学位論文

Information Theoretical Analysis of Distributed Quantum Computation
(分散型量子計算の情報論的解析)

平成 26 年 12 月博士（理学）申請

東京大学大学院理学系研究科
物理学専攻

若桑　江友里

Abstract

Distributed quantum computation is a task in which many distant parties perform a large quantum computation in collaboration with each other, by using classical communication, quantum communication or shared entanglement as resources. In this thesis, we consider one of the simplest nontrivial task of distributed quantum computation, that is, implementation of bipartite unitary by LOCC (local operations and classical communication) assisted by shared entanglement. Here, two input states are distributed over two distant parties. Contrary to previous researches, we consider an asymptotic scenario in which the two parties perform the same bipartite unitary on infinitely many independent input pairs. We analyze the minimum amount of resources required for the task by exploiting concepts and techniques developed in the field of quantum Shannon theory, which has originally been developed to analyze quantum communication.

As a tool for analyzing distributed quantum computation, we first introduce a task that we call Markovianization. Markovianization is a task in which a tripartite pure state is transformed to a state called quantum Markov chain, by a randomizing operation on one of the three subsystems. We derive the minimum cost of randomness required for Markovianization (Markovianizing cost) by exploiting a novel decomposition of a Hilbert space introduced in the context of quantum cryptography.

We then apply the obtained result to distributed quantum computation. We mainly consider implementation of bipartite unitaries by entanglement-assisted LOCC protocols consisting of three steps, called *one-round protocols*. This is the first nontrivial case because it is not possible to implement bipartite unitaries by protocols consisting of only one or two steps. Our main result is that we derive the minimum costs of entanglement and classical communication. We show that the minimum costs are given by the Markovianizing cost of a tripartite state associated with the unitary. The result indicates that information about input states of computation is divided into three components: one of which only information of amplitude in a basis is required for computation (information of phase is not), another of which both amplitude and phase information are required, and the other of which no information is involved in computation. Our construction of a protocol that requires the smallest amount of resources is based on this decomposition. We also propose a method to compute the Markovianizing cost of a state associated with the unitary.

We also consider two-round protocols, that is, protocols consisting of concatenation of two one-round protocols, for two-qubit controlled unitaries. We show that, for a particular class of two-qubit controlled unitaries, the minimum entanglement

cost in two-round protocols is strictly smaller than that in one-round protocols. This result is the first reported case that indicates a property of a trade-off relation between entanglement requirement and the number of turns in entanglement-assisted LOCC tasks.

The research in this thesis opens a possibility to build a new bridge between quantum computation theory and quantum communication theory, which are the two largest subfield of quantum information theory, by analyzing a problem in the former using concepts and techniques developed in the latter.

Contents

Chapter 1

Introduction

Recent developments in the field of quantum information theory have revealed that the fundamental limit on information processing strongly depend on the underlying laws of physics. We can perform various kinds of information processing tasks more efficiently in the framework of quantum mechanics than in that of classical physics [1]. For example, we can construct cryptographic protocols that are strongly secure against eavesdroppers within the framework of quantum mechanics, which is known to be impossible in classical physics [2]. Quantum computers can also solve problems much faster than classical computers can. One of the most famous is Shor's algorithm [3], by which we can solve in polynomial time the problem of finding the prime factors of an integer, which cannot be achieved by any known algorithms for classical computers in polynomial time.

One of the prime objects in the research field of quantum information theory is to find various information processing tasks that are implementable more efficiently within the framework of quantum mechanics than that is realizable classically, and to investigate how to realize it in the real world. However, it is not the only reason why we study quantum information theory. A more fundamental motivation behind this research field, which has recently been growing rapidly, is to understand the laws of physics from an information theoretical perspective.

As is well known, quantum mechanics predicts various counterintuitive phenomena, which cannot be understood through intuitions distilled out of our daily life. To name a few, we know the wave-particle duality, the tunneling effect, and the discreteness of energy. Most of the significant developments of quantum technology in the past half century take advantages of such quantum mechanical behavior of physical entity. However, a more essential difference of quantum mechanics from classical physics is in its probabilistic nature. In [4], J. Bell found a nonlocal property of quantum states that cannot be consistently explained within the framework of classical physics. That is, he showed that the correlation of measurement results performed independently on two or more spatially separated quantum systems can be inconsistent with a basic assumption of classical physics, that is, the assumption of local reality. This result indicates that, viewing as a theoretical framework for describing probabilities, quantum mechanics, which is constructed on an abstract mathematical framework of Hilbert spaces, is essentially different from classical physics.

It is likely that, in many cases, advantages of quantum information processing over classical counterpart originates from this probabilistic nature. Thus, by investigating what information processing tasks are possible and what are not in classical and quantum mechanics, we could acquire more profound understanding about the difference and similarity between classical and quantum mechanics, particularly from a probability theoretical aspects. In other words, we can better understand the laws of physics by investigating the fundamental limit on information processing.

Among various information processing tasks, the fundamental limits of communication are analyzed by Shannon's theory of information [5]. However, it is not clear whether Shannon's information theory, which is based on classical probability theory, can cover all cases where quantum effects are not negligible. For example, when we try to transmit classical information as efficiently as possible by using lasers, it is inevitable to consider effects like superposition principles and uncertainty principles, which are peculiar to quantum mechanics. To overcome this difficulty, a generalization of Shannon's information theory to quantum mechanics was born in [6], and has been developed. In [6], A. Holevo analyzed a fundamental limit on the capacity of classical information transmission by using quantum systems as media. Although only transmission of classical information using quantum systems is initially considered, the situation drastically changed after a seminal paper by Schumacher [7]. In [7], he considered transmission of quantum states in a framework similar to Shannon's theory, and derived the 'quantum coding theorem' analogous to Shannon's source coding theorem. Since then, this research field largely developed including transmission of quantum information, transformation of entanglement, etc. The theoretical framework of quantum generalization of Shannon's information theory is now called quantum Shannon theory [8, 9].

The major interests in quantum Shannon theory are to reveal the interconvertibility among different types of resources, such as quantum channels, noisy and noiseless entanglement, classical channels, and classical correlations. In the asymptotic limit of infinite copies, the rates at which those resources are interconvertible are derived in coding theorems.

A characteristic feature of classical and quantum Shannon theory is its universality. For efficient transmission of classical or quantum information, there are various, in principle infinite, ways of coding it. Despite the infinity of coding methods, the fundamental limit must be universally applicable regardless of a particular way of coding. Remarkably, the coding theorems teach us the fundamental limit that does not depend on a particular way of coding.

Given its universality, it would be fruitful to apply the theoretical concepts and tools developed in quantum Shannon theory beyond communication tasks. A possible direction would be its application to quantum computation. Through investigations on the possibility of performing quantum computation within the universal framework of quantum Shannon theory, we could more profoundly understand the fundamental limit on the possibility of manipulating quantum systems. In this thesis, we consider a model of quantum computation called *distributed quantum computation*, and analyze it in terms of quantum Shannon theory.

The concept of distributed quantum computation was first introduced by [26], as a method to perform a large scale quantum computation by combining small quantum computers, which are realizable by current (or near-future) experimental technologies. In distributed quantum computation, many distant parties perform a large quantum computation in collaboration with each other, by using classical communication, quantum communication and shared entanglement as resources. To exploit the full power of the small quantum computers, it is necessary to construct an efficient protocol for a given computation to be implemented. In particular, the costs of entanglement and quantum communication should be reduced as much as possible, because preparing entanglement and sending quantum information is experimentally difficult and costly.

One of the most extensively investigated tasks in distributed quantum computation is implementation of bipartite unitaries by local operations and classical communication (LOCC) assisted by shared entanglement. Here, two distant parties, say Alice and Bob, have quantum systems A and B in an unknown state $|\varphi\rangle^{AB}$, and perform a known unitary U^{AB} by LOCC using some resource entanglement shared in advance. Although this task can be implemented simply by using quantum teleportation, it was shown that the costs of resources of entanglement and classical communication can be reduced by constructing an efficient protocol, depending on the unitary to be implemented [26]. The following two questions then naturally arise: (i) How can we find efficient protocols which consume less resources for a given bipartite unitary? and (ii) What are the minimum costs of resources required for implementing that unitary? These questions are of great importance, not only from a practical point of view, but also from a fundamental viewpoint of understanding quantum computation, since bipartite unitary interaction is a core element of quantum computers. For example, in the circuit model, which is one of the most basic model of quantum computation, any computation can be described as a combination of the following three operations: (i) preparing an initial state of a multipartite quantum system in a fixed basis, (ii) applying a sequence of bipartite unitary interactions on the system, and (iii) performing a measurement on the system in a fixed basis. It is known that operations (i) and (iii) can be fixed independently of computation to be implemented. Therefore, the whole information of the computation is embedded in the sequence of the bipartite unitary interactions.

Although the two questions raised above have been addressed e.g. in [26, 27, 28, 29, 30, 31, 32, 34], most of the studies so far assume particular forms of the resource entanglement or of the bipartite unitary to be implemented. Consequently, a general method to address these problems is yet to be discovered.

In this thesis, we address the above questions by using concepts and techniques developed in quantum Shannon theory. We consider an asymptotic scenario in which the two parties perform the same bipartite unitary, simultaneously on each of a sequence of input pairs obtained from a completely random i.i.d. (independent and identically distributed) quantum information source. Our approach is complementary to previous approaches which have only dealt with single-shot cases. In an asymptotic limit of infinite pairs and vanishingly small error, we analyze the minimum costs of entanglement and classical communication per copy required for this task.

We mainly focus on protocols consisting of one-round LOCC, in which Alice first performs a measurement, sends the result to Bob, Bob performs a measurement, sends the result to Alice, and Alice performs a quantum operation. This is the first nontrivial case because any bipartite unitary cannot be implemented by protocols consisting of fewer steps. Moreover, analysis of one-round protocols is suitable for investigating bipartite unitary interaction in terms of "information flow". This is because interaction between information of Alice's input state and that of Bob's input occurs only at the step of Bob's measurement. Hence, among all pieces of information of Alice's input state, a piece of information which is involved in computation must become locally accessible for Bob after Alice's measurement and the subsequent classical communication. Therefore, by investigating one-round protocols, we could classify information of Alice's input state into two parts: one which is involved in the computation to be implemented, and the other which is not. Such a classification would be useful for constructing efficient protocols which consume less resources, as well as deriving the minimum cost of resources.

As a tool for analyzing the minimum costs of entanglement and classical communication in one-round protocols, we first introduce and analyze a task that we call Markovianization. Markovianization is a task in which a tripartite quantum state is transformed to a quantum Markov chain by a randomizing operation on one of the three subsystems. We consider cases where the initial state is a tensor product of n copies of a tripartite state ρ^{ABC}, and is transformed to a quantum Markov chain conditioned by B^n with a small error, by a random unitary operation on A^n. In an asymptotic limit of infinite copies and vanishingly small error, we analyze the

Markovianizing cost, that is, the minimum cost of randomness per copy required for Markovianization. For tripartite pure states, we derive the Markovianizing cost, that is, the minimum cost of randomness required for this task.

