Quantum information processing under local operations and classical communications

Masaki Owari
Department of Physics, Graduate School of Science
The University of Tokyo

December 20, 2006
Abstract

In this thesis, we focus on the following two quantum information processing tasks under local operations and classical communication (LOCC): LOCC discrimination (a state discrimination problem under the restriction of LOCC); and LOCC copying (a state cloning problem under the restriction of LOCC). We study the quantum non-locality concerning LOCC discrimination and LOCC copying and investigate the relationship between quantum non-locality concerning these two tasks and entanglement.

First, we concentrate on the relationship between LOCC discrimination and entanglement. I show that an upper bound of the size of a LOCC distinguishable set of states is given by entanglement measures which are defined as the distance from the set of all separable states. This result implies that the presence of entanglement guarantees a certain minimal level for non-locality concerning LOCC discrimination.

After that, we analyze the difference in LOCC discriminability for the case of permitting two-way classical communications and for the case of permitting only one-way classical communications. As a result, we derive an upper bound of the size of a one-way locally distinguishable set for bipartite states and a set of tripartite pure states. After that, in a two-qubit system, by constructing a concrete two-way local discrimination protocol, we show that two-way classical communications remarkably improves local discriminability in comparison with local discrimination by one-way classical communications.

Finally, we analyze LOCC copying of an orthogonal set of maximally entangled states and show that a necessary and sufficient condition of perfect LOCC copying is that the set is a simultaneously Schmidt diagonalizable subset of canonical Bell basis. Moreover, we show that for a set of maximally entangled states, LOCC copying is strictly more difficult than LOCC discrimination.
## Contents

1 Introduction .............................................. 4

1.1 Prologue ................................................ 4

1.1.1 Essential difference between quantum information and
       the other fields of physics .......................... 4

1.1.2 Quantum Information Processing ..................... 6

1.1.3 LOCC and quantum non-locality ...................... 8

1.1.4 Organization of the thesis ........................... 12

1.2 Historical review of Quantum Information ............... 13

1.3 Historical review of Quantum non-locality ............. 15

1.3.1 Local hidden variable model and Bell’s inequality ... 16

1.3.2 Entanglement theory ................................ 17

1.3.3 Local discrimination and local cloning .............. 22

1.4 Mathematical preparation ............................... 24

1.4.1 Space and states ...................................... 24

1.4.2 General measurement and POVM ........................ 24

1.4.3 Quantum state discrimination ....................... 26

1.4.4 Completely positive maps ........................... 27

1.4.5 Kraus representation .................................. 28

1.4.6 Steinspring representation ........................... 28

1.4.7 Definition of LOCC ................................. 29

1.4.8 Entangled states ..................................... 33
2 Bounds on multipartite local discrimination by distance-like 
entanglement monotone 40
2.1 Introduction ........................................ 41
2.2 Preliminary ........................................... 43
2.3 Proof of a new bound on local discrimination ................. 44
2.4 Application of the new bound ............................ 53
2.5 Conclusion ............................................ 55

3 Local discrimination; one-way vs two-way LOCC 60
3.1 Introduction ............................................. 60
3.2 Preliminary ............................................ 64
3.3 Local discrimination by separable POVM ...................... 66
3.4 One-way local discrimination of bipartite systems ........... 67
3.5 One-way local discrimination of Tri-partite systems .......... 71
3.6 Two-way LOCC ....................................... 73
3.7 Conclusion ............................................. 76

4 Local copying and local discrimination as a study for non-
locality of a set of states 80
4.1 Introduction ............................................. 80
4.2 The local copying problem .............................. 84
4.3 Local copying of the maximally entangled states in prime-
dimensional systems ....................................... 91
4.4 The relation between LOCC copying and LOCC discrimination 98
4.5 Application to channel copying, entanglement distillation, and 
error correction ........................................... 106
4.5.1 Channel copying .................................... 106
4.5.2 Entanglement distillation ........................... 110
4.5.3 Error correction and quantum key distribution ........... 111
4.6 Conclusion ............................................. 115
Chapter 1

Introduction

Abstract

In this chapter, after informal introduction, we give the historical reviews of quantum information and quantum non-locality. Then, we prepare mathematical tools which are used in the later chapters.

1.1 Prologue

1.1.1 Essential difference between quantum information and the other fields of physics

For people outside the field, the study of quantum information may appear strange, or even beyond the realm of physics. For example, in this work, the word “Hamiltonian” or “time” does not appear (actually, the words Hamiltonian, time, and even physical quantity never appear). Although, I myself, of course, believe that by means of the study of quantum information, we can shed light on the true nature of physics no less than by means of study of conventional fields of physics, I should make an effort to avoid some confusion on study of quantum information. Thus, at first, I would like a remark on the essential difference between quantum information and conventional fields of physics. Their most essential difference is the different notions of “nature”
which are considered in their study. In conventional fields of physics, nature is absolutely “nature in the raw”, that is, things which already exist, or happen on the universe, e.g. phase transitions of solid state, neutrino oscillations, the big bang and so on. On the other hand, in quantum information, we consider that “nature” includes our technology, especially the limit of technology which human (or other intelligent life somewhere in this universe) could “in principle” possess. In this manner, we extend the meaning of nature here. I believe this extended notion of nature is one of the most significant difference between quantum information and conventional fields of physics.

The main theme of quantum information is to know “how much” and “what kind” of information processing we can achieve under the assumption that we can do everything permitted by quantum mechanics (in other words, we can implement any unitary transformation and any measurement without refereing any Hamiltonian and any observable) and what is the gain of utilization of quantum mechanics in comparison with classical information processing. Suppose that we can do everything permitted by quantum mechanics. Then, we can completely control a given quantum system. In other words, we can freely change the Hamiltonian as we like by the control of an external field, and therefore, we can apply arbitrary unitary operations to the system. By means of quantum error correcting code, we can preserve a system in a state forever ideally. Furthermore, we can freely move a state from one system to another by the Swap operation. Finally, the only defining feature of a physical system is simply its dimension. Therefore, the Hamiltonian, time, and all (concrete) physical quantities disappear from the theory of quantum information in this most abstract setting. Since we distinguish quantum systems only by their dimension, we can treat quantum states without reference to an actual physical system. These assumptions do not only simplify our theory, but also are suitable from the viewpoint
of informatics, by which, ideally, “information” should be independent from media in the information theory. For example, the “information” of the 7 letters password to login my windows does not depend on whether it is on a piece of paper in my pocket, or it exists only in a form of electric pulse inside my brain. That is, from the viewpoint of other people, it is always a uniform classical probability distribution on $7 \times 8$ classical bits. In the quantum case, we introduce the concept of “quantum bit (qubit)” which is independent of the details of the quantum system on which it resides.

1.1.2 Quantum Information Processing

Most fundamental idea of quantum information can be written down as the words “information processing is nothing but physical process”. As we well know, the physical process on quantum systems is completely different from the physical process on classical systems. Even if we consider information processing on the most abstract setting introduced in the previous subsection, the essential difference between quantum systems and classical systems still remains. Thus, we may consider this essential difference of quantum systems and classical systems may also make a difference between information processing on quantum systems and classical systems. Therefore, we focus on the difference between information processing on quantum systems and that on classical systems by using the most abstract setting defined on the previous subsection.

When we consider “quantum information processing”, it will be better to start from classical information processing we use everyday, that is, PC and Internet. Then, our everyday life use of the internet can be summarized as follows. Spatially separated many people (call them A, B,......,Z) are sitting in front of their PCs, they calculate some problems on their PC, and communicate information through optical cables or telephone cables to each other. In other words, they send an e-mail, receive an e-mail, watch Yahoo news,
or search a recipe for cooking this evening, etc. Thus, the whole information processing systems can be written down as follows, PCs: \((A, B, \cdots, Z)\) and optical or telephone cables (called “channel”): \((C_{AB}, C_{AC}, \cdots, C_{YZ})\). Here, A computes something with his PC, and sends the results to another PC; B. Then, B computes something by using the result of A’s previous computation, and send his results to PC; C, etc. By proceeding like this, each person tries to achieve some tasks. These tasks may be numerical calculation of physics, sending of secret personal information, or searching for a recipe of today’s dinner. At least, if we omit details, information processing can be written down as the above setup. The theory of information science can be categorized into two parts, “information (Shannon) theory” and “computer science”. Computer science focuses on how fast we can solve a given problem using a PC. On the other hand, information theory focuses on how we can efficiently send information through noisy (imperfect) channels that is, how to use each channel. The extraordinary development of (classical) information science, by which we can enjoy our wonderful internet life, is based on collaboration of these two theories.

Next, we come to quantum information. In quantum information, we focus on information processing using quantum systems. Thus, each PC and each optical cable (channel) should be described by quantum mechanics. Then, we remark that, although all what we can do with classical PCs and classical channels can be done by quantum PCs and quantum channels, (this is because quantum mechanics essentially include classical mechanics), we should remain classical PCs and classical channels in our systems. This is because classical computation and classical communication are much “cheaper” than quantum computation and quantum communication. As we consider the difficulty to construct quantum computers in the real world, we do not think of using quantum PCs and channels only to derive a recipe of today’s dinner. We may use classical PCs and channels for such
easy tasks because of their cheap cost. Summarizing the above discussion, our systems of quantum information processing can be described as follows; quantum PCs \((A_q, \cdots, Z_q)\), classical PCs \((A_c, \cdots, Z_c)\), quantum channels \((Q_{AB}, Q_{AC}, \cdots, Q_{YZ})\), and classical channels \((C_{AB}, C_{AC}, \cdots, C_{YZ})\).

Here, we have types of two information processing, that is, quantum (quantum PCs + quantum channels) and classical (classical PCs + classical channels). The next question is how we can transmit information between these quantum and classical systems. Apparently, we can transmit information from classical PCs to quantum PCs by changing quantum operations depending on the computational results of classical PCs. Thus, the above question is reduced to the question: How we can transmit information from quantum PCs to classical PCs. The answer is the following: First, we apply a measurement on a quantum PC. Then, depending on a result of the measurement, we choose the classical operation applied on the classical PC in the next step. Since only the measurement can transform a quantum state to a classical probability distribution in quantum mechanics, the measurement is all what we can do in order to transmit information from quantum PCs to classical PCs. Thus, a quantum PC and a classical PC at a same place are connected by “measurement”. This is the all of the systems and operations on which we consider the theory of quantum information.

1.1.3 LOCC and quantum non-locality

In this thesis, since we are interested in what we can do by quantum PCs, we throw away all classical PCs, and compute everything on quantum PCs instead. This is because we can achieve all what we can do by classical PSs by use of quantum PCs. In this case, the results derived by the measurement on quantum PCs are directly transmitted by classical channels. Moreover, since we ignore computational cost, we can use quantum operations on quantum PCs as much as we like. Then, there remain quantum PCs
(A_q, \cdots, B_q), \text{ quantum channels } (Q_{AB}, Q_{AC}, \cdots, Q_{YZ}), \text{ and classical channels } (C_{AB}, C_{AC}, \cdots, C_{YZ}). \text{ The main theme of quantum information theory (quantum Shannon theory) is the analysis of information processing in the above systems under the assumption that the cost to transmit 1 qubit on a quantum channel is much greater than the cost to transmit 1 bit on a classical channel.}

As we have already mentioned, we generally assume that the cost of quantum communication is much greater than that of classical communication. Although we want to analyze the information processing protocol by setting of finite costs for quantum and classical communication, such an analysis is sometimes mathematically difficult. Thus, in order to make our problem simpler, we often add the following three assumptions; 1. the classical channel is “perfect” and noiseless. 2. the cost of classical communication is 0. That is, we can transmit any amount of classical information immediately. (We usually use the assumptions 1 and 2 in the discussion of quantum key distribution). 3. the cost of quantum communication is infinity, that is, we can not use a quantum channel. Under the assumptions, we do not care the individual character of “classical channels”. So, we can omit any details of classical channels. Therefore, our whole system is described by quantum PCs (A_q, B_q, \cdots, Z_q), we can apply any quantum operations on quantum PCs, and we can transmit all the results derived by measurement on a quantum PC to other quantum PCs, perfectly and immediately. These are the systems and the operations which we treat in this thesis. The whole operations which we can construct by quantum operations on quantum systems (PC) and communication of classical information as above is called “Local Operations and Classical Communications” (LOCC).

So far, we have discussed information processing. Since information processing is “states processing”, we can discuss the properties of quantum states with respect to information processing. In this thesis, since we are
physicists, we focus on the properties of quantum states related to quantum information processing. Since we have already thrown away almost all physical properties, such as the Hamiltonian, the remaining properties of our system are the following; 1. our system can be divided into several local quantum systems (PCs) \((A, B, \cdots, Z)\), and 2. we can apply LOCC to our system. Suppose each quantum systems is described by a Hilbert space \((\mathcal{H}_A, \cdots, \mathcal{H}_Z)\). Then, the whole system can be written down as \(\mathcal{H} = \mathcal{H}_A \otimes \cdots \otimes \mathcal{H}_B\). Thus, in this thesis, we consider what kind of property a state \(\rho\) or a set of states \(\{\rho_i\}\) on \(\mathcal{H}\) possesses with relation to LOCC operations. Since we have already forgotten almost all properties of a state, the property related to LOCC is a universal property which does not depend on a particular physical implementation.

Here, we add one remark. There often exist information tasks treating a state or a set of states of which we do not know the complete information; For example, tasks to discriminating, or cloning a given state, when we only know that the given state is \(\rho\) or \(\sigma\). In this case, a state which we need to treat is “an unknown state”, and these information processing treat a set of candidate states \(\{\rho, \sigma\}\).

In this thesis, we especially focus on the property of a state (or a set of states) called “non-locality”. This concept is explained as follows: For a given information processing task and a state \(\rho\) or a set of states \(\{\rho_i\}\), suppose the performance of the task is different between the case where we restrict our operations to LOCC and the case where we can apply global quantum operations for whole of our systems (this is equivalent to assuming that we can use a perfect quantum channel). Then we consider the state \(\rho\) or the set of states \(\{\rho\}\) has the non-locality related to that task.

The most famous non-locality is “(Quantum) Entanglement” which is related to a transformation of a quantum state on a system. The entanglement is defined as the non-locality which does not increase under LOCC.
Thus a state with the lower entanglement can not be convert to a state with the higher entanglement by LOCC. Since entangled states (states have the entanglement) are considered as essential resources of quantum communication, the research of entanglement is one of the central topics in quantum information.

In addition to entanglement, we consider discrimination of quantum states and cloning (copying) of quantum states as a LOCC information processing task in this thesis. These are actually tasks on a set of states. We study how well a given set of states \( \{ \rho_i \} \) can be distinguished, or cloned by only LOCC. Thus, we can consider non-locality related to LOCC discrimination and LOCC copying; note that the non-locality is actually “non-locality of a set of states”. Recently, it was shown that there exists a set of states which does not possess entanglement (that is, quantum non-locality in terms of convertibility of states), but actually possess quantum non-locality in terms of LOCC discrimination and LOCC copying [50]. That is, the quantum non-locality related to LOCC discrimination and LOCC copying are actually different from quantum non-locality in terms of entanglement. Since the study of such non-locality of a set of states are still in an early stage, it is important to study LOCC discrimination and LOCC copying from the viewpoint of quantum non-locality, and it is also important to clarify their relationship to entanglement. Therefore, the main theme of this thesis is to study of the property of these three types of non-locality (entanglement, non-locality related to LOCC discrimination, and non-locality related to LOCC copying) and the relationship among these three types of non-locality.

Although we have mentioned the motivation and importance of this thesis from the viewpoint of physics, the study of entanglement, LOCC discrimination, and LOCC copying are meaningful also from the viewpoint of informatics. First, as we will see in the next section, entanglement is considered as an essential resource of quantum communication and quantum compu-
tation. Therefore, in order to deeply understand quantum communication and quantum computation, we cannot avoid studying entanglement. LOCC discrimination is also important from the viewpoint of informatics in the following reason. In quantum communication, we encode classical information \{1, \cdots, n\} onto a set of quantum states \{\rho_i\}_{i=1}^n, and in order to decode this encoded classical information, we need to distinguish a set of states \{\rho_i\}_{i=1}^n. Therefore, information theoretically, local discriminability means how much classical information we can derive from a given set of states by LOCC. Similarly, LOCC copying also possesses the importance as a basic protocol in information processing by LOCC, and in Chapter 5, we will apply the result of LOCC copying to derive the result of many other information processing.

1.1.4 Organization of the thesis

This thesis is composed of six chapters.

In Chapter 1, we give a historical introduction of quantum information and quantum non-locality, and also of present mathematical tools which are used in the following chapters. In Chapter 2, we study the relationship between LOCC discrimination and entanglement. We present the relationship between perfect LOCC discrimination and entanglement quantities which are defined as the minimum distance from separable states (distance-like measure of entanglement). In Chapter 3, we analyze the difference in LOCC discriminability for the case of permitting two-way classical communications (two-way LOCC) and for the case of permitting only a one-way classical communication (one-way LOCC). In Chapter 4, we focus on LOCC copying of maximally entangled states, and show the relationship between LOCC copying and LOCC discrimination. In Chapter 5, we summarize the results.
1.2 Historical review of Quantum Information

Both quantum mechanics and Shannon’s information theory are considered as the key discoveries in the 20th century for the development of our present society. Quantum mechanics had been developed by many great physicists like Schrödinger, Heisenberg, Bohr, Pauli, etc in 1920’s and the discovery of quantum mechanics brought revolutionary changes not only to the field of physics, but technologies which has changed our human society. On the other hand, information theory was founded by Shannon’s revolutionary paper *A Mathematical Theory of Communication* [1] in 1948. In his paper, he treated communications of information in full mathematical ways, and proved the noiseless coding theorem and the noisy coding theorem using the entropy function of probability distributions. The noiseless coding theorem showed that the entropy can be considered as the amount of information for given probability distributions in the cases without noise. Originally, entropy had been derived in thermodynamics and also had been brought into quantum mechanics by von Neumann [2] as the von Neumann entropy. The study of quantum information was started eventually in late 1960’s by Helstrom’s quantum estimation theory [3] and by Holevo’s work for quantum channel [4]. Their works gave the foundation of quantum information theory.

The field of computer science was founded by Alan Turing in 1936 [5], and have been developed by many mathematicians including von Neumann and many engineers. Numerical calculations using computers have been an essential tool in many fields of science. Especially in physics, numerical simulations of problems which cannot be solved analytically have contributed to developments of many fields of physics, from cosmology to condensed matter physics.

The first person who considered quantum computers seems to be Richard
Feynmann. In 1982, he mentioned the possibility of revolutionary speed up of quantum simulations by using computers based on quantum mechanics [6]. A few years later, David Deutsch constructed the first quantum algorithm and showed quantum computers may have a power which strictly exceeds that of classical computers [7]. At almost the same time, it was also found that quantum mechanics may prove a significant security in the field of cryptography. Such ideas were first proposed by Wiesner [8], and then Bennet and Brassard constructed the first quantum key distribution protocol [9]. In their paper, they showed that their protocol was never eavesdropped in principle by a third party due to the no-cloning theorem, which forbids making a perfect clone of unknown quantum states [10].

In 1990’s, discovery of quantum information protocols and quantum algorithms, and their experimental realizations make the new emerging field of quantum information to be established. Bennett et al. constructed a pair of protocols which aims to send classical and quantum information utilizing EPR pair states (in a bipartite system $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$, an EPR pair states $|EPR\rangle$ is defined as $|EPR\rangle \overset{\text{def}}{=} \frac{1}{\sqrt{d}} \sum_i |ii\rangle$), respectively, namely, super dense coding [11] and quantum teleportation [12]. In their papers, they showed that if a sender and a receiver share a perfect EPR pair of $d$-dimensional systems, the sender can send a $d$-dimensional quantum state to the receiver in perfect accuracy using $\log_2 d$ bits of classical communication. Thus we can teleport an unknown quantum state from one place to another.

The most significant discovery which made the change of the position of quantum information is Shor’s factorization algorithm [14]. In 1994, Shor proposed a quantum order-finding algorithm. He also showed that the discrete logarithms and the factoring problem were able to be reduced to the order-finding problem; that is, quantum computer was able to solve the factoring problem in only a polynomial time. Because almost all public-key-type cryptographic protocols such as the RSA crypt-system [15] are based on fac-
torization, this discovery means that public-key-type secret codes, which is widely used in our internet society, are of no use in front of quantum computers. Shor’s algorithm brought on the possibility of a revolutionary change of information technology. After his discovery, the field of quantum information started more rapid growing.

We should remark that several quantum information protocols have been already realized experimentally. Quantum key distribution was first demonstrated by Bennett himself in 1992 [16]. Through many successful experiments, nowadays, systems of quantum key distribution has been already sold commercially by several companies. Quantum teleportation was also demonstrated in various systems such as the ones using the photon polarization [17], squeezed states [18] and NMR [19]. Related to quantum computer, there are also many attempts in several different systems such as liquid NMR, solid NMR, quantum dots, iron traps, cavity QED, and linear optical systems, although, at present, most of these trials are still in elementary stages.

We may need more than 30 years to make practical quantum computers and other complicated quantum information processing machines for consumers, but some quantum information protocols like the quantum key distribution protocols are already in the stage beyond just prototypes. These quantum technologies have potential to develop significantly in this century.

1.3 Historical review of Quantum non-locality

Entanglement is non-local quantum correlation in composite systems. In 1990’s, several protocols which require entangled states to perform quantum and classical communications were found: for example, dense coding, teleportation [12], quantum key distribution protocol [20] and tele-cloning [21]. Entanglement is considered as a key resource of quantum information processing to outperform classical counterparts. Entanglement theory
became one of the main areas of quantum information. We overview developments of the entanglement theory in this section.

1.3.1 Local hidden variable model and Bell’s inequality

The word of entanglement first appeared in the paper by Schödinger [13] in 1935 and became famous thanks to the memorial paper by Einstein, Podolsky and Rosen (EPR) [22]. In the EPR paper, they discussed elements of reality in theory of quantum mechanics and concluded that quantum mechanics was not complete. This paper was written by following their negative philosophy against quantum mechanics. However, to reach this conclusion, they employed certain quantum states, which became to be called EPR states later, of a composite system and showed that they did not satisfy their assumption of locality and causality. They showed that there existed states of a composite system where correlation of the subsystems were not understood as local from the classical viewpoint.

After the EPR paper, their indication of non-existence of locality and causality in quantum mechanics was developed to local hidden variable theory. Although at present, the hidden variable theory itself may be considered less significant for its original purpose, this theory is important as a tool of the study of quantum non-locality. The hidden variable theory was formulated by Bohm [23], who attempted to interpret quantum mechanics by only using classical probability theory and the assumption of the existence of “hidden variables”. In the hidden variable model, hidden variables determine measurement results completely on an individual system. However, since we do not know the value of the hidden variables of the system, the measurement results are averaged overall values of the hidden variables and we derive the each measurement result probabilistically.

The local hidden-variable model is a special case of hidden variable theory of composite systems and assumes additional locality such that the probabil-
ity distribution of the hidden variable of local measurements can be always defined as the product of two “independent” probability distributions which define each of local measurements. Then, it became apparent that there existed states which did not admit the local-hidden-variable model. In other words, quantum states admitting the local-hidden-variable model are considered having locality in a sense. In 1964, J.S. Bell presented an inequality which must be satisfied by states admitting local-hidden-variable model in a bipartite spin system [24], namely, the original Bell inequality. A few years later, Clauser, Horne, Shimony, and Holt (CHSH) generalized Bell’s result and obtained a class of inequalities (CHSH inequalities) [25], which must be satisfied by bipartite two-dimensional states admitting the local-hidden-variable model. By means of Bell’s inequality, we can consider that a state which violates the inequality more has more non-locality. In spite of its long history of study of Bell’s inequality, the problem to find all Bell inequalities for general cases is still open problem [26].

1.3.2 Entanglement theory

The clear definition of entangled states first appeared in the paper of R. Werner in 1989 [27] in the context of the local-hidden-variable model. In this paper, Werner presented separable states (he used “classically correlated states” in his paper) which can be prepared only by local measurements and classical communications, and showed that separable states always admitted a local hidden variable model and satisfied all Bell inequalities. (The Bell inequalities are defined as inequalities which states admitting a local-hidden-variable model must satisfy.)

In a finite dimensional system, separable states are defined as states which can be written in the form

$$\rho = \sum_{i=1}^{n} p_i \rho_A^i \otimes \rho_B^i \quad (1.1)$$

17
and entangled states are defined as non-separable states. On the other hands, in an infinite dimensional system, separable states are defined as the states which can be approximated in trace norm by density matrices written in the form Eq.(1.1) (related to this difference of the definitions, see [32].)

As already mentioned in the previous section, in early 1990’s, many quantum information protocols which use entanglement were discovered, i.e. dense-coding [11], teleportation [12], and quantum key distribution (E91) protocol [20]. After that, quantum non-locality has been mainly studied in relation to quantum information processing, and formulated as “entanglement theory”. In the entanglement theory, we consider that a state $\rho$ is entangled stronger than another state $\sigma$, if $\rho$ can be converted to $\sigma$ by LOCC.

Thus, in a narrow sense, the study of entanglement can be consider as a study of convertibility of quantum states under LOCC (LOCC convertibility). For bipartite pure states, the necessary and sufficient condition of LOCC convertibility is derived by Nielsen in the deterministic case [28] and Vidal in the probabilistic case [29]. However, because of its mathematical difficulty, it seems to be impossible to derive the necessary and sufficient condition for general mixed states.

If we knew the complete description of LOCC convertibility, we could derive all knowledge about entanglement. However, we cannot derive such a condition. We should study entanglement in several different ways depending on the purposes. From the viewpoint of quantum information processing, the most elementally questions about entanglement may be the following three: 1. How we can distinguish entangled states from separable states; 2. How we can generate maximally entangled states (singlets) shared between two spatially separated parties from non-maximally entangled pairs by local operations; 3. how we can quantify the entanglement.

The first question is equivalent to find the necessary and sufficient conditions that a given state is entangled (or separable). This question is imme-
diately solved for bipartite pure states by the Schmidt decomposition (i.e. polar decomposition of reduced density operator) of pure states, but is non-trivial for mixed states. The necessary conditions of separable states are called separability criteria.

Research of the separability criteria has been remarkably progressed by Peres and the Horodecki family (the father and two sons) in 1996. Peres showed that “If $\rho$ is separable, then $T \otimes I \rho$ is positive,” where $T$ is a transpose operator of a subsystem and $T \otimes I \rho$ is called the partial transpose of $\rho$ [30]. Soon after the discovery of Peres, the Horodecki family presented a necessary and sufficient condition of separability. The theorem is the following [31]: “Let $\rho$ act on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $\rho$ is separable if and only if for any positive map $\Lambda : T(\mathcal{H}_A) \rightarrow T(\mathcal{H}_B)$, the operator $(I \otimes \Lambda) \rho$ is positive” where $T(\mathcal{H})$ is a space of trace class operators (and therefore coincides with the space of linear operators on $\mathcal{H}$ in this case). Usually we cannot use this theorem to determine the separability of specific states concretely, but they proved that Peres’s criterion is necessary and sufficient for 2 × 2 or 2 × 3 dimensional systems. Therefore the separability problem was completely solved in these low dimensional systems even for mixed states. On the other hand, P. Horodecki found a counterexample in a higher dimensional system, that is, states whose partial transpose is not positive (in the following discussion, we call the states giving positive partial transpose as PPT states), but they are still entangled [32].

