrm{Tr}[P_+\mathcal{T}_U(O_-)]\end{aligned} \tag{4.23}$$

The third equality holds because $I_{\mathcal{H}} = P_+ + P_-$, and the last equality holds because $\mathcal{T}_U$ is trace-preserving. Now, it holds that

$$\begin{aligned} \mathrm{Tr}[P_-\mathcal{T}_U(O_+)] + \mathrm{Tr}[P_+\mathcal{T}_U(O_-)] &= |\mathrm{Tr}[P_-\mathcal{T}_U(O_+)] + \mathrm{Tr}[P_+\mathcal{T}_U(O_-)]| \\ &= |\mathrm{Tr}[P_+\mathcal{T}_U(O)] - \mathrm{Tr}[O_+]| \\ &= |\mathrm{Tr}[P_+(\mathcal{T}_U(O) - O)]| \\ &\leqq \|P_+\|_\infty \|\mathcal{T}_U(O) - O\|_1 \\ &\leqq \epsilon. \end{aligned} \tag{4.24}$$

The first inequality follows from Proposition 2.38. Since inequalities $\mathrm{Tr}[P_-\mathcal{T}_U(O_+)] \geqq 0$ and $\mathrm{Tr}[P_+\mathcal{T}_U(O_-)] \geqq 0$ hold, we have

$$\mathrm{Tr}[P_-\mathcal{T}_U(O_+)] \leqq \epsilon. \tag{4.25}$$

Now, observe that

$$\begin{aligned} \mathrm{Tr}[P_-\mathcal{T}_U(O_+)] &= \mathrm{Tr}[(P_- \otimes I_E)U(O_+ \otimes \omega_E)U^\dagger] \\ &= \|(P_- \otimes I_E)U(\sqrt{O_+} \otimes \omega_E)\|_2^2, \end{aligned} \tag{4.26}$$

which implies that

$$\|(P_- \otimes I_E)U(\sqrt{O_+} \otimes \omega_E)\|_2 \leqq \sqrt{\epsilon}. \tag{4.27}$$

By using Lemma 4.4, we have

$$\|(P_- \otimes I_E)U(P_+ \otimes \omega_E)\|_2 \leqq \frac{\|(P_- \otimes I_E)U(\sqrt{O_+} \otimes \omega_E)\|_2}{s_{\min}(\sqrt{O_+})} \leqq \sqrt{\frac{\epsilon}{s_{\min}(O_+)}}. \tag{4.28}$$

Therefore, noticing that $P_- = I_{\mathcal{H}} - P_+$, we have

$$\|[P_+ \otimes \mathbb{1}_E, U](P_+ \otimes \omega_E)\|_2 \leqq \sqrt{\frac{\epsilon}{s_{\min}(O_+)}}. \tag{4.29}$$

□

We now prove the third lemma, which shows a property of quantum channels that almost preserve a state on a given quantum system consisting of two orthogonal systems. The lemma states that if a Stinespring representation of the channel has an approximate block diagonal structure with respect to one of the two orthogonal systems, it must be also approximately block diagonal with respect to the other system.

Lemma 4.7. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be quantum systems. Define a quantum system $\mathcal{H} \coloneqq \mathcal{H}_1 \oplus \mathcal{H}_2$. Let ρ be a density operator and suppose that $\mathrm{supp}(\rho) = \mathcal{H}$. Suppose that a quantum channel $\mathcal{T}_U$ with the Stinespring representation U almost preserves ρ; that is, for a fixed positive number ϵ, it holds that*

$$\|\mathcal{T}_U(\rho) - \rho\|_1 \leqq \epsilon. \tag{4.30}$$

Suppose also that

$$\|[P_1 \otimes I_E, U](P_1 \otimes \omega_E)\|_2 \leqq \delta. \tag{4.31}$$

Then, it holds that

$$\|[P_2 \otimes I_E, U](P_2 \otimes \omega_E)\|_2 \leqq \sqrt{\frac{\epsilon + 2\delta + \delta^2}{s_{\min}(P_2 \rho P_2)}}. \tag{4.32}$$

Proof. First, observe that Eq. (4.31) implies that

$$\|(P_2 \otimes I_E)U(P_1 \otimes \omega_E)\|_2 \leqq \delta \tag{4.33}$$

because $P_1 + P_2 = I_{\mathcal{H}}$. Now, it holds that

$$\begin{aligned}
|\mathrm{Tr}[P_1 \mathcal{T}_U(P_2 \rho P_2)]| &= |\mathrm{Tr}[\mathcal{T}_U(P_2 \rho P_2) - P_2 \mathcal{T}_U(P_2 \rho P_2)]| \\
&= |\mathrm{Tr}[P_2 \rho - P_2 \mathcal{T}_U(P_2 \rho P_2)]| \\
&= |\mathrm{Tr}[P_2(\rho - \mathcal{T}_U(\rho))] + \mathrm{Tr}[P_2 \mathcal{T}_U(P_1 \rho P_1 + P_1 \rho P_2 + P_2 \rho P_1)]| \\
&\leqq |\mathrm{Tr}[P_2(\rho - \mathcal{T}_U(\rho))]| \\
&\quad + |\mathrm{Tr}[P_2 \mathcal{T}_U(P_1 \rho P_1)]| + |\mathrm{Tr}[P_2 \mathcal{T}_U(P_1 \rho P_2)]| + |\mathrm{Tr}[P_2 \mathcal{T}_U(P_2 \rho P_1)]|.
\end{aligned} \tag{4.34}$$