We then apply the obtained results and show that any two-round protocol for implementing a bipartite unitary can be described as a combination of Markovianization of a tripartite state associated with the unitary, followed by a protocol callled *quantum state merging* [16, 17]. Consequently, the minimum costs of resources in two-round protocols are obtained as the sum of ones in those two subroutines. Our main result is that we derive the minimum costs of resources of entanglement and classical communication required in one-round protocols are given by the Markovianizing cost of a state associated with the unitary.

We also analyze a particular example of two-qubit controlled-unitaries. First, by applying the obtained results, we derive the minimum entanglement cost for implementing this class of unitary by one-round protocols. Second, we consider two-round protocols, that is, protocols consisting of concatenation of two one-round protocols. We show that, for a particular class of two-qubit controlled unitaries, the minimum entanglement cost in two-round protocols is strictly smaller than that in one-round protocols.

This thesis is organized as follows. In Chapter 2, we review basic mathematical tools that will be used in the following chapters. In Chapter 3, we explain core elements of quantum Shannon theory that will also be used in the following chapters. In Chapter 4, we introduce and analyze the task of Markovianization. In Chapter 5, we derive the optimal rates of costs of entanglement and classical communication required in any protocol composed of minimal number of steps. In Chapter 6, we introduce two methods to compute the Markovianizing cost of a tripartite state associated with a unitary. In Chapter 7, we consider a particular example of two-qubit controlled-unitaries. We close with a summary in Chapter 8.

Notations and Abbreviations

- $\mathcal{H}^A$: Hilbert space associated with a quantum system labeled by A
- d_A, $\dim\mathcal{H}^A$: The dimension of the Hilbert space $\mathcal{H}^A$
- ρ^A : A density operator that represent a state of quantum system A
- $\mathcal{L}(\mathcal{H})$: The set of linear operators on a Hilbert space $\mathcal{H}$
- $\mathcal{B}(\mathcal{H})$: The set of bounded linear operators on a Hilbert space $\mathcal{H}$
- $\mathcal{P}(\mathcal{H})$: The set of positive semidefinite linear operators on a Hilbert space $\mathcal{H}$
- $\mathcal{S}(\mathcal{H})$: The set of positive semidefinite linear operators on a Hilbert space $\mathcal{H}$, the trace of which are equal to 1
- $\mathcal{E}^A$: A quantum operation on quantum system A
- $H(X)$, $H(\{p_x\}_x)$: The Shannon entropy of a random variable X with probability distribution $p(x) = \Pr\{X = x\}$
- $I(X : Y)$: The (classical) mutual information between two random variables X and Y
- $S(\rho^A)$, $S(A)_\rho$: The von Neumann entropy of system A in state ρ^A
- $I(A : B)_\rho$: The quantum mutual information between system A and B in the state ρ^{AB}
- $I(A : B|C)_\rho$: The quantum conditional mutual information between system A and B conditioned by C, in the state ρ^{ABC}
- Φ_d : A maximally entangled state with Schmidt rank d
- $|\Psi(U)\rangle^{ABR_AR_B}$: A four-partite pure state associated with a bipartite unitary U, defined by $|\Psi(U)\rangle^{ABR_AR_B} = U^{AB}|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$
- $M^{A\to A'}$: A linear map from $\mathcal{H}^A$ to $\mathcal{H}^{A'}$

- A^n, $\bar{A}$: A system composed of n identical systems A

- $F(\rho, \sigma)$: The fidelity of two states ρ and σ, defined as $F(\rho, \sigma) = (\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})^2$

- $\eta(\epsilon)$: A proper continuous concave function of $\epsilon > 0$ that satisfy $\lim_{\epsilon \to 0} = 0$ and does not depend on n or the dimension of Hilbert spaces.

- $M_{A|B}(\Psi^{ABC})$: The Markovianizing cost of a tripartite state Ψ^{ABC} on A conditioned by B, defined by Definition 19

- $M(U)$: The Markovianizing cost of a bipartite unitary U, defined by Definition 29

We abbreviate $(M^A \otimes I^B)|\psi\rangle^{AB}$ as $M^A|\psi\rangle^{AB}$ and $(M^A \otimes I^B)\rho^{AB}(M^A \otimes I^B)^\dagger$ as $M^A\rho^{AB}M^{A\dagger}$. For $|\phi\rangle^{AB}$, ϕ^A and $\mathrm{Tr}_B[|\phi\rangle]$ represents $\mathrm{Tr}_B[|\phi\rangle\langle\phi|]$. When $\mathcal{E}^A$ is a quantum operation on A, we abbreviate $(\mathcal{E}^A \circ \mathrm{id}^B)(\rho^{AB})$ as $\mathcal{E}^A(\rho^{AB})$. $\mathcal{E}(|\phi\rangle)$ represents $\mathcal{E}(|\phi\rangle\langle\phi|)$. We abbreviate $F(\rho, |\phi\rangle\langle\phi|)$ as $F(\rho, |\phi\rangle)$.

Chapter 2

Mathematical Preliminaries

In this chapter, we review several mathematical tools used in quantum information theory. First, we introduce a mathematical formalism to describe operations and measurements, and review basic properties of bipartite states (Section 2.1). We then discuss two functions to quantify 'how different' two states are, and their properties as well (Section 2.2). In Section 2.3, we describe properties of entropic functions such as the von Neumann entropy and the quantum mutual information, which are extensively used in the following chapters. In Section 2.4, we introduce a method of decomposing a Hilbert space in such a way that it completely characterizes a set of operations that does not change a given set of states. The decomposition plays a central role in the following chapters. Contents in this chapter are mainly based on [1, 10, 18].

2.1 Quantum States, Operations and Measurements

A quantum system is described by a Hilbert space $\mathcal{H}$. In this thesis, we only consider cases where the dimension of the Hilbert space is finite. A Hilbert space corresponding to a system labeled by A is denoted by $\mathcal{H}^A$.

A quantum operation on system A is described by a linear map $\mathcal{E} : \mathcal{S}(\mathcal{H}^A) \rightarrow \mathcal{S}(\mathcal{H}^{A'})$ that satisfies the following conditions,

$$\begin{aligned}
&\forall T \in \mathcal{L}(\mathcal{H}^A);\ \ \mathrm{Tr}[\mathcal{E}(T)] = \mathrm{Tr}[T] \quad \text{(trace-preserving)}\\
&\forall\, \text{system } B,\ \forall \rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B);\ \ (\mathcal{E}^A \circ \mathrm{id}^B)(\rho) \geq 0 \quad \text{(completely positive).}
\end{aligned}$$

A' is the output system which is in general different from the input system A. A linear map $\mathcal{E}$ satisfying these two conditions is called a *CPTP map.* Any CPTP map is decomposed as

$$\mathcal{E}(\tau) = \sum_k E_k \tau E_k^{\dagger}, \tag{2.1}$$

where E_k are linear operators from $\mathcal{H}^A$ to $\mathcal{H}^{A'}$ that satisfy

$$\sum_k E_k^\dagger E_k = I. \tag{2.2}$$

The decomposition given by (2.1) is called the *Kraus representation* of $\mathcal{E}$. Any CPTP map is also described as

$$\mathcal{E}(\tau) = \mathrm{Tr}_E[V\tau V^\dagger], \tag{2.3}$$

where V is an isometry from $\mathcal{H}^A$ to $\mathcal{H}^{A'} \otimes \mathcal{H}^E$. We can construct V from (2.1) as

$$V = \sum_k |k\rangle^E \otimes E_k^{A\to A'}. \tag{2.4}$$

It is easy to check that V is an isometry, namely,

$$V^\dagger V = \sum_{k,k'} \langle k|k'\rangle E_k^\dagger E_{k'} = \sum_k E_k^\dagger E_k = I^A, \tag{2.5}$$

and that V satisfies (2.3) as

$$\mathrm{Tr}_E[V\tau V^\dagger] = \mathrm{Tr}_E\left[\sum_{k,k'} |k\rangle\langle k'| \otimes E_k\tau E_{k'}^\dagger\right] = \sum_{k,k'} \delta_{k,k'} E_k \tau E_{k'}^\dagger = \sum_k E_k\tau E_k^\dagger.$$

A measurement on A is described by a set $\{\mathcal{E}_k\}_{k=1}^K$ of linear maps $\mathcal{E}_k : \mathcal{S}(\mathcal{H}^A) \to \mathcal{S}(\mathcal{H}^{A'})$ satisfying the following conditions,

$$\begin{aligned}&\sum_{k=1}^K \mathcal{E}_k \text{ is trace-preserving}\\ &\mathcal{E}_k \text{ is completely positive for each } k.\end{aligned}$$

Index k represents a measurement result, and K is the number of measurement results. $\mathcal{E}_k$ is decomposed as

$$\mathcal{E}_k(\tau) = \sum_{j=1}^{J_k} M_{kj}\tau M_{kj}^\dagger, \tag{2.6}$$

where E_{kj} are linear operators from $\mathcal{H}^A$ to $\mathcal{H}^{A'}$ satisfying

$$\sum_{k=1}^K \sum_{j=1}^{J_k} M_{kj}^\dagger M_{kj} = I. \tag{2.7}$$

In this thesis, we only consider cases where $J_k = 1$ for all k. Then a measurement on A can be identified with a set of linear operators $\{M_k^{A\to A'}\}_{k=1}^K$ that satisfy

$$\sum_{k=1}^K M_k^\dagger M_k = I^A. \tag{2.8}$$

For any measurement $\{M_k^{A\to A'}\}_k$, there exists an isometry $W : \mathcal{H}^A \to \mathcal{H}^{A'} \otimes \mathcal{H}^E$ such that $M_k = \langle k|^E W$, where $\{|k\rangle\}_k$ is an orthonormal basis of $\mathcal{H}^E$. Indeed, let $W = \sum_k |k\rangle^E \otimes M_k^{A\to A'}$. This is an isometry because

$$W^\dagger W = \sum_{k,k'} \langle k|k'\rangle M_k^\dagger M_{k'} = \sum_k M_k^\dagger M_k = I^A, \tag{2.9}$$

and satisfies $M_k = \langle k|^E W$. The correspondence from $\{M_k^{A\to A'}\}_k$ to $\{|k\rangle\}_k$ and W is called the *Naimark extension* of $\{M_k^{A\to A'}\}_k$.

An example of quantum operations is a *random unitary operation* defined by

$$\mathcal{E}(\tau) = \sum_x p_x U_x \tau U_x^\dagger, \tag{2.10}$$

where U_x are unitaries and $\{p_x\}_x$ is a probability distribution. A useful property of random unitary operations is that it preserves the completely mixed state as

$$\mathcal{E}\left(\frac{1}{d}I\right) = \frac{1}{d}I, \tag{2.11}$$

where d denotes the dimension of the Hilbert space of the system. Another example is the *complete dephasing channel* defined by

$$\mathcal{E}(\tau) = \sum_{j=1}^{d} |j\rangle\langle j|\tau|j\rangle\langle j| \quad (\forall \tau \in \mathcal{S}(\mathcal{H})), \tag{2.12}$$

where $\{|j\rangle\}_j$ is an orthonormal basis of $\mathcal{H}$. This channel can also be described as a random unitary operation as

$$\mathcal{E}(\tau) = \frac{1}{d}\sum_{k=1}^{d} U_k \tau U_k^\dagger, \tag{2.13}$$

where

$$U_k = \sum_{j=1}^{d} e^{\frac{2\pi jki}{d}} |j\rangle\langle j|. \tag{2.14}$$

Any bipartite pure state $|\psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is written as

$$|\psi\rangle = \sum_i \sqrt{p_i}|e_i\rangle^A |e_i'\rangle^B, \tag{2.15}$$

where $\{|e_i\rangle\}_i$ and $\{|e_i'\rangle\}_i$ are orthonormal bases of $\mathcal{H}^A$ and $\mathcal{H}^B$, respectively, and $\{p_i\}_i$ is a probability distribution. $\{|e_i\rangle\}_i$ and $\{|e_i'\rangle\}_i$ are called the *Schmidt basis* of $|\psi\rangle$. We omit the symbol of the tensor product, and denote $|e_i\rangle^A \otimes |e_i'\rangle^B$ as $|e_i\rangle^A |e_i'\rangle^B$.