The answer of the second question (that is, how we can generate maximally entangled states (singlets) shared between two spatially separated parties from non-maximally entangled states by local operations) is obtained by Bennett et. al. in 1996. In their paper [33], they showed that if two parties share infinitely many two-dimensional partially entangled states, then we are able to transform them into the maximally entangled states by using only local operations (generalized measurements and interactions to ancilla systems
to local systems) and classical communications of the measurement results. This protocol is called \textit{entanglement distillation}.

Bennett et al. also first tried to answer the third question (that is, the quantification of entanglement), in the paper [34] also in 1996. They showed that for a pure bipartite state $|\Psi\rangle$, both of the number of singlets (two-dimensional maximally entangled states) which can be transformed from $|\Psi\rangle$ by local operations and classical communications (LOCC), called distillable entanglement ($E_d(|\Psi\rangle)$), and the number which requires to create $|\Psi\rangle$ by only LOCC, called entanglement cost ($E_c(|\Psi\rangle)$), tend to

$$E_d(|\Psi\rangle) = E_c(|\Psi\rangle) = S(\text{Tr}_B(|\Psi\rangle \langle \Psi|))$$ (1.2)

where $S(\rho)$ is the von Neumann entropy of $\rho$, in the limit of infinitely many copies of $|\Psi\rangle$ [34]. Thus, $S(\text{Tr}_B(|\Psi\rangle \langle \Psi|))$ is called the amount of entanglement. In [35], they also estimated the distillable entanglement and the entanglement cost of special classes of mixed states. However, the general procedure of calculating such quantities for mixed states have not been obtained so far.

In 1997, Vedral and Plenio presented an axiomatic approach to quantization of entanglement [36]. In the paper, they defined an entanglement measure as a function which satisfies the following three conditions: 1. monotonically decreasing by LOCC transformations; 2. continuity for the trace norm; 3. additivity for the tensor product of states. They also presented, as an example of the entanglement measure, a function called the relative entropy of entanglement:

$$E_r(\rho) = \inf_{\sigma} S_{rel}(\rho|\sigma),$$ (1.3)

where

$$S_{rel}(\rho|\sigma) \equiv \text{Tr}\rho \log_2 \rho - \text{Tr}\rho \log_2 \sigma$$ (1.4)

is the quantum relative entropy, and the infimum is taken over all separable states $\sigma$. The additivity and continuity in their definition are for considering
their entanglement measure as a measure of entanglement, which is a resource of quantum information processing. Of course, the distillable entanglement and the entanglement cost are also entanglement measures.

In order to explain the importance of this axiomatic approach to quantification entanglement, we should note that, in a bipartite system, there exists the uniqueness theorem of entanglement measures for pure bipartite states [37]. The theorem states that any entanglement measure must be upper-bounded by the entanglement cost, and lower-bounded by the distillable entanglement [38, 39]. Hence, any entanglement measures $E(\rho)$, that is a function of states satisfying the above three conditions, satisfy $E_d(\rho) \leq E(\rho) \leq E_c(\rho)$ for all $\rho$. Moreover, since $E_d(|\Psi\rangle)$ and $E_c(|\Psi\rangle)$ coincide with $S(\text{Tr} |\Psi\rangle \langle \Psi|)$ for all bipartite pure states $|\Psi\rangle$, all entanglement measures coincide with $S(\text{Tr} |\Psi\rangle \langle \Psi|)$. Thus, the axiom of entanglement measure is not only natural from the viewpoint of physics, but it is also meaningful from an information theoretic viewpoint as a lower bound of the entanglement cost and an upper bound of the distillable entanglement. In this sense, we can say, at least for a bipartite system, that the theory of entanglement measure is the most successful way of quantifying entanglement so far.

Although the axioms of entanglement measure succeeded in a bipartite system, it does not work well in a multi-partite system. The most important property of entanglement in a multi-partite system is that there exist many different types of entangled states and generally, there is no maximally entangled states in multi-partite systems [41, 42]. Thus, there is no unique way of extending the definition of the entanglement cost and the distillable entanglement. That is, the theorems related to entanglement measure (such as the uniqueness theorem of entanglement measures) fail and people have not succeed to efficiently extend the theory of entanglement measure to multi-partite systems [43].

There also exists another way of quantifying entanglement proposed by
G. Vidal [40]. Vidal gave the axiom of entanglement monotone without the additivity condition. Thus, the axiom of entanglement monotone is weaker than the axiom of entanglement measure and there are many functions which satisfy the axiom of entanglement monotone, but do not satisfy the axiom of entanglement measure; e.g. robustness of entanglement[44], and geometric measure of entanglement[45]. In multi-partite systems, since the theory of entanglement measure does not work well, we can say that this axiom of entanglement monotone is more useful in a multi-partite system.

1.3.3 Local discrimination and local cloning

Recently, several new directions of research have been done as studies of quantum non-locality. Research of local discrimination (a state-discrimination problem under the restriction of LOCC) and local cloning (a state-cloning problem under the restriction of LOCC) are two of such new direction of study of quantum non-locality. We mainly focus on these topics in this thesis.

From a very early stage of quantum information, since the necessity of including the effect of measurement is one of the most significant differences between quantum information theory and classical information theory, state discrimination, a study of “how well we can determine a given unknown state from a set of candidate states”, was one of the central topics of fields of quantum information. Depending on its purpose, state discrimination has been studied in several different forms as quantum hypothesis testing, quantum statistical estimation [3, 46], and quantum channel coding[4, 47]; for more detail of the state discrimination, see [49, 48].

Since the situation in which we can only use LOCC is natural and common in quantum information processing, discussions about the restriction of LOCC have been sometimes included in several studies of quantum estimation and quantum channel coding. Bennett et al. first discussed local
discrimination in the context of quantum non-locality in 1999 [50]. They showed that the following bases \( \{ |\Psi_i\rangle \}_{i=1}^{9} \) in \( \mathbb{C}^3 \otimes \mathbb{C}^3 \),

\[
\begin{align*}
|\Psi_1\rangle &= |1\rangle \otimes |1\rangle \\
|\Psi_{2(3)}\rangle &= \frac{1}{\sqrt{2}} |0\rangle \otimes (|0\rangle \pm |1\rangle) \\
|\Psi_{4(5)}\rangle &= \frac{1}{\sqrt{2}} |2\rangle \otimes (|1\rangle \pm |2\rangle) \\
|\Psi_{6(7)}\rangle &= \frac{1}{\sqrt{2}} (|1\rangle \pm |2\rangle) \otimes |0\rangle \\
|\Psi_{8(9)}\rangle &= \frac{1}{\sqrt{2}} (|0\rangle \pm |1\rangle) \otimes |2\rangle,
\end{align*}
\]

where \( \{ |i\rangle \}_{i=0}^{2} \) is an orthonormal basis set of each local space, cannot be distinguished by only LOCC. Apparently, all \( \{ |\Psi_i\rangle \} \) are product states, and do not have any entanglement. However, from the fact of local indistinguishability of this basis set, this basis set (not each individual state) seems to possess a kind of non-locality, and Bennett et al. called this non-locality “quantum non-locality without entanglement”. In this sense, quantum non-locality related to local discrimination cannot be explained only by entanglement theory, and the local-discrimination problem is worth studying to derive deep understanding of quantum non-locality. After the work of Bennett et al., many researches of local discrimination have been done; however, the theory of local discrimination is still on the way of developing [51].

Similarly to local discrimination, generally speaking, if a quantum information processing task for bi(multi)-parties becomes difficult to be performed by adding the restriction of LOCC, we can consider this difficulty arises from quantum non-locality concerning this quantum information processing task. Cloning is also one of the most elemental quantum information processing. After the discovery of the non-cloning theorem [10] which prohibits to make a perfect copy of a completely unknown quantum state, many researches of cloning have been done [52]. A few years ago, cloning problem under the
restriction of LOCC was also proposed [53]. However, the study of local cloning is just starting now, and there are only few papers.

In this thesis, we investigate the local discrimination in Chapter 2 and 3, and local cloning in Chapter 4.

1.4 Mathematical preparation

In this section, we present several definitions and mathematical tools which are foundations for the following chapters of this thesis. We refer to the articles [39, 48, 54, 55, 56] and do not give the proof of each theorem in this section. Elementary knowledge of functional analysis is assumed.

1.4.1 Space and states

In this thesis, we always construct the theory on the separable Hilbert space composed of more than one Hilbert spaces $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots$. The Banach space of bounded operators on $\mathcal{H}$ is denoted by $\mathfrak{B}(\mathcal{H})$ and the Banach space of trace class operators is denote by $\mathcal{T}(\mathcal{H})$. The states are positive trace class operators with a unit trace and a set of states is denoted by $S(\mathcal{H})$. By definition, $S(\mathcal{H})$ is a closed convex set and the set of extremal points of $S(\mathcal{H})$ coincides with the set of one-dimensional projectors on $\mathcal{H}$ called pure states. We always identify $|\Psi\rangle\langle\Psi| \in S(\mathcal{H})$ with $|\Psi\rangle \in \mathcal{H}$ without any remark.

1.4.2 General measurement and POVM

In a Hilbert space $\mathcal{H}$, a measurement is described by a set of (bounded) linear operators $\{A_i\}_{i=1}^N$ which satisfies

$$\sum_i A_i^\dagger A_i = 1. \quad (1.6)$$

For a state $\rho \in S(\mathcal{H})$, by applying a measurement described by $\{A_i\}_{i=1}^N$, we obtain one of the possible results indexed by classical information $i \in \cdots$
\{1 \cdots N\} \text{ with the probability given by } \text{Tr} \rho A_i^\dagger A_i. \text{ After we obtain a result } i, \text{ the state } \rho \text{ changes to another state given by } \rho' = \frac{A_i \rho A_i^\dagger}{\sqrt{\text{Tr} \rho A_i^\dagger A_i}}. \text{ On the other hand, if we do a measurement } \{A_i\}_{i=1}^N \text{ on a state } \rho, \text{ and if we do not see the result of the measurement, the state after the measurement is given as }

\[
\sum_{i=1}^N A_i \rho A_i^\dagger.
\]

(1.7)

Note that since, by choosing \(N = 1\), a set that consists of only one unitary operator \(\{U\}\) also satisfies Eq.\((1.6)\), we can consider a unitary operator as a measurement.

In quantum information, we often have a situation where we do not have interests in the states after measurements. In such a situation, since the output probability only depends on a set of positive operators \(\{A_i^\dagger A_i\}_{i=1}^N\) instead of \(\{A_i\}_{i=1}^N\), the following definition of positive-operator-valued measure (POVM) is useful.

**Definition 1.1 (POVM)** A set of (bounded) linear operator \(\{M_i\}_{i=1}^N \in \mathfrak{B(H)}\) is called POVM, if it satisfies

\[
M_i \geq 0 \ (\forall i)
\]

\[
\sum_{i=1}^N M_i = I,
\]

where \(\geq\) is defined as a matrix inequality.

In this case, since we discard information of the states after the measurement, we can consider POVM \(\{M_i\}_{i=1}^N\) as a tool to derive a classical probability distribution \(\{p_i\}_{i=1}^N = \{\text{Tr} \rho M_i\}_{i=1}^N\) from a quantum state \(\rho\), or a tool to change quantum information of \(\rho\) to classical information \(\{p_i\}_{i=1}^N = \{\text{Tr} \rho M_i\}_{i=1}^N\).

We give a simple example of POVM. For a spin \(\frac{1}{2}\) system \(\mathbb{C}^2 = \text{span}\{\left|\uparrow\right\rangle, \left|\downarrow\right\rangle\}\), the measurement of the spin in the Z-direction is described by the POVM \(\{\left|\uparrow\right\rangle \left\langle \uparrow\right|, \left|\downarrow\right\rangle \left\langle \downarrow\right|\}\) and the measurement of the spin in the X-direction is described by the POVM \(\left\{\frac{1}{2} \left(\left|\uparrow\right\rangle + \left|\downarrow\right\rangle\right) \left(\left\langle \uparrow\right| + \left\langle \downarrow\right|\right), \frac{1}{2} \left(\left|\uparrow\right\rangle - \left|\downarrow\right\rangle\right) \left(\left\langle \uparrow\right| - \left\langle \downarrow\right|\right)\right\}.\)
1.4.3 Quantum state discrimination

By using the definition of POVM, the discrimination problem of quantum states (the quantum state discrimination) is given by the following: Now, we have an unknown state $\rho$ and know that $\rho$ is one of $\{\rho_i\}_{i=1}^N$ with a priori probabilities $\{p_i\}_{i=1}^N$. The question is how well we can determine $\rho$ by using the information above; we call this problem “discrimination of $\rho$ from a set of candidates $\{\rho\}_{i=1}^N$”. Suppose that we use a POVM $\{M_i\}_{i=1}^N$ for this problem and measure the outcome “$i$”. We then decide that “the given state is $\rho_i$”. From the discussion in the previous subsection, if the given state is $\rho_i$, we measure the outcome “$j$” with the probability $\text{Tr}\rho_i M_j$. Thus, the success probability (the probability that our decision is right) is $\text{Tr}\rho_i M_i$, and the error probability (the probability that our decision is wrong) is $\sum_{j\neq i} \text{Tr}\rho_i M_j$. In this case, we say that “we detect $\rho_i$ by the POVM $M_i$ with probability $\text{Tr}\rho_i M_i$”. By taking the average for all $\rho_i$ of the a priori probability $p_i$, we can write down the mean success probability $p_{\text{ave}}$ as follow,

$$p_{\text{ave}} \triangleq \sum_{i=1}^N p_i \text{Tr}\rho_i M_i.$$  

Similarly, the mean error probability $e_{\text{ave}}$ is given as follows,

$$e_{\text{ave}} \triangleq \sum_{i=1}^N p_i (\sum_{j\neq i} \text{Tr}\rho_i M_j)$$

$$= 1 - \sum_{i=1}^N \rho_i Tr \rho_i M_i$$

$$= 1 - p_{\text{ave}}.$$

For the state-discrimination problem, our main aim is to derive the maximum success probability $p_{\text{ave}}$ (the minimum error probability $e_{\text{ave}}$) and the POVM which gives the maximum success probability for a given a set of states $\{\rho_i\}_{i=1}^N$ and a priori probability $\{p_i\}_{i=1}^N$. 

26
1.4.4 Completely positive maps

In quantum information, we often treat a (mixed) state $\rho \in S(\mathcal{H}) \subset T(\mathcal{H})$, not only a pure state $|\Psi\rangle \in \mathcal{H}$, since we may not have full information of our state. Hence, it is important to consider what kind of map $\Lambda$ on $T(\mathcal{H})$ is “physical”, and what is not. Apparently, a physical map should map the set of all state $S(\mathcal{H})$ onto $S(\mathcal{H})$ itself; thus, it should be a positive map in the following meaning:

**Definition 1.2 (Positive map)** If a bounded linear operator $\Lambda(\rho)$ on $T(\mathcal{H})$ satisfies

$$\rho \geq 0 \implies \Lambda(\rho) \geq 0,$$

(1.8)

where $\rho \geq 0$ means $\rho$ is a positive operator, then $\Lambda$ is called a positive map.

If $\Lambda$ is a positive map, then, for a state $\rho$, $\Lambda(\rho)/\|\Lambda(\rho)\|_{\text{Tr}}$ is also a state, where $\|\|_{\text{Tr}}$ denotes the trace norm; in the following, we usually omit Tr in this section.

Adding to the positivity, we sometimes need to consider an open system, that is, a Hilbert space $\mathcal{H}$ may be a part of larger space $\mathcal{H} \otimes \mathcal{H}'$; for example, $\mathcal{H}'$ denotes the space of an environment. Thus, a physical map $\Lambda$ should also satisfy the condition of the positivity for all such extensions of the Hilbert space; that is, it should satisfy the following condition of the complete positivity:

**Definition 1.3 (Completely positive map)** If a bounded operator $\Lambda(\rho)$ on $T(\mathcal{H})$ satisfies for all $d \in \mathbb{N}$,

$$\rho \geq 0 \implies \Lambda \otimes I(\rho) \geq 0$$

(1.9)

where $\rho \in S(\mathcal{H} \otimes \mathbb{C}^d)$, then $\Lambda(\rho)$ is called complete positive.
1.4.5 Kraus representation

We call a complete positive map on $T(\mathcal{H})$ that does not increase the trace norm, that is, $\|\Lambda(\rho)\| \leq \|\rho\|$ for all $\rho \in T(\mathcal{H})$, “a quantum operation” or “a quantum channel”. If a quantum operation $\Lambda$ transforms $\rho \in S(\mathcal{H})$ to $\Lambda(\rho)$, then we interpret that $\Lambda$ transforms $\rho$ to $\Lambda(\rho)/\|\Lambda(\rho)\| \in S(\mathcal{H})$ with the probability $\|\Lambda(\rho)\|$. This interpretation is justified by the following theorem.

**Theorem 1.1 (Kraus Representation)** For an arbitrary quantum operation $\Lambda$, there exists operators $\{A_k\}, k \in K$ (a finite or countably infinite index set) on $\mathcal{H}$, satisfying $\sum_{k \in K_0} A_k^\dagger A_k \leq 1$ for all finite subsets $K_0 \subset K$, such that, the mapping $\Lambda$ is given for an arbitrary state $\rho \in T(\mathcal{H})$ by the following:

$$\Lambda(\rho) = \sum_{k \in K} A_k \rho A_k^\dagger. \quad (1.10)$$

The operators $\{A_k\}_{k \in K}$ in this theorem are called Kraus operators. We note that they are not unique. The decomposition of operations in this theorem is called the Kraus representation. By comparing Eq.(1.7) and Eq.(1.10), any quantum operation always can be considered as a generalized measurement. Thus, the generalized measurements describe all what we can perform to a quantum state.

1.4.6 Steinspring representation

Deterministic (trace preserving) quantum operations, or complete positive trace preserving maps (CPTP maps) are characterized as linear maps that can be decomposed into the following elementary operations:

1. Adding an uncorrelated ancilla:
   $$\Lambda_1 : T(\mathcal{H}_1) \rightarrow T(\mathcal{H}_1 \otimes \mathcal{K}_1), \quad \Lambda_1(\rho) \equiv \rho \otimes \sigma,$$
   where $\mathcal{H}_1$ and $\mathcal{K}_1$ denote the Hilbert spaces of the original quantum system and of the ancilla, respectively, where $\sigma \in S(\mathcal{K}_1)$.
2. Tracing out a part of the system:
\[ \Lambda_2 : \mathcal{T}(\mathcal{H}_2 \otimes \mathcal{K}_2) \rightarrow \mathcal{T}(\mathcal{H}_2), \Lambda_2(\rho) \equiv \text{Tr}_{\mathcal{K}_2}(\rho) \] where \( \mathcal{H}_2 \) and \( \mathcal{K}_2 \) denote the Hilbert spaces of the original quantum system and of the discarded part respectively, where \( \text{Tr}_{\mathcal{K}_2} \) denotes the partial trace over \( \mathcal{K}_2 \).

3. Unitary transformations:
\[ \Lambda_3 : \mathcal{T}(\mathcal{H}_3) \rightarrow \mathcal{T}(\mathcal{H}_3), \Lambda_3(\rho) = U\rho U^\dagger \] where \( U \) is a unitary operation on \( \mathcal{H}_3 \).

That is, if we consider the set of ancilla spaces \( \mathcal{K}_1 \) and \( \mathcal{K}_2 \), an ancilla state \( \sigma \) and a unitary operator \( U \) on the composite system are equivalent to a CPTP map \( \Lambda \) on \( \rho \). The above representation of a CPTP map is called the Steinspring representation.

### 1.4.7 Definition of LOCC

In this thesis, we consider the situation where there are several parties in spatially separated laboratories, and they are only allowed to perform quantum operations in their own laboratories and to communicate classical information of measurement results among different laboratories. Therefore, the following definition of LOCC plays an essential role for all discussions in this thesis. Here, we only give a definition of LOCC for bi-partite (two-party) systems; however, the extension for multi-party systems is trivial.

Consider the case \( \mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B \), that is, the case where our system consists of two spatially separated subsystems \( \mathcal{H}^A \) and \( \mathcal{H}^B \). \( \mathcal{H}^A \) and \( \mathcal{H}^B \) are called "local systems". A local operation is a quantum operation (generalized measurements including unitary operations) on individual subsystems; that is, \( \Lambda \) can be written in the Kraus representation by
\[ \Lambda(\rho) = \sum_{k \in \mathcal{K}} (A_k^\dagger \otimes B_k^\dagger)\rho(A_k \otimes B_k). \quad (1.11) \]
We consider a class of operations that can be performed by local operations and classical communications of measurement results. At first, we consider a situation allowing classical communications in only one direction and define the one-way LOCC operation.

**Definition 1.4 (One-way LOCC)** A completely positive map $\Lambda : T(\mathcal{H}_A^1 \otimes \mathcal{H}_B^1) \rightarrow T(\mathcal{H}_A^2 \otimes \mathcal{H}_B^2)$ is called a one-way LOCC operation from $A$ to $B$ if it can be written in the Kraus representation as

$$\Lambda(\rho) = \sum_{i,j \in K,L} (1_A^1 \otimes W_{ji}^B)(V_i^A \otimes 1_B^1)\rho(V_i^{A\dagger} \otimes 1_B^1)(1_A^2 \otimes W_{ji}^{B\dagger})$$

for all $\rho \in T(\mathcal{H}_A^1 \otimes \mathcal{H}_B^1)$ and some sequences of operators $(V_i^A : \mathcal{H}_1^A \rightarrow \mathcal{H}_2^A)_i$ and $(W_{ji}^A : \mathcal{H}_1^B \rightarrow \mathcal{H}_2^B)_{ji}$ with $\sum_{i \in K} V_i^{A\dagger}V_i^A = 1_A^1$ and $\sum_{j \in L} W_{ji}^{B\dagger}W_{ji}^B = 1_B^1$ for each $i$, where $1_A^1, 1_B^1$ and $1_A^2$ are the unit operations acting on the Hilbert space $\mathcal{H}_1^A$, $\mathcal{H}_1^B$, and $\mathcal{H}_2^A$, respectively.

If $V_i^A$ and $W_{ji}^B$ only satisfy the two conditions

$$\sum_{i \in K} V_i^{A\dagger}V_i^A \leq 1_A^1,$$  

$$\sum_{j \in L} W_{ji}^{B\dagger}W_{ji}^B \leq 1_B^1,$$

then, we call $\Lambda$ a one-way Stochastic LOCC (SLOCC) operation, or a non-deterministic one-way LOCC operation.

More generally, we consider a situation where classical communications are allowed in both directions, and use the following definition of the two-way LOCC.

**Definition 1.5 (Two-way LOCC)** A completely positive map $\Lambda : T(\mathcal{H}^A \otimes \mathcal{H}^B) \rightarrow T(\mathcal{K}^A \otimes \mathcal{K}^B)$ is called an LOCC operation if there exist $n > 0$ and sequences of Hilbert spaces $(\mathcal{H}_k^A)_{k=1}^{n+1}$ and $(\mathcal{H}_k^B)_{k=1}^{n+1}$ with $\mathcal{H}_1^{A(B)} = \mathcal{H}^{A(B)}$ and...
$H_{n+1}^{A(B)} = K^{A(B)}$, such that $\Lambda$ can be written in the Kraus representation as

$$\Lambda (\rho) = \sum_{i_1, \ldots, i_{2n} \in K_{1}, \ldots, K_{2n}} V_{i_1, \ldots, i_{2n}} \rho V_{i_1, \ldots, i_{2n}}^{\dagger}$$

(1.15)

for all $\rho \in T(\mathcal{H}^{A} \otimes \mathcal{H}^{B})$, where

$$V_{i_1, \ldots, i_{2n}} : \mathcal{H}^{A} \otimes \mathcal{H}^{B} \rightarrow K^{A} \otimes K^{B}$$

(1.16)

is given by

$$V_{i_1, \ldots, i_{2n}} = (1^{A}_{n+1} \otimes W_{2n}^{i_2, \ldots, i_1})(V_{2n-1}^{i_2, \ldots, i_1} \otimes 1^{B}_{n})(1^{A}_{n} \otimes W_{2n-2}^{i_2, \ldots, i_1}) \cdots (1^{A}_{2} \otimes W_{2}^{i_2, i_1})(1^{B}_{1})$$

(1.17)

using families of operators

$$(V_{2k-1}^{i_2, \ldots, i_1} : \mathcal{H}^{A}_{k} \rightarrow \mathcal{H}^{A}_{k+1})_{k=1}^{n}$$

(1.18)

$$(W_{2k}^{i_2, \ldots, i_1} : \mathcal{H}^{B}_{k} \rightarrow \mathcal{H}^{B}_{k+1})_{k=1}^{n}$$

(1.19)

such that for $k = 0, \ldots, n - 1$ and each sequence of indices $(i_{2k}, \ldots, i_1)$, we have

$$\sum_{i_{2k+1} \in K_{2k+1}} (V_{2k+1}^{i_2, \ldots, i_1})^{\dagger} V_{2k+1}^{i_2, \ldots, i_1} = 1^{A}_{k+1}$$

(1.20)

and for $k = 1, \ldots, n$ and each sequence of indices $(i_{2k-1}, \ldots, i_1)$,

$$\sum_{i_{2k} \in K_{2k}} (W_{2k}^{i_2, \ldots, i_1})^{\dagger} W_{2k}^{i_2, \ldots, i_1} = 1^{B}_{k}$$

(1.21)

where for all $k > 0$, $1^{A}_{k}$ and $1^{A}_{k}$ denote the unit operator on $\mathcal{H}^{A}_{k}$ and $\mathcal{H}^{B}_{k}$, respectively.

Obviously, the two-way LOCC includes the one-way LOCC. We can also define SLOCC, which maps a state to another state with non-zero probability under LOCC (but not necessary the unit probability as in the deterministic LOCC case), similarly to the case of one-way LOCC.

Since the definition of LOCC is complicated and mathematically intractable, we define a simpler local operations, a separable operation.
**Definition 1.6 (Separable operations)** A CPTP map $\Lambda : \mathcal{T}(\mathcal{H}^A \otimes \mathcal{H}^B) \to \mathcal{T}(\mathcal{K}^A \otimes \mathcal{K}^B)$ is called a separable operation, if $\Lambda$ can be written in the Kraus representation

$$\Lambda(\rho) = \sum_{i=1}^N (M_i \otimes N_i)\rho(M_i^\dagger \otimes N_i^\dagger), \quad (1.22)$$

where $\{M_i\}_{i=1}^N \subset \mathcal{B}(\mathcal{H}^A, \mathcal{K}^A)$ and $\{N_i\}_{i=1}^N \subset \mathcal{B}(\mathcal{H}^B, \mathcal{K}^B)$ satisfy $\sum_{i=1}^N M_i^\dagger M_i \otimes N_i^\dagger N_i = I_{AB}$.

Separable operations belong to a slightly larger set than the two-way LOCC. For example, the POVM which distinguishes the special product bases given by (1.5) is a separable operation, but is not a two-way LOCC operation [50].