Then, as for the first term of Eq. (4.34), by Proposition 2.38, we have that

$$|\mathrm{Tr}[P_2(\rho - \mathcal{T}_U(\rho))]| \leqq \|P_2\|_\infty \|\rho - \mathcal{T}_U(\rho)\|_1 \leqq \epsilon. \tag{4.35}$$

In addition, we can upper-bound the second term of Eq. (4.34).

$$\begin{aligned}
|\mathrm{Tr}[P_2 \mathcal{T}_U(P_1 \rho P_1)]| &= |\mathrm{Tr}[(P_2 \otimes I_E)U(P_1 \otimes \omega_E)(\rho \otimes \omega_E)(P_1 \otimes \omega_E)U^\dagger(P_2 \otimes I_E)]| \\
&= \|(P_2 \otimes I_E)U(P_1 \otimes \omega_E)(\sqrt{\rho} \otimes \omega_E)\|_2^2 \\
&\leqq \|(P_2 \otimes I_E)U(P_1 \otimes \omega_E)\|_2^2 \|\sqrt{\rho} \otimes \omega_E\|_2^2 \\
&\leqq \delta^2.
\end{aligned} \tag{4.36}$$

Moreover, we can upper-bound the third term of Eq. (4.34) in the following way.

$$\begin{aligned}
|\mathrm{Tr}[P_2 \mathcal{T}_U(P_1 \rho P_2)]| &= |\mathrm{Tr}[(P_2 \otimes I_E)U(P_1 \otimes \omega_E)(\rho \otimes \omega_E)(P_2 \otimes \omega_E)U^\dagger]| \\
&= \|(P_2 \otimes I_E)U(P_1 \otimes \omega_E)\|_2 \|U(P_2 \otimes \omega_E)(\rho \otimes \omega_E)\|_2 \\
&\leqq \delta \|(P_2 \otimes \omega_E)(\rho \otimes \omega_E)\|_2 \\
&\leqq \delta.
\end{aligned} \tag{4.37}$$

Using similar argument, we have

$$\mathrm{Tr}[P_2\mathcal{T}_U(P_2\rho P_1)]| \leqq \delta. \tag{4.38}$$

Therefore, we have

$$\begin{aligned}\|(P_1\otimes I_E)U(\sqrt{P_2\rho P_2}\otimes\omega_E)\|_2^2 &= |\mathrm{Tr}[P_1\mathcal{T}_U(P_2\rho P_2)]| \\ &\leqq \epsilon + 2\delta + \delta^2.\end{aligned} \tag{4.39}$$

Using Lemma 4.4, we have

$$\|(P_1\otimes I_E)U(P_2\otimes\omega_E)\|_2 \leqq \frac{\|(P_1\otimes I_E)U(\sqrt{P_2\rho P_2}\otimes\omega_E)\|_2}{s_{\min}(\sqrt{P_2\rho P_2})} \leqq \sqrt{\frac{\epsilon+2\delta+\delta^2}{s_{\min}(P_2\rho P_2)}}. \tag{4.40}$$

□

Using these three lemmas, we can prove our main theorems. First, we show our first theorem, which shows that a quantum channel has an approximate block diagonal structure determined by the given states when it preserves two quantum states approximately. The statement of the following theorem is considered as an approximate version of Theorem 4.2; hence, we successfully approximate the structure shown in Eq. (4.22).

Theorem 4.8. *Let $\mathcal{H}$ be a quantum system. Let ρ and ρ' be quantum states, and suppose that $\mathrm{supp}(\rho+\rho')=\mathcal{H}$. Define $O := \rho-\rho'$, and consider the decomposition of $O = O_+ - O_-$, where O_+ is the positive operator corresponding to the positive eigenvalues of O. Let P_+ be the projector onto the space spanned by the eigenstates of O_+ and P_+ be the projector onto the space spanned by the eigenstates of O_-. In addition, define a quantum state*

$$\sigma := \frac{\rho+\rho'}{2} \tag{4.41}$$

Suppose that a quantum channel $\mathcal{T}_U$ almost preserves the two given states; that is,

$$\|\mathcal{T}_U(\rho)-\rho\|_1 \leqq \epsilon, \tag{4.42}$$

$$\|\mathcal{T}_U(\rho')-\rho'\|_1 \leqq \epsilon'. \tag{4.43}$$

Then, it holds that

$$\begin{aligned}&\|U-(P_+\otimes I_E)U(P_+\otimes\omega_E)\oplus(P_-\otimes I_E)U(P_-\otimes\omega_E)\|_2 \\ &\leqq \sqrt{\frac{\epsilon+\epsilon'}{s_{\min}(O_+)}} + \frac{1}{\sqrt{s_{\min}(P_-\sigma P_-)}}\sqrt{\left(\frac{1}{2}+\frac{1}{s_{\min}(O_+)}\right)(\epsilon+\epsilon')+2\sqrt{\frac{\epsilon+\epsilon'}{s_{\min}(O_+)}}}.\end{aligned} \tag{4.44}$$