(2.15) is called the *Schmidt decomposition* of $|\psi\rangle$. Reduced states of $|\psi\rangle$ on either subsystem A and B are given by

$$\psi^A = \sum_i p_i |e_i\rangle\langle e_i|, \quad \psi^B = \sum_i p_i |e'_i\rangle\langle e'_i|, \tag{2.16}$$

respectively. A bipartite pure state $|\psi\rangle^{AB}$ is called a *purification* of ρ^A if it satisfies $\psi^A = \rho^A$. For any ensemble of pure states $\{q_k, |\phi_k\rangle\}_k$ on A such that

$$\psi^A = \sum_k q_k |\phi_k\rangle\langle \phi_k|, \tag{2.17}$$

where $\{|\phi_k\rangle\}_k$ are not necessarily orthogonal, there exists a measurement $\{M_k\}_k$ on B such that

$$\mathrm{Tr}_B[M_k^B |\psi\rangle\langle\psi| M_k^{\dagger B}] = q_k |\phi_k\rangle\langle\phi_k|. \tag{2.18}$$

This can be shown by the following manner. Let $c_{j|k} = \langle j|\phi_k\rangle$. Then $\{M_k\}_k$ defined by

$$M_k := |\tilde{\phi}_k\rangle\langle\tilde{\phi}_k|, \quad |\tilde{\phi}_k\rangle := \sqrt{p_k} \sum_j \sqrt{q_j^{-1}} c_{j|k}^* |j\rangle \tag{2.19}$$

satisfies (2.18) and $\sum_k M_k^\dagger M_k = I$.

Any two bipartite pure states which have the same reduced state on one subsystem are interconvertible by an isometry on the other subsystem. Indeed, consider states $|\psi\rangle^{AB}$ and $|\varphi\rangle^{AC}$ such that

$$\psi^A = \varphi^A = \sum_i p_i |e_i\rangle\langle e_i| \tag{2.20}$$

with an orthonormal basis $\{|e_i\rangle\}_i$. The Schmidt decomposition of the states are given by

$$|\psi\rangle = \sum_i \sqrt{p_i} |e_i\rangle^A |e'_i\rangle^B, \quad |\varphi\rangle = \sum_i \sqrt{p_i} |e_i\rangle^A |e''_i\rangle^C, \tag{2.21}$$

respectively. There exists an isometry $V : \mathcal{H}^B \to \mathcal{H}^C$ such that $V|e'_i\rangle^B = |e''_i\rangle^C$ for all i, and thus $V|\psi\rangle = |\varphi\rangle$.

A bipartite pure state $|\Phi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is called a *maximally entangled state* if its Schmidt decomposition takes the form of

$$|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |e_i\rangle^A |e'_i\rangle^B, \tag{2.22}$$

where $d = \dim \mathcal{H}^A = \dim \mathcal{H}^B$. A useful property of the maximally entangled states is that any unitary on one subsystem can be replaced by another unitary on the other subsystem. Indeed, for a unitary $U = \sum_{ji} u_{ji} |e_j\rangle\langle e_i|$, we have

$$U^A|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i,j=1}^{d} u_{ji} |e_j\rangle^A |e'_i\rangle^B = \frac{1}{\sqrt{d}} \sum_{i,j=1}^{d} |e_i\rangle^A u_{ij} |e'_j\rangle^B = (U^T)^B |\Phi\rangle, \tag{2.23}$$

where the superscript T denotes transposition with respect to the basis $\{|e_i\rangle\}_i$. The maximally entangled state with $d = 2$ is called a *Bell pair.*

A bipartite state $\rho \in \mathcal{S}(\mathcal{H}^X \otimes \mathcal{H}^A)$ is called a *classical-quantum state* between X and A if it is written as

$$\rho^{XA} = \sum_i p_i |i\rangle\langle i| \otimes \rho_i^A. \tag{2.24}$$

2.2 Distance Measures

In this section, we introduce two functions that measure the 'distance' between two quantum states. We review properties of those functions which are used in the following chapters.

2.2.1 Trace Distance

The trace norm of $A \in \mathcal{L}(\mathcal{H})$ is defined as

$$\|A\|_1 := \mathrm{Tr}\sqrt{A^\dagger A}. \tag{2.25}$$

For $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, the trace distance is defined as

$$d(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1. \tag{2.26}$$

This function satisfies the conditions for distance, i.e,

(i) *nonnegativity*: $d(\rho, \sigma) \geq 0$,

(ii) *coincidence axiom*: $d(\rho, \sigma) = 0$ if and only if $\rho = \sigma$,

(iii) *symmetry*: $d(\rho, \sigma) = d(\sigma, \rho)$,

(iv) *triangle inequality*: $d(\rho, \sigma) \leq d(\rho, \tau) + d(\tau, \sigma)$.

The trace distance is invariant under any isometric operation $U : \mathcal{H} \to \mathcal{H}'$ as

$$\left\|U\rho U^\dagger - U\sigma U^\dagger\right\|_1 = \|\rho - \sigma\|_1,$$

and is nonincreasing under any CPTP map $\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ as

$$\left\| \rho - \sigma \right\|_1 \geq \left\| \mathcal{E}(\rho) - \mathcal{E}(\sigma) \right\|_1 .$$

In particular, it is nonincreasing under discarding one of the composite systems as

$$\left\| \rho^{AB} - \sigma^{AB} \right\|_1 \quad \geq \quad \left\| \rho^{A} - \sigma^{A} \right\|_1 \tag{2.27}$$

for $\rho, \sigma \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$. When two states $\rho, \sigma \in \mathcal{S}(\mathcal{H}^X \otimes \mathcal{H}^A)$ are classical-quantum states, which are given by

$$\rho^{XA} = \sum_i q_i |i\rangle\langle i|^X \otimes \rho_i^A, \;\; \sigma^{XA} = \sum_i q_i |i\rangle\langle i|^X \otimes \sigma_i^A, \tag{2.28}$$

the trace distance is written as

$$\left\| \rho^{XA} - \sigma^{XA} \right\|_1 = \sum_i q_i \left\| \rho_i^A - \sigma_i^A \right\|_1 , \tag{2.29}$$

Combining this with (2.27) yields joint convexity of trace distance as

$$\sum_i q_i \left\| \rho_i - \sigma_i \right\|_1 \geq \left\| \sum_i q_i \rho_i - \sum_i q_i \sigma_i \right\|_1 , \tag{2.30}$$

which, as a particular case of $\sigma_i = \sigma$, implies

$$\sum_i q_i \left\| \rho_i - \sigma \right\|_1 \geq \left\| \sum_i q_i \rho_i - \sigma \right\|_1 . \tag{2.31}$$

Consider two states $\rho, \sigma \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$ such that $\rho^B = \sigma^B$ and a measurement $\{M_k\}_k$ on B. Let the probability to obtain a result k of the measurement to be

$$p_k := \mathrm{Tr}[M_k^B \rho^{AB} M_k^{\dagger A}] = \mathrm{Tr}[M_k^B \sigma^{AB} M_k^{\dagger B}], \tag{2.32}$$

and the state after the measurement to be

$$\rho_k^{AB} := p_k^{-1} M_k^B \rho^{AB} M_k^{\dagger B}, \;\; \sigma_k^{AB} := p_k^{-1} M_k^B \sigma^{AB} M_k^{\dagger B}. \tag{2.33}$$

Then we have

$$\sum_k p_k \left\| \rho_k^A - \sigma_k^A \right\|_1 \leq \left\| \rho^{AB} - \sigma^{AB} \right\|_1 . \tag{2.34}$$

To show this relation, define a CPTP map $\mathcal{E} : B \to B'B$ by

$$\mathcal{E}(\tau) = \sum_k |k\rangle\langle k|^{B'} \otimes M_k \tau M_k^{\dagger B}. \tag{2.35}$$

Then

$$\mathcal{E}(\rho^{AB}) = \sum_k p_k |k\rangle\langle k|^{B'} \otimes \rho_k^{AB}, \;\; \mathcal{E}(\sigma^{AB}) = \sum_k p_k |k\rangle\langle k|^{B'} \otimes \sigma_k^{AB}, \tag{2.36}$$

and thus

$$\begin{aligned} \left\| \rho^{AB} - \sigma^{AB} \right\|_1 \quad &\geq \quad \left\| \mathcal{E}(\rho^{AB}) - \mathcal{E}(\sigma^{AB}) \right\|_1 = \sum_k p_k \left\| \rho_k^{AB} - \sigma_k^{AB} \right\|_1 \\ &\geq \quad \sum_k p_k \left\| \rho_k^A - \sigma_k^A \right\|_1 . \end{aligned} \tag{2.37}$$

2.2.2 Fidelity

The fidelity of two states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ is defined as

$$F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1^2. \tag{2.38}$$

The function is symmetric on ρ and σ, satisfies $0 \leq F(\rho, \sigma) \leq 1$, and quantifies 'how close' the two states are. $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. When one of the two states is pure, the fidelity has a simple form as

$$F(\rho, |\psi\rangle) = \langle\psi|\rho|\psi\rangle. \tag{2.39}$$

Consequently, we have

$$F\left(\sum_i q_i\rho_i, |\psi\rangle\right) = \sum_i q_i F(\rho_i, |\psi\rangle). \tag{2.40}$$

Similarly to the trace distance, the fidelity is invariant under any isometric operation $U : \mathcal{H} \to \mathcal{H}'$ as

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho, \sigma), \tag{2.41}$$

and is nondecreasing under any CPTP map $\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ as

$$F(\rho, \sigma) \leq F(\mathcal{E}(\rho), \mathcal{E}(\sigma)).$$

It is nondecreasing under discarding one of the composite systems as

$$F(\rho^{AB}, \sigma^{AB}) \leq F(\rho^A, \sigma^A)$$

for $\rho, \sigma \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$.

The trace distance and the fidelity are related by simple inequalities.

$$1 - \sqrt{F(\rho, \sigma)} \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)} \tag{2.42}$$

Consequently, if $\|\rho - \sigma\|_1 \leq \epsilon$ then

$$F(\rho, \sigma) \geq \left(1 - \frac{1}{2}\|\rho - \sigma\|_1\right)^2 \geq 1 - \epsilon + \frac{1}{4}\epsilon^2 = 1 - \eta(\epsilon), \tag{2.43}$$

and conversely, if $F(\rho, \sigma) \geq 1 - \epsilon$ then

$$\|\rho - \sigma\|_1 \leq 2\sqrt{1 - F(\rho, \sigma)} \leq 2\sqrt{\epsilon} = \eta(\epsilon). \tag{2.44}$$

We also have

$$F(\rho, \sigma) = \max_\phi F(|\psi\rangle, |\phi\rangle) = \max_\phi |\langle\psi|\phi\rangle|^2, \tag{2.45}$$

where ψ is an arbitrary purification of ρ, and the maximization is taken over all possible purifications of σ.