Since we consider a state-discrimination problem in the one-way LOCC, the two-way LOCC and the separable operation in the later chapter, we give a definition of a POVM which can be implemented by these operations.

**Definition 1.7** We say that a POVM $\{M_i\}_{i=1}^N$ can be implemented by a one-way LOCC, if there exists a one-way LOCC operation $\Lambda : \mathcal{T}(\mathcal{H}^A_1 \otimes \mathcal{H}^B_1) \to \mathcal{T}(\mathcal{H}^A_2 \otimes \mathcal{H}^B_2)$ such that $\Lambda$ satisfies the following three conditions:

1. $\Lambda$ can be written in the Kraus representation

$$\Lambda(\rho) = \sum_{i,j \in K,L} (1^A \otimes W^B_{ji})(V^A_i \otimes 1^B_1)\rho(V^A_i \otimes 1^B_1)(1^A \otimes W^B_{ji}^\dagger) \quad (1.23)$$

for all $\rho \in \mathcal{T}(\mathcal{H}^A_1 \otimes \mathcal{H}^B_1)$ and some sequences of operators $(V^A_i : \mathcal{H}^A_1 \to \mathcal{H}^A_2)_i$ and $(W^B_{ji} : \mathcal{H}^B_1 \to \mathcal{H}^B_2)_j$ with $\sum_{i \in K} V^A_i V^A_i = 1^A_1$ and $\sum_{j \in L} W^B_{ji} W^B_{ji} = 1^B_1$ for each $i$.

2. $K \times L$ can be decomposed as $K \times L = \bigcup_{l=1}^N X_l$ with $X_l \cap X_{l'}$ for all $l \neq l'$.

3. $\sum_{(i,j) \in X_l} V^A_i V^A_i \otimes W^B_{ji} W^B_{ji} = M_l$.

We define a POVM which can be implemented by a two-way LOCC and a separable operation in a similar fashion.
1.4.8 Entangled states

Finally, we give the definition of entangled states and several theorems used later. Now, we consider an \( n \)-partite system \( \mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i \), and assume \( \dim \mathcal{H}_i < \infty \) for all \( i \). As we have already mentioned in the previous section, the definitions of separable states and entangled states are given as follows.

**Definition 1.8 (Separable state)** A state \( \rho \) on \( \bigotimes_{i=1}^{n} \mathcal{H}_i \) is called a separable state, if \( \rho \) can be written as follows:

\[
\rho = \sum_{k=1}^{K} p_k \rho_1^k \otimes \cdots \otimes \rho_n^k,
\]

(1.24)

where \( \{p_k\}_k \) is a probability distribution \( \{p_k \geq 0 \text{ and } \sum_k p_k = 1\} \) and \( \rho_i^k \) is a state on \( \mathcal{H}_i \) for all \( i \) and \( k \).

**Definition 1.9 (Entangled state)** A state \( \rho \) on \( \bigotimes_{i=1}^{n} \mathcal{H}_i \) is called an entangled state, if \( \rho \) is not a separable state.

In a bipartite system \( \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \), all pure states have the following standard form.

**Theorem 1.2 (Schmidt decomposition)** For a bipartite pure state \( |\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \), there exists a basis of \( \mathcal{H}_A: \{|e_i\rangle\}_{i=1}^{\dim \mathcal{H}_A} \), a basis of \( \mathcal{H}_B: \{|f_i\rangle\}_{i=1}^{\dim \mathcal{H}_B} \), and a probability distribution \( \{\lambda_i\}_{l=1}^{\min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}} \) such that \( |\Psi\rangle \) can be written down as follows:

\[
|\Psi\rangle = \sum_{l=1}^{\min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}} \sqrt{\lambda_l} |e_i\rangle \otimes |f_i\rangle,
\]

(1.25)

where \( \lambda_i \leq \lambda_{i+1} \) for all \( i \).

The above decomposition is called the Schmidt decomposition of a composite system of the systems \( A \) and \( B \), and \( \{\lambda_i\} \) and \( \{|e_i\rangle \otimes |f_i\rangle\}_{ll} \) are called the Schmidt coefficient and the Schmidt basis, respectively. The number of
the non-zero terms is called the “Schmidt rank”. We can easily see that the Schmidt rank of a bipartite pure state $|\Psi\rangle$ is equal to the rank of its reduced density operator $\text{rank} \text{Tr}_B |\Psi\rangle \langle \Psi|$, where $\text{Tr}_B \rho$ is called the reduced density operator of the system $A$, and $\text{Tr}_B$ is a partial trace with respect to the system $B$ defined by the equation $\text{Tr}_B \rho \overset{\text{def}}{=} \sum_{i=1}^{\dim \mathcal{H}_B} \langle i | \rho | i \rangle \{ | i \rangle \}_{i=1}^{\dim \mathcal{H}_B}$ is an arbitrary orthonormal basis set of $\mathcal{H}_B$. It is known that the Schmidt rank is a SLOCC monotone, that is, a monotonic function of SLOCC convertibility.

As we can easily see from Eq.(1.25), a bipartite pure state $|\Psi\rangle$ can be transformed to another bipartite pure state $|\Phi\rangle$ by a local unitary transformation, if and only if their Schmidt coefficients are equal. Since the entanglement is the quantum non-locality in terms of the LOCC convertibility of states, if $|\Psi\rangle$ can be converted to $|\Phi\rangle$ by LOCC, and if $|\Phi\rangle$ can be also converted to $|\Psi\rangle$, we consider $|\Psi\rangle$ and $|\Phi\rangle$ to have the same entanglement. Thus, the bipartite states with the same values of Schmidt coefficients have the same entanglement and the entanglement property of bipartite states only depend on their Schmidt coefficients.

As we mentioned in Section 1.1, the only defining features of the systems are their dimension and their tensor product structure. That is, in $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$, all basis sets of a local system $\mathcal{H}_i$ have the same property with respect to information processing and non-locality which we discuss in this thesis, although, for convenience, we sometimes refer to a “computational basis”, a basis which we can arbitrarily choose as a reference basis.
Bibliography


Chapter 2

Bounds on multipartite local discrimination by distance like entanglement monotone

Abstract

In this chapter, we concentrate on the relation between the LOCC discrimination and the entanglement. I give a bound of the LOCC discriminability which is determined only by the value of the entanglement function of each candidate state. That is, I show that upper bounds of the size of an LOCC distinguishable set of states are given by the geometric measures of entanglement, the relative entropy of entanglement, and the global robustness of entanglement, which are entanglement measures defined as the distance from the set of all separable states (“distance-like measures of entanglement”). This result implies that the presence of entanglement guarantees at least a certain minimal level of non-locality concerning the LOCC discrimination. Since this result is valid for all finite sets of multipartite states, this result also gives operational meaning for such entanglement monotonic functions in terms of the LOCC discrimination task.


2.1 Introduction

As we discussed in Chapter 1, the main researches of quantum non-locality so far have been concentrated on the study of entanglement theory. It is known, however, that there is quantum non-locality which cannot be explained only by the entanglement of quantum states. In this chapter, we focus on the local discriminability and investigate its relation with the entanglement.

Non-locality of a state \( \rho \) or a set of states \( \{ \rho_i \} \) is generally defined in an operational way as follows. We can achieve an arbitrary information processing task for a state \( \rho \) or a set of states \( \{ \rho_i \} \) in the case where we can apply coherent quantum operations to our entire systems; this is equivalent to assuming that we can use a perfect quantum channel. If we restrict our operations to LOCC, however, the task may become impossible, or exhibit a difficulty in achieving the task. In this case, we consider that a state \( \rho \) or a set of states \( \{ \rho_i \} \) has non-locality related to that task and that non-locality is greater when the difficulty of achieving task is larger. The discrimination problem of a set of states \( \{ \rho \} \) apparently satisfies the above condition; the discriminability of a set of states is different between between the case where we are allowed to apply all measurements and the case where we are only permitted to apply measurements implemented by LOCC. Thus we can consider this difficulty as non-locality of the set of states.

As mentioned in the subsection 1.3.3 of the previous chapter, Bennett et al. first proved that the non-locality of a set of states related to the local discrimination cannot be explained only by the entanglement of the states. As an example they presented a set of locally indistinguishable orthonormal product states that apparently does not possess any non-locality in terms of the entanglement but possesses non-locality in terms of the local discriminability [11]. After this Bennett et al.'s study, many researches of local discrimination have been done. For example, it was shown that any two pure
states can be discriminated optimally by LOCC no matter how entangled they are [12]. There have been several results since then on various LOCC settings [13], and connections have been made to bipartite the entanglement distillation and formation [14]. However, the exact relation between the local discrimination and the entanglement is still unclear. In particular, there are no general quantitative results which connect the local discriminability to the entanglement.

One of the main reasons of the difficulty in giving the clear connection between the local discriminability and the entanglement is that the general (two-way) LOCC is too difficult to treat in a mathematically rigorous way. Thus, in this chapter, we focus on separable operations (see Definition 1.6), which is a slightly generalized operation compared to the two-way LOCC and is easier to treat than the two-way LOCC. I consider the relation between the perfect local discrimination by separable operations and three entanglement monotones, namely, the robustness of entanglement [3], the relative entropy of entanglement [4], and the geometric measure of entanglement [5], which are defined as the distance from the set of all separable states (“distance-like measure of entanglement”). As a result, we show that an upper bound of the size of a locally distinguishable set of states is given by these three entanglement monotones. Therefore, although we know the local discriminability of a set of states cannot be explained only by the entanglement of the states, we can conclude that the entanglement describes the difficulty of local discrimination by this upper bound of the size of the locally distinguishable set. On the other hand, since this result is valid for all finite sets of multipartite states, this result gives an operational meaning to these entanglement monotones in terms of local discrimination task.

By using this new bounds of local discrimination, we also compute the size of a locally distinguishable set of states for GHZ-states and W-states. These two states are considered as important multi-partite states, since they
are natural multi-partite extensions of a maximally entangled bipartite state and are known to be entangled in a different manners [6]. As a result, we prove that a W-state is more non-local than a GHZ-state in the view point of local discrimination.

2.2 Preliminary

In this section, as a preparation for the next section, we give the problem setting of the local discrimination and the definition of the distance-like measures of entanglement.

In a multi-partite system $H = \bigotimes_{i=1}^{m} H_i$, the local discriminability of a set of states is defined as follows:

**Definition 2.1** A set of states $\{\rho_i\}_{i=1}^{N} \subset S(H)$ is called locally distinguishable by a one-way LOCC, a two-way LOCC, or a separable operation, if there exists a POVM $\{M_i\}_{i=1}^{N}$ such that it can be implemented by a one-way LOCC, a two-way LOCC, or a separable operation, respectively, that satisfies

$$\forall i \quad \text{Tr} \rho_i M_i = 1. \quad (2.1)$$

We have already given the definition of a POVM which can be implemented by a one-way LOCC, a two-way LOCC, and a separable operation in the subsection 1.4.7 of the previous chapter.

Then, we define three entanglement monotones used in this section, namely, the geometric measure of entanglement $E_g(\rho)$, the relative entropy of entanglement $E_R(\rho)$, and the global robustness of entanglement $R_g(\rho)$. For a pure state on $\mathcal{H} = \bigotimes_{i=1}^{m} \mathcal{H}_i$, the definition of the geometric measure of entanglement is defined as the minimum fidelity from the set of all product states [5], that is,

$$E_g(|\Psi\rangle) = -\log_2 \max_{|\Phi\rangle \in \text{Pro}(\mathcal{H})} |\langle \Phi | \Psi \rangle|^2, \quad (2.2)$$

43
where $\text{Pro}(\mathcal{H})$ is the set of all pure product states on $\mathcal{H} = \otimes_{i=1}^{m} \mathcal{H}_i$. Usually, the definition of the geometric measure of entanglement is extended to mixed states by convex roof extension [5]; however, we do not treat this extension of the geometric measure of entanglement in this thesis. The relative entropy of entanglement $E_r(\rho)$ is defined as the minimum relative entropy to the nearest separable state,

$$E_R(\rho) \overset{\text{def}}{=} \min_{\sigma \in \text{Sep}} D(\rho \parallel \sigma),$$

(2.3)

where $\rho$ is a state on $\mathcal{H}$, Sep is the set of all separable states, $D(\rho \parallel \sigma) \overset{\text{def}}{=} \text{Tr}(\rho \log \rho - \rho \log \sigma)$ is the quantum relative entropy [4]. Similarly, the global robustness of entanglement $R_g(\rho)$ is defined as the minimum noise which we need to destroy the entanglement of $\rho$ as follows [3]:

$$R_g(\rho) \overset{\text{def}}{=} \min \left\{ t \in [0, 1] \mid \exists \sigma \in S(\mathcal{H}), \right. \left. s.t. \right\| \frac{1}{1 + t} (\rho + t\sigma) \in \text{Sep} \left\}, \right.$$

(2.4)

where $S(\mathcal{H})$ denotes the set of all states on $\mathcal{H}$. The above three functions, the geometric measure of entanglement, the relative entropy of entanglement, and the robustness of entanglement, are shown to be monotonically increasing under the LOCC operation. Therefore, we can consider these functions as measures of entanglement. Since all the above three functions are defined as the minimums of “distance like” functions from the set of all separable states, we call these three functions as “distance-like measures of entanglement”.

### 2.3 Proof of a new bound on local discrimination

Our new bound of the local discrimination can be written down as follows.

**Theorem 2.1** If a set of $N$ states $\{\rho_i\}_{i=1}^{N}$ is locally distinguishable by sep-
arable operations, the number of states $N$ is bounded by

$$N \leq D/d(\rho_i) \leq D/r(\rho_i) \leq D/2^{E_R(\rho_i)+S(\rho_i)} \leq D/2^G(\rho_i),$$

(2.5)

where $\pi_i := 1/N \sum_{i=1}^{N} x_i$ denotes the average, and the functions $d(\rho)$, $r(\rho)$, $s(\rho)$, and $G(\rho)$ are defined as follows:

$$d(\rho) \overset{\text{def}}{=} \min \left\{ \frac{1}{\text{Tr}(\rho \omega)} \left| 0 \leq \frac{\omega}{\text{Tr}(\rho \omega)} \leq I, \omega \in \text{Sep} \right. \right\},$$

(2.6)

$$r(\rho) \overset{\text{def}}{=} \text{Tr}P \left[ 1 + R_g \left( \frac{P}{\text{Tr}P} \right) \right],$$

(2.7)

$$s(\rho) \overset{\text{def}}{=} -\text{Tr}\rho \log \rho,$$

(2.8)

$$G(\rho) \overset{\text{def}}{=} -\log_2 \left\{ \max_{\omega \in \text{Sep}} \text{tr}\{\rho \omega\} \right\},$$

(2.9)

where $P$ is the support of the state $\rho$, $s(\rho)$ is the von Neumann entropy, and we call $G(\rho)$ the geometric measure.

Thus, if all states are pure, we derive the bounds as follows:

**Corollary 2.1** If a set of $N$ pure states $\{|\Psi_i\rangle\}_{i=1}^{N}$ is locally distinguishable by separable operations, the number of states $N$ is bounded by

$$N \leq D/d(|\Psi_i\rangle) \leq D/1 + R_g(|\Psi_i\rangle) \leq D/2^{E_R(|\Psi_i\rangle)} \leq D/2^{E_g(|\Psi_i\rangle)}.$$

(2.10)

Hence, in a case of pure state, where the bounding quantities are reduced to the geometric measure of entanglement, the relative entropy of entanglement and the robustness of entanglement (from right to left), we can interpret these three distance-like entanglement measures as bounds on the number of pure states that we can discriminate perfectly by LOCC (see Fig. 2.1). Apparently, Theorem 2.1 and Corollary 2.1 are also valid if we swap the word “separable operations” to “two-way LOCC” or “one-way LOCC” on
To discriminate the pure states \( \{\varphi_i\}_{i=1}^{N} \) perfectly under LOCC, the sum of the entanglement “distances” \( E(|\varphi_i\rangle) \) must be less than the total dimension \( D \) (Theorem 2.1 and Corollary 2.1), thus \( N \leq D/E(|\varphi_i\rangle) \).

Figure 2.1: To discriminate the pure states \( \{\varphi_i\}_{i=1}^{N} \) perfectly under LOCC, the sum of the entanglement “distances” \( E(|\varphi_i\rangle) \) must be less than the total dimension \( D \) (Theorem 2.1 and Corollary 2.1), thus \( N \leq D/E(|\varphi_i\rangle) \).

We start proving Theorem 2.1 by the following lemma.

**Lemma 2.1** A set of states \( \{\rho_i | i = 1..N\} \) is distinguishable by separable operations, if the following inequality holds:

\[
\sum_i d(\rho_i) \leq D, 
\]  

(2.11)

where \( D \) is the total dimension of the system.

**Proof** To prove Theorem 1, we begin by listing some conditions that the POVMs (positive operator value measures) must satisfy. The task of the state discrimination is to perform a measurement (in our case by LOCC) on a system to find out which one of a set of states the system is in. If it is possible to perfectly discriminate among a set of density matrices \( \mathcal{S} := \{\rho_i | i = 1..N\} \)
by separable operations, then it is necessary that there exists a POVM \( \{ M_i \} \) (Subsection 1.4.2 and 1.4.3) satisfying the following conditions:

\[
\sum_i M_i = I \quad (2.12)
\]

\[
I \geq M_i \geq 0 \quad (2.13)
\]

\[
\forall i \quad \text{Tr}(M_i \rho_i) = 1 \quad (2.14)
\]

\[
\forall i \quad M_i / \text{Tr} M_i \in \text{Sep} \quad (2.15)
\]

The conditions (2.12) and (2.13) are simply the conditions which mean that \( \{ M_i \} \) is a POVM. The condition (2.14) says that, given a state \( \rho_i \), the result corresponding to the outcome \( M_i \) occurs with probability 1, i.e. the discrimination is deterministic. The condition (2.15) is a necessary and sufficient condition that the POVM \( \{ M_i \} \) is to be implementable by separable operations; see the definition of separable operations (Definition 1.6).

To make a connection to distances between states, we first notice that any POVM element \( M_i \) can be expressed as a positive number \( s_i = \text{Tr}(M_i) \) times a density matrix \( \omega_i \), \( M_i = s_i \omega_i \). We can then use this to immediately rewrite (2.12)-(2.15). The condition (2.14) is rewritten as \( s_i = 1 / \text{Tr}(\rho_i \omega_i) \). The condition (2.15) means that \( \omega_i \) is separable. For pure states, \( s_i \) now looks like a distance-like quantity between a state \( \rho_i \) and a separable state \( \omega_i \), the remaining conditions are satisfied (that is, \( \sum_i s_i \omega_i = I, \ I \geq s_i \omega_i \geq 0 \)).

If we then minimize \( s_i \) such that the conditions (2.13), (2.14) and (2.15) are satisfied for each \( i \) independently, we get exactly the definition of \( d(\rho_i) \) in Theorem 1, (2.6). From the condition (2.12), we derive a lower bound of
as follows,

\[ D = \sum_i \text{Tr}(M_i) \]
\[ = \sum_i s_i \]
\[ = \sum_i \frac{1}{\text{Tr} \rho_i \omega_i} \]
\[ \geq \sum_i d(\rho_i), \quad (2.16) \]

completing the proof. \( \square \)

At this point \( d(\rho) \) cannot be considered as an entanglement measure of the ‘distance to the closest separable state’. It turns out that condition \( \frac{\rho}{\text{Tr}(\rho \omega)} \leq I \) in (2.6) complicates things a lot, and indeed, even without this condition it is not an entanglement monotone for mixed states; see a comment below the definition of the geometric measure (2.26). Hence the connection to the entanglement is not immediate. However, we can use this quantity \( d(\rho) \) to relate the problem of the state discrimination to other distance-like entanglement monotones, as in the following theorem.

**Lemma 2.2** The following bounds hold for all states \( \rho \):

\[ d(\rho) \geq r(\rho) \geq 2^{E_R(\rho) + S(\rho)} \geq 2^{G(\rho)}, \quad (2.17) \]

where \( G(\rho) \) is the geometric measure, \( E_R(\rho) \) is the relative entropy of entanglement, \( S(\rho) \) is the von Neumann entropy, and \( r(\rho) := |P|(1 + R_G(\rho/|P|)) \), where \( P \) is the support of the state \( \rho \) (the support of a state \( \rho \), with eigen-decomposition \( \rho = \sum_i \alpha_i |i\rangle\langle i| \) is given by \( P = \sum_i |i\rangle\langle i| \), \( |P| := \text{Tr}P \), and \( R_G(\rho) \) is the robustness of entanglement of the state \( \rho \).

In the case of a pure state, \( S(\rho) = 0 \) and \( P = \rho \), and hence these quantities \( G(\rho), E_R(\rho) + S(\rho) \), and \( r(\rho) \) become exactly (up to log) the geometric
measure of entanglement, the relative entropy of entanglement and the robustness of entanglement, respectively. In the case of a mixed state, they include some quantification of how mixed the state is. This makes sense in the problem of the state discrimination, since the more mixed the states are, the fewer orthogonal states there can be for a given Hilbert space dimension $D$; that is, if a set of states $\{\rho_i\}$ is an orthogonal set, then, $\sum_i \text{rank}(\rho_i) \leq D$.

We will later show that the quantities in Eq.(2.17) are equivalent for GHZ states (GHZ states are multipartite states defined originally in [15]).

(Proof of Lemma 2.2): To prove the first inequality of Eq.(2.17), we must first write $d(\rho)$ in a more convenient form. We can rewrite the condition $\frac{\omega}{\text{Tr}(\rho \omega)} \leq I$ in (2.6) as

$$\langle \psi | \omega | \psi \rangle \leq \text{Tr}(\rho \omega)$$

(2.18)

for all $|\psi\rangle$.

Suppose $\omega$ can be decomposed as follows:

$$\omega = \lambda |P'| \frac{P'}{|P'|} + (1 - \lambda |P'|) \Delta'$$

(2.19)

where $\lambda$ is the maximal eigenvalue of $\omega$, $P'$ is the projection onto the eigenspace of $\lambda$, and $\Delta'$ is a states satisfying $\text{Tr} P' \Delta' = 0$. Then, since $\lambda$ is the maximally eigenvalue and $\rho$ is a positive operator with the unit trace norm, $\text{Tr} \rho \omega \leq (\text{Tr} \rho) \|\omega\|_\text{op} = \lambda$, where $\|\omega\|_\text{op} \overset{\text{def}}{=} \max_{\|\Psi\| = 1} \langle \Psi | \omega | \Psi \rangle$ is the operator norm of $\omega$. Moreover, from the inequality (2.18) and the well known equation $\lambda = \|\omega\|_\text{op} = \max_{\|\Psi\| = 1} \langle \Psi | \omega | \Psi \rangle$, we also have $\text{Tr} (\rho \omega) \geq \lambda$; that is, $\text{Tr} (\rho \omega) = \lambda$. This equality immediately means $P' \geq P$, where $P$ is a projection onto the support of $\rho$. Thus there exists a projection $P''$ such that
\(P' = P + P''\) and \(\text{Tr}PP' = 0\). Therefore, we can decompose \(\omega\) as follows:

\[
\omega = \lambda P' + (1 - \lambda|P'|)\Delta' \\
= \lambda(P + P'') + (1 - \lambda|P'|)\Delta' \\
= \lambda P + \{\lambda P'' + (1 - \lambda|P'|)\Delta\} \\
= \lambda|P|\frac{P}{|P|} + (1 - \lambda|P|)\Delta,
\]

where \(\Delta\) is a state defined as \(\Delta \overset{\text{def}}{=} \frac{\lambda}{(1 - \lambda|P'|)}P'' + \frac{(1 - \lambda|P'|)}{(1 - \lambda|P'|)}\Delta'.\) Since both \(\Delta'\) and \(P''\) are orthogonal to \(P\), \(\Delta\) is also orthogonal to \(P\). Thus, \(\omega\) can always be rewritten in the form \(\omega = \lambda|P|\frac{P}{|P|} + (1 - \lambda|P|)\Delta\) with the additional conditions \(\text{tr}\{P\Delta\} = 0\) and \(\lambda \geq \langle \psi |\omega|\psi \rangle\) for all \(|\psi\rangle\).

By substituting the above formula of \(\omega\) for the definition of \(d(\rho)\) (2.6), we can rewrite \(d(\rho)\) as

\[
d(\rho) = \min(1/\lambda) \\
such that \exists\ a state \Delta,\ satisfying \\
\omega = \lambda|P|\frac{P}{|P|} + (1 - \lambda|P|)\Delta \in \text{Sep}, \\
\text{tr}\{P\Delta\} = 0,\ \lambda \geq \langle \psi |\omega|\psi \rangle\ for\ all\ |\psi\rangle.
\]

(2.20)

We can now compare the above formula of \(d(\rho)\) to the **global robustness of entanglement** \(R_g(\rho)\) [3].

\[R_g(\rho) := \min t\]

such that \(\exists\ a state \Delta,\ satisfying\)

\[
\frac{1}{1+t}(\rho + t\Delta) \in \text{Sep}.\]

(2.21)

We can understand this as the minimum (arbitrary) noise \(\Delta\) that we need to add to make the state separable.

We can see that the global robustness of entanglement of the support of state \(\rho\), \(R_g(P/|P|)\), is very similar in definition to \(d(\rho)\) in (2.20); that is, we...
can easily see the following correspondences,

\[
\begin{align*}
\lambda |P| & \leftrightarrow \frac{1}{1 + t} \\
1 - \lambda |P| & \leftrightarrow \frac{t}{1 + t} \\
\frac{1}{d(\rho)} |P| & \leftrightarrow \frac{1}{1 + R_g(\rho)} \\
d(\rho) & \leftrightarrow |P|(1 + R_g(\rho))
\end{align*}
\]

The crucial difference is the removal of the two conditions in the last line of (2.20). Since relaxing conditions can only lead to a lower minimum, we can see that

\[d(\rho) \geq r(\rho) := |P| [1 + R_g(P/|P|)]\],

proving the first inequality of Lemma 2.2.