Proof. From Eqs. (4.42) and (4.43), we have

$$\|\mathcal{T}_U(O) - O\|_1 \leqq \epsilon + \epsilon', \tag{4.45}$$

$$\|\mathcal{T}_U(\sigma) - \sigma\|_1 \leqq \frac{\epsilon + \epsilon'}{2}. \tag{4.46}$$

Then, by Lemma 4.6, we have

$$\|[P_+ \otimes \mathbb{1}_E, U](P_+ \otimes \omega_E)\|_2 \leqq \sqrt{\frac{\epsilon + \epsilon'}{s_{\min}(O_+)}}. \tag{4.47}$$

Therefore, by Lemma 4.7, we also have

$$\|[P_- \otimes \mathbb{1}_E, U](P_- \otimes \omega_E)\|_2 \leqq \frac{1}{\sqrt{s_{\min}(P_2 \sigma P_2)}} \sqrt{\left(\frac{1}{2} + \frac{1}{s_{\min}(O_+)}\right)(\epsilon + \epsilon') + 2\sqrt{\frac{\epsilon + \epsilon'}{s_{\min}(O_+)}}}. \tag{4.48}$$

Combining these two equations, we have

$$\begin{aligned}
&\|U - (P_+ \otimes I_E)U(P_+ \otimes \omega_E) \oplus (P_- \otimes I_E)U(P_- \otimes \omega_E)\|_2 \\
&= \|U(P_+ \otimes \omega_E) + U(P_- \otimes \omega_E) - (P_+ \otimes I_E)U(P_+ \otimes \omega_E) + (P_- \otimes I_E)U(P_- \otimes \omega_E)\|_2 \\
&\leqq \|U(P_+ \otimes \omega_E) - (P_+ \otimes I_E)U(P_+ \otimes \omega_E)\|_2 + \|U(P_- \otimes \omega_E) - (P_- \otimes I_E)U(P_- \otimes \omega_E)\|_2 \\
&= \|[P_+ \otimes \mathbb{1}_E, U](P_+ \otimes \omega_E)\|_2 + \|[P_- \otimes \mathbb{1}_E, U](P_- \otimes \omega_E)\|_2 \\
&\leqq \sqrt{\frac{\epsilon + \epsilon'}{s_{\min}(O_+)}} + \frac{1}{\sqrt{s_{\min}(P_- \sigma P_-)}} \sqrt{\left(\frac{1}{2} + \frac{1}{s_{\min}(O_+)}\right)(\epsilon + \epsilon') + 2\sqrt{\frac{\epsilon + \epsilon'}{s_{\min}(O_+)}}}.
\end{aligned} \tag{4.49}$$

□

Next, we show our second theorem, which shows that a quantum channel with a block structure almost preserving a state with off block diagonal elements has a finer approximate block structure. The following theorem is considered as a tool to refine the structure obtained through Theorem 4.8, which is essential to achieve a sophisticated approximate block diagonal structure of quantum channels.

Theorem 4.9. *Let $\mathcal{H}$ be a quantum system with decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Let ρ be a density operator with $\mathrm{supp}(\rho) = \mathcal{H}$ and $P_2 \rho P_1 \neq 0$, where P_i is the projection onto $\mathcal{H}_i$ for $i = 1, 2$. Let $\mathcal{G}_1$ be the support of $P_2 \rho P_1$ and $\mathcal{G}_2$ be the image of $P_2 \rho P_1$, and define*

$\mathcal{G}_i^{\perp} \coloneqq \mathcal{H}_i - \mathcal{G}_i$ *for* $i = 1, 2$. *Consider a polar decomposition* $P_2 \rho P_1 \coloneqq WQ$, *where* W *is a unitary operator* Q *is a positive operator. Note that* $W^\dagger W$ *is the projection onto* $\mathcal{G}_1$ *and that* $WW^\dagger$ *is the projection onto* $\mathcal{G}_2$. *Here, we have* $W^2 = 0$ *because* P_1 *and* P_2 *are orthogonal.*

Suppose that a pair of unitary operators U_1 *and* U_2 *satisfies*

$$\|\mathcal{T}_{U_1 \oplus U_2}(\rho) - \rho\|_1 \leqq \epsilon. \tag{4.50}$$

Then, it holds that

$$\begin{aligned} &\|U_1 - (W^\dagger W \otimes I_E) U_1 (W^\dagger W \otimes \omega_E) \oplus ((P_1 - W^\dagger W) \otimes I_E) U_1 ((P_1 - W^\dagger W) \otimes \omega_E)\|_2 \\ &\leqq \alpha_O \sqrt{2\epsilon} + \sqrt{\frac{2(1+\alpha_O^2)\epsilon + 2\alpha_O \sqrt{2\epsilon}}{s_{\min}((P_1 - W^\dagger W)\rho(P_1 - W^\dagger W))}}, \end{aligned} \tag{4.51}$$

where

$$\alpha_O \coloneqq \frac{1}{\sqrt{s_{\min}(O_+)}} + \frac{1}{\sqrt{s_{\min}(O_-)}} \tag{4.52}$$