2.2.3 Uhlmann's Theorem

If two bipartite pure states $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB'}$ satisfy $\|\psi^A - \phi^A\|_1 \leq \epsilon$, where ψ^A and ϕ^A are reduced state of $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB'}$ in A, there exists an isometry $V : \mathcal{H}^{B'} \to \mathcal{H}^B$ such that

$$\left\| |\psi\rangle\langle\psi|^{AB} - V|\phi\rangle\langle\phi|^{AB'}V^\dagger \right\|_1 \leq \eta(\epsilon). \tag{2.46}$$

By the condition we have $F(\psi^A, \phi^A) \geq 1 - \eta(\epsilon)$, and thus there exists $|\phi^*\rangle^{AB}$ such that $\phi^{*A} = \phi^A$ and $F(|\psi\rangle, |\phi^*\rangle) \geq 1 - \eta(\epsilon)$. There exists an isometry $V : \mathcal{H}^{B'} \to \mathcal{H}^B$ such that $|\phi^*\rangle = V|\phi\rangle$, and hence $F(|\psi\rangle, V|\phi\rangle) \geq 1 - \eta(\epsilon)$ and (2.46). Existence of such an isometry is know as *Uhlmann's theorem* [11, 12].

2.2.4 Distance between Operations

It is not common but we can also consider distance between two operations on a particular input state. Consider two CPTP maps $\mathcal{E}$, $\mathcal{F}$ on $\mathcal{S}(\mathcal{H})$, and a state $\rho \in \mathcal{S}(\mathcal{H})$. The trace distance between $\mathcal{E}$ and $\mathcal{F}$ with respect to ρ can be defined as

$$d_\rho(\mathcal{E}, \mathcal{F}) := d(\mathcal{E}(\rho), \mathcal{F}(\rho)) = \frac{1}{2} \|\mathcal{E}(\rho) - \mathcal{F}(\rho)\|_1 . \tag{2.47}$$

From (2.18) and (2.34), for any state $|\varphi\rangle^{AB}$ and ensemble $\{p_k, |\psi_k\rangle^A\}_k$ such that $\sum_k p_k |\psi_k\rangle\langle\psi_k| = \varphi^A$, we have

$$\sum_k p_k \|\mathcal{E}(|\psi_k\rangle\langle\psi_k|) - \mathcal{F}(|\psi_k\rangle\langle\psi_k|)\|_1 \leq \left\|(\mathcal{F}^A \circ \mathrm{id}^B)(|\varphi\rangle\langle\varphi|) - (\mathcal{E}^A \circ \mathrm{id}^B)(|\varphi\rangle\langle\varphi|)\right\|_1 .$$

Consequently, if

$$F\left((\mathcal{F}^A \circ \mathrm{id}^B)(|\varphi\rangle\langle\varphi|), (\mathcal{E}^A \circ \mathrm{id}^B)(|\varphi\rangle\langle\varphi|)\right) \geq 1 - \epsilon \tag{2.48}$$

then

$$\sum_k p_k F\left(\mathcal{E}(|\psi_k\rangle\langle\psi_k|), \mathcal{F}(|\psi_k\rangle\langle\psi_k|)\right) \geq 1 - \eta(\epsilon). \tag{2.49}$$

The L.H.S. in (2.48) and (2.49) are generalizations of *entanglement fidelity* and *ensemble fidelity* introduced in Ref.[13], respectively.

2.3 Von Neumann Entropy and Quantum Mutual Information

In this section, we introduce information theoretical functions such as entropy and mutual information, and review their properties that are used in the following chapters. We do not give detailed proofs for all properties. For more detail, see e.g. Refs.[1, 10, 8].

Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability distribution $p(x) = \Pr\{X = x\}$. The Shannon entropy of X, denoted by $H(X)$, is defined by

$$H(X) := -\sum_{x\in\mathcal{X}} p(x)\log p(x). \tag{2.50}$$

To clarify that $H(X)$ is a function of probability distribution $\{p(x)\}_{x\in\mathcal{X}}$, we also denote this quantity by $H(\{p(x)\}_x)$. Let $\{q(x)\}_{x\in\mathcal{X}}$ be another probability distribution. The relative entropy between $\{p(x)\}_x$ and $\{q(x)\}_x$ is defined as

$$D(\{p(x)\}\|\{q(x)\}) = \sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}, \tag{2.51}$$

and is known to be nonnegative. In particular, by letting $q(x) = 1/|\mathcal{X}|$, we have

$$\begin{aligned} \log|\mathcal{X}| - H(X) &= \sum_{x\in\mathcal{X}} p(x)\left(\log|\mathcal{X}| + \log p(x)\right) \\ &= \sum_{x\in\mathcal{X}} p(x)\left(\log p(x) - \log q(x)\right) \\ &\geq D(\{p(x)\}\|\{q(x)\}) \geq 0. \end{aligned} \tag{2.52}$$

Consider two random variables X and Y with a joint probability distribution $p(x,y)$ and marginal probability distributions $p(x) = \sum_y p(x,y)$ and $p(y) = \sum_x p(x,y)$. The joint entropy of X and Y is defined as

$$H(X,Y) := -\sum_{x,y} p(x,y)\log p(x,y), \tag{2.53}$$

and the conditional entropy of Y, conditioned by X, is defined as

$$H(Y|X) := \sum_x p(x)H(Y|X=x) = -\sum_x p(x)\sum_y p(y|x)\log p(y|x), \tag{2.54}$$

where $p(y|x)$ is a conditional probability distribution defined as $p(y|x) = p(x,y)/p(x)$. The mutual information between X and Y is defined as

$$I(X:Y) := D(\{p(x,y)\}\|\{p(x)p(y)\}) = H(X) + H(Y) - H(X,Y). \tag{2.55}$$

For three random variables X, Y and Z with a joint probability distribution $p(x,y,z)$, the conditional mutual information between X and Y, conditioned by Z, is defined as

$$I(X:Y|Z) := \sum_z p(z)I(X:Y|Z=z) = \sum_z p(z)D(\{p(x,y|z)\}\|\{p(x|z)p(y|z)\}) \tag{2.56}$$

All entropic functions defined here are nonnegative:

$$H(X), H(X,Y), H(Y|X), I(X:Y), I(X:Y|Z) \geq 0. \tag{2.57}$$

The quantum mechanical counterpart of the Shannon entropy, called the von Neumann entropy of a state ρ, is defined by

$$S(\rho) := -\mathrm{Tr}[\rho \log \rho] \tag{2.58}$$

for $\rho \in \mathcal{S}(\mathcal{H})$. By using the eigenvalues $\{\lambda_x\}_x$ of ρ, $S(\rho)$ is rewritten as

$$S(\rho) := -\sum_x \lambda_x \log \lambda_x. \tag{2.59}$$

Analogously to (2.52), we have

$$S(\rho) \leq \log \mathrm{rank}\rho \leq \log \dim\mathcal{H}. \tag{2.60}$$

$S(\rho) = 0$ if and only if ρ is a pure state. For $\rho \in \mathcal{S}(\mathcal{H}^A)$, we also denote $S(\rho)$ as $S(A)_\rho$. The entropy for a composite system AB in the state $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$ is straightforwardly defined as

$$S(AB)_\rho := -\mathrm{Tr}[\rho \log \rho]. \tag{2.61}$$

The quantum conditional entropy is defined by

$$S(B|A)_\rho := S(AB)_\rho - S(A)_\rho.$$

For two states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, the quantum relative entropy is defined as

$$D(\rho\|\sigma) := \mathrm{Tr}[\rho \log \rho - \rho \log \sigma], \tag{2.62}$$

which is nonnegative analogously to the classical relative entropy. The quantum mutual information is defined by

$$I(A:B)_\rho := D(\rho^{AB}\|\rho^A \otimes \rho^B) = S(A)_\rho + S(B)_\rho - S(AB)_\rho,$$

which is nonnegative, and $I(A:B)_\rho = 0$ if and only if $\rho^{AB} = \rho^A \otimes \rho^B$. The quantum conditional mutual information for a tripartite state $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C)$ is defined as

$$\begin{aligned} I(A:C|B) &:= S(A|B) + S(C|B) - S(AC|B) \\ &= S(AB) + S(BC) - S(B) - S(ABC). \end{aligned}$$

Contrary to the classical conditional mutual information, the quantum conditional entropy can be negative. For example, let $\rho^{AB} = |\psi\rangle\langle\psi|$, where the Schmidt decomposition of $|\psi\rangle^{AB}$ is given by

$$|\psi\rangle^{AB} = \sum_i \sqrt{p_i} |e_i\rangle^A |e_i'\rangle^B \tag{2.63}$$

Then the reduced states are

$$\rho^A = \sum_i p_i |e_i\rangle\langle e_i|^A, \quad \rho^B = \sum_i p_i |e_i'\rangle\langle e_i'|^B, \tag{2.64}$$

and thus

$$S(A)_\rho = S(B)_\rho = H(\{p_i\}_i). \tag{2.65}$$

Since ρ^{AB} is a pure state, we have $S(AB)_\rho = 0$. Thus

$$S(B|A)_\rho = S(AB)_\rho - S(A)_\rho = -H(\{p_i\}_i). \tag{2.66}$$

The von Neumann entropy of the reduced state of $|\psi\rangle^{AB}$ given by (2.65) is also called the *entanglement entropy* of $|\psi\rangle^{AB}$.