Next, we consider the second inequality in Eq. (2.17). The relative entropy of entanglement is defined as [4]

\[E_R(\rho) := \min_{\omega \in SEP} S(\rho||\omega),\] (2.23)

where \(S(\rho||\omega) = -S(\rho) - \text{Tr}(\rho \log_2 \omega)\) is the relative entropy and \(S(\rho)\) is the von Neumann entropy. From the definition of \(R_g(\rho)\), we know that for a state \(\Delta\), the state given by \(\omega_i := [P_i/|P_i| + R_g(P_i/|P_i|)\Delta]/[1 + R_g(P_i/|P_i|)]\) is a separable state. Hence the following inequalities must hold:

\[
\begin{align*}
E_R(\rho_i) + S(\rho_i) & = -\min_{\omega \in SEP} \text{Tr}(\rho_i \log_2 \omega) \\
& \leq -\text{Tr}(\rho_i \log_2 \omega_i) \\
& = -\text{Tr} \left[ \rho_i \log_2 \left( \frac{P_i/|P_i| + R_g(P_i/|P_i|)\Delta}{1 + R_g(P_i/|P_i|)} \right) \right] \\
& \leq -\text{Tr} \left[ \rho_i \log_2 \left( \frac{P_i/|P_i|}{1 + R_g(P_i/|P_i|)} \right) \right] \\
& = -\left[ \log_2 \left( \frac{1/|P_i|}{1 + R_g(P/|P_i|)} \right) \right] \text{Tr}(\rho_i P_i) \\
& = \log_2 \left[ |P_i| (1 + R_g(P_i/|P_i|)) \right],
\end{align*}
\]

51
where the fourth line follows from the monotonicity of the logarithm, which states that $\log(A + B) \geq \log(A)$ whenever $B \geq 0$, for two operators $A, B$ [16]. In the fifth line, we use the equality $\log(pP) = (\log p)P$ for a positive coefficient $p \geq 0$ and a projection $P$, where we substituted $\frac{1}{1+R_\rho(P_i/P_i)}$ and $P_i$ for $\lambda$ and $P$, respectively. The last line is true even if $\rho_i$ is any state in the span of $P_i$. Hence:

\[ 2^{E_R(\rho_i) + S(\rho_i)} \leq r(\rho_i). \]  

Finally, we consider the last inequality in Eq.(2.17). We call the geometric measure $G(\rho)$

\[ G(\rho) := -\log_2 \left[ \max_{\omega \in SEP} \text{Tr}(\rho \omega) \right]. \]  

In the case of pure states, this is reduced to the geometric measure of entanglement [5]. However, for mixed states, this is not an entanglement monotone. For example, it is maximized by the maximally mixed state. We immediately see that this would be equivalent (up to log) to $d(\rho)$ in (2.6) if we were to drop the condition $\frac{\omega}{\text{Tr}(\rho \omega)} \leq I$. Hence we have $d(\rho) \geq 2^{G(\rho)}$. However, it is possible to show a stronger bound. In [17] it was shown that in the case of a pure state, $G(\rho)$ is bounded from above by the relative entropy of entanglement. We use the same simple concavity arguments now for the case of a mixed state. By definition $E_R(\rho_i) + S(\rho_i) = -\max_{\omega \in SEP} \text{Tr}(\rho \log_2 \omega)$. By concavity of the logarithm, we have for all $\rho$ and $\omega$, $\text{Tr}(\rho \log_2 \omega) \leq \text{Tr}(\rho \omega)$. Thus

\[ E_R(\rho) + S(\rho) \geq G(\rho) \]  

Combining (2.22), (2.25) and (2.27) we complete the proof of the lemma. □

Finally, by means of the above two lemmas, we can immediately derive our main theorem.

**Proof** (Proof of Theorem 2.1) Combining Lemma 2.1 and Lemma 2.2, and dividing the equation by $N$, we immediately obtain Eq.(2.5). □
2.4 Application of the new bound

Given the hierarchy of the bounds (2.5), we can apply known results from entanglement theory to find bounds on $N$, one of which we will show is tight. Firstly, the robustness of entanglement is completely solved for pure bipartite states [3]. For a state with the Schmidt decomposition $|\Psi\rangle = \sum_i \sqrt{\lambda_i} |ii\rangle$, the robustness was found to be

$$R_{g}(|\Psi\rangle) = (\sum_{i} \sqrt{\lambda_i})^2 - 1. \quad (2.28)$$

We can immediately put this into (2.10). For instance, if we have a set of pure bipartite states $\{|\Psi_k\rangle\}_{k=1}^{N}$ all with the same entanglement of $|\Psi\rangle$ (in other words, each $|\Psi_k\rangle$ can be transformed to $|\Psi\rangle$ by a local unitary transformation), we have

$$N \leq d_1 d_2 / (\sum_i \sqrt{\lambda_i})^2, \quad (2.29)$$

where $d_1$ and $d_2$ are the dimensions of the Hilbert spaces, respectively, and $\{\lambda_i\}_{i=1}^{\min\{d_1,d_2\}}$ are the Schmidt coefficients for any one of the states in the set. In particular, in the case where all $\{|\Psi\rangle\}$ is a maximally entangled state, we derived the following consequence: it is impossible to distinguish more than $d$ maximally entangled states (where $d$ is the dimension of one subspace, then $(\sum \sqrt{\lambda_i})^2 = d$). This fact about a set of maximally entangled states has been already known [18, 19] (see also chapter 4).

In an $m$-qubit system $\mathcal{H} \equiv (\mathbb{C}^2)^\otimes m$, the W-state and the GHZ-state are define as follows: the W state $|W\rangle \equiv \frac{1}{\sqrt{m}}(|00\ldots01\rangle + |00\ldots10\rangle + \ldots + |01\ldots00\rangle + |10\ldots00\rangle)$, and the GHZ state $|\text{GHZ}\rangle := \frac{1}{\sqrt{2}}(|0\rangle^\otimes m + |1\rangle^\otimes m)$. These two states, the W-state and the GHZ-state, are considered as important states, since these are natural extensions of the maximally entangled states of the two-qubit system and are known to be inconvertible under SLOCC transformations. (The SLOCC inconvertible states do not exist in a bipartite system.
and are considered as one of the most important features of the multi-partite entanglement [6]. We know from Wei et al. [17] that for the \( m \)-party W state \(|W⟩\) and GHZ state \(|GHZ⟩\), the relative entropy of entanglement and the geometric measure coincide with each other and are given by \( E_R(|GHZ⟩) = E_g(|GHZ⟩) = 1 \) and \( E_R(|W⟩) = E_G(|W⟩) = \log_2(m/(m - 1))^{(m-1)} \). Therefore, suppose sets of states \(|Ψ_k⟩\) \(_{k=1}^{N(GHZ)}\) and \(|Φ_l⟩\) \(_{l=1}^{N(W)}\) are locally distinguishable and each state \( |Ψ_k⟩\) and \( |Φ_l⟩\) satisfy \( E_g(|Ψ_k⟩) = E_g(|GHZ⟩) \) and \( E_g(|Φ_l⟩) = E_g(|W⟩) \), respectively. Then, we have

\[
N(GHZ) \leq 2^{m-1} \\
N(W) \leq 2^m ((m - 1)/m)^{(m-1)} .
\] (2.30)

Then, we focus on sets which is made of GHZ-type states and W-type states; GHZ-type states are states which can be transformed from GHZ states by local unitary transformations and W-type states are states which can be transformed from W states by local unitary transformations. The maximum number of states in locally distinguishable sets of GHZ-type states \( N(S_{GHZ}) \) is defined as follows:

\[
N(S_{GHZ}) \overset{\text{def}}{=} \max(N) \\
\text{such that } \exists \{ |Ψ_k⟩ \}_{k=1}^{N} \\
\{ |Ψ_k⟩ \}_{k=1}^{N} \text{ is locally distinguishable} \\
\exists \text{ local unitary } U_k \text{ satisfying } |Ψ_k⟩ = U_k |GHZ⟩
\]

We also define the maximum number of states in locally distinguishable sets of W-type states \( N(S_W) \) in the same manner. Then we can show \( N(S_{GHZ}) = 2^{m-1} \) by an explicit construction. We form a set of states \( \{ |Ψ_k⟩ := I \otimes U_k |GHZ⟩ \}_{k=1}^{m-1} \) by local unitaries \( U_k \) over \( m - 1 \) parties. The unitaries \( U_k \) are formed from all the possible combinations of products of the identity and \( σ_x \) Pauli operations, e.g. \( U_1 = I^\otimes (m-2) \otimes σ_x \), giving a set of \( \sum_{j=0}^{m-1} \binom{m-1}{k} = 2^{m-1} \),

54
states. It is easy to check that these can be discriminated by making local $\sigma_z$ measurements. Calling $\mathcal{S}_W$ a set of states equal to the W state up to local unitary transformations, with (2.30) gives

$$N(\mathcal{S}_W) < N(\mathcal{S}_{GHZ}).$$  

This result guarantees that the W-state is more non-local than the GHZ-state in terms of local discriminability.

We also note that if we can find such a bound by any of the entanglement measures in (2.5) and show that it is tight, the measures below it in the hierarchy are equal. The GHZ case is such an example giving $R_g(\vert\text{GHZ}\rangle) = 1$, and is one of the few cases where the global robustness of entanglement is known for multiparty systems. We round off the examples by showing another simple known result. If even one state in a complete basis is entangled, then (2.5) shows that the basis cannot be discriminated perfectly [20].

2.5 Conclusion

In this chapter, we have focused on the relation between the local discrimination and the entanglement for general multi-partite states and have shown that the distance-like measures of the entanglement (that is, the geometric measure of entanglement, the relative entropy of entanglement, and the robustness of entanglement) give upper bounds for the size of a set of locally distinguishable set of states (Theorem 2.1 and Corollary 2.1 in Section 2.3). Therefore, we have concluded that the entanglement guarantees a certain minimum level of the difficulty of the local discrimination. This difficulty is non-locality in terms of local discrimination. On the other hand, this result also gives an operational meaning for the distance like measures of entanglement in terms of the local discrimination task. By using known results on the entanglement, we have also given examples of existing and new LOCC discrimination bounds in a unified manner (Section 2.4).
Finally, we give the following two remarks. First, the simplicity of the basis of the proofs of the main results here allows it to be used with other necessary conditions on LOCC measurements. The condition of the separability (2.15), for example, may be changed to more tractable conditions such as the positivity of partial transpose or bi-separability [21]. It can easily be seen that these conditions would lead to analogous bounds to those derived in this chapter. In the case of biseparability, Eq.(2.29) shows that, for pure states, it is always possible to give some easily computable bound. Second, the difficulty of LOCC state discrimination is an important concept in various quantum information tasks (e.g. quantum data hiding [10]), which may give more usages of these results. This is a topic of ongoing investigations. In this direction, it is also possible to extend Lemma 2.1 to the case of imperfect discrimination. This leads to bounds on the LOCC accessible information as in [19, 22], which will be presented in a separate paper.
Bibliography


Chapter 3

Local discrimination; one-way vs two-way LOCC

Abstract

In this chapter, we analyze the difference in the LOCC discriminability for the case of permitting two-way classical communications (two-way LOCCs) and for the case of permitting only one-way classical communications (one-way LOCCs). As a result, I derive an upper bound of the size of a one-way locally distinguishable set for bipartite states and a set of three partite pure states. After that, in a two-qubit system, by constructing a concrete two-way local discrimination protocol, we show that two-way classical communications remarkably improves the local discriminability in comparison with the local discrimination by one-way classical communications.

3.1 Introduction

In the fields of quantum information, as we mentioned in section 1.1, we often consider performing distributed information processing. In distributed quantum information processing, two or more spatially separated parties “communicate” each other to perform a task. In such cases, since transmission of quantum states over long distance is technologically much more
difficult than transmission of classical information, the consideration as to which kind of information processing tasks can be achieved without communications of quantum states, or only by local (quantum) operations and classical communication (LOCC) is useful [4].

In the LOCC scenario, we consider two different cases related to the classical communication, that is, the case where all parties are allowed to communicate classically with each other as much as they like, and the case where the classical communication is restricted to one direction (for example, parties are labelled as $A, B, \cdots, Z$, and they are allowed to communicate classically only in the alphabetical order.) The former is called the two-way LOCC, and the latter is called the one-way LOCC. Since, by definition, the two-way LOCC apparently includes the one-way LOCC, we may think that we can achieve information tasks by the two-way LOCC more than the one-way LOCC. For example, the quantum capacity (the size of a quantum state that can be reliably transmitted over a quantum channel) and the distillable entanglement (the number of maximally entangled states which can be derived from a given state by LOCC) with two-way classical communications are proven to be greater than those with one-way classical communications [1]. However, so far, the above are rare examples, and we hardly know the information tasks which we achieve by two-way LOCCs more than by one-way LOCCs. The first reason is that the mathematical definition of two-way LOCC (Definition 1.5) is too complicated, and it is mathematically hard to rigorously evaluate the performance of a given two-way LOCC protocol. The second reason is that for several simple tasks, the performance of the two-way LOCC is actually shown to be the same as the performance of the one-way LOCC. For example, it was shown that the LOCC convertibility of bipartite pure states does not depend on the choice of the two-way LOCC and the one-way LOCC; that is, in bipartite systems, if a pure state can be transformed to another pure state by the two-way LOCC, then it is always possible by
the one-way LOCC [2]. In this chapter, we present a new example of the “local discrimination” where two-way classical communications enhance the performance of a task, compared to one-way classical communications.

The local discrimination (a problem to discriminate an unknown states by only LOCC [3]) is an important task in the LOCC scenario. In quantum communication, we encode classical information \{1, \cdots, n\} onto a set of quantum states \{\rho_i\}_{i=1}^n, and in order to decode this encoded classical information, we need to distinguish the set of states \{\rho_i\}_{i=1}^n. That is, information theoretically, the local discriminability means how much classical information we can derive from an unknown state by LOCC. Therefore, if the performance of local discrimination with two-way classical communications is better than the performance with one-way classical communications, we can conclude that the two-way classical communication enhances the power to derive a classical information from an unknown state compared to the one-way classical communication. In this chapter, we will show the above statement is true.

In the previous chapter, we derived an upper bound on the size of locally distinguishable sets with separable operations in terms of entanglement monotones (Theorem 2.1). Since the two-way LOCC is included by separable operations, this upper bound is also an upper bound for two-way LOCC. However, no upper bound has been known for the case of local discrimination with the one-way LOCC. In order to compare the one-way LOCC with the two-way LOCC, we first derive an upper bound of the size of locally distinguishable states in the case of the one-way LOCC. Then, by presenting an example of the two-way LOCC discrimination protocol, we show that two-way classical communications really enhance the performance of LOCC discrimination compared to one-way classical communications.

This chapter is organized as follows: In Section 3.2, we consider the discrimination problem between an arbitrary given states \rho and a completely
mixed state $I/\dim\mathcal{H}$ on a multi-partite system $\mathcal{H}$ under the condition that the given state is detected perfectly, and let $t_{\rightarrow}(\rho)$, $t_{\leftarrow}(\rho)$, $t_{\text{sep}}(\rho)$ denote the minimum error probability to detect $\rho$ by the one-way LOCC, the two-way LOCC, and the separable operation, respectively. Then, we show that these functions are considered as appropriate measures of the local discriminability, since they are not only the minimum probability of the above problem, but also give upper bounds of the size of locally distinguishable sets in general local discrimination problems. In section 3.3, we show that the function $d(\rho)$ defined in the previous chapter gives a lower bound for $t_{\text{sep}}(\rho)$. In section 3.4, we show that, in the bipartite case, the measure of the one-way LOCC discriminability $t_{\rightarrow}$ is same as the Schmidt rank of the states. In Section 3.5, we prove that, in three qubits systems, a set of W-type states (where a W-type state is a state written as $\lambda_1 |a_0b_0c_1\rangle + \lambda_2 |a_0b_1c_0\rangle + \lambda_3 |a_1b_0c_0\rangle$ with complex coefficients $\lambda_1, \lambda_2, \lambda_3$ and an arbitrary orthonormal basis set of local systems $\{|a_i\rangle_i\}, \{|b_j\rangle_j\}, \{|c_k\rangle_k\}$) cannot be discriminated more than three by the one-way LOCC. Finally, in section 3.6, we consider a two-way LOCC discrimination protocol for two-qubits systems. We derive an upper bound for $t_{\leftarrow}(\rho)$. By means of this upper bound, we show that $t_{\leftarrow}(\rho)$ is strictly greater than $t_{\rightarrow}(\rho)$ for entangled states except for the maximally entangled states. Moreover, we show that since $d(\rho)$ gives a lower bound for $t_{\text{sep}}(\rho)$, $t_{\rightarrow}(\rho)$ and $t_{\text{sep}}(\rho)$ gives almost the same value for two-qubits states (see FIG. 3.6). Since $t_{\leftarrow}(\rho)$ itself has operational meaning of local discrimination presented in the previous sections, this is, a piece of evidence which shows that the two-way classical communication enhances the performance of the local discrimination compared to the one-way classical communication.
3.2 Preliminary

In this section, we consider a system which is composed of several subsystems. We do not restrict the number of the subsystems. We denote the Hilbert space of the composite system as \( \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots \), and define the dimension of the Hilbert space as \( d = \text{dim} \mathcal{H} = \text{dim} \mathcal{H}_A \times \text{dim} \mathcal{H}_B \times \cdots \). In the composite system \( \mathcal{H} \), we call \( T \) with \( 0 \leq T \leq I \) a POVM element, where \( I \) is an identity operator on \( \mathcal{H} \), if the two-valued POVM \( \{ T, I - T \} \) can be implemented by the one-way LOCC, the two-way LOCC, and the separable operations, respectively [5]. That is, if \( T \) is a one-way LOCC (two-way LOCC, or separable) POVM element, then there exists a one-way LOCC (two-way LOCC, or separable, respectively) POVM which includes \( T \) as its element. We write a set of one-way LOCC, two-way LOCC, and separable POVM elements as \( T_{\rightarrow}, T_{\leftrightarrow}, T_{\text{sep}} \). Obviously, they satisfy the relation \( T_{\rightarrow} \subseteq T_{\leftrightarrow} \subseteq T_{\text{sep}} \). We note the following fact, if \( T \in T_c \), then \( I - T \in T_c \), where \( c \) can be either \( \rightarrow, \leftrightarrow, \text{sep} \).

In this chapter, we are interested in how the power of the one-way LOCC, the two-way LOCC, and the separable operations are different for the local discrimination. Although there are many problem settings for the local discrimination, we especially focus on one of the simplest problem settings as follows: We consider local discrimination of \( \rho \) and \( \rho_{\text{mix}} = \frac{I_{AB}}{d} \), and investigate how well we can detect \( \rho_{\text{mix}} \) under the assumption that we do not make any error to detect \( \rho \), namely, how well we can distinguish \( \rho \) from the completely random mixed state \( \rho_{\text{mix}} \) without making any error.

Our problem can be written down rigorously as follows. We detect \( \{ \rho, \rho_{\text{mix}} \} \) by the POVM \( \{ T, I - T \} \), where \( T \in T_{\rightarrow}, T_{\leftrightarrow}, \text{or} T_{\text{sep}} \), that is, if we get the result corresponding to \( T \), then we decide that the state is in \( \rho \), and if we get the result corresponding to \( I - T \), then we decide that the state is in \( \rho_{\text{mix}} \). We consider two kinds of error probability as follows: the type 1
error probability $\text{Tr}\rho(I - T)$, and the type 2 error probability $\text{Tr}\rho_{\text{mix}}T$. In this case, the type 1 error probability corresponds to the error probability that the real state is $\rho$ and our decision is $\rho_{\text{mix}}$, and the type 2 error probability corresponds to the error probability that the real state is $\rho_{\text{mix}}$ and our decision is $\rho$. Under the above condition, we try to minimize the type 2 error probability $\text{Tr}\rho_{\text{mix}}T$ under the assumption that the type 1 error probability $\text{Tr}\rho(I - T)$ is 0. Thus, we focus on the following minimum of the type 2 error probability:

$$\beta_c(\rho\|\rho_{\text{mix}}) = \min\{\text{Tr}(\rho_{\text{mix}}T) | T \in T_c, \text{Tr}T = 1\},$$

(3.1)

where $c = \rightarrow, \leftrightarrow, \text{and sep}$. Then, we define $t_c(\rho)$ as

$$t_c(\rho) = \min\{\text{Tr}T | T \in T_c, \text{Tr}T \rho = 1\}.$$

(3.2)

Hence it satisfies

$$t_c(\rho) = d\beta_c(\rho\|\rho_{\text{mix}}),$$

(3.3)

where $d$ is the dimension of the whole system $\mathcal{H}$. Therefore, $t_c(\rho)$ is in proportion to the minimum of the type 2 error probability $\beta_c(\rho\|\rho_{\text{mix}})$ of one-way LOCC, two-way LOCC, and separable POVM in the case where $c = \rightarrow, \leftrightarrow, \text{and sep}$, respectively. Obviously, $t_c(\rho)$ satisfies $t_{\text{sep}}(\rho) \leq t_{\leftrightarrow}(\rho) \leq t_{\rightarrow}(\rho)$.

Moreover, suppose that a set of states $\{\rho_i\}_{i=1}^{N_c}$ is locally distinguishable by one-way ($c = \rightarrow$), two-way LOCC ($c = \leftrightarrow$), or separable ($c = \text{sep}$) POVM. From the result obtain in the previous chapter, $t_c(\rho_i)$ (which corresponds to $d(\rho)$) gives an upper bound of $N_c$ as [7],

$$N_c \leq d/t_c(\rho_i),$$

(3.4)

where $t_c(\rho_i)$ is the average of $\{t_c(\rho_i)\}_{i=1}^{N_c}$ [7]. Thus, $t_c(\rho)$ is an appropriate measure of local discriminability in an operational sense, and also as a function whose average gives an upper bound for the locally distinguishable sets of states. Therefore, we investigate the difference of local discriminability of
\( \rho \) by one-way LOCC POVM, two-way LOCC POVM, and separable POVM in terms of \( t_c(\rho) \) in the following sections.

### 3.3 Local discrimination by separable POVM

In this section, we investigate the local discrimination by separable POVM and give a formula of a slight generalization of \( T_{\text{sep}} \) for bipartite systems.

We can define a set of POVM elements \( T_{\tilde{\text{sep}}} \) by,

\[
T_{\tilde{\text{sep}}} \overset{\text{def}}{=} \{ T \mid T \leq I_{AB}, T = \sum_i N_i \otimes M_i \otimes \cdots, \forall i, N_i \geq 0, M_i \geq 0, \cdots \}.
\]

\( T_{\tilde{\text{sep}}} \) is a set of POVM whose POVM elements can be decomposed into separable forms; we say a positive linear operator \( M \) is in a separable form, if \( M/\text{Tr}M \) is a separable state; See Definition 1.8. Since the definition of \( T_{\tilde{\text{sep}}} \) is the definition of \( T_{\text{sep}} \) without the condition \( I - T \in T_{\text{sep}} \), we have \( T_{\text{sep}} \subset T_{\tilde{\text{sep}}} \).

Note that even if \( T \in T_{\tilde{\text{sep}}} \), \( I - T \) does not necessarily satisfy \( I - T \in T_{\tilde{\text{sep}}} \), that is \( T_{\text{sep}} \neq T_{\tilde{\text{sep}}} \). Similarly, we can define \( t_{\tilde{\text{sep}}} \) as,

\[
T_{\tilde{\text{sep}}} = \{ T \mid T \leq I_{AB}, T = \sum N_i \otimes M_i \otimes \cdots, \forall i, N_i \geq 0, M_i \geq 0, \cdots \}.
\]

By definition, \( t_{\tilde{\text{sep}}}(\rho) \) apparently gives a lower bound of \( t_{\text{sep}}(\rho) \), that is, for all \( \rho \in S(\mathcal{H}) \),

\[
t_{\tilde{\text{sep}}}(\rho) \leq t_{\text{sep}}(\rho).
\]

Then, we can easily see that \( t_{\tilde{\text{sep}}} \) is actually equal to \( d(\rho) \) which we defined in Eq. (2.6) of the previous chapter. From the result of the previous chapter, for bipartite pure state, \( t_{\tilde{\text{sep}}} \) is represented by

\[
t_{\tilde{\text{sep}}}(|\Psi\rangle) = d(|\Psi\rangle) = 1 + R_g(|\Psi\rangle) = \sum_{i,j} \sqrt{\lambda_i \lambda_j},
\]

where \( R_g(\rho) \) is the global robustness of entanglement [6] and \( \{\lambda_i\}_{i=0}^{d-1} \) the Schmidt coefficients of \( |\Psi\rangle \). We can easily see the first equality by comparing

66
the definition of $t_{\text{sep}}(|\Psi\rangle)$ and $d(\rho)$. We can check the second equality by using the fact that the optimal states of $R_G(\rho)$ which derived in [6] satisfies the condition of $d(\rho)$; the optimal state of $R_G(\rho)$ is also an optimal state of $d(\rho)$. The last equality is found in [6].

3.4 One-way local discrimination of bipartite systems

In this section, we give a formula for $t_\rightarrow(\rho)$ (which gives an upper bound of the one-way locally distinguishable set) for bipartite states, and show that $t_\rightarrow(\rho)$ is equal to the Schmidt rank of a state.

As a preparation for calculating $t_\rightarrow(\rho)$, we prove several lemmas about one-way LOCC POVM elements. First, we see the fact that there are several equivalent representations of the definition of one-way LOCC POVM elements. We start from the following three representations which we can see immediately from the definition; that is, in a bipartite system $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, if $T \in T_\rightarrow$, there exist sets of positive operators $\{M_i\}_i$ and $\{N^i_j\}_j$ such that

$$T = \sum_{ij} M_i \otimes N^i_j,$$

(3.8)

$$\sum_i M_i \leq I_B,$$ and $$\sum_j N^i_j \leq I_A,$$ where $\{M_i\}_i$ is the POVM of the local measurement on $\mathcal{H}_1$ and $\{N^i_j\}_j$ is the POVM of the local measurement on $\mathcal{H}_2$ depending on the first measurement result $i$. Other representations for a one-way POVM element are as follows:
Lemma 3.1

\[ T \in T_\rightarrow \]
\[ \iff \exists \{M_i\}_i \text{ and } \{N_i\}_i, \]
\[ s.t. \forall i, 0 \leq M_i, 0 \leq N_i \leq I_B, \sum_i M_i \leq I_A, \text{ and } T = \sum_i M_i \otimes N_i \]
\[ \iff \exists \{p_i\}_i, \{q_{ij}^i\}_ij, \{|u_i\}_i, \{|v_{ij}^i\}_ij \]
\[ s.t. \forall i, j, 0 \leq p_i, 0 \leq q_{ij}^i, \sum_i p_i \langle u_i | v_{ij}^i \rangle \leq I_A, \sum_j q_{ij}^i \langle v_{ij}^i | v_{ij}^i \rangle \leq I_B \]
\[ \text{and } T = \sum_{ij} p_i q_{ij}^i |u_i\rangle \langle u_i| \otimes |v_{ij}^i\rangle \langle v_{ij}^i| \]  

(3.10)

Proof  In Eq.(3.8), by redefining \( N_{i,j} \) as
\[ N_{i,j} = \sum_j N_{i,j} \]
we derive the first representation (3.9). For the second representation (3.10), by making spectral-decomposition of \( M_i \) and \( N_{i,j} \) as
\[ M_i = \sum_k p_{ik} |u_{ik}\rangle \langle u_{ik}| \]
\[ N_{i,j} = \sum_l q_{ij,l} |v_{ij,l}\rangle \langle v_{ij,l}| \]
then, we have
\[ T = \sum_{ijkl} p_{ik} q_{ij,l} |u_{ik}\rangle \langle u_{ik}| \otimes |v_{ij,l}\rangle \langle v_{ij,l}| \]
Reframing \( ik \) and \( jl \) to \( i \) and \( j \), respectively, we derive the second representation. \( \square \)

In the following proofs, we choose suitable one from the above three representations of one-way LOCC elements depending on the situations.

The following is important to calculate \( t_\rightarrow (\rho) \).