with a Hermitian operator $O \coloneqq P_1 \rho P_2 + P_2 \rho P_1$. *Here,* $O_\pm \coloneqq P_\pm O P_\pm$ *where* $P_\pm$ *is the projection operator onto the space spanned by the eigenstates corresponding to the nonzero positive/negative eigenvalues of* O. *It also holds that*

$$\begin{aligned} &\|U_2 - (WW^\dagger \otimes I_E) U_2 (WW^\dagger \otimes \omega_E) \oplus ((P_2 - WW^\dagger) \otimes I_E) U_1 ((P_1 - WW^\dagger) \otimes \omega_E)\|_2 \\ &\leqq \alpha_O \sqrt{2\epsilon} + \sqrt{\frac{2(1+\alpha_O^2)\epsilon + 2\alpha_O \sqrt{2\epsilon}}{s_{\min}((P_2 - WW^\dagger)\rho(P_2 - WW^\dagger))}}. \end{aligned} \tag{4.53}$$

Proof. First, it holds that

$$\begin{aligned} &\|[(P_+ + P_-) \otimes I_E, U_1 \oplus U_2](P_+ + P_- \otimes \omega_E)\|_2 \\ &\leqq \|((P_+ + P_-) \otimes I_E)(U_1 \oplus U_2)((P_+ + P_-) \otimes \omega_E) - (U_1 \oplus U_2)((P_+ + P_-) \otimes \omega_E)\|_2 \\ &\leqq \|((P_+ + P_-) \otimes I_E)[(U_1 \oplus U_2)((P_+ + P_-) \otimes \omega_E) \\ &\qquad - (P_+ \otimes I_E)(U_1 \oplus U_2)(P_+ \otimes \omega_E) - (P_+ \otimes I_E)(U_1 \oplus U_2)(P_+ \otimes \omega_E)] \\ &\quad + (P_+ \otimes I_E)(U_1 \oplus U_2)(P_+ \otimes \omega_E) + (P_+ \otimes I_E)(U_1 \oplus U_2)(P_+ \otimes \omega_E) \\ &\qquad - (U_1 \oplus U_2)((P_+ + P_-) \otimes \omega_E)\|_2 \\ &= \|((I_{\mathcal{H}} - P_+ - P_-) \otimes I_E)[[P_+ \otimes I_E, U_1 \oplus U_2](P_+ \otimes \omega_E) + [P_- \otimes I_E, U_1 \oplus U_2](P_- \otimes \omega_E)]\|_2 \\ &\leqq \|[P_+ \otimes I_E, U_1 \oplus U_2](P_+ \otimes \omega_E) + [P_- \otimes I_E, U_1 \oplus U_2](P_- \otimes \omega_E)\|_2 \\ &\leqq \|[P_+ \otimes I_E, U_1 \oplus U_2](P_+ \otimes \omega_E)\|_2 + \|[P_- \otimes I_E, U_1 \oplus U_2](P_- \otimes \omega_E)\|_2. \end{aligned} \tag{4.54}$$

On the other hand, it holds that

$$\begin{aligned}\|\mathcal{T}_{U_1\oplus U_2}(P_i\rho P_j) - P_i\rho P_j\|_1 &= \|P_i\mathcal{T}_{U_1\oplus U_2}(\rho)P_j - P_i\rho P_j - P_i\rho P_j\| \\ &\leqq \|P_i\|_\infty\|\mathcal{T}_{U_1\oplus U_2}(\rho) - \rho\|_1\|P_j\|_\infty \\ &= \|\mathcal{T}_{U_1\oplus U_2}(\rho) - \rho\|_1 \\ &\leqq \epsilon\end{aligned} \tag{4.55}$$

for $i, j = 1, 2$. Therefore, we have

$$\|\mathcal{T}_{U_1\oplus U_2}(O) - O\|_1 \leqq 2\epsilon. \tag{4.56}$$

By Lemma 4.6, it holds that

$$\|[P_+ \otimes I_E, U_1 \oplus U_2](P_+ \otimes \omega_E)\|_2 \leqq \sqrt{\frac{2\epsilon}{s_{\min}(O_+)}} \tag{4.57}$$

$$\|[P_- \otimes I_E, U_1 \oplus U_2](P_- \otimes \omega_E)\|_2 \leqq \sqrt{\frac{2\epsilon}{s_{\min}(O_-)}}. \tag{4.58}$$

Then, we have

$$\|[(P_+ + P_-) \otimes I_E, U_1 \oplus U_2]((P_+ + P_-) \otimes \omega_E)\|_2 \leqq \alpha_O\sqrt{2\epsilon}. \tag{4.59}$$

Since $P_+ + P_- = W^\dagger W + WW^\dagger$, we have that

$$\|[(W^\dagger W + WW^\dagger) \otimes I_E, U_1 \oplus U_2]((W^\dagger W + WW^\dagger) \otimes \omega_E)\|_2 \leqq \alpha_O\sqrt{2\epsilon}, \tag{4.60}$$

which implies that

$$\|[(W^\dagger W) \otimes I_E, U_1]((W^\dagger W) \otimes \omega_E)\|_2 \leqq \alpha_O\sqrt{2\epsilon} \tag{4.61}$$

$$\|[(WW^\dagger) \otimes I_E, U_2]((WW^\dagger) \otimes \omega_E)\|_2 \leqq \alpha_O\sqrt{2\epsilon}. \tag{4.62}$$