One of the most important results in quantum information theory is the property of von Neumann entropy called the *strong subadditivity*, which is expressed as

$$S(ABC) + S(B) \leq S(AB) + S(BC) \tag{2.67}$$

for an arbitrary tripartite state ρ^{ABC}. This inequality was proved first in [14] in a highly mathematical way, and more intuitive proofs are later given in [17, 15]. As direct consequences of this inequality, we have the nondecreasing property of the quantum conditional entropy under discarding part of conditional systems:

$$S(A|BC) \leq S(A|B), \tag{2.68}$$

the nonnegativity of the quantum conditional mutual information:

$$I(A:C|B) \geq 0,$$

and the nonincreasing property of the quantum mutual information under discarding part of the system:

$$I(A:BC) \geq I(A:B).$$

Entropies and mutual informations are invariant under local isometric operations. Let $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C)$, let $V : \mathcal{H}^A \to \mathcal{H}^{A'}$ be any isometry, and let $\sigma^{A'BC} := V\rho V^\dagger$. We have

$$S(A)_\rho = S(A')_\sigma, \quad S(B|A)_\rho = S(B|A')_\sigma, \quad I(A:B)_\rho = I(A':B)_\sigma \tag{2.69}$$

and so forth. Consequently, under any CPTP map $\mathcal{E} : B \to B'$, we have

$$\begin{aligned} I(A:B)_\rho &\geq I(A:B')_{\rho'} \\ S(A|B)_\rho &\leq S(A|B')_{\rho'} \end{aligned} \tag{2.70}$$

where $\rho' := \mathcal{E}(\rho)$. Let $V : \mathcal{H}^B \to \mathcal{H}^{B'} \otimes \mathcal{H}^E$ be an isometry such that the Stinespring dilation of $\mathcal{E}$ is given by $\mathcal{E}(\tau) = \mathrm{Tr}_E[V\tau V^\dagger]$. Let $\tilde{\rho}^{AB'E} := V\rho V^\dagger$. From (2.68) and (2.69), we have

$$S(A|B)_\rho = S(A|B'E)_{\tilde{\rho}} \leq S(A|B')_{\tilde{\rho}} = S(A|B')_{\rho'}. \tag{2.71}$$

Inequalities (2.70) are called the *data processing inequality.*

The quantum conditional entropies and the quantum conditional mutual informations satisfy simple equalities as follows.

$$\begin{aligned} &S(BC|A) = S(B|A) + S(C|AB) \\ &I(A:C|B) = I(AB:C) - I(B:C) = I(A:BC) - I(A:B) \end{aligned} \tag{2.72}$$

These equalities are shown as

$$\begin{aligned} S(BC|A) &= S(ABC) - S(A) \\ &= S(AB) - S(A) + S(ABC) - S(AB) \\ &= S(B|A) + S(C|AB) \end{aligned}$$

and

$$\begin{aligned} I(A:C|B) &= S(A|B) + S(C|B) - S(AC|B) \\ &= S(AB) + S(BC) - S(B) - S(ABC) \\ &= S(AB) + S(C) - S(ABC) - S(B) - S(C) - S(BC) \\ &= I(AB:C) - I(B:C). \end{aligned} \tag{2.73}$$

The symmetry of $I(A:C|B)$ in A and C implies (2.72). These properties are called the *chain rule.*

The quantum conditional entropy and the quantum conditional mutual information can be represented by simple forms when the conditioning part is a classical system. Namely, for ρ and σ defined as

$$\begin{aligned} \rho^{XA} &:= \sum_x p(x)|x\rangle\langle x|^X \otimes \rho_x^A \\ \sigma^{XAB} &:= \sum_x p(x)|x\rangle\langle x|^X \otimes \sigma_x^{AB}, \end{aligned}$$

we have

$$S(A|X) = \sum_x p(x) S(A)_{\rho_x} \tag{2.74}$$

and

$$I(A:B|X) := \sum_x p(x) I(A:B)_{\sigma_x}. \tag{2.75}$$

Consequently, we have

$$0 \leq I(A:X)_\rho = S(A)_\rho - \sum_x p(x) S(A)_{\rho_x}, \tag{2.76}$$

which expresses the concavity of the von Neumann entropy. This property implies that the von Neumann entropy is nondecreasing under random unitary operations. For $\rho \in \mathcal{S}(\mathcal{H}^A)$, let

$$\rho' = \sum_x p_x U_x \rho U_x^\dagger. \tag{2.77}$$

Then we have

$$S(A)_{\rho'} \geq \sum_x p_x S(A)_{U_x \rho U_x^\dagger} = \sum_x p_x S(A)_\rho = S(A)_\rho. \tag{2.78}$$

When ρ is a pure state, we also have

$$S(A)_{\rho'} \leq H(\{p_x\}_x). \tag{2.79}$$

To see this, let $\rho = |\psi\rangle\langle\psi|$, $|\psi_x\rangle := U_x|\psi\rangle$ and $|\varphi\rangle^{AB} := \sum_x \sqrt{p_x}|\psi_x\rangle^A|x\rangle^B$. We have

$$S(B)_\varphi = S(A)_\varphi = S\left(\sum_x p_x |\psi_x\rangle\langle\psi_x|\right) = S(A)_{\rho'}. \tag{2.80}$$

and also have

$$\mathcal{D}(\varphi^B) = \sum_x p_x |x\rangle\langle x|, \tag{2.81}$$

where $\mathcal{D}$ is the complete dephazing map on B with respect to the basis $\{|x\rangle\}_x$. Note that $\mathcal{D}$ can be described by a random unitary operation as (2.13). Hence we have

$$H(\{p_x\}_x) = S\left(\mathcal{D}(\varphi^B)\right) \geq S(B)_\varphi = S(A)_{\rho'}. \tag{2.82}$$

As indicated in (2.74), the quantum conditional entropy is nonnegative if the conditioning part is a classical system. Generally, the quantum conditional entropy is nonnegative if the state is separable. Indeed, suppose $\rho^{AB} = \sum_x p_x \rho_x^A \otimes \sigma_x^B$, and let

$$\rho^{XAB} := \sum_x p(x) |x\rangle\langle x|^X \otimes \rho_x^A \otimes \sigma_x^B. \tag{2.83}$$

Then we have

$$\begin{aligned} S(B|A)_\rho &\geq S(B|AX)_\rho = S(AB|X)_\rho - S(A|X)_\rho \\ &= \sum_x p(x) \left(S(\rho_x^A \otimes \sigma_x^B) - S(\rho_x^A)\right) \\ &= \sum_x p(x) S\left(\sigma_x^B\right) \geq 0. \end{aligned} \tag{2.84}$$

The von Neumann entropy is a continuous function of states. If $|\rho^A - \sigma^A| \leq \epsilon$, then

$$|S(A)_\rho - S(A)_\sigma| \leq \epsilon \log d_A + \eta(\epsilon) \leq \eta(\epsilon) \log d_A. \tag{2.85}$$

Also, if $|\rho^{AB} - \sigma^{AB}| \leq \epsilon$, then

$$|S(A|B)_\rho - S(A|B)_\sigma| \leq 4\epsilon \log d_A + 2h(\epsilon) \leq \eta(\epsilon) \log d_A, \tag{2.86}$$

where $h(\epsilon) := -\epsilon \log \epsilon - (1-\epsilon) \log(1-\epsilon)$. Note that the upper bound in (2.86) depends only on the dimension of $\mathcal{H}^A$. Consequently, we have

$$|I(A:B)_\rho - I(A:B)_\sigma| \leq |S(A)_\rho - S(A)_\sigma| + |S(A|B)_\rho - S(A|B)_\sigma| \leq \eta(\epsilon) \log d_A.$$

2.4 Koashi-Imoto decomposition

Ref.[18] introduces a method to decompose a Hilbert space in such a way that it completely characterizes the set of operations which do not disturb a given set of states. We call it as the *Koashi-Imoto decomposition*, or the *KI decomposition* for short. The decomposition provides a way to classify and quantify "information" generated by an ensemble of quantum states, by dividing it into the classical part, the quantum part and the redundant part. Moreover, it is also useful for characterizing (ir)reversibility of quantum operations. In particular, Ref.[21] shows that the decomposition characterize the structure of states satisfying strong subadditivity of von Neumann entropy with equality. Moreover, it is important in this thesis due to its relation to the quantum Markov chain (Section 4.1).

Theorem 1 Associated to any set of states $\mathfrak{S} := \{\rho_k\}_k$ in a finite dimensional quantum system B, there exists a unitary isomorphism $\Gamma : \mathcal{H}^B \to \mathcal{H}^{b_0} \otimes \mathcal{H}^{b_L} \otimes \mathcal{H}^{b_R}$ such that the following two properties hold.

1. The states in $\mathfrak{S}$ are decomposed as

   $$\Gamma \rho_k \Gamma^\dagger = \sum_{j \in J} p_{j|k} |j\rangle\langle j| \otimes \rho_{j|k} \otimes \sigma_j \tag{2.87}$$

   with some probability distribution $\{p_{j|k}\}_{j \in J}$, orthonormal basis $\{|j\rangle\}_{j \in J}$ of $\mathcal{H}^{b_0}$, states $\rho_{j|k} \in \mathcal{S}(\mathcal{H}^{b_L})$ and $\sigma_j \in \mathcal{S}(\mathcal{H}^{b_R})$.

2. Any CPTP map $\mathcal{E}$ on B which leaves all ρ_k invariant has a Stinespring dilation of the form

   $$\mathcal{E}(\rho) = \mathrm{Tr}_E[U \rho U^\dagger].$$

Here, $U : B \to BE$ is an isometry that is decomposed as

$$\Gamma U \Gamma^\dagger = \sum_{j \in J} |j\rangle\langle j|^{b_0} \otimes I_j^{b_L} \otimes U_j^{b_R},$$

where I_j is the identity operator on $\mathcal{H}_j^{b_L} := \operatorname{supp} \sum_k \rho_{j|k}$, and $U_j^{b_R} : \mathcal{H}_j^{b_R} \to \mathcal{H}_j^{b_R} \otimes \mathcal{H}^E$ are isometries that satisfy

$$\mathrm{Tr}_E[U_j \sigma_j U_j^\dagger] = \sigma_j$$

for all j.

Proof. See Ref.[18]. ■

We call Γ as the KI decomposition of $\mathcal{H}^B$ with respect to a set of states $\mathfrak{S}$. The KI decomposition is uniquely determined from $\mathfrak{S}$, up to trivial relabelings of j and changes of the basis. As indicated in (2.87), $\mathcal{H}^{b_0}$ stores the classical part of the information on ρ_k, $\mathcal{H}^{b_L}$ the quantum part, and $\mathcal{H}^{b_R}$ stores no information. An algorithm for obtaining the KI decomposition is proposed in [18]. We informally denote the composite system $b_0 b_L b_R$ by B when there is no fear of confusion. There are several derivative types of the KI decomposition as follows.

Definition 2 The KI decomposition of $\mathcal{H}^B$ with respect to an ensemble of states $\mathfrak{E} := \{p_k, \rho_k\}_k$, for which $p_k > 0$ for all k, is defined as the KI decomposition of $\mathcal{H}^B$ with respect to $\{\rho_k\}_k$.

Definition 3 The KI decomposition of $\mathcal{H}^B$ with respect to a bipartite state ρ^{AB} is defined as the KI decomposition of $\mathcal{H}^B$ with respect to the following set of states.

$$\mathfrak{S}_{\rho^{A \to B}} := \{\varphi \in \mathcal{S}(\mathcal{H}^B) | \exists M \in \mathcal{P}(\mathcal{H}^A) \text{ s.t. } \varphi = \mathrm{Tr}_A[M^A \rho^{AB}]\}$$

Here, $\mathcal{P}(\mathcal{H}^A)$ denotes the set of positive semidefinite operators on $\mathcal{H}^A$.

By definition, it is straightforward that there exists a POVM $\{M_\mu\}_\mu$ on A, i.e., the set of positive semidefinite operators on $\mathcal{H}^A$ satisfying $\sum_\mu M_\mu = I$, such that the KI decomposition of $\mathcal{H}^B$ with respect to a bipartite state ρ^{AB} is equivalent to the KI decomposition of an ensemble $\{p_\mu, \rho_\mu\}_\mu$, which is defined by $p_\mu = \mathrm{Tr}[M_\mu^A \rho^{AB}]$ and $\rho_\mu = p_\mu^{-1} \mathrm{Tr}_A[M_\mu^A \rho^{AB}]$. Moreover, the following property is proved in [21].