Lemma 3.2 A one-way LOCC POVM element \( T = \sum_{ij} M_i \otimes N_{i,j} \in T_\rightarrow \) satisfies \( \text{Tr} \rho T = 1 \) if and only if \( \text{Tr}(\rho_A \sum_i M_i) = 1 \) and \( \text{Tr}(\rho_{B,M_i} \sum_j N_{i,j}) = 1 \)
for all \( i \), where \( \rho_A \stackrel{\text{def}}{=} \text{Tr}_B \rho \) and \( \rho_{B,M_i} \stackrel{\text{def}}{=} \text{Tr}_A \rho M_i \otimes I_B / \text{Tr} \rho M_i \otimes I_B \).

Proof
\[ \text{Tr} \rho T = \sum_{ij} \text{Tr} \rho M_i \otimes N_{i,j} = \sum_{ij} \text{Tr} \{(\text{Tr}_A \rho (M_i \otimes I_B)) N_{i,j}\} \]
\[ = \sum_i \text{Tr} \rho_A M_i \cdot \text{Tr} \rho_{B,M_i}(\sum_j N_{i,j}) = 1. \]

Since \( \sum_i \text{Tr} \rho_A M_i \leq 1 \) and \( \text{Tr} \rho_{B,M_i}(\sum_j N_{i,j}) \leq 1 \) for all \( i \), we derive \( \sum_i \text{Tr} \rho_A M_i = 1 \) and \( \text{Tr} \rho_{B,M_i}(\sum_j N_{i,j}) = 1 \). The opposite direction is trivial. \( \square \)
The above lemma means that to detect a state perfectly, we need to detect the reduced density operators of a local system $A$, $\rho_A$, and that of a local system $B$, $\rho_{B,M}$, perfectly in each step.

For bipartite states, we can further derive the following Theorem for $t_-$ by means of Lemma 3.1 and Lemma 3.2.

**Theorem 3.1**

\[
t_-(\rho) = \min_{(p_i, |u_i\rangle)} \left\{ \sum_i p_i \text{rank}(\rho_{B, |u_i\rangle}) \mid \forall i, 0 \leq p_i, \sum_i p_i |u_i\rangle \langle u_i| \leq I_A, \text{ and } \text{Tr}_{A}(\sum_i p_i |u_i\rangle \langle u_i|) = 1 \right\},
\]

where $\rho_A = \text{Tr}_B \rho$ and $\rho_{B, |u_i\rangle} \overset{\text{def}}{=} \text{Tr}_A \rho |u_i\rangle \langle u_i| \otimes I_B$.

**Proof** Suppose $T \in T_-$ and $\text{Tr} \rho T = 1$. Then, from Lemma 3.1, without losing generality, we can write $T$ as $T = \sum_i p_i |u_i\rangle \langle u_i| \otimes N^i$, where $0 \leq \sum_i p_i |u_i\rangle \langle u_i| \leq I_A$ and $0 \leq N^i \leq I_B$. Then, from Lemma 3.2, $\text{Tr}_{A} \sum_i p_i |u_i\rangle \langle u_i| = 1$, and for all $i$, $\text{Tr}_{B, |u_i\rangle} N^i = 1$.

Suppose the state $\rho_{B, |u_i\rangle}$ have a spectral decomposition form $\rho_{B, |u_i\rangle} = \sum_l q_l |\phi_l\rangle \langle \phi_l|$, where $q_l \geq 0$ and $\sum_l q_l = 1$. Then from $\text{Tr}_{B, |u_i\rangle} N^i \leq 1$, we have $\sum_l q_l |\phi_l\rangle N^i \langle \phi_l| = 1$. Since $\sum_l q_l = 1$ and $N^i = 1$, the above equality holds only if $N^i \geq \{\rho_{B, |u_i\rangle} > 0\}$, where $\{\rho_{B, |u_i\rangle} > 0\}$ is a projection onto the support of $\rho_{B, |u_i\rangle}$. Therefore, we have

\[
T = \sum_i p_i |u_i\rangle \langle u_i| \otimes N^i \geq \sum_i p_i |u_i\rangle \langle u_i| \otimes \{\rho_{B, |u_i\rangle} > 0\} \quad (3.11)
\]

By taking the trace, we obtain

\[
\text{Tr} T \geq \sum_i p_i \text{rank}(\rho_{B, |u_i\rangle}).
\]

From this theorem, we can easily derive the following theorem for pure states.
**Corollary 3.1** For a bipartite pure state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$,

$$t_- (|\Psi\rangle) = \text{rank} \text{Tr}_B (|\Psi\rangle \langle \Psi|)$$  \hfill (3.12)

**Proof** Since $\rho_{B,|u_i\rangle}$ is a pure state, $\text{rank} \rho_{B,|u_i\rangle} = 1$. Then,

$$t_- (|\Psi\rangle) = \left\{ \sum_i p_i \mid \forall i, 0 \leq p_i, \sum_i p_i |u_i\rangle \langle u_i| \leq I_A, \text{ and } \text{Tr}_A (\sum_i p_i |u_i\rangle \langle u_i|) = 1 \right\}$$

Since $\text{Tr}_A (\sum_i p_i |u_i\rangle \langle u_i|) = 1$, $\sum_i p_i |u_i\rangle \langle u_i| \geq \{ \rho_A > 0 \}$. Therefore,

$$\sum_i p_i = \text{Tr} \sum_i p_i |u_i\rangle \langle u_i| \geq \text{rank} \rho_A = \text{rank} \text{Tr}_B (|\Psi\rangle \langle \Psi|).$$

That is, we can conclude that $t_- (|\Psi\rangle) = \text{rank} \text{Tr}_B (|\Psi\rangle \langle \Psi|).$ \hfill $\Box$

Therefore, for a bipartite pure state, $t_- (|\Psi\rangle)$ is equal to the Schmidt rank (see subsection 1.4.8). The optimal way of detecting a state is the following. Both Alice and Bob measure their local states in the Schmidt basis, (of course, they only need to detect the support of the local states) and then Alice informs her measurement result to Bob. Therefore, in this case, local discriminability of the one-way LOCC depends only on the Schmidt rank and does not depend on the Schmidt coefficients. In other words, by only one-way LOCC, we cannot utilize the information of the Schmidt coefficients to discriminate a given state.

By comparing the previous result of $t_{\text{sep}}$ of Eq.(3.7) and Corollary 3.1, we can easily see that if $|\Psi\rangle$ is not a maximally entangled state nor a product state, then, the strict inequality $t_{\text{sep}} (|\Psi\rangle) < t_- (|\Psi\rangle)$ holds. However, since the original local discriminability for separable operation is given by $t_{\text{sep}}$, and $t_{\text{sep}} \leq t_{\text{sep}}$, we can not conclude whether there is a gap between the one-way local discriminability and the separable local discriminability from this argument.
3.5 One-way local discrimination of Tri-partite systems

On the basis of the results of bipartite systems, we analyze tri-partite systems. First, we show that the optimization for tri-partite pure states can be reduced to the optimization for the bipartite mixed states represented by reduced density matrixes. Then, in the case of three qubits, we calculate $t_\rightarrow(|\Psi\rangle)$ for a special kind of tri-partite states, called W-states [9].

For a tri-partite pure state, we can derive the following Lemma.

**Lemma 3.3** For a tri-partite pure state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, $t_\rightarrow(|\Psi\rangle) = t_\rightarrow(\text{Tr}_C |\Psi\rangle \langle \Psi|)$

**Proof** Suppose $T = \sum_{ij} p_i |u_i\rangle \langle u_i| \otimes q_j^i |v_j^i\rangle \langle v_j^i| \otimes L^ij$. Since $\text{Tr}T |\Psi\rangle \langle \Psi| = 1$, by substituting the systems $AB$ and $C$ in this lemma for the systems $A$ and $B$ in Lemma 3.2, respectively, $\text{Tr}\rho_{AB} \sum_{ij} p_i |u_i\rangle \otimes q_j^i |v_j^i\rangle \langle v_j^i| = 1$ and $\text{Tr}\rho_{C,|u_i\rangle \otimes |v_j^i\rangle} L_{ij} = 1$. Thus, $L_{ij} \geq \{\rho_{C,|u_i\rangle \otimes |v_j^i\rangle} > 0\}$. Then, we have

$$T = \sum_{ij} p_i |u_i\rangle \otimes q_j^i |v_j^i\rangle \langle v_j^i| \otimes L^ij \geq \sum_{ij} p_i |u_i\rangle \otimes q_j^i |v_j^i\rangle \langle v_j^i| \otimes \{\rho_{C,|u_i\rangle \otimes |v_j^i\rangle} > 0\}.$$ 

Therefore, $\text{Tr}T \geq \sum_{ij} p_i q_j^i$, where $p_i$ and $q_j^i$ satisfy $\text{Tr}\rho_{AB} \sum_{ij} p_i |u_i\rangle \otimes q_j^i |v_j^i\rangle \langle v_j^i| = 1$. Thus, we obtain the following equality:

$$t_\rightarrow(|\Psi\rangle) = \min_{p_i, q_j^i, |u_i\rangle, |v_j^i\rangle} \left\{ \text{Tr}T \right\} \text{ s.t. } \begin{cases} \forall i,j, \ 0 \leq p_i, 0 \leq q_j^i, \\ T = \sum_{ij} p_i |u_i\rangle \langle u_i| \leq I_A, \sum_j q_j^i |v_j^i\rangle \langle v_j^i| \leq I_B, \text{ and } \text{Tr}\rho_{AB} T = 1 \end{cases}.$$

By Lemma 3.2, we have $t_\rightarrow(|\Psi\rangle) = t_\rightarrow(\text{Tr}_C |\Psi\rangle \langle \Psi|)$.

This lemma means that the value of $t_\rightarrow$ of a tri-partite pure state is the same as the value of $t_\rightarrow$ of its bipartite reduced density matrix.
By means of the above lemma, we calculate $t_-$ for W-type states in three-qubit systems.

**Theorem 3.2** In $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$, for a tri-partite pure state $|\Psi\rangle \equiv |001\rangle \otimes |010\rangle \otimes |100\rangle$ with $|\lambda_1|^2 + |\lambda_2|^2 + |\lambda_3|^2 = 1$, we have $t_- (|\Psi\rangle) = 3$. Here, we assume $|\lambda_i| > 0$ for all $i$.

**Proof** By Lemma 3.3, $t_- (|\Psi\rangle) = t_- (\rho_{AB})$, where $\rho_{AB}$ is the reduced density matrix of $|\Psi\rangle$ as follows,

$$
\rho_{AB} \overset{\text{def}}{=} \text{Tr}_c |\Psi\rangle \langle \Psi|,
$$

where $\lambda_2^\ast$ and $\lambda_3^\ast$ are the complex conjugate of $\lambda_2$ and $\lambda_3$, respectively. We minimize $\sum_i p_i \text{rank} \rho_{B,u_i}$ under the restrictions of $p_i > 0$, $\sum_i p_i \langle u_i | \rho_A | u_i \rangle = 1$, and $\sum_i p_i \leq \dim \mathcal{H}_A$. Since $\rho_A = (|\lambda_1|^2 + |\lambda_2|^2) |00\rangle \langle 00| + |\lambda_3|^2 |11\rangle \langle 11|$, \(\sum_i p_i \langle u_i | \rho_A | u_i \rangle = 1\) is reduced to $\sum_i p_i |u_i\rangle \langle u_i| \geq \{\rho_A > 0\} = I_A$. That is, we have $\sum_i p_i |u_i\rangle \langle u_i| = I_A$. Therefore, we obtain

$$
t_- (|\Psi\rangle) = \min \{ \sum_i p_i \text{rank} \rho_{B,u_i} | \sum_i p_i |u_i\rangle \langle u_i| = I_A, p_i > 0 \}. \quad (3.13)
$$

Suppose $|u_i\rangle = \alpha |0\rangle + \beta |1\rangle$, then $\rho_{B,u_i} = (|\lambda_1|^2 |\alpha|^2 + |\lambda_3|^2 |\beta|^2) |0\rangle \langle 0| + \lambda_2 \lambda_3 \alpha \beta |1\rangle \langle 0| + \lambda_2 \lambda_3 \alpha \beta |0\rangle \langle 1| + |\lambda_2|^2 |\alpha|^2 |1\rangle \langle 1|$, and we have $\det \rho_{B,u_i} = \lambda_2^2 |\alpha|^4$. That is, $\det \rho_{B,u_i} = 0$ if and only if $\alpha = 0, |\beta| = 1$. Therefore, if $|u_i\rangle = |1\rangle$, $\text{rank} \rho_{B,u_i} = 1$, otherwise $\text{rank} \rho_{B,u_i} = 2$.

At first, we assume $|u_i\rangle \neq |1\rangle$ for all $i$. Then, since $\text{Tr} (\sum_i p_i |u_i\rangle \langle u_i|) = \sum_i p_i = \text{Tr} I_A = 2$, we have $\sum_i p_i \text{rank} \rho_{B,u_i} = 2 \sum_i p_i = 4$. On the other hand, suppose that there exists $i_0$ such that $|u_{i_0}\rangle = |1\rangle$. In this case, we can assume $|u_0\rangle = |1\rangle$ and $|u_i\rangle \neq |1\rangle$ for all $i \leq 1$. Since $\sum_{i=0}^n p_i = 2$, we obtain $\sum_{i=1}^n p_i = 2 - p_0$. Therefore,

$$
\sum_i p_i \text{rank} \rho_{B,u_i} = p_0 + 2 \sum_{i=1}^n p_i \geq 3,
$$

72
where in the third line, we use $p_0 \leq 1$. The condition of the equality is $p_0 = p_1 = 1$, $|u_0\rangle = |0\rangle$, $|u_1\rangle = |1\rangle$. Finally, we conclude $t_- (|\Psi\rangle) = 3$. □

By using the above theorem, we can show that the number of one-way LOCC distinguishable states $N_- \leq d/t_- (|\Psi\rangle) = 8/3 < 3$. Therefore, we can not distinguish any set of three $W$-type states by one-way LOCC in $C^2 \otimes C^2 \otimes C^2$.

### 3.6 Two-way LOCC

In this section, by deriving an upper bound of the two-way local discriminability $t_\leftrightarrow$ in two-qubit systems by presenting a concrete protocol, we show that there is a gap between the one-way and two-way local discriminabilities at least in two-qubit systems.

We consider the following example of a three-step two-way LOCC POVM which can detect a bipartite entangled state $|\Psi\rangle = \sqrt{\lambda} |00\rangle + \sqrt{1-\lambda} |11\rangle$ in a two-qubits system. Suppose that Alice has the system A, and Bob has the system B. Then, first, Alice measures her system by POVM $\{M, I_A - M\}$, where $M \overset{\text{def}}{=} |0\rangle \langle 0| + (1 - \delta) |1\rangle \langle 1|$. In the case where Alice obtains the result $I_A - M$, the states after the measurement is $|1\rangle \otimes |1\rangle$. Then, they detect the state after Alice’s first measurement by measuring each local system in computational basis, or arbitrary chosen basis (the definition of computational basis is in Subsection 1.4.8 of Chapter 1). If the result of both are $|1\rangle$, they conclude the first state is $|\Psi\rangle$, and otherwise they conclude not $|\Psi\rangle$. In the case where Alice obtains the result $M$ in the first step, then Bob measures his system in the basis $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle \overset{\text{def}}{=} \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$. In the last step, if Bob’s result is $|+\rangle$, Alice measures her system in $\{|\uparrow\rangle, |\downarrow\rangle\}$, and if Bob’s result is $|-\rangle$, she measures her system in $\{|\overline{\uparrow}\rangle, |\overline{\downarrow}\rangle\}$, where $|\uparrow\rangle \overset{\text{def}}{=} \sqrt{p} |0\rangle + \sqrt{1 - p} |1\rangle$, $|\downarrow\rangle \overset{\text{def}}{=} \sqrt{1 - p} |0\rangle - \sqrt{p} |1\rangle$, $|\overline{\uparrow}\rangle \overset{\text{def}}{=} \sqrt{p} |0\rangle - \sqrt{1 - p} |1\rangle$, $|\overline{\downarrow}\rangle \overset{\text{def}}{=} \sqrt{1 - p} |0\rangle + \sqrt{p} |1\rangle$, and $p$ is defined by $\frac{\sqrt{M \otimes I_B |\Psi\rangle}}{\|M \otimes I_B |\Psi\rangle\|} = \sqrt{p} |00\rangle +$
$|\sqrt{1-p}11\rangle$. If she obtains $|\uparrow\rangle$ or $|\bar{\uparrow}\rangle$, they conclude that the first state is $|\Psi\rangle$ and otherwise they conclude it is not $|\Psi\rangle$.

We define a POVM element $T(\lambda, \delta)$ in two-qubit systems $\mathcal{H}_A \otimes \mathcal{H}_B$ by the above measurement process. $T(\lambda, \delta)$ can be written by,

$$T(\lambda, \delta) = (\sqrt{M} \otimes I_B)(|\uparrow\rangle \langle \uparrow| \otimes |+\rangle \langle +|)(\sqrt{M} \otimes I_B))$$

$$+ (\sqrt{M} \otimes I_B)(|\bar{\uparrow}\rangle \langle \bar{\uparrow}| \otimes |-\rangle \langle -|)(\sqrt{M} \otimes I_B))$$

$$+ (I_A - M) \otimes |1\rangle \langle 1|.$$ (3.14)

The POVM element $T(\lambda, \delta)$ can be obviously implemented by the three-step two-way LOCC, and moreover, since Bob measures in the mutually unbiased basis set $\{|+\rangle, |-\rangle\}$ [10], this measurement process can not be reduced to two steps. Therefore, it cannot be one-way LOCC.

We can calculate $\langle \Psi | T | \Psi \rangle$ as follows:

$$\langle \Psi | T | \Psi \rangle = (\langle \Psi | \sqrt{M} \otimes I_B)(|\uparrow\rangle \langle \uparrow| \otimes |+\rangle \langle +| + |\bar{\uparrow}\rangle \langle \bar{\uparrow}| \otimes |-\rangle \langle -|)(\sqrt{M} \otimes I_B |\Psi\rangle))$$

$$+ \langle \Psi | ((I_A - M) \otimes |1\rangle \langle 1|) |\Psi\rangle$$

$$= (1 - \delta + \delta \lambda)(\sqrt{I} |00\rangle + \sqrt{1-p} |11\rangle)$$

$$|\uparrow\rangle \langle \uparrow| \otimes |+\rangle \langle +| + |\bar{\uparrow}\rangle \langle \bar{\uparrow}| \otimes |-\rangle \langle -|)(\sqrt{I} |00\rangle + \sqrt{1-p} |11\rangle|$$

$$+ (1 - \lambda)\delta$$

$$= (1 - \delta + \delta \lambda) + (1 - \lambda)\delta$$

$$= 1.$$ (3.15)

Thus, the detection of $|\Psi\rangle$ by the above measurement process is perfectly possible. Therefore, we derive a lower bound of $t_\rightarrow(|\Psi\rangle)$ by means of $T(\lambda, \delta)$, and we show that $t_\rightarrow(|\Psi\rangle)$ is actually less than $t_\rightarrow(|\Psi\rangle)$.

We show the following upper bound of $t_\rightarrow(|\Psi\rangle)$ by means of $T(\lambda, \delta)$.

**Lemma 3.4** In two-qubit systems,

$$t_\rightarrow(|\Psi\rangle) \leq 2 - \frac{(1 - \sqrt{2\lambda})^2}{1 - \lambda}$$ (3.15)

74
Proof Thanks to the Schmidt decompositions, we can represent a general pure bipartite state in the form $|\Psi\rangle = \sqrt{\lambda} |00\rangle + \sqrt{1-\lambda} |11\rangle$ with $0 < \lambda \leq 1/2$, by introducing suitable basis set for each qubit, without losing generality.

From Eq.(3.14), we derive

$$\text{Tr}T(\lambda, \delta) = \text{Tr}(M |\uparrow\rangle \langle \uparrow|) \cdot \text{Tr}(|+\rangle \langle +|) + \text{Tr}(M |\uparrow\rangle \langle \uparrow|) \cdot \text{Tr}(-) \langle -|) + \text{Tr}(I - M)$$

$$= \text{Tr}M(|\uparrow\rangle \langle \uparrow| + |\uparrow\rangle \langle \uparrow|) + \text{Tr}(I - M)$$

$$= 2\text{Tr}M^2\rho_A + \text{Tr}(I_A - M)$$

$$= 2 - \frac{\delta\{(1 - 2\lambda) - (1 - \lambda)\delta\}}{1 - \delta(1 - \lambda)}$$

where $\rho_A \overset{\text{def}}{=} \text{Tr}_B |\Psi\rangle \langle \Psi|$, we used $|\uparrow\rangle \langle \uparrow| + |\uparrow\rangle \langle \uparrow|$ in the third line, and we used $\text{Tr}M\rho_A = 1 - \delta + \lambda\delta$ and $\text{Tr}M^2\rho_A = 1 - 2\delta(1 - \lambda) + \delta^2(1 - \lambda)$ in the fourth line. Then, since

$$\frac{\partial\text{Tr}T(\lambda, \delta)}{\partial\delta} = -\{(1 - \lambda)\delta - (1 - \sqrt{2\lambda})\}\{(1 - \lambda)\delta - (1 + \sqrt{2\lambda})\},$$

(3.16)

for fixed $0 < \lambda < 1/2$, the value $\text{Tr}T(\lambda, \delta)$ attains its minimum when $\delta = \frac{1 - \sqrt{2\lambda}}{1 - \lambda}$.

Therefore, we derive

$$\min_{0 \leq \delta \leq 1} \text{Tr}T(\lambda, \delta) = 2 - \frac{(1 - \sqrt{2\lambda})^2}{1 - \lambda}.$$  

(3.17)

Since $T(\lambda, \delta)$ is a two-way LOCC POVM which detects $|\Psi\rangle$ perfectly, $t_{\leftrightarrow}(|\Psi\rangle) \leq \min \text{Tr}T(\lambda, \delta)$. Therefore, we derive the inequality (3.15).

□

Since $2 - \frac{(1 - \sqrt{2\lambda})^2}{1 - \lambda} < 2$ for all $0 < \lambda < 1/2$, for all non-maximally pure entangled states in two-qubit systems,

$$t_{\leftrightarrow}(|\Psi\rangle) \leq 2 - \frac{(1 - \sqrt{2\lambda})^2}{1 - \lambda} < 2 = t_{\leftrightarrow}(|\Psi\rangle).$$  

(3.18)

On the other hand, from Eq.(3.7) and Eq.(3.6), we have a lower bound of $t_{\leftrightarrow}(|\Psi\rangle)$ as

$$1 + 2\sqrt{\lambda(1 - \lambda)} = t_{\text{sep}}(|\Psi\rangle) \leq t_{\text{sep}}(|\Psi\rangle) \leq t_{\leftrightarrow}(|\Psi\rangle).$$
We present the graph of these bounds in Fig. 3.1. From this figure, we can see that there is a big gap between $t_{\rightarrow}(|\Psi\rangle)$ and $t_{\leftrightarrow}(|\Psi\rangle)$ and the difference between $t_{\rightarrow}(|\Psi\rangle)$ and $t_{\text{sep}}(|\Psi\rangle)$ is (if the difference exists) relatively small. Thus, for any non-maximally entangled pure states, there is a gap between the one-way and two-way local distinguishabilities at least for two-qubit systems in terms of $t_{\rightarrow}(|\Psi\rangle)$. That is, the two-way classical communication remarkably improves the local discriminability compared to the local discrimination by the one-way classical communication.

3.7 Conclusion

In this chapter, in bipartite systems, we have given a formula for $t_{\rightarrow}$, which is an appropriate measure to local discriminability. We also have shown that $t_{\rightarrow}(|\Psi\rangle)$ for tri-partite states $|\Psi\rangle$ can be reduced to $t_{\rightarrow}$ of bipartite mixed states $\text{Tr}_{C} |\Psi\rangle \langle \Psi|$ and we have shown that $t_{\rightarrow}(|\Psi\rangle) = 3$ for W-type states. Since $t_{\rightarrow}(|\Psi\rangle)$ gives an upper bound of the size of the locally distinguishable set, we show that a set of three W-type states (three different states which
can be transformed W-type states by local unitary transformations) can not be discriminated by the one-way LOCC. Finally, we investigate the two-way LOCC and show that $t_{\rightarrow}(|\Psi\rangle)$ is strictly less than $t_{\leftarrow}(|\Psi\rangle)$ for non-maximally entangled states in two-qubit systems. This implies that there is a gap between the one-way and two-way local distinguishabilities, although our result of the two-way LOCC is a very restricted case compared to the separable and the one-way LOCC.
Bibliography


Chapter 4

Local copying and local discrimination as a study for non-locality of a set of states

Abstract

In this chapter, we focus on the LOCC copying and its relation to the LOCC discrimination. I analyze the LOCC copying for an orthogonal set of maximally entangled states and show that for such a set in a system with prime-dimensional local systems, a necessary and sufficient condition of the perfect LOCC copying is that the set is a simultaneous Schmidt diagonalizable subset of canonical Bell bases. Moreover, I show that for a set of maximally entangled states, the LOCC copying is strictly more difficult than the LOCC discrimination by using the fact that a simultaneous Schmidt diagonal set of states is always locally distinguishable.

4.1 Introduction

As mentioned in Chapter 1, the study of entanglement theory is the study of convertibility between entangled states under locality restrictions for operations, (e.g. LOCC, separable operations, and PPT (positive partial trans-
pose) operations \([1, 2, 3, 4, 5, 6, 7]\)). On the other hand, there are problems of non-locality which cannot be explained by one-to-one convertibility of states. One of such problems is the “\textit{local discrimination}” (a problem to discriminate an unknown state by only LOCC) \([8, 9, 10, 11, 12, 13, 14]\). The starting point was the discovery of a product basis which cannot be perfectly discriminated by LOCC \([15]\) by Bennett \textit{et al.}, namely “\textit{Non-locality without entanglement}”. In \([15]\), they proposed a locally indistinguishable product basis and regarded its impossibility for perfect discrimination under LOCC as its non-locality.

The study by Bennett \textit{et al.} suggests a new kind of non-locality, namely \textit{non-locality of a set of states}. At first, in analogy to the non-locality discussion in their paper, we can expand the concept of non-locality as follows. If the local (LOCC) restriction causes difficulty for a task concerning a set of states, \textit{e.g.} discrimination, copying etc., then we consider that this set has non-locality and regard the degree of this difficulty as \textit{non-locality of the set}. This concept of non-locality is not unnatural, since it is consistent with the conventional entanglement theory because of the following reasons. In entanglement theory, the entanglement cost \([2]\) is one of the most established measures of entanglement, and can be regarded as a kind of difficulty of a task, \textit{i.e.} the difficulty of the entanglement dilution \([2]\). Moreover, if we consider the task to approximate a given state by separable states, we derive the relative entropy of entanglement \([6]\) by measuring this difficulty in terms of the accuracy of the approximation, using the relative entropy. These can be regarded as the degrees of difficulty of tasks with the local restrictions.

Indeed, the local discrimination can be regarded as a task for a set of states with the local restrictions, because these problems are usually treated based on a set of candidates of unknown states. Hence, we can measure the “\textit{non-locality of a set of states}” by the degree of difficulty of the local discrimination. We should note that this kind of difficulty cannot be often
characterized only by entanglement of states of the given set. A typical example is the impossibility of the local discrimination of the product basis of Bennett et al. mentioned above. In addition to the local discrimination, similar non-locality also appears commonly in various different fields of quantum information, e.g. quantum capacity, quantum estimation, etc. [16, 17].