Using Lemma 4.7, we have the inequalities (4.51) and (4.53). □

From Theorem 4.8 and Theorem 4.9, we can give an approximate block diagonal structure of a quantum channel that approximately preserves a given quantum ensemble. Theorem 4.8 leads to a rough structure, and then Theorem 4.9 allows us to refine the structure. On the other hand, one may notice that approximate Stinespring operators obtained through approximations, shown in Eqs. (4.44), (4.51), and (4.53), are not necessarily isometries because for a unitary operator U and a projection operator P, PUP is not necessarily

a unitary operator. To resolve this problem, we show the following theorem, which states that when a unitary operator can be approximated by a block diagonal operator, the unitary can indeed be approximated by a block diagonal unitary operator. The following theorem allows us to approximate a quantum channel almost preserving a quantum ensemble as another quantum channel with a block diagonal structure.

Theorem 4.10. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be quantum systems, and let P_i be the projection onto $\mathcal{H}_i$ for $i = 1, 2$. Suppose that a unitary operator $U \in \mathrm{U}(\mathcal{H}_1 \oplus \mathcal{H}_2)$ satisfies*

$$\|U - P_1 U P_1 \oplus P_2 U P_2\|_2 \leqq \epsilon \tag{4.63}$$

for a fixed positive real number $0 < \epsilon < 1$. Then, there exists a pair of unitary operators $U_1 \in \mathrm{U}(\mathcal{H}_1)$ and $U_2 \in \mathrm{U}(\mathcal{H}_2)$ such that

$$\|U - U_1 \oplus U_2\|_2 \leqq \max\left\{\frac{\epsilon}{\sqrt{1-\epsilon}}, 2\epsilon\right\}. \tag{4.64}$$

Proof. Consider singular value decomposition of $P_1 U P_1$ and $P_2 U P_2$

$$P_1 U P_1 \coloneqq W_1 S_1 V_1, \tag{4.65}$$
$$P_2 U P_2 \coloneqq W_2 S_2 V_2, \tag{4.66}$$

where S_i is a diagonal matrix with the singular values of $P_i U P_i$, and W_i and V_i are unitary matrices for the singular value decomposition, for $i = 1, 2$. By Eq. (4.63), we have

$$\|(W_1 \oplus W_2)^\dagger U (V_1 \oplus V_2)^\dagger - S_1 \oplus S_2\|_2 \leqq \epsilon \tag{4.67}$$

since the 2 norm does not change if we apply unitary operators. Here we define

$$U' \coloneqq (W_1 \oplus W_2)^\dagger U (V_1 \oplus V_2)^\dagger, \tag{4.68}$$
$$S \coloneqq S_1 \oplus S_2, \tag{4.69}$$

leading to

$$\|U' - S\|_2 \leqq \epsilon. \tag{4.70}$$

Note that U' is also a unitary matrix by definition and that S is a diagonal matrix. Then, considering a matrix representation

$$U' = \sum_{a,b \in \Sigma} U'_{ab} |a\rangle \langle b|, \tag{4.71}$$
$$S = \sum_{a \in \Sigma} S_{aa} |a\rangle \langle a| \tag{4.72}$$

with an alphabet Σ such that $\mathcal{H}_1 \oplus \mathcal{H}_2 = \mathbb{C}^\Sigma$, we have

$$\sum_{a\in\Sigma} |U'_{aa} - S_{aa}|^2 + \sum_{\substack{a,b\in\Sigma \\ a\neq b}} |U'_{ab}|^2 \leqq \epsilon^2. \tag{4.73}$$

There exists $\delta_1 \geqq 0$ and $\delta_2 \geqq 0$ such that $\delta_1 + \delta_2 \leqq \epsilon^2$ and

$$\sum_{a\in\Sigma} |U'_{aa} - S_{aa}|^2 = \delta_1, \tag{4.74}$$

$$\sum_{\substack{a,b\in\Sigma \\ a\neq b}} |U'_{ab}|^2 = \delta_2. \tag{4.75}$$

Then, it holds that

$$\begin{aligned} \|S - I_{\mathcal{H}_1\oplus\mathcal{H}_2}\|_2^2 &= \sum_{a\in\Sigma} |S_{aa} - 1|^2 \\ &= \sum_{a\in\Sigma} |S_{aa}|^2 - 2\sum_a S_{aa} + D_{\mathcal{H}_1\oplus\mathcal{H}_2} \\ &= \delta_1 + \delta_2 - 2\sum_{a\in\Sigma} (1 - \mathrm{Re}(U'_{aa}))\, S_{aa} \\ &\leqq \epsilon^2 \end{aligned} \tag{4.76}$$

because $S_{aa} \geqq 0$ and $1 - \mathrm{Re}(U'_{aa}) \geqq 0$ for all a. Then, we first consider the following chain of inequalities:

$$\|U - W_1V_1 \oplus W_2V_2\|_2 = \|U' - I_{\mathcal{H}_1\oplus\mathcal{H}_2}\|_2 \tag{4.77}$$

$$\leqq \|U' - S\|_2 + \|S - I_{\mathcal{H}_1\oplus\mathcal{H}_2}\|_2 \tag{4.78}$$

$$\leqq 2\epsilon. \tag{4.79}$$

On the other hand, we consider the following chain of inequalities:

$$\|U - W_1V_1 \oplus W_2V_2\|_2^2 = \|U' - I_{\mathcal{H}_1\oplus\mathcal{H}_2}\|_2^2 \tag{4.80}$$

$$= \sum_{a\in\Sigma} |U'_{aa} - 1|^2 + \sum_{\substack{a,b\in\Sigma \\ a\neq b}} |U'_{ab}|^2 \tag{4.81}$$

$$= 2D_{\mathcal{H}_1\oplus\mathcal{H}_2} - 2\sum_{a\in\Sigma} \mathrm{Re}(U'_{aa}) \tag{4.82}$$

$$= 2\sum_{a\in\Sigma} (1 - \mathrm{Re}(U'_{aa}))\,. \tag{4.83}$$

Now, since $\|S - I_{\mathcal{H}_1 \oplus \mathcal{H}_2}\|_2 \leqq \epsilon$, letting $S_{\min}$ denote the smallest (diagonal) element of S, we have

$$(S_{\min} - 1)^2 \leqq \|S - I_{\mathcal{H}_1 \oplus \mathcal{H}_2}\|_2^2 \leqq \epsilon^2, \tag{4.84}$$

which implies that

$$1 - \epsilon \leqq S_{\min} \leqq 1 + \epsilon. \tag{4.85}$$

Since it holds that

$$0 \leqq \|S - I_{\mathcal{H}_1 \oplus \mathcal{H}_2}\|_2^2 \leqq \epsilon^2 - 2 \sum_{a \in \Sigma} (1 - \mathrm{Re}(U'_{aa}))\, S_{aa}, \tag{4.86}$$

it follows that

$$2 S_{\min} \sum_{a \in \Sigma} (1 - \mathrm{Re}(U'_{aa})) \leqq 2 \sum_{a \in \Sigma} (1 - \mathrm{Re}(U'_{aa}))\, S_{aa} \leqq \epsilon^2, \tag{4.87}$$

which implies that

$$2 \sum_{a \in \Sigma} (1 - \mathrm{Re}(U'_{aa})) \leqq \frac{\epsilon^2}{S_{\min}} \leqq \frac{\epsilon^2}{1 - \epsilon}. \tag{4.88}$$

Therefore, we have

$$\|U - W_1 V_1 \oplus W_2 V_2\|_2 \leqq \frac{\epsilon}{\sqrt{1 - \epsilon}}. \tag{4.89}$$

Combining two inequalities, we have

$$\|U - W_1 V_1 \oplus W_2 V_2\|_2 \leqq \max\left\{ \frac{\epsilon}{\sqrt{1 - \epsilon}}, 2\epsilon \right\}. \tag{4.90}$$

Therefore, we can take $U_1 \coloneqq W_1 V_1$ and $U_2 \coloneqq W_2 V_2$. □

Thus, combining Theorems 4.8, 4.9, and 4.10, we have that a Stinespring representation of a quantum channel can be approximated by an isometry with a block diagonal structure when the channel approximately preserves some given set of states. Then, by repeatedly applying these theorems, we can find a fine approximate structure of a given quantum channel within allowed errors. Finally, we should note that the upper bounds obtained in Eqs. (4.44), (4.51), and (4.53) depend on the smallest singular values of some operators. While Remark 4.5 implies that we cannot remove this factor of the smallest singular values in a straightforward way, further investigating the tightness of the bounds would be interesting and needed to fully understand the approximate structure.

4.3 Summary and Discussions

In this chapter, we investigated a structure of quantum channels that almost preserve a given quantum ensemble. As noted in previous research [2,3], an approximate KI structure of quantum ensembles is hard to analyze because even an extremely small approximation causes a significant difference of the KI structure. At the same time, an approximate variation of KI decomposition is essential for further error analysis of the optimal rate of blind quantum data compression. To investigate an approximate KI structure, we consider a structure of quantum channel approximately preserving a quantum ensemble. We first proved that when a quantum channel approximately preserves two quantum states, the isometry corresponding to its Stinespring representation has an approximate block diagonal structure induced by the two states. The extent of approximation is determined by how well the channel preserves the given two states and the eigenvalues of the operator defined as the difference of the two states. The bound is tight in the sense that when the given two states are similar, the approximation becomes worse. Then, we also proved that when a quantum channel with a block diagonal Stinespring representation almost preserves a state with off block-diagonal elements, we can approximately refine the block diagonal structure. Finally, considering the fact that approximated operators appearing in these theorems are not necessarily isometries, we show a theorem ensuring that Stinespring representations in our theorems are indeed close to block diagonal *isometries*.