Lemma 4 When Γ_ρ is the KI decomposition of $\mathcal{H}^B$ with respect to ρ^{AB}, the state ρ^{AB} is decomposed as

$$\rho_{KI}^{AB} := \Gamma_\rho^B \rho^{AB} \Gamma_\rho^{\dagger B} = \sum_{j \in J} p_j |j\rangle\langle j|^{b_0} \otimes \rho_j^{Ab_L} \otimes \sigma_j^{b_R}, \tag{2.88}$$

where $\{p_j\}_{j \in J}$ is a probability distribution, $\rho_j^{Ab_L} \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^{b_L})$, $\sigma_j^{b_R} \in \mathcal{S}(\mathcal{H}^{b_R})$ and $\langle j | j' \rangle = \delta_{jj'}$.

Proof. See Ref.[21]. ■

We call (2.88) as the KI decomposition of ρ^{AB} on B.

Chapter 3

Elements of Quantum Shannon Theory

In this chapter, we introduce *quantum state merging* and the *decoupling theorem*, which are at the core of the recent developments in the field of quantum Shannon theory. Results and techniques reviewed in this chapter will be used in Chapter 4, 5 and 6.

Quantum state merging is a quantum communication task, first introduced and analyzed in [16, 17]. It has been shown that many of the central coding theorems in quantum Shannon theory are systematically derived from quantum state merging and its variant known as the fully-quantum Slepian Wolf theorem [20]. The core of their usefulness is that they approach problems of quantum communication from the viewpoint of decoupling, that is, destroying correlation between two quantum systems by locally acting on one of the two subsystems.

The minimum amount of randomness required for decoupling is revealed by the decoupling theorem. Depending on the types of operations applied to destroy the correlation, there are several formulations of the decoupling theorem, such as the one based on the partial trace [20], random unitary operations [15], and projective measurements [17]. In this chapter, we consider decoupling by random unitary operations.

We start with explaining the idea and properties of typicality [1, 8, 22], which is one of the most basic tools in classical and quantum Shannon theory.

3.1 Typical sequences and subspaces

Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability distribution $p(x) = \Pr\{X = x\}$. We assume $|\mathcal{X}| < \infty$, where $|\mathcal{X}|$ denotes the number of elements in $\mathcal{X}$. Consider a source that generates a sequence $X_1, \cdots, X_n \in \mathcal{X}^n$ with a certain probability distribution. If probabilities of each sequence is given by

$$\Pr\{X_1 = x_1, \cdots, X_n = x_n\} = p(x_1) \cdots p(x_n) \tag{3.1}$$

for arbitrary n, the source is said to be independent and identically distributed (i.i.d.).

Definition 5 A sequence $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathcal{X}^n$ is called ϵ-weakly typical with respect to $p(x)$ if it satisfies

$$2^{-n(H(X)+\epsilon)} \le p(\boldsymbol{x}) \le 2^{-n(H(X)-\epsilon)}. \tag{3.2}$$

The ϵ-weakly typical set $T_{n,\epsilon}$ is defined as the set of all sequences $\boldsymbol{x} \in \mathcal{X}^n$ that are ϵ-weakly typical.

Definition 6 A sequence $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathcal{X}^n$ is called ϵ-strongly typical with respect to $p(x)$ if it satisfies

$$\left| \frac{1}{n} N(x|\boldsymbol{x}) - p(x) \right| < \frac{\epsilon}{|\mathcal{X}|} \tag{3.3}$$

for all $x \in \mathcal{X}$ and $N(x|\boldsymbol{x}) = 0$ if $p(x) = 0$. Here, $N(x|\boldsymbol{x})$ is the number of occurrences of the symbol x in the sequence $\boldsymbol{x}$. The ϵ-strongly typical set $T^*_{n,\epsilon}$ is defined as the set of all sequences $(x_1, \cdots, x_n) \in \mathcal{X}^n$ that are ϵ-strongly typical.

Theorem 7

1. The number of elements in $T_{n,\epsilon}$, denoted by $|T_{n,\epsilon}|$, is bounded as $|T_{n,\epsilon}| \le 2^{n(H(X)+\epsilon)}$.

2. For any $\epsilon, \delta > 0$ and for sufficiently large n,

$$\sum_{\boldsymbol{x} \in T_{n,\epsilon}} p(\boldsymbol{x}) > 1 - \delta. \tag{3.4}$$

3. For any $\epsilon, \delta > 0$ and for sufficiently large n,

$$\sum_{\boldsymbol{x} \in T^*_{n,\epsilon}} p(\boldsymbol{x}) > 1 - \delta. \tag{3.5}$$

Proof.

1. From (3.2), we have

$$1 = \sum_{\boldsymbol{x} \in \mathcal{X}^n} p(\boldsymbol{x}) \ge \sum_{\boldsymbol{x} \in T_{n,\epsilon}} p(\boldsymbol{x}) \ge \sum_{\boldsymbol{x} \in T_{n,\epsilon}} 2^{-n(H(X)+\epsilon)} = |T_{n,\epsilon}| \cdot 2^{-n(H(X)+\epsilon)}. \tag{3.6}$$

2. The expectation value of a random variable $-\log p(X)$ is

$$\mathbb{E}\left[-\log p(X)\right] = -\sum_x p(x)\log p(x) = H(X). \tag{3.7}$$

From the weak law of large numbers, we have

$$\Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n}(-\log p(X_i)) - H(X)\right| \le \epsilon\right\} > 1-\delta \tag{3.8}$$

for sufficiently large n. On the other hand, from Condition (3.2), $\boldsymbol{x} \in T_{n,\epsilon}$ if and only if

$$\left|\frac{1}{n}\sum_{i=1}^{n}(-\log p(x_i)) - H(X)\right| \le \epsilon. \tag{3.9}$$

Thus we obtain (3.4).

3. The expectation value of a random variable $\delta_{x,X}$ for any fixed $x \in \mathcal{X}$ is $\mathbb{E}\left[\delta_{x,X}\right] = p(x)$. From the weak law of large numbers, we have

$$\Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\delta_{x,X_i} - p(x)\right| \le \frac{\epsilon}{|\mathcal{X}|}\right\} > 1-\delta' \tag{3.10}$$

for all $x \in \mathcal{X}$ and for sufficiently large n. Hence

$$\begin{aligned} & \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\delta_{x,X_i} - p(x)\right| \le \frac{\epsilon}{|\mathcal{X}|} \quad (\forall x \in \mathcal{X})\right\} \\ = \;& 1 - \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\delta_{x,X_i} - p(x)\right| > \frac{\epsilon}{|\mathcal{X}|} \quad (\exists x \in \mathcal{X})\right\} \\ \ge \;& 1 - \sum_x \Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\delta_{x,X_i} - p(x)\right| > \frac{\epsilon}{|\mathcal{X}|}\right\} \\ = \;& 1 - \delta'|\mathcal{X}|. \end{aligned} \tag{3.11}$$

Since

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{x,X_i} = \frac{1}{n}N(x|X_1,\cdots,X_n), \tag{3.12}$$

we obtain (3.5) by letting $\delta' = \delta/|\mathcal{X}|$. ■

Definition 8 Suppose the spectral decomposition of $\rho \in \mathcal{S}(\mathcal{H})$ is given by $\rho = \sum_x p(x)|x\rangle\langle x|$. The ϵ-weakly typical subspace $\mathcal{H}_{n,\epsilon}^{typ} \subset \mathcal{H}^{\otimes n}$ with respect to ρ is defined as

$$\mathcal{H}_{n,\epsilon}^{typ} := \mathrm{span}\{|x_1\rangle\cdots|x_n\rangle \in \mathcal{H}^{\otimes n}|(x_1,\cdots,x_n) \in T_{n,\epsilon}\}, \tag{3.13}$$

and the ϵ-strongly typical subspace $\mathcal{H}^{typ*}_{n,\epsilon} \subset \mathcal{H}^{\otimes n}$ is defined as

$$\mathcal{H}^{typ*}_{n,\epsilon} := \mathrm{span}\{|x_1\rangle \cdots |x_n\rangle \in \mathcal{H}^{\otimes n} | (x_1, \cdots, x_n) \in T^*_{n,\epsilon}\}, \tag{3.14}$$

where $T_{n,\epsilon}$ and $T^*_{n,\epsilon}$ are typical sets with respect to $p(x)$.

Theorem 9

1. The dimension of the typical subspace is bounded as $\dim \mathcal{H}^{typ}_{n,\epsilon} \leq 2^{n(S(\rho)+\epsilon)}$.

2. For any $\epsilon, \delta > 0$ and for sufficiently large n, let $\Pi_{n,\epsilon}$ be the projection onto $\mathcal{H}^{typ}_{n,\epsilon}$ with respect to ρ. Then for any state ψ^{AR} such that $\psi^A = \rho$,

$$\left\| \Pi^A_{n,\epsilon} (\psi^{AR})^{\otimes n} \Pi^A_{n,\epsilon} - (\psi^{AR})^{\otimes n} \right\|_1 < \eta(\delta). \tag{3.15}$$

3. For any $\epsilon, \delta > 0$ and for sufficiently large n, let $\Pi^*_{n,\epsilon}$ be the projection onto $\mathcal{H}^{typ}_{n,\epsilon}$ with respect to ρ. Then for any state ψ^{AR} such that $\psi^A = \rho$,

$$\left\| \Pi^{*A}_{n,\epsilon} (\psi^{AR})^{\otimes n} \Pi^{*A}_{n,\epsilon} - (\psi^{AR})^{\otimes n} \right\|_1 < \eta(\delta). \tag{3.16}$$

Proof.

1. $\dim \mathcal{H}^{typ}_{n,\epsilon} = |T_{n,\epsilon}| \leq 2^{n(H(X)+\epsilon)} = 2^{n(S(\rho)+\epsilon)}$.

2. Let $|\psi\rangle^{ARR'}$ be a purification of ψ^{AR}. The Schmidt decomposition of the state is given by $|\psi\rangle^{ARR'} = \sum_x \sqrt{p_x} |x\rangle^A |e_x\rangle^{RR'}$, where $\langle e_x | e_{x'} \rangle = \delta_{xx'}$. We have

$$(\langle \psi |^{ARR'})^{\otimes n} \Pi^A_{n,\epsilon} (|\psi\rangle^{ARR'})^{\otimes n} = \sum_{\boldsymbol{x}, \boldsymbol{x}' \in T_{n,\epsilon}} \sqrt{p_{\boldsymbol{x}} p_{\boldsymbol{x}'}} \langle e_{\boldsymbol{x}} | e_{\boldsymbol{x}'} \rangle = \sum_{\boldsymbol{x} \in T_{n,\epsilon}} p_{\boldsymbol{x}} \geq 1 - \delta,$$

and hence

$$\mathrm{Tr}[\Pi^A_{n,\epsilon} (\psi^{AR})^{\otimes n}] = \mathrm{Tr}[\Pi^A_{n,\epsilon} (|\psi\rangle\langle\psi|^{ARR'})^{\otimes n}] \geq 1 - \delta. \tag{3.17}$$

From Lemma 10 given below, we obtain (3.15).