Recently, a similar problem to the local discrimination, namely “local copying”, was also raised [18, 19], as a problem to study cloning of unknown entangled states under the LOCC restriction with only the minimum entanglement resource. Local copying is also defined for a set of states with the local restriction; therefore we can consider the non-locality of a set of states based on the local copying. Moreover, this non-locality cannot be also characterized only by the entanglement of states of the given set [18].

In this chapter, as a study of the non-locality of a set of states, we focus on the local copying and the local discrimination. Specifically, we concentrate on a set of orthogonal maximally entangled states in a prime-dimensional system for mathematical simplicity and investigate the relation between their local copiability and local distinguishability. As a result, we completely characterize the local copiability of such a set; that is, we prove that such a set is locally copiable if and only if it has a canonical Bell form and is simultaneously Schmidt decomposable. Using this result, we prove the following two facts. First, the maximal size of locally copiable sets is equal to the dimension of the local space which is equal to the maximal size of local distinguishable sets. Second, we also show that if such a set is locally copiable, then it is locally distinguishable by one-way communications. Thus, in this case, the local copying is strictly more difficult than the one-way local discrimination. The relation of the local copiability and the local discriminability is summarized in Fig. 4.1. From this relation, we derive the conclusion related to the non-locality of a set concerning the local copying and the local discrimination: A simultaneous Schmidt decomposable state does not have non-locality
Figure 4.1: The hierarchy of the non-locality of sets of maximally entangled states. In this figure, LD, SSD, and c.c. mean locally distinguishable, simultaneously Schmidt decomposable, and classical communication, respectively.
of a set of states beyond individual entanglement due to the local discrimination, since it is locally distinguishable. However, even if a set is simultaneous Schmidt decomposable, if such a set does not have a canonical Bell form, such a set still has non-locality due to the local copying.

Although we mainly concentrate on the aspect of the local copying and the local discrimination as the study of non-locality in this chapter, the local copying and the local discrimination themselves are worth to investigate as basic protocols of quantum information processing with two parties. In the last part of this chapter, we show that there are many important relation between our local copying protocol and other quantum information protocols. These results give many other interpretations for the local copying.

This chapter is organized as follows. In Section 4.2, in preparation of our analysis, we review a necessary and sufficient condition for a locally copiable set, which is the main result of the paper [18]. In Section 4.3, we give an example of a locally copiable set of $D$ maximally entangled states, and then, prove that, in a prime-dimensional local system, the above example is the only case where local copying is possible for maximally entangled states. In Section 4.4, we discuss the relation between the local copying and the local discrimination by means of simultaneous Schmidt decomposition. In Section 4.5, we present other protocols which are strongly related to our theory of local copying, i.e., the channel copying, the entanglement distillation protocol, the error correction, and the quantum key distribution. And then, we extend our results of local copying to these protocols. Finally, we summarize and discuss our results in Section 4.6.

### 4.2 The local copying problem

In this section, in preparation for our analysis, we introduce formulation and known results of the local copying from [18].
Many researchers treated approximated cloning, for example, universal cloning \cite{22}, asymmetric cloning \cite{23}, tele-cloning \cite{24}. This is because the perfect cloning, or copying, is impossible without prior knowledge (no-cloning theorem) \cite{25}. That is, the possibility of copying depends on prior knowledge about the state to be copied, or, in other words, a set of candidates for the unknown target state, where we call the state to be copied the ‘target’ state. If we know that an unknown state to be copied is contained by a set of orthogonal states, which is called the copy set, we can copy the given state. However, if our operation is restricted to local operations and classical communication (LOCC)\cite{1}, we cannot necessarily copy the given quantum state even with the above orthogonal assumption, perfectly. Thus, it is interesting from both viewpoints of entanglement theory and cloning theory to extend cloning problem to the bipartite entangled states. This is the original motivation of cloning problems with the LOCC restriction \cite{18, 19}.

Recently, F. Anselmi \textit{et al.} \cite{18} focused on the perfect cloning of bipartite systems under the following assumptions;

1. Our operation is restricted to LOCC.

2. It is known that the unknown state to be copied is contained in a set of known orthogonal entangled states (the copy set).

3. A known entangled state of the same size is shared.

They called this problem local copying and they characterized copied sets which we can locally copy in special cases. In the following, for simplicity, we say that the set is locally copiable if local copying is possible with the prior knowledge as to which set the given state belongs to.

The problem of local copying can be phrased as follows. We assume two players at a long distance, namely Alice and Bob in this protocol. They have two quantum systems $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, each of which is composed
by two $D$-dimensional systems, i.e., the systems $\mathcal{H}_A$ and $\mathcal{H}_B$ are described by $\mathcal{H}_A = \mathcal{H}_1 \otimes \mathcal{H}_3$, $\mathcal{H}_B = \mathcal{H}_2 \otimes \mathcal{H}_4$. Since the dimensions of these systems is the same, by defining “fixed” computational bases in each system, we can mathematically identify these systems $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$, and $\mathcal{H}_4$. In our problem, they try to copy an unknown state $|\Psi\rangle$ on the initial system $\mathcal{H}_1 \otimes \mathcal{H}_2$ to the target system $\mathcal{H}_3 \otimes \mathcal{H}_4$ with the prior knowledge that $|\Psi\rangle$ belongs to the copy set $\{|\Psi_j\rangle\}_{j=0}^{N-1}$. Moreover, we assume that they implement copying only by LOCC between them. Since LOCC operations do not increase the entanglement of the whole states, they cannot copy any entangled states by LOCC without any entanglement resources. Thus, we also assume that they share a blank entangled state $|b\rangle$ in the target system $\mathcal{H}_3 \otimes \mathcal{H}_4$; see Fig. 4.2. Therefore, a set of states $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ is called locally copiable with a blank state $|b^{34}\rangle \in \mathcal{H}_3 \otimes \mathcal{H}_4$, if there exists a LOCC operation $\Lambda$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ which satisfies the following condition for all $j = 0, \cdots , N - 1$:

$$
\Lambda(|\Psi_j^{12}\rangle \otimes |b^{34}\rangle \langle \Psi_j^{12}| \otimes \langle b^{34}|) = |\Psi_j^{12}\rangle \otimes |\Psi_j^{34}\rangle \langle \Psi_j^{12}| \otimes \langle \Psi_j^{34}|,
$$

(4.1)

where $|\Psi^{12}\rangle$ and $|\Psi^{34}\rangle$ are the state $|\Psi\rangle$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\mathcal{H}_3 \otimes \mathcal{H}_4$, respectively.
and we treat $\mathcal{H}_A = \mathcal{H}_1 \otimes \mathcal{H}_3$ and $\mathcal{H}_B = \mathcal{H}_2 \otimes \mathcal{H}_4$ as local spaces with respect to the LOCC operation $\Lambda$. In this chapter, we keep using the above notation of systems and states in the local copying problem.

Even in the most simple case $N = 1$, it is very hard to completely characterize which state can be used as a blank state $|b\rangle$ for a given state $|\Psi_0\rangle$ in Eq. (4.1). This is because the transformation in Eq. (4.1) is the entanglement transformation from $|b\rangle$ to $|\Psi_0\rangle$ using $|\Psi_0\rangle$ as an entanglement catalysis [3]; that is, even if $|b\rangle$ cannot be transformed to $|\Psi_0\rangle$, it may be possible for $|b\rangle$ to be transformed to $|\Psi_0\rangle$ with the help of the catalysis $|\Psi_0\rangle$. It is very hard to characterize this catalytic transformation (for detail see Sec. II B of Ref. [18]). Thus, it is very hard to derive a necessary and sufficient condition for a general setting of local copying. It is well known that no maximally entangled state works as an entanglement catalysis [3]. In this chapter, to avoid the difficulty of the entanglement catalysis, we restrict our analysis to the case where all of $|\Psi_j\rangle$ are maximally entangled states, where a maximally entangled state is defined as a state whose Schmidt coefficients, that is, eigenvalues of the reduced density matrix, are all $1/D$.

If we restrict a copy set $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ to a set of maximally entangled states, we can simplify the problem setting as follows. First, because of the monotonicity of entanglement under LOCC, a blank state $|b\rangle$ also must be maximally entangled and we can always choose this blank state $|b\rangle = |\Psi_0\rangle$. This is because, before we implement an operation $\Lambda$ in Eq. (4.1), we can always perform a local unitary operation for $|b\rangle$ to change $|b\rangle$ to an arbitrary maximally entangled state. Therefore, by the assumption $|b\rangle = |\Psi_0\rangle$ the problem does not lose generality. Second, if a set of states $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ is locally copiable, we can easily see that the set of states $\{U \otimes V |\Psi_j\rangle\}_{j=0}^{N-1}$ is also locally copiable for any local unitary operation $U \otimes V$. In other words,\footnote{The paper [18] showed that if a copy set $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ has at least one maximally entangled state and is locally copiable, then all of states $|\Psi_j\rangle$ in the copied set must be maximally entangled.}

87
local copiability is invariant under the action of local unitary operations for a set of states. Thus, by means of this freedom of the action of local unitaries, we can fix $|b\rangle$ to one state in $\{ |\Psi_j\rangle \}_{j=0}^{N-1}$. Here, we choose the standard maximally entangled state as $|\Psi_0\rangle$, that is, $|\Psi_0\rangle = \frac{1}{\sqrt{D}} \sum_{i=0}^{D-1} |i\rangle \otimes |i\rangle$, where $\{ |i\rangle \}_{i=0}^{D-1}$ is a computational basis in each local space. Finally, all we need to consider is the following condition:

$$
\Lambda (|\Psi_{12}^j\rangle \otimes |\Psi_{34}^j\rangle \langle \Psi_{12}^j| \otimes \langle \Psi_{34}^j|) = |\Psi_{12}^j\rangle \otimes |\Psi_{34}^j\rangle \langle \Psi_{12}^j| \otimes \langle \Psi_{34}^j|, \tag{4.2}
$$

where $|\Psi_0\rangle$ is the standard maximally entangled state. In the following discussion, we always assume the above condition.

In paper [18], Anselmi et al. derived a necessary and sufficient condition for a specific locally copiable set (Lemma 4.1). They also completely characterized the local copiability of maximally entangled states in the case of $N = 2$. In the following, in preparation for our analysis, we shortly summarize Anselmi et al.’s results for local copying of maximally entangled states.

First, Anselmi et al. showed the following theorem.

**Theorem 4.1 (Anselmi et al.)** If a set of maximally entangled states $\{ |\Psi_j\rangle \}_{j=0}^{N-1}$ is locally copiable, then it is copiable only by local unitary transformation.

That is, when all $|\Psi_j\rangle$ are maximally entangled states, we always choose a local unitary operation as the copying operation $\Lambda$ in Eq. (4.1). In [18], by means of the above result, Anselmi et al. derived a necessary and sufficient condition of copiability for a set of maximally entangled states. Here, we do not present this necessary and sufficient condition in the same form they gave, but we present a modified version of their statement for the benefit of a detailed analysis later in Section 4.3.

**Lemma 4.1 (Anselmi et al.)** A set of maximally entangled states $\{ |\Psi_j\rangle \}_{j=0}^{N-1}$ is locally copiable if and only if there exists a unitary operator $A$ on $\mathcal{H}_1 \otimes \mathcal{H}_3$. 

88
and unitary operations \( \{U_j\}_{j=0}^{N-1} \) on \( \mathcal{H}_1 \) such that

\[
A(U_j \otimes I)A^\dagger = U_j \otimes U_j,
\]

and

\[
|\Psi_j\rangle \langle \Psi_j| = (U_j \otimes I)|\Psi_0\rangle \langle \Psi_0| (U_j^\dagger \otimes I).
\]

**Proof**

We first look at the necessary condition. Suppose that \( \{|\Psi_j\rangle\}_{j=0}^{N-1} \) is locally copiable. Then, by Theorem 4.1, there exists a local unitary operation \( A \otimes B \) and real numbers \( 0 \leq \theta_j < 2\pi \) such that, for all \( j \),

\[
A^{13} \otimes B^{24} |\Psi_j^{12}\rangle \otimes |\Psi_0^{34}\rangle = e^{i\theta_j} |\Psi_j^{12}\rangle \otimes |\Psi_j^{34}\rangle,
\]

where \( e^{i\theta_j} \) are phase factors. Then, we define \( A' \) and \( \theta'_j \) as \( A' \defeq e^{-i\theta_0} A \) and \( \theta'_j \defeq \theta_j - \theta_0 \), respectively. From Eq. (4.5), we derive

\[
A'^{13} \otimes B^{24} |\Psi_j^{12}\rangle \otimes |\Psi_0^{34}\rangle = e^{i\theta'_j} |\Psi_j^{12}\rangle \otimes |\Psi_j^{34}\rangle,
\]

where \( \theta'_0 = 0 \). Since all \( |\Psi_j\rangle \) are maximally entangled states, there exists a set of unitary operations \( \{U_j\}_{j=0}^{N-1} \) such that

\[
|\Psi_j\rangle = e^{-i\theta'_j} U_j \otimes I |\Psi_0\rangle,
\]

where we can choose \( U_0 = I \). Then, we can easily see \( |\Psi_j\rangle \langle \Psi_j| = U_j \otimes I |\Psi_0\rangle \langle \Psi_0| U_j^\dagger \otimes I \) and

\[
(A'^{13} \otimes B^{24})(U_j^1 \otimes I^{234}) |\Psi_0^{12}\rangle \otimes |\Psi_0^{34}\rangle = U_j^1 \otimes U_j^3 \otimes I^{24} |\Psi_0^{12}\rangle \otimes |\Psi_0^{34}\rangle.
\]

By using the symmetry of the standard maximally entangled state \( (I \otimes V |\Psi_0\rangle = V^T \otimes I |\Psi_0\rangle \) for all unitary \( V \), we derive

\[
\{A'^{13}(U_j^1 \otimes I^{3})B^{T13}\} \otimes I^{24} |\Psi_0^{12}\rangle \otimes |\Psi_0^{34}\rangle = U_j^1 \otimes U_j^3 \otimes I^{24} |\Psi_0^{12}\rangle \otimes |\Psi_0^{34}\rangle.
\]
where $B^T$ is transpose of $B$ in the computational basis. Then, by projecting Eq. (4.9) to a state in computational basis $|k^1\rangle \otimes |l^2\rangle \otimes |m^3\rangle \otimes |n^4\rangle$, we have
\[
\langle k^1 | \langle l^3 | A^{13}(U^1_j \otimes I^3)B^{T13} | m^1 \rangle \otimes |n^3\rangle = \langle k^1 | \langle l^3 | U^1_j \otimes U^3_j | m^1 \rangle \otimes |n^3\rangle.
\]
(4.10)
Since the above equation is valid for all states in the computational basis, we derive $A' B^T = U^1_j \otimes U^3_j$. By substituting $U^0_j = I$ in the above formula, we derive $A' B^T = U^1_j \otimes U^3_j$. Therefore, finally, we derive Eq. (4.3).

We now look at the sufficient condition. Suppose that Eq. (4.3) and Eq. (4.4) are valid. Then, by defining $B = A^*$, we can directly check Eq. (4.5).

Here, we remark on the following two facts. First, multiplying Eq. (4.3) by the complex conjugate of Eq. (4.3), we derive $A((U^1_j \otimes I)A^\dagger) = (U^1_j \otimes U^1_j)^\dagger$. By taking the trace, we derive
\[
D \text{Tr} U^1_j U^1_{j'} = (\text{Tr} U^1_j U^1_{j'})^2.
\]
Thus, Tr$U^1_j U^1_{j'} = 0$ or Tr$U^1_j U^1_{j'} = D$. In the case Tr$U^1_j U^1_{j'} = D$, we can easily see $U_j = U_{j'}$. Since $U_j \neq U_{j'}$ for all $j \neq j'$, we derive Tr$U^1_j U^1_{j'} = \delta_{j,j'}$. Thus, for a locally copiable set $\{|\Psi_j\rangle\}_{j=0}^{N-1}$, $|\Psi_j\rangle$ must be orthogonal to each other. Second, by the proof of the above lemma, the local copying operation $\Lambda$ is explicitly represented as a local unitary transformation $A^{13} \otimes A^{24}$.

In the paper [18], Anselmi et al. also obtained an explicit expression of $U_j$ in Eq. (4.3) in the case of $N = 2$ (the case where the copy set can be written as $\{|\Psi_0\rangle , |\Psi_1\rangle\}$); that is, they completely characterized local copiability under the assumption that the copy set consists of two states. In this case, since $U_0 = I$, there is only one independent equation $A(U_1 \otimes I)A^\dagger = U_1 \otimes U_1$. The following theorem is the conclusion of their analysis of Eq. (4.3) for $N = 2$. 

90
**Theorem 4.2 (Anselmi et al.)** There exists a unitary operator $A$ satisfying

$$A(U \otimes I)A^\dagger = U \otimes U,$$

(4.11)

if and only if the unitary operator $U$ satisfies the following two conditions:

1. The spectrum of $U$ is the set of the powers of $M$th roots of unity, where $M$ is a factor of $D$.

2. The distinct eigenvalues of $U$ have equal degeneracy.

Thus, if $D$ is a prime and $U \neq I$, then, the set of eigenvalues of $U$ is completely determined to be \(\{\omega^a\}_{a=0}^{D-1}\), where $\omega \overset{\text{def}}{=} \exp(2\pi i/D)$. We use this notation of $\omega$ throughout the following discussion.

Here, we should remark on the number of maximally entangled states as the resource. If we allow the use of three entangled states as a resource, we can always locally copy any orthogonal set of maximally entangled states by performing quantum teleportation [20]. In the case where we share two entangled states as resources, see [19].

### 4.3 Local copying of the maximally entangled states in prime-dimensional systems

Our main purpose is developing the relation between the local copiability and the local discriminability and understanding the non-locality of a set of maximally entangled states concerning local copiability and local discriminability. For this purpose, the necessary and sufficient condition given in Lemma 4.1 is rather abstract and we need a simpler criterion by which we can easily determine whether a given set is locally copiable or not. Therefore, in this section, we construct such a simple criterion for local copiability and completely characterize local copiability of a set of maximally entangled states. That is, we obtain an explicit expression of $U_j$ in Eq. (4.3) and get
a simpler necessary and sufficient condition of the local copiability for all \( N \) in the case of prime-\( D \)-dimensional local systems (Theorem 4.4). That is, by this Theorem 4.4, the form of \( A \) is completely determined. As a consequence, we show that \( D \) is the maximum size of a locally copiable set.

In the first step, we construct an example of a locally copiable set of \( D \) maximally entangled states.

**Theorem 4.3** When the set of maximally entangled states \( \{ |\Psi_j\rangle \}_{j=0}^{N-1} \) is defined by

\[
|\Psi_j\rangle = (U_j \otimes I) |\Psi_j\rangle
\]  

and

\[
U_j = \sum_{j=0}^{D-1} \omega^{jk} |k\rangle \langle k|,
\]

where \( \{ |k\rangle \}_{k=0}^{D-1} \) is an orthonormal basis of \( \mathcal{H}_1 \), then the set \( \{ |\Psi_j\rangle \}_{j=0}^{N-1} \) can be locally copied.

**Proof** We define the unitary operator \( A \) by

\[
A = \text{CNOT} \overset{\text{def}}{=} \sum_{a,b} |a \oplus b\rangle |b\rangle \langle a| \langle b|,
\]

where CNOT is an extension of Control-NOT gate represented in \( \{ |k\rangle \}_{k=0}^{D-1} \) for \( D \) dimensional systems, and \( \oplus \) is subtraction modulo \( D \). Then, we can easily
verify Equation (4.3) as

\[
A(U_j \otimes I)A^\dagger
= \text{CNOT}(U_j \otimes I)\text{CNOT}^\dagger
= \sum_{a_1,b_1} |a_1 \oplus b_1\rangle \langle b_1| (|a_3\rangle \langle a_3|)
|a_2\rangle \langle a_2 \oplus b_2| \langle b_2|
= \sum_{a_1,b_1} \omega^{ja_1} |a_1 \oplus b_1\rangle \langle a_1 \oplus b_1| \otimes |b_1\rangle \langle b_1|
= \sum_{c,b_1} \omega^{j(b_1 \oplus c)} |c\rangle \otimes |b_1\rangle \langle b_1|
= U_j \otimes U_j,
\]

where \(\oplus\) is addition modulo \(D\) and we set \(c = a_1 \oplus b_1\). Therefore, Lemma 4.1 guarantees that the set \(\{|\Psi_j\rangle\}_j\) \(D^{-1}\) can be locally copied. \(\square\)

As we already mentioned in the previous section, the copying operation can be chosen to be the local unitary operation \(A^{12} \otimes A^{34*} = \text{CNOT}^{12} \otimes \text{CNOT}^{34*}\). Therefore, this protocol of local copying used in the above proof is represented as Fig 4.3. Here, we should remark that \(U_1\) is the generalized Pauli’s \(Z\) operator which is one of the generators of the Weyl-Heisenberg group, and another \(U_j\) is the \(j\)th power of \(U_1 = Z\). Hence, in the case of
non-prime-dimensional local systems, the spectrum of $U_j$ is different from that of $U_1$ if $j$ is a non-trivial factor of $D$.

Moreover, the property of the Weyl-Heisenberg group not only guarantees that the above example satisfies (4.3), but also is essential for the condition (4.3). That is, as is proved below, any locally copiable set of maximally entangled states is restricted exclusively to the above example. Therefore, our main theorem can be written down as follows.

**Theorem 4.4** For systems whose local spaces are prime-dimensional, the set of maximally entangled states $\{U_j \otimes I | \Psi_0 \rangle \langle \Psi_0 | U_j^\dagger \otimes I \}_{j=0}^{N-1}$ can be locally copied if and only if there exists an orthonormal basis set $\{ |a\rangle \}_{a=0}^{D-1}$ and a set of integers $\{n_j\}_{j=0}^{N-1}$ such that the unitary $U_j$ can be written as

$$U_j = \sum_{a=0}^{D-1} \omega^{n_j a} |a\rangle \langle a|,$$

where $\omega$ is the $D$th root of unity.

From Eq. (4.15), we can easily see that the number of different candidates of $U_j$ is at most $D$. Thus, $D$, which is equal to the dimension of the local space, is the maximum size of a locally copiable set of maximally entangled states with prime-dimensional local systems. In comparison with the case without the LOCC restriction, where we can copy $D^2$ orthogonal states, this is actually the square root.

The proof of Theorem 4.4 is as follows.

**Proof** (Sufficient Condition) We have already proven that $\{U_j \otimes I | \Psi_j \rangle \}_{j=0}^{D-1}$ can be copied by LOCC in Theorem 4.3. Therefore, any subset of them can be copied by LOCC.

(Necessary Condition) Assume that a unitary operator $A$ satisfies the condition (4.3) for all $j$. By applying Theorem 4.2 to $U_1$, we can choose an
orthonormal basis set \( \{|a\rangle\}_{a=0}^{D-1} \) such that

\[
U_1 = \sum_{a=0}^{D-1} \omega^a |a\rangle \langle a| ,
\]

where \( \omega \) is \( D \)th root of unity. Then, we focus on Eq. (4.3) in the case of \( j = 1 \). In this equation, \(|a\rangle \langle a| \otimes \mathcal{H} \) is the eigenspace of the corresponding eigenvalue \( \omega^a \) of \( U_1 \otimes I \) in the left-hand side, and \( \text{span}\{|a \ominus c\rangle \langle a \ominus c| \otimes |c\rangle \langle c|\}_{c=0}^{D-1} \) is the eigenspace of the corresponding eigenvalue \( \omega^a \) of \( U_1 \otimes U_1 \) in the right-hand side. Since \( U_1 \otimes I \) is transformed to \( U_1 \otimes U_1 \) by the action of the unitary \( A \) in Eq. (4.3), the unitary \( A \) should transform the subspace \(|a\rangle \langle a| \otimes \mathcal{H} \) to the subspace \( \text{span}\{|a \ominus c\rangle \langle a \ominus c| \otimes |c\rangle \langle c|\}_{c=0}^{D-1} \) and the remaining freedom of \( A \) is unitary transformations between these subspaces. That is, \( A \) is expressed as

\[
A = \sum_{a,b,c} \xi_{b,c}^a |a \ominus c\rangle |c\rangle \langle a| \langle b| ,
\]

where \( \xi_{b,c}^a \) is a unitary matrix for \( b \) and \( c \) for a fixed \( a \); that is, \( \sum_{c=0}^{D-1} \xi_{b,c}^a \bar{\xi}_{b',c}^a = \delta_{b,b'} \) and \( \sum_{b=0}^{D-1} \xi_{b,c}^a \xi_{b',c}^a = \delta_{c,c'} \). For every \( a, \xi_{b,c}^a \) determines a unitary transformation from \(|a\rangle \otimes \mathcal{H} \) to \( \text{span}\{|a \ominus c\rangle \langle a \ominus c| \otimes |c\rangle \langle c|\}_{c=0}^{D-1} \). Thus, based on the basis set \( \{|a\rangle\}_{a=0}^{D-1} \), the matrix elements of Eq. (4.3) for all \(|a_1\rangle \otimes |b_1\rangle \langle a_2| \otimes |b_2| \) is written down as

\[
\langle a_1| \langle b_1| A(U_j \otimes 1) A^\dagger |a_2\rangle \langle b_2| = \langle a_1| U_j |a_2\rangle \langle b_1| U_j |b_2\rangle .
\]

Therefore, substituting Eq. (4.17) to Eq. (4.18) for any integer \( j \), we obtain

\[
\langle a_1| U_j |a_2\rangle \langle b_1| U_j |b_2\rangle = \sum_{a,b,c,a',b',c'} \xi_{b,c}^a \bar{\xi}_{b',c'}^{a'} \delta_{a_1,a \ominus c} \delta_{a_2,a' \ominus c'} \delta_{b_1,c} \delta_{b_2,c'} \delta_{b,b'} \langle a| U_j |a'\rangle
\]

\[
= \sum_{b=0}^{D-1} \xi_{b,b_1}^{a_1 \oplus b_1} \bar{\xi}_{b,b_2}^{a_2 \oplus b_2} \langle a_1 \oplus b_1| U_j |a_2 \oplus b_2\rangle ,
\]

for all \( a_1, a_2, b_1 \) and \( b_2 \).
To see that $U_1$ and $U_j$ can be simultaneously diagonalized, we need to prove the following lemma.

**Lemma 4.2** If a non-zero $D \times D$ matrix $U_{ab}$ satisfies the equation

$$
\Xi^{c_1 \oplus b_1, a_2 \oplus b_2}_{b_1 b_2} U_{a_1 \oplus b_1, a_2 \oplus b_2} = U_{a_1 a_2} U_{b_1 b_2},
$$

(4.20)

where $\Xi^{cci}_{b_1 b_2} = \delta_{b_1 b_2}$ and all indices have their values between 0 and $D-1$, then $U_{ab}$ is a diagonal matrix.