Although it is still hard to obtain a complete form of the approximate structure, we think our work makes progress towards analyzing such an approximate KI structure. It is essential to extend our results to obtain a more general approximate KI structure of quantum channels, leading to an approximate KI decomposition of quantum ensembles. We believe this analysis is also vital for a good approximate protocol for blind compression, which is discussed in Chapter 3.

Chapter 5

Conclusion

In this thesis, we investigated the quantum blind compression with finite local approximations and an approximate structure of quantum channels.

We constructed a novel compression protocol that works remarkably well with the allowance of approximations; that is, the protocol showed a substantial reduction of the compression rate for specific examples. Furthermore, our numerical experiments revealed that the arithmetic mean method performs well over the geometric mean method for the compression of classical ensembles with a local approximation. In addition, we successfully extended the key theorems shown in Ref. [4] to the approximate setup; that is, we proved that the Stinespring isometry corresponding to a quantum channel that approximately preserves a quantum ensemble can be approximated by a block diagonal isometry. We also discussed the tightness of the extent of the approximation with an example.

Thus, our results shed light on the instability of the rate of blind quantum data compression against approximations. Thanks to our new compression protocol, the much more efficient data transmission is achieved in a realistic setup with the allowance of approximations. Furthermore, our results on the approximation of KI decomposition not only contributes to understandings of restrictions for quantum operations that cause a small disturbance, which is vital in terms of fundamental quantum mechanics, but also provides us with insights into an approximate KI decomposition of quantum states that is essential to conduct rigorous error analysis of blind data compression and to construct better and more general protocol with approximations. We have not completely revealed the general approximate structure of quantum ensembles due to the difficulties of general analysis caused by the instability of KI decomposition; nevertheless, our protocol works surprisingly

better for some examples, and we showed a structure of quantum channels. We believe that our work is an essential first step of the general investigation of approximate KI decomposition, which is demanded for the study of blind quantum data compression with finite approximations and the further error analysis of blind compression.

References

[1] M. Koashi and N. Imoto, "Compressibility of quantum mixed-state signals," *Phys. Rev. Lett.*, vol. 87, p. 017902, 2001.

[2] A. Anshu, D. Leung, and D. Touchette, "Incompressibility of classical distributions," 2019. arXiv:1911.09126.

[3] Z. B. Khanian and A. Winter, "General mixed state quantum data compression with and without entanglement assistance," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1852–1857, 2020.

[4] M. Koashi and N. Imoto, "Operations that do not disturb partially known quantum states," *Phys. Rev. A*, vol. 66, p. 022318, 2002.

[5] E. Wakakuwa, A. Soeda, and M. Murao, "Markovianizing cost of tripartite quantum states," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1280–1298, 2017.

[6] H. Yamasaki and M. Murao, "Quantum state merging for arbitrarily small-dimensional systems," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3950–3972, 2019.

[7] P. Hayden, R. Jozsa, D. Petz, and A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality," *Commun. Math. Phys.*, vol. 246, no. 2, pp. 359–374, 2004.

[8] B. Ibinson, N. Linden, and A. Winter, "Robustness of quantum markov chains," *Commun. Math. Phys.*, vol. 277, no. 2, pp. 289–304, 2008.

[9] D. Sutter, *Approximate Quantum Markov Chains*, pp. 75–100. Cham: Springer International Publishing, 2018.

[10] A. Jenčová and D. Petz, "Sufficiency in quantum statistical inference," *Commun. Math. Phys.*, vol. 263, no. 1, pp. 259–276, 2006.

[11] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, "Information-preserving structures: A general framework for quantum zero-error information," *Phys. Rev. A*, vol. 82, p. 062306, 2010.

[12] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.

[13] C. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[14] C. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[15] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[16] B. Schumacher, "Quantum coding," *Phys. Rev. A*, vol. 51, pp. 2738–2747, 1995.

[17] R. Jozsa and B. Schumacher, "A new proof of the quantum noiseless coding theorem," *J. Mod. Opt.*, vol. 41, no. 12, pp. 2343–2349, 1994.

[18] H. Barnum, C. A. Fuchs, R. Jozsa, and B. Schumacher, "General fidelity limit for quantum channels," *Phys. Rev. A*, vol. 54, pp. 4707–4711, 1996.

[19] H.-K. Lo, "Quantum coding theorem for mixed states," *Opt. Commun.*, vol. 119, no. 5, pp. 552–556, 1995.

[20] R. Jozsa, M. Horodecki, P. Horodecki, and R. Horodecki, "Universal quantum information compression," *Phys. Rev. Lett.*, vol. 81, pp. 1714–1717, 1998.

[21] R. Jozsa and S. Presnell, "Universal quantum information compression and degrees of prior knowledge," *Proc. R. Soc. Lond. A*, vol. 459, no. 2040, pp. 3061–3077, 2003.

[22] C. H. Bennett, A. W. Harrow, and S. Lloyd, "Universal quantum data compression via nondestructive tomography," *Phys. Rev. A*, vol. 73, p. 032336, 2006.

[23] S. Braunstein, C. Fuchs, D. Gottesman, and H.-K. Lo, "A quantum analog of huffman coding," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1644–1649, 2000.

[24] M. Horodecki, "Limits for compression of quantum information carried by ensembles of mixed states," *Phys. Rev. A*, vol. 57, pp. 3364–3369, 1998.