3. Similarly, we have

$$\mathrm{Tr}[\Pi^{*A}_{n,\epsilon} (\psi^{AR})^{\otimes n}] \geq 1 - \delta, \tag{3.18}$$

and thus we obtain (3.16). ■

Lemma 10 (Lemma 9 in [19]) Let ρ be a subnormalized state, i.e. $\rho \geq 0$ and $\mathrm{Tr}[\rho] \leq 1$, and let $0 \leq X \leq I$. If $\mathrm{Tr}[\rho X] \geq 1 - \delta$, then

$$\left\| \sqrt{X} \rho \sqrt{X} - \rho \right\|_1 \leq 2\sqrt{2\delta}. \tag{3.19}$$

Here, $0 \leq X \leq I$ is an operator inequality stating that X and $I - X$ are positive semidefinite.

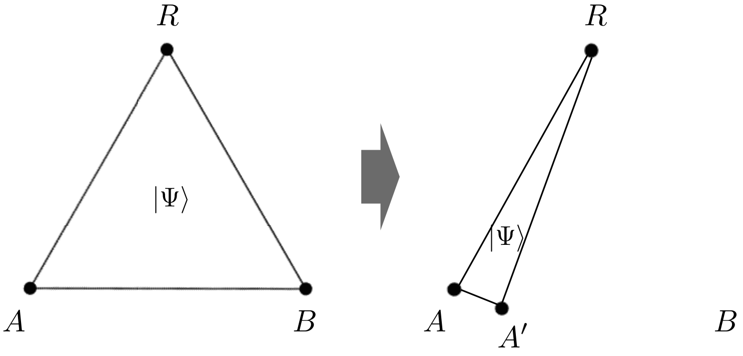


Figure 3.1: Quantum state merging is a task in which Bob transfers his share of $|\Psi\rangle^{ABR}$ to Alice. R is an inaccessible reference system. For later convenience, we consider Bob as the sender and Alice as the receiver.

Proof. Let $\rho = \sum_i p_i |i\rangle\langle i|$ be the spectral decomposition of ρ. From (2.42), (2.30) and $1 - x^2 \leq 2 - 2x$, we have

$$\begin{aligned}
\left\| \sqrt{X}\rho\sqrt{X} - \rho \right\|_1^2 &\leq \left(\sum_i p_i \left\| \sqrt{X}|i\rangle\langle i|\sqrt{X} - |i\rangle\langle i| \right\|_1 \right)^2 \\
&\leq \sum_i p_i \left\| \sqrt{X}|i\rangle\langle i|\sqrt{X} - |i\rangle\langle i| \right\|_1^2 \\
&\leq 4 \sum_i p_i \left(1 - \langle i|\sqrt{X}|i\rangle\langle i|\sqrt{X}|i\rangle \right) \\
&\leq 8 \sum_i p_i \left(1 - \langle i|\sqrt{X}|i\rangle \right) \\
&= 8 \left(1 - \mathrm{Tr}[\rho\sqrt{X}] \right) \leq 8 \left(1 - \mathrm{Tr}[\rho X] \right) \leq 8\delta.
\end{aligned}$$

■

3.2 Quantum State Merging

Suppose Alice and Bob share a tripartite pure state $|\Psi\rangle^{ABR}$ with an inaccessible reference system R. Quantum state merging [16, 17] is a task in which Bob sends his share of Ψ to Alice so that Alice has both A and B parts of Ψ (Figure 3.1). A trivial way is that Bob directly sends B to Alice. However, there are cases where Bob can reduce the cost of quantum communication by exploiting correlation in Ψ^{AB}. If Bob is not allowed to perform quantum communication but is only allowed to send classical information, it would be in general impossible to accomplish the task unless assistance of shared entanglement is available. But if correlation in Ψ^{AB} is strong enough and contains sufficient amount of entanglement, it might be possible to accomplish state merging without consuming additional entanglement. Moreover, it would even be possible to gain a certain amount of pure entanglement between

Alice and Bob after accomplishing state merging. Thus our question is as follows: For a given state $|\Psi\rangle^{ABR}$,

1. How much classical communication is required for state merging?
2. How much entanglement must be consumed for state merging, or can be gained after state merging?

It is proved in [17] that, in an asymptotic situation where $\Psi = \psi^{\otimes n}$ and $n \to \infty$, the optimal rate of the classical communication cost and entanglement cost per copy can be expressed by simple functions, namely, the quantum mutual information $I(A : R)_\psi$ and the quantum conditional entropy $S(B|A)_\psi$, respectively. In particular, it was shown that if $S(B|A)_\psi < 0$, then Alice and Bob can gain entanglement which amounts to $-S(B|A)_\psi$ after state merging. Thereby a rigorous operational meaning of negative quantum conditional entropy was uncovered for the first time. In the following, we review the rigorous definition of quantum state merging and two main theorems regarding its entanglement cost and classical communication cost.

Definition 11 Consider a tripartite pure state $|\Psi\rangle^{ABR}$. Let Alice and Bob have quantum systems A_0, A_1 and B_0, B_1, respectively. The following protocol $\mathcal{M}$ consisting of a sequence of quantum operations is called state merging of Ψ with error ϵ, entanglement cost $\log K - \log L$ and classical communication cost C. Here, $\mathcal{M} : AA_0BB_0 \to AA'A_1B_1$ is a LOCC and

$$F(\rho(\mathcal{M}), |\Psi\rangle^{AA'R}|\Phi_L\rangle^{A_1B_1}) \geq 1 - \epsilon \tag{3.20}$$

for $\rho(\mathcal{M}) = \mathcal{M}(|\Psi\rangle^{AA'R}|\Phi_K\rangle^{A_0B_0})$. C is the total amount of classical communication transmitted from Bob to Alice in $\mathcal{M}$, measured by bits.

The following theorem states that there always exists state merging with an error determined by the state and the amount of entanglement obtained after state merging.

Theorem 12 Let $D := (\mathrm{Tr}[(\Psi^A)^2])^{-1}$. For any $L \leq d_B$, there exists state merging of Ψ with error $\epsilon = \eta(\sqrt{Ld_R/D} + L/d_B)$, entanglement cost $-\log L$ and classical communication cost $C = \log(d_B/L)$.

Proof. See Ref.[17]. ■

The following theorem reveals the necessary amount of entanglement and classical communication required for state merging with a small error. Our version is essentially the same, but is technically different from that of [17]. We give a rigorous proof for completeness.

Theorem 13 Let $\mathcal{M}$ be state merging of Ψ with error ϵ. Entanglement cost and classical communication cost of $\mathcal{M}$ are bounded below as

$$\begin{aligned} \log K - (1-\eta(\epsilon))\log L &\geq S(B|A)_\Psi - \eta(\epsilon)\log(d_A d_B d_C) \qquad (3.21)\\ C &\geq I(B:R)_\Psi - \eta(\epsilon)\log(d_A d_B d_C). \qquad (3.22)\end{aligned}$$

Proof. See Appendix.

3.3 Decoupling Theorem

Suppose Alice and Bob share n copies of a bipartite state ρ^{AB}. By performing a random unitary operation on A, Alice destroys the correlation between A and B, that is, she turns $(\rho^{AB})^{\otimes n}$ close to a product state of the form $\rho_n'^{A^n} \otimes (\rho^B)^{\otimes n}$. The main problem is to minimize the amount of randomness required for this task. A precise definition is given as follows.

Definition 14 We say that ρ^{AB} is decoupled with the randomness cost R on A if, for any $\epsilon > 0$ and for sufficiently large n, there exists a random unitary operation $\mathcal{T}_n^{\bar{A}} : \tau \mapsto 2^{-nR}\sum_{k=1}^{2^{nR}} V_k \tau V_k^\dagger$ on $\bar{A}$ such that

$$\left\|(\mathcal{T}_n^{\bar{A}} \otimes \mathrm{id}^{\bar{B}})(\rho^{\otimes n})^{\bar{A}\bar{B}} - \mathcal{T}_n^{\bar{A}}(\rho^{\otimes n})^{\bar{A}} \otimes (\rho^{\otimes n})^{\bar{B}}\right\|_1 \leq \epsilon. \qquad (3.23)$$

The following theorem states that, in the asymptotic limit of infinite copies, the minimum randomness cost per copy required for destroying the correlation between two quantum systems is given by the quantum mutual information.

Theorem 15 ρ^{AB} is decoupled with the randomness cost R on A if and only if $R \geq I(A:B)_\rho$.

Proof. We only consider cases where ρ is a pure state, i.e, $\rho = |\varphi\rangle\langle\varphi|$. For the proof of mixed-state cases, see [15]. Proof techniques presented here will be applied in Chapter 4.

Suppose that a random unitary operation $\mathcal{T}_n^{\bar{A}} : \tau \mapsto 2^{-nR}\sum_{k=1}^{2^{nR}} V_k \tau V_k^\dagger$ satisfies Condition (3.23). The state after the operation,

$$\varphi_n'^{\bar{A}\bar{B}} := (\mathcal{T}_n^{\bar{A}} \otimes \mathrm{id}^{\bar{B}})(|\varphi\rangle\langle\varphi|^{\otimes n}), \qquad (3.24)$$

satisfies

$$\left\|\varphi_n'^{\bar{A}\bar{B}} - \varphi_n'^{\bar{A}} \otimes \varphi_n'^{\bar{B}}\right\|_1 \leq \epsilon. \qquad (3.25)$$

From (2.78) and (2.85), we have

$$\begin{aligned} nR &\geq S(\bar{A}\bar{B})_{\varphi'} \geq S(\bar{A})_{\varphi'} + S(\bar{B})_{\varphi'} - n\eta(\epsilon)\log(d_A d_B) \\ &= S(\bar{A})_{\varphi'} + S(\bar{B})_{\varphi} - n\eta(\epsilon)\log(d_A d_B) \\ &\geq S(\bar{A})_{\varphi^{\otimes n}} + S(\bar{B})_{\varphi^{\otimes n}} - n\eta(\epsilon)\log(d_A d_B) \\ &= n(S(A)_\varphi + S(B)_\varphi) - n\eta(\epsilon)\log(d_A d_B) \\ &= nI(A:B)_\varphi - n\eta(\epsilon)\log(d_A d_B). \end{aligned} \tag{3.26}$$

Since ϵ can be arbitrarily small and $\eta(\epsilon) \to 0$ when $\epsilon \to 0$, we obtain the "only if" part of the theorem.