**Proof** See Appendix A.

We apply Lemma 4.2 to the case when $U_{ab} = \langle a | U_j | b \rangle$ and $\Xi_{b_1 b_2}^{a_1 a_2} = \sum_{b=0}^{D-1} c_{bb_1}^{a_1} \Xi_{bb_2}^{a_2}$. Then this lemma shows that $\langle a | U_j | b \rangle$ is a diagonal matrix for all $j$, and therefore the unitaries $\{U_j\}_{j=0}^{N-1}$ are simultaneously diagonalizable in the eigenbasis $\{|a\rangle\}_{a=0}^{D-1}$ of $U_1$. Then we can get the formula of $\{U_j\}_{j=0}^{N-1}$ explicitly as follows. From the diagonal element of (4.19), we derive

$$
\langle a \oplus b | U_j | a \oplus b \rangle = \langle a | U_j | a \rangle \langle b | U_j | b \rangle.
$$

(4.21)

From Theorem 4.2 and Lemma 4.1, all $U_j$ have the same eigenvalues $\{\omega^n\}_{n=0}^{D-1}$, and from the above discussion, all $U_j$ also have the same eigenbasis $\{|a\rangle\}_{a=0}^{D-1}$. Therefore, we can express $U_j$ as

$$
U_j = \sum_{a=0}^{D-1} \omega^{P_j(a)} |a\rangle \langle a|,
$$

(4.22)

where $P_j(a)$ is a bijection from $\{a\}_{a=0}^{D-1}$ to themselves. Then Eq. (4.21) guarantees that $P_j(a)$ is a self-isomorphism of the cyclic group $\{a\}_{a=0}^{D-1}$. Since a self-isomorphism of a cyclic group is identified as the image of the generator [27], we derive the formula (4.15) with $P_j(1) = n_j$.

This theorem completely characterizes the local copiability of maximally entangled states in the case of prime-dimensional local spaces.
So far, we have solved the LOCC copying problem only for a prime-dimensional local space. Now, we discuss about the case of a non-prime-dimensional local space. Since Theorem 4.3 is valid in the non-prime case, the sufficient condition of Theorem 4.4 is also valid in non-prime dimensional local systems. However, the necessary condition is extended straightforwardly, if the set \( \{U_j\}_{j=0}^{N-1} \) contains at least one unitary whose eigenvalues are generated by \( \omega \); in other words, in the set of \( \{U_j\}_{j=0}^{N-1} \), there exists \( U_k \) whose eigenspace is not degenerate. We can only show the following statement for the necessary condition.

**Theorem 4.5** If a set of maximally entangled states \( \{U_j \otimes I | \Psi_0 \rangle \langle \Psi_0 | U_j^\dagger \otimes I \} \) is locally copiable and if there exists \( k \) such that \( \omega \) is an eigenvalue of \( U_k \), then there exist an orthonormal basis set \( \{ |a\rangle \}_{D-1}^{a=0} \) and a set of integers \( \{n_j\}_{N-1}^{j=0} \) such that, for all \( j \), \( U_j \) can be written down in the form of Eq. (4.15).

**Proof** The proof does not lose generality under the assumption that \( U_1 \) has eigenvalue \( \omega \). Then, by Theorem 4.2, Eq. (4.16) holds. By the same procedure of the prime-dimensional case, Eq. (4.17) and Eq. (4.18) hold. Thus we obtain Eq. (4.19) in the same way of the prime-dimensional case. Then Lemma 4.2 implies that all \( U_j \) can be diagonalized and also implies Eq. (4.21) for all \( U_j \). By writing \( U_j \) as (4.22), we get the equation \( P_j(a \oplus b) = P_j(a) \oplus P_j(b) \) and hence \( P_j(a) = aP_j(1) \). Hence, Theorem 4.2 guarantees the same representation of \( U_j \) as (4.15). \( \square \)

Therefore, we can solve the problem of the local copying in non-prime-dimensional local spaces as the direct extension of Theorem 4.4, only in the case where eigenspace of one of \( U_j \) is not degenerate. On the other hand, if eigenvalues of all \( U_j \) are degenerate, our proof of the necessary condition does not hold.
4.4 The relation between LOCC copying and LOCC discrimination

If we have no LOCC restriction, the possibility of the deterministic copying is equivalent to that of the perfect discriminability. However, we can easily see that under the restriction of LOCC, this relation is nontrivial at all. As we already mentioned in the introduction, these two problems share a common feature; that is, their difficulty can be regarded as the non-locality of a set of states and this non-locality cannot be explained only by the entanglement convertibility. Therefore, the study of their relation is really important to understand the non-locality of a set of states. In this section, we compare the local discriminability and the local copiability for a set of orthogonal maximally entangled states. Thus, by introducing the simultaneous Schmidt decomposition, we show the relation between these two problems of the non-locality.

At first, we review the definition of a locally distinguishable set and then mention several known and new results of the local discriminability. A set of states \( \{ |\Psi_j\rangle \}_{j=0}^{N-1} \) is called two-way (one-way) classical communication locally distinguishable, if there exists a POVM \( \{ M_j \}_{j=0}^{N-1} \) which can be performed by two-way (one-way) LOCC and also satisfies the following conditions:

\[
\forall i, j, \quad \langle \Psi_i | M_j | \Psi_i \rangle = \delta_{ij}. \quad (4.23)
\]

In order to compare the local copying and the local discrimination, we should take care of the following point: we assume an extra maximally entangled state only in the local copying case. This is because local copying of a set of maximally entangled states is trivially impossible without a blank entangled state. This fact is contrary to local discrimination; that is, we do not allow the parties to share maximally entangled in the local discrimination problem, since if we allow, they can always discriminate by teleportation.
In the previous section, we already proved that $D$ is the maximum size of the locally copiable set of maximally entangled states. In the case of the local discrimination, we can also prove that $D$ is the maximum size of a locally distinguishable set of maximally entangled states. This statement was proved by the paper [14] only when the set of maximally entangled states $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ consists of canonical-form Bell states, where a canonical-form Bell state $|\Psi_{nm}\rangle$ is defined as

$$
|\Psi_{nm}\rangle \overset{\text{def}}{=} Z^n X^m \otimes I |\Psi_{00}\rangle
$$

$$
|\Psi_{00}\rangle \overset{\text{def}}{=} \sum_{k=0}^{d-1} |k\rangle \otimes |k\rangle
$$

$$
X \overset{\text{def}}{=} \sum_{k=1}^{d} |k\rangle \langle k \oplus 1|
$$

$$
Z \overset{\text{def}}{=} \sum_{k=1}^{d} \omega^k |k\rangle \langle k|.
$$

Such a set is a special case of a set of maximally entangled states. Here, we give a simple proof of this statement for a general set of maximally entangled states by the same technique as [28].

**Theorem 4.6** If an orthogonal set of maximally entangled states $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ is locally distinguishable, then $N \leq D$.

**Proof** Suppose that $\{|\Psi_j\rangle\}_{j=0}^{N-1}$ is locally distinguishable by LOCC POVM $\{M_j\}_{j=0}^{N-1}$. Then, since LOCC operation is always separable, $\{M_j\}_{j=0}^{N-1}$ can be decomposed as $M_i = \sum_{k=0}^{L} p_{ik} |\psi_k\rangle \langle \psi_k| \otimes |\phi_k\rangle \langle \phi_k|$, where $p_{ik}$ is a positive
coefficient. Then, we can derive an upper bound of $\langle \Psi_j | M_i | \Psi_j \rangle$ as follows,

\[
\langle \Psi_j | M_i | \Psi_j \rangle = \sum_{k=0}^{L} p_{ik} \langle \Psi_j | (| \psi_k \rangle \otimes | \phi_k \rangle) | \Psi_j \rangle \\
\leq \sum_{k=0}^{L} p_{ik} \langle \psi_k | (\frac{1}{D} I) | \psi_k \rangle \\
\leq \frac{\text{Tr} M_i}{D},
\]

where the first inequality comes from the monotonicity of the fidelity under the partial trace operations with respect to the system $B$. Since $\langle \Psi_j | M_j | \Psi_j \rangle = 1$, we have $1 \leq \text{Tr}(M_j)/D$. Finally, taking the summation over the inequality for $j$, we obtain $N \leq D^2/D = D$, since $\sum_{j=0}^{D-1} \text{Tr}(M_j) = D^2$. \hfill \Box

Therefore, in this case, the maximal size of both locally copiable and locally distinguishable sets is equal to the dimension of the local space.

When we consider the relation between the local discrimination and the local copying of a set of maximally entangled states, it is quite useful to introduce the simultaneous Schmidt Decomposition [9, 21]. A set of states $\{ | \Psi_\alpha \rangle \}_{\alpha \in \Gamma} \subset \mathcal{H}_1 \otimes \mathcal{H}_2$ is called simultaneously Schmidt decomposable, if they can be written as

\[
| \Psi_\alpha \rangle = \sum_{k=0}^{d-1} b_k^{(\alpha)} | e_k \rangle | f_k \rangle,
\]

(4.24)

where $\Gamma$ is a parameter set, $\{ | e_k \rangle \}_{k=0}^{d-1}$ and $\{ | f_k \rangle \}_{k=0}^{d-1}$ are orthonormal basis of the local spaces (the simultaneous Schmidt basis) and $b_k^{(\alpha)}$ is a complex coefficient. For a set of orthogonal maximally entangled states, the simultaneous Schmidt decomposability is a sufficient condition for the one-way local discriminability [9] and a necessary condition for the local copiability. Moreover, the simultaneous Schmidt decomposability is not a necessary and sufficient condition for the both cases. Therefore, the family of locally copiable sets of maximally entangled states is strictly included by the family of
one-way locally distinguishable sets of maximally entangled states. In the following, we prove this relation.

First, we explain the relation between the local discrimination and the simultaneous Schmidt decomposition which has been already obtained by the paper [9]. If an unknown state $|\Psi_\alpha\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is in a simultaneously Schmidt decomposable set of states $\{|\Psi_\alpha\rangle\}_{\alpha \in \Gamma}$, such a state can be transformed onto a single local space $\mathcal{H}_A$ or $\mathcal{H}_B$ by LOCC. Rigorously speaking, there exists an LOCC $\Lambda$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ which transforms $|\Psi_A B_1 B_2^\alpha\rangle \otimes |0_{B_2}\rangle$ to $\sigma_A \otimes |\Psi_B^{A_1 B_1^{\alpha}}\rangle$ for all $\alpha \in \Gamma$, and there also exists an LOCC $\Lambda'$ on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2} \otimes \mathcal{H}_B$ which transforms $|0^{A_1}\rangle \otimes |\Psi_A^{A_2 B_2^\alpha}\rangle$ to $|\Psi_A^{A_1 A_2^\alpha}\rangle \otimes \sigma_B$ for all $\alpha \in \Gamma$, where $|0\rangle$ and $\sigma_A$ and $\sigma_B$ are an arbitrary blank state, and unimportant garbage states of Alice and Bob, respectively. This LOCC transformation can be written in the following Kraus representation [9]:

$$\rho \mapsto \sum_{k=0}^{D-1} F_k \rho F_k^*,$$

where

$$F_k \overset{\text{def}}{=} (I_A \otimes \text{CNOT})(U_k \otimes I_{A_2 B_2})(P_k \otimes I_{B_1 B_2})$$

$$P_k \overset{\text{def}}{=} 1/D(\sum_i \omega^{k_i} |e_i\rangle)(\sum_l \omega^{k_l} \langle e_l|)$$

$$U_k \overset{\text{def}}{=} \sum_i \omega^{k_i} |f_i\rangle \langle f_i|$$

$$\text{CNOT} \overset{\text{def}}{=} \sum_{kl} |e_k\rangle \otimes |f_k \oplus l\rangle \langle f_k| \otimes |l|.$$

In the above formula, both $\{|e_k\rangle\}_{k=0}^{D-1}$ and $\{|f_i\rangle\}_{k=0}^{D-1}$ are simultaneous Schmidt basis set of $\{|\Psi_\alpha\rangle\}_{\alpha \in \Gamma}$, and $\{|l\rangle\}_{l=0}^{D-1}$ is the standard computational basis. This protocol can be depicted in Fig 4.4, where $|G_k\rangle$ is a garbage state with no information. Using the above protocol, if a set $\{|\Psi_\alpha\rangle\}$ is simultaneously Schmidt decomposable, there exists a one-way-LOCC POVM $M' = \{M'_i\}$ for a given arbitrary POVM $M = \{M_i\}$ such that

$$\langle \Psi_\alpha | M_i | \Psi_\alpha \rangle = \langle \Psi_\alpha | M'_i | \Psi_\alpha \rangle, \quad \forall i, \forall \alpha.$$
Figure 4.4: A set of simultaneous Schmidt decomposable states can be sent to Bob’s space by LOCC

That is, any POVM can be essentially realized by one-way LOCC. Therefore, “a simultaneously Schmidt decomposable set of orthogonal maximally entangled states is one-way locally distinguishable.”

On the other hand, there exists a set of orthogonal maximally entangled states which is not simultaneously Schmidt decomposable, but locally distinguishable. For example, a set of maximally entangled states \{\ket{\Psi_0}, X \otimes I \ket{\Psi_0}, Z \otimes I \ket{\Psi_0}\} is not simultaneously Schmidt decomposable in any prime-dimensional systems, since \([X, Z] \neq 0\). However, this set is one-way locally distinguishable in \(D \geq 3\) dimensional systems, since all sets of canonical Bell states with \(l\) elements are one-way locally distinguishable, if \(l(l-1)/2 \leq D\) \cite{13}. There are also other types of examples of a locally distinguishable, but not locally copiable set of maximally entangled states in non-prime dimensional systems \cite{18}. In this case, they are simultaneously Schmidt decomposable, however not canonical Bell states, for example, \{\ket{\Psi_0}, I \otimes U \ket{\Psi_0}\} where \(U \overset{\text{def}}{=} e^{\frac{\pi}{3}} \ket{0}\bra{0} + e^{\frac{2\pi}{3}} \ket{1}\bra{1} + e^{\frac{4\pi}{3}} \ket{2}\bra{2} + e^{\frac{5\pi}{3}} \ket{3}\bra{3}\) in \(4 \times 4\) dimensional systems. Thus, the family of simultaneously Schmidt decomposable sets of maximally entangled states is strictly included by the family of locally
distinguishable sets of maximally entangled states.

On the other hand, the relation between the simultaneous Schmidt decomposability and the local copiability can be described by the following theorem.

**Theorem 4.7** In prime-dimensional local systems, an orthogonal set of maximally entangled states \( \{ |\Psi_j\rangle \}^{N-1}_{j=0} \) is locally copiable, if and only if it is a simultaneously Schmidt decomposable subset of canonical-form Bell states under the same local unitary operation.

**Proof** We can easily see the necessary condition of the above statement from Theorem 4.4 as follows. In Theorem 4.4, since Eq.(4.15) means that each \( U_j \) is equivalent to \( Z^{n_j} \) under the same unitary operation, a set of maximally entangled states \( \{ U_j \otimes I |\Psi_0\rangle \}^{N-1}_{j=0} \) is local unitary equivalent to \( \{ Z^{n_j} \otimes I |\Psi_0\rangle \}^{N-1}_{j=0} \), which is a simultaneous Schmidt decomposable subset of canonical-form Bell states. The “if” part can be showed as follows. The paper [21] shows that the states \( |\Psi_{n_a m_\alpha}\rangle (\alpha = 1, 2, \cdots, l) \) are simultaneously Schmidt decomposable, if and only if there exist integers \( p, q \) and \( r \) (\( p \neq 0 \) or \( q \neq 0 \)) satisfying \( p m_{\alpha} + q m_{\alpha} = r \) for all \( \alpha \). Since the ring \( \mathbb{Z}_p \) is a field if \( p \) is a prime number, the above condition is reduced to the existence of \( f \) and \( g \) such that \( m_{\alpha} = fn_{\alpha} + g \). Then we get

\[
|\Psi_{n_a m_\alpha}\rangle = |\Psi_{n_a(fn_{\alpha}+g)}\rangle = C_\alpha(ZX^f)^{n_a}X^g \otimes I |\Psi_0\rangle.
\] (4.25)

Since \( ZX^f \) is unitary equivalent to \( Z \) [13], the state \( |\Psi_{n_a m_\alpha}\rangle \) is locally unitary equivalent with \( U_j \otimes I |\Psi_0\rangle \) in Theorem 4.4. \( \Box \)

We add a remark here. Under the assumption of the simultaneous Schmidt decomposition, a set has canonical Bell form, if and only if the set of corresponding unitary operators is a cyclic group, that is, the group with only one
generator. Therefore, we can rephrase this necessary and sufficient condition as follows, the set is simultaneously Schmidt decomposable and satisfies the following condition by renumbering:

$$U_1^D = I, \ U_k = \overrightarrow{U_1 \cdots U_1}.$$ (4.26)

Finally, we derive Fig.4.1 and therefore for maximally entangled states, the family of locally copiable sets is strictly included by the family of simultaneously Schmidt decomposable sets. In other words, the local copying is more difficult than the local discrimination.

In the last part of this section, we discuss our main results in Fig 4.1, from the viewpoint of non-locality of a set of states.

In the case of bipartite pure states, all information of a bipartite state $|\Psi\rangle = \sum_{i=0}^{D-1} \lambda_i |e_i\rangle \otimes |f_i\rangle$ can be separated to two parts, namely the Schmidt coefficients $\lambda_i$ and the Schmidt basis set $\{|e_i\rangle, |f_i\rangle\}_{i=0}^{D-1}$, where $\lambda_i \geq 0$. Because of local unitary equivalence, the Schmidt coefficients completely determine the entanglement convertibility [1]. On the other hand, Schmidt bases determine the non-locality which is different from the entanglement. In general, for different states, their Schmidt basis are also different. Thus, conversely, we can regard the non-locality coming from interrelationship among the Schmidt basis set for the different states as the non-locality purely beyond the entanglement of individual states. In the following discussion, we try to separate the non-locality which depends on the Schmidt coefficients and the Schmidt basis.

At first, all sets in Fig. 4.1 are sets of maximally entangled states; that is, they have the same Schmidt coefficients and the same amount of entanglement. Thus the structure of non-locality in Fig.4.1 is determined only by the interrelationship of the Schmidt bases for different states, and the effect of the Schmidt coefficients do not appear directly in this figure. On the other hand, when we calculate the maximal sizes of local distinguishable
and copiable sets, we need to optimize all possible choices of the Schmidt bases. That is, the maximal sizes depend only on the Schmidt coefficients. Therefore, the Schmidt coefficients may affect only the maximal size of local distinguishable and copiable sets.

The interrelationship of Schmidt bases is determined by the unitary operator \( U = \sum_{i=0}^{D-1} |e_i \rangle \langle f_i| \). In Fig. 4.1, the two properties of the interrelationship of Schmidt basis, that is, such unitary operators, are related to the non-locality of a set. That is, the simultaneous Schmidt decomposability and canonical Bell form seems to reduce non-locality of a set. For simultaneous Schmidt decomposable sets, we can explain their lack of non-locality as follows. As we well know, in the case of pure bipartite states, one person can always apply the local operation which causes the same transformation for a given state as another person’s local operation causes (Lo-Popescu's theorem [29]). The simple structure of the entanglement convertibility originates in the above symmetry between local systems. This symmetry is caused by the existence of a Schmidt decomposition. Similarly, in the case of local discrimination, the protocol Fig. 4.4 seems to utilize this kind of symmetry between local systems. Therefore, the existence of simultaneous Schmidt decomposable basis can give the symmetry between the local systems and this fact may decrease the non-locality of the sets of states.

In the case of canonical Bell form, an interesting fact is that the algebraic property of the Weyl-Heisenberg group is related to the local copiability, and not to the local discriminability. As we have already seen, since a simultaneous Schmidt decomposable set can be transformed to a single local space by LOCC, we can use any global discrimination protocols to such a set by only LOCC. Therefore, with respect to the local discrimination, the sets of simultaneous Schmidt decomposable states do not seem to possess any non-locality which originates in interrelationship between their Schmidt basis. However, if such a set does not have a canonical Bell form, it is not
that is, a set has extra non-locality beyond individual entanglement with respect to the local copying, if it has no algebraic structure given in (4.26), even if it is simultaneous Schmidt decomposable. Finally, we can conclude that, from the viewpoint of problems of the non-locality beyond individual entanglement, the above algebraic non-locality is the most remarkable difference between the local copying and the local discrimination.

4.5 Application to channel copying, entanglement distillation, and error correction

So far, we have treated the local copying mainly in the context of the non-locality of a set of states. On the other hand, we now consider how the local copying itself is benefitting to information processing. In this last section, we apply our results, especially Theorem 4.7 to different contexts and give several other interpretations of our results, such as the channel copying, the entanglement distillation, the error correction, and the quantum key distribution (QKD). These many connections imply the fruitfulness of the local copying problem as a fundamental two-party protocol. Moreover, seeing the local copying from these various points of view, we may also derive some clue which helps us to construct further development of understanding of the non-locality beyond the entanglement convertibility.

4.5.1 Channel copying

In the analysis of the local copying in Section 4.3, we treated not directly maximally entangled states, but unitary operators which represent the maximally entangled states based on the standard maximally entangled states. This method is a kind of an operator-algebraic method, or equivalent to the Heisenberg picture [30]. Therefore, we can interpret our results as directly the results for these unitary operators themselves. As a result, we can look
at the problem of “unitary channel copying”.

Here, we consider a problem of “channel copying”, that is, a problem in which we ask a question as follows: in the case where we do not have a complete description of a channel, “can we simulate two copies of an unknown channel by using the unknown channel only once and also using a known blank channel once.” For example, such a question may occur in the following case. There exists an unknown and rare quantum operation, for which we would like to have as many outputs (results of the operation) as possible. However, we cannot restrict inputs of the channel; therefore the inputs might be arbitrary states. Under the above condition, we would like to decrease the frequency of use of the operation. Such a situation may occur in query complexity problems, for example [31]. In this case, an unknown channel is a query which is represented by a unitary operation.

As we will see in the following discussion, the channel copying with the help of one-way classical communications between the sender and the receiver is equivalent to the local copying of corresponding entangled states with help of one-way classical communication between Alice and Bob. The problem setting of the channel copying can be written as follows:

**Definition 4.1** We say that a set of channel \( \{ \Lambda_i \}_{i=1}^{N} : \mathcal{T}(\mathcal{H}_{A1}) \rightarrow \mathcal{T}(\mathcal{H}_{B1}) \) is copiable with one-way classical communications and a blank channel \( \Lambda_b \), if for all \( i \), there exists sets of Kraus’s operators \( \{ A_k \}_{k=1}^{K} \subset \mathcal{T}(\mathcal{H}_{A1} \otimes \mathcal{H}_{A2}) \) and \( \{ B^k_l \}_{l=1}^{L} \subset \mathcal{T}(\mathcal{H}_{B1} \otimes \mathcal{H}_{B2}) \) such that \( \sum_{k=1}^{K} A^\dagger_k A_k = I_A \), \( \sum_{l=1}^{L} B^k_l B^k_l = I_B \) for all \( l \), and

\[
\sum_{kl} B^k_l [\Lambda_i \otimes \Lambda_b(A_k \rho A^\dagger_k)] B^k_l = \Lambda_i \otimes \Lambda_i(\rho), \tag{4.27}
\]

for all \( i \) and \( \rho \) on \( \mathcal{H}_{A1} \otimes \mathcal{H}_{A2} \), where \( \mathcal{T}(\mathcal{H}) \) is the Banach space of all trace class operators on \( \mathcal{H} \).

The meaning of the above definition can be sketched as the Fig 4.5; by an encoding operation \( \{ A_k \}_{k=1}^{K} \) and a decoding operation \( \{ B^k_l \}_{l=1}^{L} \), one copy
of unknown channel $\Lambda_i$ with one copy of an arbitrary fixed blank channel $\Lambda_b$ works as two copies of unknown channels $\Lambda_i \otimes \Lambda_i$. For simplicity, we always assume $\dim \mathcal{H}_{A1} = \dim \mathcal{H}_{A2} = \dim \mathcal{H}_{B1} = \dim \mathcal{H}_{B2}$ in the following discussion.

![Figure 4.5: Definition of channel copying with one way classical communication: By suitable encoding $\{A_k\}$ and decoding $\{B_l^k\}$ operations, each $\Lambda_i$ works as $\Lambda_i \otimes \Lambda_i$.](image)

Then, by means of Choi-Jamiolkowski’s isomorphism which is the isomorphism between the channel $\Lambda$ and the entangled state $\Lambda \otimes I(\ket{\Psi} \bra{\Psi})$ [32], we can easily show that the channel copying problem with one-way classical communications is exactly the same as the local copying of corresponding entangled states with one-way classical communication.

**Theorem 4.8** A set of channel $\{\Lambda_i\}_{i=0}^{N-1}$ is copiable with one-way classical communications and a blank channel $\Lambda_b$, if and only if a set of entangled states $\{\Lambda_i \otimes I(\ket{\Psi} \bra{\Psi})\}_{i=N}$ is locally copiable with one-way classical communication and a blank states $\Lambda_b \otimes I(\ket{\Psi} \bra{\Psi})$, where $\ket{\Psi}$ is an arbitrarily maximally entangled state.

**Proof** Suppose that $\{\Lambda_i\}_{i=0}^{N-1}$ is copiable with one-way classical communications and a blank channel $\Lambda_b$. Consider four systems $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \mathcal{H}_4$, where Alice has $\mathcal{H}_1 \otimes \mathcal{H}_3$ and Bob has $\mathcal{H}_2 \otimes \mathcal{H}_4$, and prepare two copies of
maximally entangled states $|\Psi\rangle \langle \Psi|$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\mathcal{H}_3 \otimes \mathcal{H}_4$, respectively. Then, by applying a channel copying protocol for $\mathcal{H}_1 \otimes \mathcal{H}_3$, we derive the following calculations:

$$
\sum_{kl} B_{13}^k \otimes I_{24}^{24} [\Lambda_i^1 \otimes \Lambda_b^3 \otimes I_{24}^{24} (A_{13}^{1} \otimes I_{24}^{24} |\Psi^{12}\rangle \langle \Psi^{12}|] B_{13}^k \otimes I_{24}^{24}
$$

$$
= \sum_{kl} B_{13}^k \otimes I_{24}^{24} [\Lambda_i^1 \otimes \Lambda_b^3 \otimes I_{24}^{24} (I_{13} \otimes A_{24}^{12} |\Psi^{12}\rangle \langle \Psi^{12}|] B_{13}^k \otimes I_{24}^{24}
$$

$$
= \sum_{kl} B_{13}^k \otimes A_{24}^{12} [(\Lambda_i^1 \otimes I_{24}^{24} (|\Psi^{12}\rangle \langle \Psi^{12}|] B_{13}^k \otimes I_{24}^{24}
$$

$$
= \Lambda_i^1 \otimes I_{24}^{24} (|\Psi^{12}\rangle \langle \Psi^{12}|] \otimes \Lambda_b^3 \otimes I_{24}^{24} (|\Psi^{34}\rangle \langle \Psi^{34}|),
$$

where the last equality comes from Eq. (4.27). Therefore, $\Lambda_i \otimes I(\langle \Psi \rangle \langle \Psi|)$ is locally copiable with one-way classical communications. We can also easily check the opposite direction of the proof. □

The above theorem shows that the channel copying problems can be always identified with corresponding local copying problems of entangled states in the case of one-way classical communications. On the other hand, since not all states can be written as $\Lambda \otimes I(\langle \Psi \rangle \langle \Psi|)$ for a maximally entangled state $|\Psi\rangle$, not all local discrimination problems can be considered as a channel copying problem.