[25] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "On quantum coding for ensembles of mixed states," *J. Phys. A*, vol. 34, no. 35, pp. 6767–6785, 2001.

[26] W. Dür, G. Vidal, and J. I. Cirac, "Visible compression of commuting mixed states," *Phys. Rev. A*, vol. 64, p. 022308, 2001.

[27] G. Kramer and S. A. Savari, "Quantum data compression of ensembles of mixed states with commuting density operators," 2001. arXiv:quant-ph/0101119.

[28] M. Horodecki, "Optimal compression for mixed signal states," *Phys. Rev. A*, vol. 61, p. 052309, 2000.

[29] M. Hayashi, "Optimal visible compression rate for mixed states is determined by entanglement of purification," *Phys. Rev. A*, vol. 73, p. 060301, 2006.

[30] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, "Remote state preparation," *Phys. Rev. Lett.*, vol. 87, p. 077902, 2001.

[31] C. Bennett, P. Hayden, D. Leung, P. Shor, and A. Winter, "Remote preparation of quantum states," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 56–74, 2005.

[32] A. Abeyesinghe, P. Hayden, G. Smith, and A. Winter, "Optimal superdense coding of entangled states," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3635–3641, 2006.

[33] R. Jain, J. Radhakrishnan, and P. Sen, "Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states," in *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pp. 429–438, 2002.

[34] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2926–2959, 2014.

[35] A. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, 1998.

[36] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, pp. 131–138, 1997.

[37] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, pp. 3081–3084, 1999.

[38] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, 2002.

[39] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, 2005.

[40] N. Cai, A. Winter, and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *Probl. Inf. Transm.*, vol. 40, no. 4, pp. 318–336, 2004.

[41] M.-H. Hsieh, Z. Luo, and T. Brun, "Secret-key-assisted private classical communication capacity over quantum channels," *Phys. Rev. A*, vol. 78, p. 042306, 2008.

[42] M. M. Wilde, "Comment on "secret-key-assisted private classical communication capacity over quantum channels"," *Phys. Rev. A*, vol. 83, p. 046303, 2011.

[43] P. Hayden, M. Horodecki, A. Winter, and J. Yard, "A decoupling approach to the quantum capacity," *Open Syst. Inf. Dyn.*, vol. 15, no. 01, pp. 7–19, 2008.

[44] P. Hayden, P. W. Shor, and A. Winter, "Random quantum codes from gaussian ensembles and an uncertainty relation," *Open Syst. Inf. Dyn.*, vol. 15, no. 01, pp. 71–89, 2008.

[45] M. Horodecki, J. Oppenheim, and A. Winter, "Partial quantum information," *Nature*, vol. 436, no. 7051, pp. 673–676, 2005.

[46] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," *Commun. Math. Phys.*, vol. 269, no. 1, pp. 107–136, 2007.

[47] I. Devetak and J. Yard, "Exact cost of redistributing multipartite quantum states," *Phys. Rev. Lett.*, vol. 100, p. 230501, 2008.

[48] J. T. Yard and I. Devetak, "Optimal quantum source coding with quantum side information at the encoder and decoder," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5339–5351, 2009.

[49] J. Oppenheim, "State redistribution as merging: introducing the coherent relay," 2008. arXiv:0805.1065.

[50] I. Devetak, A. W. Harrow, and A. Winter, "A family of quantum protocols," *Phys. Rev. Lett.*, vol. 93, p. 230504, 2004.

[51] I. Devetak, "Triangle of dualities between quantum communication protocols," *Phys. Rev. Lett.*, vol. 97, p. 140503, 2006.

[52] C. Ahn, A. Doherty, P. Hayden, and A. Winter, "On the distributed compression of quantum information," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4349–4357, 2006.

[53] I. Devetak, A. W. Harrow, and A. J. Winter, "A resource framework for quantum shannon theory," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4587–4618, 2008.

[54] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, "The mother of all protocols: restructuring quantum information’s family tree," *Proc. R. Soc. A*, vol. 465, no. 2108, pp. 2537–2563, 2009.

[55] A. Anshu, R. Jain, and N. A. Warsi, "A generalized quantum slepian–wolf," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1436–1453, 2018.

[56] M.-H. Hsieh and S. Watanabe, "Fully quantum source compression with a quantum helper," in *2015 IEEE Information Theory Workshop - Fall (ITW)*, pp. 307–311, 2015.

[57] E. Wakakuwa, Y. Nakata, and M.-H. Hsieh, "One-shot trade-off bounds for state redistribution of classical-quantum sources," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1864–1869, 2020.

[58] E. Wakakuwa and Y. Nakata, "Randomized partial decoupling unifies one-shot quantum channel capacities," 2020. arXiv:2004.12593.

[59] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[60] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, 2 ed., 2017.

[61] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[62] M. Hayashi, "Exponents of quantum fixed-length pure-state source coding," *Phys. Rev. A*, vol. 66, p. 032321, 2002.

[63] N. Datta and F. Leditzky, "Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 582–608, 2015.

[64] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theor. Comput. Sci.*, vol. 560, pp. 7–11, 2014.