Choosing all $\Lambda_i$ as unitary channels, we derive maximally entangled states for the corresponding entangled states. Therefore, our results in Section 4.3 and 4.4 give also results for unitary channel copying as follows.

**Corollary 4.1** In prime-dimensional systems, a set of unitary channels $\{\Lambda_i\}_{i=0}^{N-1}$, where $\Lambda_i(\rho) = U_i \rho U_i^\dagger$, is copiable with a blank noiseless channel $\Lambda_b = I$ and one-way classical communications, if and only if $\{U_i\}_{i=0}^{N-1}$ is a simultaneous
diagonalizable subset of the Weyl-Heisenberg (Generalized Pauli) group.

**Proof**  We can easily see the proof from Theorem 4.7 and 4.8. \(\square\)

As we will see later, the above analysis of the channel copying can also be used in the context of the error correction of quantum channel.

### 4.5.2 Entanglement distillation

Since our local copying protocol consists of local unitary operations, we can also consider the opposite direction of our protocol. This inverse of the local copying protocol is actually an entanglement distillation protocol by local unitary operations. Since we usually use measurements in entanglement distillation protocols \([2, 26, 4]\), this is a really rare example of entanglement distillation by unitary transformation. As we can see in Fig 4.3, the inverse of our protocol operates on states of the form \(|\Psi_i\rangle \otimes |\Psi_i\rangle\) by a local unitary operation.

Therefore, if we consider a mixed state as

\[
\rho = \sum_{ij} a_{ij} |\Psi_i\rangle \otimes |\Psi_i\rangle \langle \Psi_j| \otimes \langle \Psi_j|
\]  

(4.28)

and apply our local copying protocol, where \(\{|\Psi_i\rangle\}_{i=0}^{D-1}\) is a set of simultaneous Schmidt decomposable subset of canonical Bell states, then we derive

\[
A^{13} \otimes A^{24} \rho^{1234} A^{13} \otimes A^{24} = \sum_{ij} a_{ij} \langle A^{13} \otimes A^{24} |\Psi_i^{12}\rangle \otimes |\Psi_i^{34}\rangle \langle \Psi_j^{12}| \otimes \langle \Psi_j^{34}| A^{13} \otimes A^{24} \rangle
\]

\[
= \sum_{ij} a_{ij} |\Psi_i^{12}\rangle \langle \Psi_j^{12}| \otimes |\Psi_i^{34}\rangle \langle \Psi_j^{34}|,
\]

where \(A\) is a local unitary operator defined in (4.14) and we used \(A^{13} \otimes A^{24} |\Psi_i^{12}\rangle \otimes |\Psi_i^{34}\rangle = |\Psi_i^{12}\rangle \otimes |\Psi_0^{34}\rangle\) in the second equality. This protocol is an actually entanglement distillation protocol deriving one maximally entangled state for all mixed states which satisfy the above condition. Moreover, in the
case \( a_{ij} = \delta_{ij}/D \), since \( \sum_{i=0}^{D-1} \frac{1}{D} |\Psi_i\rangle \langle \Psi_i| \) is a separable state, this distillation protocol by the local unitary is optimal. The state (4.28) belongs to a class of states called “maximally correlated states”, and the simple formula of distillable entanglement for maximally correlated states has been already known [4]. However, the above protocol is deterministic and moreover unitary, which is actually an important point. Generally speaking, deterministic distillable entanglement (distillable entanglement by means of the deterministic distillation protocols) is strictly less than the usual asymptotic one [5]. Therefore, this is a very rare case where we can derive the lower bound of deterministic distillable entanglement for mixed states.

### 4.5.3 Error correction and quantum key distribution

As another application of our results of the local copying problem, we apply our result to the error correction and quantum key distribution (QKD) with the following specific noisy channel in this subsection.

The error correction is the protocol in which we derive noiseless channels from noisy channels by applying encoding and decoding operations. The error correction of quantum channels is described mathematically as follows: For given quantum channel \( \Lambda \) on \( \mathcal{H} \), we say that the error on the quantum channel \( \Lambda \) is corrected by means of an encoding operation \( \Lambda_e \) and a decoding operation \( \Lambda_d \) on \( \mathcal{H} \otimes n \), if we drive the following equality in the limit \( n \to \infty \):

\[
\Lambda_e \circ \Lambda \otimes n \circ \Lambda_d = I^{\otimes m} \otimes \Lambda_g,
\]

where \( \Lambda_g \) is an important garbage (noisy) channel on \( \mathcal{H}^{\otimes (n-m)} \). In the above case, the channel capacity of \( \lambda \) is defined by \( \lim_{n \to \infty} \frac{m}{n} \).

Now, we consider the inverse of the channel copying protocol in subsection 4.5.1 and we derive the error correcting protocol which corresponds to the above distillation protocol. Consider a channel \( \Lambda(\rho) = \sum_{k=1}^{N} E_k \rho E_k^\dagger \) on \( \mathcal{H}_1 \otimes \mathcal{H}_2 \), where the error operators \( E_k \) satisfies \( \sum_{k=1}^{N} E_k^\dagger E_k = I \) and can be written...
as \( E_k = \sum_{i=0}^{D-1} c_k U_i \otimes U_i \) in terms of a simultaneous diagonalized subset \( \{U_i\}_{i=0}^{D-1} \) of Generalized Pauli’s Group. In particular, when a channel can be decomposed by a set of Kraus operators which have a form \( E_k = p_k U_k \otimes U_k \), the channel is called the collective noise. Such a noise may occur, for example, in the case where we send two photonic qubits simultaneously through an optical fibre or a free space [34]. Since the whole dimension of the operator space \( \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) \) is \( D^4 \), these error operators \( \{E_k\}_k \) have very limited forms. However, in the present problem, the inverse of our copying protocol gives one noiseless channel as follows. If the channel \( \Lambda \) satisfies the above condition, the channel \( \Lambda \) can be written in the form \( \Lambda(\rho) = \sum_{ij} a_{ij} U_i \otimes U_i \rho U^\dagger_j \otimes U^\dagger_j \), where \( a_{ij} \) are appropriate complex coefficients. Then, by the inverse of the channel copying operation, there exists an encoding operation \( A \) satisfying the following relation:

\[
A^\dagger [\Lambda(A \rho A^\dagger)] A = \sum_{ij} a_{ij} A^\dagger (U_i \otimes U_i) A \rho A^\dagger (U^\dagger_j \otimes U^\dagger_j) A
\]

\[
= \sum_{ij} a_{ij} (U_i \otimes I) \rho (U^\dagger_j \otimes I).
\]

Thus, using an ancilla \( \sigma_0 \), the encoding operation \( A \) and the decoding operation \( A^\dagger \), we derive a noiseless channel in \( \mathcal{H}_2 \) as follows:

\[
\text{Tr}_1 A^\dagger [\Lambda(A(\sigma_0 \otimes \sigma) A^\dagger)] A = \sigma,
\]

where \( \sigma \) is an arbitrary input state. Similarly to the distillation case, as is shown later, when \( a_{ij} = \delta_{ij} / D \), this error correcting protocol attains the asymptotic optimal rate of transmitting a quantum state through the channel \( \Lambda \). That is, the transmission rate of this protocol is equal to the quantum capacity of this rate.

This fact can be seen from the correspondence between the quantum capacity and the distillable entanglement given in [2]. Thus, for a generalized Pauli’s channel, the quantum capacity coincides with the distillable entanglement of the corresponding state, which is the state derived as
the output state when we input a part of a maximally entangled state, i.e. \[ \sum_{i=0}^{D-1} \frac{1}{D} |\Psi_i\rangle \otimes |\Psi_i\rangle \langle \Psi_i| \otimes \langle \Psi_i| \]. Since our protocol is the optimal distillation protocol for this states, this channel coding protocol is also optimal.

Next, we apply this error correcting protocol to QKD. In the \( D \)-dimensional case, we apply the above encoding and decoding operations for \( D \)-dimensional version of the BB84 protocol [33]. Here, we fix noise operators as \( U_i \equiv X_i = FZ_iF^\dagger \) and encoding operation as \( A \equiv F \otimes F(CNOT)F^\dagger \otimes F^\dagger \), where \( F \equiv \sum_i |\tilde{i}\rangle \langle i| \) is the Fourier transformation and \( |\tilde{i}\rangle \equiv \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} \omega^{ij} |j\rangle \) is a Fourier-transformed basis. Then, we can easily see that these \( U_i \) and \( A \) satisfy Eq.(4.3) as follows:

\[
AU_i \otimes I A^\dagger = F \otimes F(CNOT)Z^i \otimes I(CNOT)^\dagger F^\dagger \otimes F^\dagger \\
= (FZ^i F^\dagger) \otimes (FZ^i F^\dagger) \\
= U_i \otimes U_i,
\]

where we use \( (CNOT)Z^i \otimes I(CNOT) = Z^i \otimes Z^i \) in the second line. For simplicity, we choose an ancilla \( \sigma_0 = |\tilde{0}\rangle \langle \tilde{0}| \). Then, applying the error correcting code to the BB84 protocol, we derive the following protocol.

**Protocol 1: State Preparation.** Alice randomly chooses a basis from \( \{\sum_k \omega^{-ik} |\tilde{k}\rangle \otimes |k\rangle\}_{i=0}^{D-1} \) and \( \{|\tilde{i}\rangle \otimes |i\rangle\}_{i=0}^{D-1} \), and randomly generate one state from the chosen basis.

**3: Decoding.** Bob randomly chooses a measurement basis from \( \{\sum_{j=0}^{D-1} |a\rangle \langle a| \otimes |a \oplus i\rangle \langle a \oplus i|\}_{i=0}^{D-1} \) and \( \{|I\otimes |\tilde{i}\rangle \langle i|\}_{i=0}^{D-1} \), and measure the states.

**4: Basis announcement.** Alice and Bob discard any digits where they prepared and measured in a different basis.

The step 1 can be decomposed into the following step 1A and 1B, and the step 3 can be decomposed into the following step 3A and 3B:

**1A: Preparation of the BB84 states.** Alice randomly chooses a basis from \( \{|i\rangle\}_{i=0}^{D-1} \) and \( \{|\tilde{i}\rangle\}_{i=0}^{D-1} \), and create a qubits from the chosen basis.

**1B: Encoding.** Alice encodes the qubits by our error correcting code, that is, applies the encoding
operation \( A = F \otimes F(\text{CNOT})F^\dagger \otimes F^\dagger \) with an ancilla qudit \(|\tilde{0}\rangle\).

3A: Error correction. Bob applies the decoding operation \( F \otimes F(\text{CNOT}^\dagger)F^\dagger \otimes F^\dagger \), and throw away the ancilla qudit. 3B: Detection. Bob randomly chooses a basis from \( \{|i\rangle\}_{i=0}^{D-1} \) and \( \{\tilde{i}\rangle\}_{i=0}^{D-1} \), and measure the decoded qudits. We can check the above reduction of the protocol as follows: From 1A and 1B to 1,

\[
\begin{align*}
A |\tilde{0}\rangle \otimes |\tilde{i}\rangle &= \sum_{ab} \begin{pmatrix} a - b \end{pmatrix} \begin{pmatrix} \bar{a} \end{pmatrix} \begin{pmatrix} \tilde{b} \end{pmatrix} \begin{pmatrix} |\tilde{0}\rangle \otimes \sum_{k} \omega^{-ik} |\tilde{k}\rangle \end{pmatrix} \\
&= \sum_{k} \omega^{-ik} |\tilde{k}\rangle \otimes |\tilde{k}\rangle,
\end{align*}
\]

and

\[
\begin{align*}
A |\tilde{0}\rangle \otimes |\tilde{i}\rangle &= \sum_{ab} \begin{pmatrix} a - b \end{pmatrix} \begin{pmatrix} \bar{a} \end{pmatrix} \begin{pmatrix} \tilde{b} \end{pmatrix} \begin{pmatrix} |\tilde{0}\rangle \otimes |\tilde{i}\rangle \end{pmatrix} \\
&= |\tilde{i}\rangle \otimes |\tilde{i}\rangle ;
\end{align*}
\]

From 3A and 3B to 3

\[
\begin{align*}
AI \otimes |i\rangle \langle i| A^\dagger &= \sum_{ckk'} A |\tilde{c}\rangle \langle \tilde{c}| \otimes |\tilde{k}\rangle \langle \tilde{k}'| \\
&= \frac{1}{D} \sum_{ckk'} \omega^{-i(k-k')} \begin{pmatrix} c - k \end{pmatrix} \begin{pmatrix} \bar{c} \end{pmatrix} \begin{pmatrix} \tilde{k} \end{pmatrix} \begin{pmatrix} \tilde{k}' \end{pmatrix} \\
&= \frac{1}{D^2} \sum_{ckk'ab} \omega^{-i(k-k')-(ak-bk')+c(a-b)} |a\rangle \langle b| \otimes |\tilde{k}\rangle \langle \tilde{k}'| \\
&= \frac{1}{D} \sum_{ckc'k/a} \omega^{-(a+i)(k-k')} |a\rangle \langle k| \otimes |\tilde{k}\rangle \langle \tilde{k}'| \\
&= \sum_{a} |a\rangle \langle a| \otimes |a \ominus i\rangle \langle -a \ominus i| ,
\end{align*}
\]

and

\[
\begin{align*}
AI \otimes |\tilde{i}\rangle \langle \tilde{i}| A^\dagger &= \sum_{c} \begin{pmatrix} c - \tilde{i}\end{pmatrix} \begin{pmatrix} \bar{c} \end{pmatrix} \begin{pmatrix} \tilde{i}\end{pmatrix} \langle \tilde{i}| \\
&= \frac{1}{D} \sum_{cab} \omega^{(a-b)(c-i)} |a\rangle \langle b| \otimes |\tilde{i}\rangle \langle \tilde{i}| \\
&= I \otimes |\tilde{i}\rangle \langle \tilde{i}| .
\end{align*}
\]
Therefore, if the noise of the quantum channel satisfies our assumption (that is, the noise operator $E_k$ can be written down as $E_k = \sum_{i=0}^{D-1} c_{ki} X^i \otimes X^i$) by means of the error correction code, we can realize the noiseless QKD by the above protocol.

4.6 Conclusion

In this chapter, we focused on a set consisting of several maximally entangled states in a prime-dimensional system. In this case, we completely characterized the local copiability and showed the relation between the locally copiability and the local discriminability. In sections 4.3 and 4.4, we proved that such a set is locally copiable, if and only if it has a canonical Bell form and simultaneous Schmidt decomposable (Theorem 4.7). This theorem yields the following two conclusions. First, as well as the maximal size of local distinguishable sets, the maximal size of locally copiable sets is equal to the dimension of the local space $D$. This maximal size is the square root of the maximal size without the LOCC restriction. Second, as we can see in Fig.4.1, when such a set is locally copiable, it is also one-way locally distinguishable but the opposite direction is not true. In other words, at least in prime-dimensional systems, the local copying is more difficult than the one-way local discrimination for a set of maximally entangled states.

In the case of the local discrimination, a simultaneous Schmidt decomposable set is locally distinguishable. However, if such a set of states does not have a canonical Bell form, the set is not locally copiable. We can interpret the above fact as follows. A simultaneous Schmidt decomposable set does not possess non-locality beyond individual entanglement with respect to the local discrimination. On the other hand, if such a set does not have canonical Bell form, such a set still has non-locality with respect to the local copying. In other words, we can conclude that the lack of algebraic structure
causes extra non-locality of a set with respect to the local copying.

Although we only treated orthogonal sets of maximally entangled states in this chapter, our result of Fig.4.1 is also regarded as a classification of sets of the Schmidt basis by their non-locality. Therefore, in the case of a set of general entangled states, the structure of non-locality of sets of Schmidt basis may be similar to Fig.4.1, though it possesses additional non-locality which originates in various Schmidt coefficients. Therefore, our result may be useful as the base for more general discussion of non-locality problems of the Schmidt basis, especially for general discussion of local copying problems.

In Section 4.4, we showed that our results and protocol of local copying can be interpreted as results of several different and closely related forms of quantum information processing, namely, the channel copying, the entanglement distillation, the error correction, and the quantum key distribution. These close relations with many other protocols suggest the importance of the local copying as a fundamental protocol of non-local quantum information processing.

Finally, we should mention a remaining open problem. In this chapter, we showed the necessity of the form of states (4.15) for LOCC copying only in prime-dimensional local systems. However, we restrict this dimensionality only by the technical reason and this restriction has no physical meaning. The validity of Theorem 4.7 for non-prime-dimensional systems still remains as an open question.

After finishing the first draft [35], the authors found a related paper [36] which contains a different approach to Theorem 4.6, and also found a paper which extends our result to a set of non-maximally entangled states based on our main theorem (Theorem 4.4) [37].
Appendix: Proof of Lemma 4.2

In this appendix, we prove Lemma 4.2 by induction.

**Proof**

First, in Eq. (4.20) by choosing \( c = a_1 \oplus b_1 = a_2 \oplus b_2 \), we have

\[
\delta_{b_1 b_2} U_{cc} = U_{c \oplus b_1} \oplus U_{c \oplus b_2} U_{b_1 b_2}.
\]  

(4.30)

In addition, choosing \( b_1 \neq b_2 \), we derive

\[
U_{c \oplus b_1} \oplus U_{c \oplus b_2} U_{b_1 b_2} = 0
\]  

(4.31)

for all \( c \). The above equation means

\[
b_1 \neq b_2 \implies U_{b_1 b_2} = 0 \quad \text{or} \quad \forall c, U_{c \oplus b_1} \oplus U_{c \oplus b_2} = 0.
\]  

(4.32)

By means of the above fact, we prove

\[
U_{b_b \oplus n} = U_{b_b \ominus n} = 0
\]  

(4.33)

for all \( b \) and for all \( 0 \leq n \leq D - 1 \) by induction with respect to the integer \( n \).

First, we prove \( U_{b_b \ominus 1} = 0 \) and \( U_{b_b \oplus 1} = 0 \) for all \( b \) by contradiction, and then, under the assumption of \( U_{b_b \ominus k} = U_{b_b \oplus k} = 0 \) for all \( b \) and for all \( k \leq n - 1 \), we prove \( U_{b_b \ominus n} = U_{b_b \oplus n} = 0 \) for all \( b \).

(Proof of \( U_{b_b \ominus 1} = 0 \) for all \( b \).)

We prove \( U_{b_b \ominus 1} = 0 \) for all \( b \) by contradiction. We assume that there exists \( b_1 \) such that \( U_{b_1 \ominus b_1 \ominus 1} \neq 0 \). Then, Eq. (4.32) implies \( U_{b_b \ominus 1} = 0 \) for all \( b \). In order to show the contradiction, we prove \( U_{b_b \ominus k} = 0 \) for all \( b \) and \( k \) by induction concerning \( k \).

We assume \( U_{b_b \ominus k} = 0 \) for all \( b \) and show \( U_{b_b \ominus k \ominus 1} = 0 \) for all \( b \). By substituting \( a_1 = b \ominus b_1, a_2 = b \ominus b_1 \ominus k \ominus 1, \) and \( b_2 = b_1 \ominus 1 \), Eq. (4.20) guarantees that

\[
\Box_{b_1 b_1 \ominus 1} U_{b_b \ominus k} = U_{b_b \ominus b_1, b_b \ominus k \ominus 1} U_{b_1 b_1 \ominus 1}.
\]  

(4.34)
Then, by substituting $U_b b \oplus k = 0$, Eq. (4.34) guarantees that

$$U_{b \oplus b_1 b \oplus k \oplus 1} U_{b_1 b_1 \oplus 1} = 0$$

(4.35)

for all $b$. Since we assumed $U_{b_1 b_1 \oplus 1} \neq 0$, Eq. (4.35) guarantees that $U_{b \oplus b_1 b \oplus k \oplus 1} = 0$ for all $b$, that is, $U_b b \oplus k \oplus 1 = 0$ for all $b$.

Therefore, by induction with respect to $k$, we derive $U_b b \oplus k = 0$ for all $b$ and $k$ by induction. This is contradiction to the assumption that $U_{b_1 b_1 \oplus 1} \neq 0$. Thus, we have $U_b b \oplus 1 = 0$ for all $b$.

(Proof of $U_b b \oplus 1 = 0$ for all $b$)

Similarly, we can prove $U_b b \oplus 1 = 0$ for all $b$ by contradiction as follows. Suppose that there exists $b_1$ such that $U_{b_1 b_1 \oplus 1} = 0$. Then, Eq. (4.32) implies $U_b b \oplus 1 = 0$ for all $b$. In order to show the contradiction, we prove $U_b b \oplus k = 0$ for all $b$ and $k$ by induction with respect to $k$. Equation (4.20) implies

$$\Xi_{b_1 b_1 \oplus 1} U_b b \oplus k = U_{b \oplus b_1 b \oplus k \oplus 1} U_{b_1 b_1 \oplus 1}.$$  

(4.36)

Thus, if $U_b b \oplus k = 0$ for all $b$, we derive $U_b b \oplus k \oplus 1 = 0$ for all $b$. Therefore, by induction, we have $U_b b \oplus k = 0$ for all $k$ and $b$. This is contradiction to the assumption $U_{b_1 b_1 \oplus 1} \neq 0$. Therefore, $U_b b \oplus 1 = 0$ for all $b$.

(Proof of $U_b b \oplus n = U_b b \oplus n = 0$ for all $b$ and $0 \leq n \leq D - 1$)

As the final step, in order to prove $U_b b \oplus n = U_b b \oplus n = 0$ for all $b$ and $1 \leq n \leq D - 1$, we use induction with respect to $n$ again. We assume $U_b b \oplus k = U_b b \oplus k = 0$ for all $k \leq n - 1$ and show $U_b b \oplus n = U_b b \oplus n = 0$ for any $b$ by contradiction. Assume that there exists $b_1$ such that $U_{b_1 b_1 \oplus n} \neq 0$. Then Eq. (4.32) implies that $U_b b \oplus n = 0$ for all $b$. To show the contradiction, under the above assumption, we prove $U_b b \oplus k = 0$ for all $b$, all $0 \leq k \leq ln - 1$, and all $l$, (that is, for all $k$) by induction with respect to $l$. We assume $U_b b \oplus k = 0$ for all $b$ and all $(l - 1)n \leq k \leq ln - 1$, and show $U_b b \oplus k = 0$ for all $b$ and all
ln \leq k \leq (l + 1)n - 1. By substituting \( a_1 = b \ominus b_1 \), \( a_2 = b \ominus b_1 \oplus n \oplus k \), and \( b_2 = b_1 \ominus n \), Eq. (4.20) implies

\[
\Xi_{b_1}^{b} b_{b_1}^{k} U_{b_1} b_{b_1}^{k} = U_{b_1} b_{b_1}^{k} \ominus b_{b_1}^{n} \oplus k U_{b_1} b_{b_1}^{n}.
\]  

(4.37)

By substituting \( U_{b_1} b_{b_1}^{k} = 0 \) for all \( b \) and \((l - 1)n \leq k \leq ln - 1\), we have \( U_{b_1} b_{b_1}^{k} = 0 \) for all \( ln \leq k \leq (l + 1)n - 1 \) and \( b \). Therefore, by induction with respect to \( l \), we derive \( U_{b_1} b_{b_1}^{k} = 0 \) for all \( b \) and all \( 1 \leq k \). This contradicts \( U_{b_1} b_{b_1}^{n} \neq 0 \). Therefore, \( U_{b_1} b_{b_1}^{n} = 0 \) for all \( b \).

By means of the same discussion for \( U_{b_1} b_{b_1}^{n} \), we can prove \( U_{b_1} b_{b_1}^{n} = 0 \) for all \( b \); what we only have to do is changing \( \ominus \) to \( \oplus \) in the above proof. Finally, by the mathematical induction with respect to \( n \), we prove \( U_{b_1} b_{b_1}^{n} = U_{b_1} b_{b_1}^{n} = 0 \) for all \( b \) and all \( 0 \leq n \leq D - 1 \). Therefore, \( U_{ab} \) is a diagonal matrix. \( \square \)
Bibliography


[27] S. Lang, *Algebra*, Section 4 of Chapter 1, Addison-Wesley.


[34] X.B. Wang, quant-ph/0406100

122


Chapter 5

Conclusion

In this thesis, we have concentrated on the three types of quantum non-locality, namely, the entanglement, the non-locality with respect to LOCC discrimination, and the non-locality with respect to LOCC copying. We have studied the property of these three types of non-locality and the relation among them.

In Chapter 2, we concentrated on the relation between the LOCC discrimination and the entanglement. We gave a bound of the LOCC discrim- inability which is determined only by the value of entanglement function of each candidate state. That is, we showed that upper bounds of the size of the LOCC distinguishable set of states are given by the geometric measures of entanglement, the relative entropy of entanglement, and the global robustness of entanglement, which are entanglement measures defined as the distance from the set of all separable states ("distance-like measure of entanglement"). This result implies that the presence of entanglement guarantees at least the minimal level of non-locality with respect to the LOCC discrimination. Since this result is valid for all finite sets of multipartite states, this result also gives an operational meaning for the distance-like measures of entanglement in terms of the LOCC discrimination task.

In Chapter 3, we analyzed the difference in the LOCC discriminability in
the case of permitting two-way classical communications (two-way LOCC) and in the case of permitting only one-way classical communications (one-way LOCC). As a result, we derived an upper bound of the size of a one-way locally distinguishable set for bipartite states and a set of three partite pure states. After that, in a two-qubit system, by constructing a concrete two-way local discrimination protocol, we showed that two-way classical communications remarkably improves the local discriminability in comparison with the local discrimination by one-way classical communications.

In Chapter 4, we focused on the LOCC copying and its relation to the LOCC discrimination. We analyzed the LOCC copying for an orthogonal set of maximally entangled states and shown that for such a set in a system with prime-dimensional local systems, a necessary and sufficient condition of the perfect LOCC copying is that the set is a simultaneously Schmidt diagonalizable subset of canonical Bell bases. Moreover, we showed that for a set of maximally entangled states, the LOCC copying is strictly more difficult than the LOCC discrimination by using the fact that a simultaneously Schmidt diagonal set of states is always locally distinguishable.
Publications

Main results of the dissertation have been published as the following list of papers, and also presented in both international and domestic conference.


Acknowledgment

First, I would like to thank my supervisor Mio Murao for all discussions and comments throughout this research and for her support and encouragement throughout all my life in the doctoral course. Second, I would like to thank Prof. Masahito Hayashi (ERATO, JST) and Dr. Damian Markham for discussions and checking the paper through our research and for their encouragement. I would like to thank Dr. Shashank Virmani (Imperial College) for discussions throughout this research. These four persons are the coworkers of this research and without one of them I could not finish this thesis.

I am grateful to Prof. Keiji Matsumoto (NII, ERATO) and Prof. Martin Plenio (College) for useful discussions. I would like to thank Mr. Akihito Soeda for checking English. I would like to thank Prof. Hiroshi Imai for his encouragement. I would like to acknowledge assistance and supports from my colleagues, Yu Tanaka, Yoshifumi Nakata, and Keiichi Shigechi. Finally, I thank my family and all my friends for their continuous supports and encouragements.

This work has been supported by the Asahi Glass Foundation and JSPS Research Fellowships for Young Scientists.