For the "if" part, let $|\varphi\rangle = \sum_x \sqrt{p_x}|x\rangle^A|x\rangle^B$ be the Schmidt decomposition of $|\varphi\rangle$. Take arbitrary $\epsilon > 0$, choose sufficiently large n, and let $\Pi_{n,\epsilon}^A$ be the projection onto the ϵ-weakly typical subspace $\mathcal{H}_{n,\epsilon}^{A_{typ}}$ with respect to φ^A. Define a subnormalized state $|\varphi_{n,\epsilon}\rangle := \Pi_{n,\epsilon}^A(|\varphi\rangle)^{\otimes n}$. From Theorem 9, we have

$$\left\| |\varphi_{n,\epsilon}\rangle\langle\varphi_{n,\epsilon}|^{A^nB^n} - (|\varphi\rangle\langle\varphi|^{AB})^{\otimes n} \right\|_1 < \eta(\epsilon). \tag{3.27}$$

For any unitary V acting on $\mathcal{H}_{n,\epsilon}^{A_{typ}}$, define $|\varphi_{n,\epsilon}(V)\rangle^{A^nB^n} := V|\varphi_{n,\epsilon}\rangle^{A^nB^n}$. Suppose V is randomly chosen from an ensemble $\{p(dV), V\}$ such that for all $F \in \mathcal{L}(\mathcal{H}_{n,\epsilon}^{A_{typ}})$,

$$\int_V p(dV) VFV^\dagger = \mathrm{Tr}[F] \cdot \pi^{\bar{A}}, \tag{3.28}$$

where $\pi^{\bar{A}} := \Pi_{n,\epsilon}^{A_{typ}}/\mathrm{Tr}\Pi_{n,\epsilon}^{A_{typ}}$. Existence of such unitary ensembles will be discussed in Page 32. As an ensemble average, we have

$$\bar{\varphi}_{n,\epsilon} := \mathbb{E}\left[|\varphi_{n,\epsilon}(V)\rangle\langle\varphi_{n,\epsilon}(V)|\right] = \pi^{\bar{A}} \otimes \varphi_{n,\epsilon}^{\bar{B}}. \tag{3.29}$$

By Theorem 9 and the definition of the typical subspace, the nonzero eigenvalues of $\pi^{\bar{A}}$ are equal to $(\mathrm{dim}\mathcal{H}_{n,\epsilon}^{A_{typ}})^{-1} \geq 2^{-n(S(\varphi^A)+\epsilon)}$, and those of $\varphi_{n,\epsilon}^{\bar{B}}$ is not smaller than $2^{-n(S(\varphi^A)+\epsilon)}$.

Suppose $V_1, \cdots, V_N$ are unitaries that are randomly and independently chosen from the ensemble $\{p(dV), V\}$. From Lemma 16 given below, we have

$$\Pr\left\{ \frac{1}{N}\sum_{k=1}^N \varphi_{n,\epsilon}(V_k) \notin [(1-\epsilon)\bar{\varphi}_{n,\epsilon}, (1+\epsilon)\bar{\varphi}_{n,\epsilon}] \right\} \leq 2d_A^n d_B^n \exp\left(-\frac{N\lambda\epsilon^2}{2}\right),$$

where $\lambda \geq 2^{-n(2S(\varphi^A)+2\epsilon)}$. The R.H.S. of this inequality is smaller than 1 for sufficiently large n if $N = 2^{nR}$ and $R > 2S(\varphi^A) + 2\epsilon = I(A:B)_\varphi + 2\epsilon$. Thus there exists a set of unitaries $\{V_k\}_{k=1}^{2^{nR}}$ such that

$$(1-\epsilon)\bar{\varphi}_{n,\epsilon} \leq \frac{1}{N}\sum_{k=1}^N \varphi_{n,\epsilon}(V_k) \leq (1+\epsilon)\bar{\varphi}_{n,\epsilon}, \tag{3.30}$$

which implies

$$\left\| \frac{1}{N} \sum_{k=1}^{N} \varphi_{n,\epsilon}(V_k) - \bar{\varphi}_{n,\epsilon} \right\|_1 \leq \epsilon. \tag{3.31}$$

Combining with (3.27) and (3.29), we obtain (3.23). ■

The following lemma, called the *Operator Chernoff Bound*, plays a central role in the proof of Theorem 15 as described above. It will also be used in Chapter 4 and Chapter 6.

Lemma 16 (*Operator Chernoff Bound*: Lemma 3 in Ref.[15]) Let $X_1, \cdots X_N$ be i.i.d. random variables taking values in the operator interval $[0 : I] \subset \mathcal{B}(\mathcal{H})$ and with expectation $M = \mathbb{E}X_i \geq \lambda I$. Then, for $0 \leq \epsilon \leq 1$, and denoting $\bar{X} = (1/N)\sum_{i=1}^{N} X_i$,

$$\Pr[\bar{X} \not\leq (1+\epsilon)M] \leq d \exp\left(-\frac{N\lambda\epsilon^2}{2}\right),$$
$$\Pr[\bar{X} \not\geq (1-\epsilon)M] \leq d \exp\left(-\frac{N\lambda\epsilon^2}{2}\right).$$

Here, $\bar{X} \not\leq (1+\epsilon)M$ and $\bar{X} \not\geq (1-\epsilon)M$ represent that $(1+\epsilon)M - \bar{X} \geq 0$ and $\bar{X} - (1-\epsilon)M \geq 0$ do not hold, respectively.

Let us consider a random unitary ensemble $\{p(dV), V\}$ on a d-dimensional Hilbert space $\mathcal{H}$ that satisfy

$$\int p(dV) V F V^\dagger = \mathrm{Tr}[F] \cdot \frac{1}{d} I \tag{3.32}$$

for all $F \in \mathcal{L}(\mathcal{H})$.

An example of a random unitary ensemble satisfying (3.32) is that of generalized Pauli operators. The generalized Pauli operators on d-dimensional Hilbert space is defined as

$$\sigma_{pq} := X^p Z^q \quad (0 \leq p, q \leq d-1), \tag{3.33}$$

where $X := \sum_{t=1}^{d} |t-1\rangle\langle t|$ and $Z := \sum_{t=1}^{d} e^{2\pi i t/d} |t\rangle\langle t|$ with a fixed basis $\{|t\rangle\}_{t=1}^{d}$. Here, subtraction is taken with mod d. The random generalized Pauli ensemble $\{d^{-2}, \sigma_{pq}\}_{p,q}$ satisfies Condition (3.32). Indeed, consider the random generalized Pauli operation defined by

$$\mathcal{R}(\tau) = \frac{1}{d^2} \sum_{p,q} \sigma_{pq} \tau \sigma_{pq}^\dagger. \tag{3.34}$$

By using $XZ = e^{2\pi i/d} ZX$, $\mathcal{R}$ is described by

$$\mathcal{R}(\tau) = \frac{1}{d} \sum_{q} Z^q \left(\frac{1}{d} \sum_{p} X^p \tau X^{\dagger p} \right) Z^{\dagger q} = \frac{1}{d} \sum_{p} X^p \left(\frac{1}{d} \sum_{q} Z^q \tau Z^{\dagger q} \right) X^{\dagger p}. \tag{3.35}$$

Thus we have $[\mathcal{R}(F), X] = [\mathcal{R}(F), Z] = 0$ for any $F \in \mathcal{L}(\mathcal{H})$. Hence we have

$$\mathcal{R}(F) = \frac{\mathrm{Tr}[F]}{d} \cdot I \tag{3.36}$$

by Schur's lemma.

Another example, which is widely used in quantum Shannon theory, is the Haar distributed random unitary ensemble. The Haar distribution on $\mathcal{U}(\mathcal{H})$ is defined as the unique distribution which is invariant under unitary transformations. For this ensemble $\{p(dV), V\}$, for any $G, H \in \mathcal{U}(\mathcal{H})$ we have

$$G \int p(dV) V F V^\dagger = \int p(d(GV)) G V F V^\dagger G^\dagger \cdot G = \int p(dV) V F V^\dagger G. \tag{3.37}$$

Thus the ensemble satisfies (3.32) by Schur's lemma. Due to the unitary invariance, the ensemble also satisfies

$$\int p(dV) V = 0. \tag{3.38}$$

Chapter 4

Markovianizing Cost

Tripartite states for which the quantum conditional mutual information are zero are called *quantum Markov chains* or *Markov states* for short. In this chapter, we define a task (Markovianization) of transforming a tripartite state into a state sufficiently close to a quantum Markov chain, by performing a random operation on one of the three systems. We derive the minimal cost of randomness (Markovianizing cost) required to accomplish this task. Results obtained in this chapter will be used in Chapter 5.

4.1 Quantum Markov Chains and the Markovianizing Cost

Tripartite quantum states for which the quantum conditional mutual information is *exactly* zero are called quantum Markov chain, or Markov states for short. For clarity, we say that a tripartite state Υ^{ABC} is a Markov state conditioned by B if it satisfies $I(A:C|B)=0$. It is proved in [21, 23] that the following three conditions are equivalent:

1. Υ^{ABC} is a Markov state conditioned by B,

2. There exists a unitary isomorphism $\Gamma:\mathcal{H}^B\rightarrow\mathcal{H}^{b_0}\otimes\mathcal{H}^{b_L}\otimes\mathcal{H}^{b_R}$, such that Υ^{ABC} is decomposed as

$$\Gamma^B\Upsilon^{ABC}\Gamma^{\dagger B}=\sum_{j\in J}p_j|j\rangle\langle j|^{b_0}\otimes\sigma_j^{Ab_L}\otimes\phi_j^{b_RC}, \tag{4.1}$$

3. Υ^{ABC} is decomposed as

$$\Gamma^B\Upsilon^{ABC}\Gamma^{\dagger B}=\sum_{j\in J}p_j|j\rangle\langle j|^{b_0}\otimes\sigma_j^{Ab_L}\otimes\phi_j^{b_RC},$$

where Γ is the KI decomposition of $\mathcal{H}^B$ with respect to Υ^{AB} and Υ^{BC}.

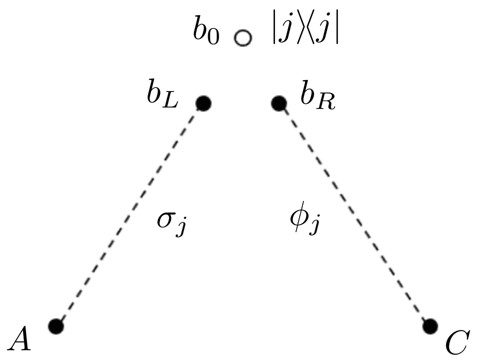


Figure 4.1: A graphical representation of the Markov decomposition of Markov states (4.1). The white circle represents the 'classical' system in the sense of (2.28). The dotted lines represent mixed states. The whole state is the probabilistic mixture of the above state with probability p_j, namely, $\sum_{j\in J} p_j |j\rangle\langle j|^{b_0} \otimes \sigma_j^{Ab_L} \otimes \phi_j^{b_R C}$.

4. Υ^{ABC} satisfies

$$\begin{aligned}\Upsilon^{ABC} &= (\Upsilon^{AB})^{\frac{1}{2}}(\Upsilon^{B})^{-\frac{1}{2}}\Upsilon^{BC}(\Upsilon^{B})^{-\frac{1}{2}}(\Upsilon^{AB})^{\frac{1}{2}} \\ &= (\Upsilon^{BC})^{\frac{1}{2}}(\Upsilon^{B})^{-\frac{1}{2}}\Upsilon^{AB}(\Upsilon^{B})^{-\frac{1}{2}}(\Upsilon^{BC})^{\frac{1}{2}}.\end{aligned}$$

We call $\Gamma : \mathcal{H}^B \to \mathcal{H}^{b_0} \otimes \mathcal{H}^{b_L} \otimes \mathcal{H}^{b_R}$ in (4.1) as the Markov decomposition of $\mathcal{H}^B$ with respect to a Markov state Υ^{ABC}, which is equivalent to the KI decomposition of $\mathcal{H}^B$ with respect to Υ^{AB} and Υ^{BC}. We call (4.1) as the Markov decomposition of a Markov state Υ^{ABC} (Figure 4.1). If a tripartite state Υ^{ABC} is decomposed as (4.1) by some Γ, we call Υ^{ABC} as a Markov state with respect to Γ.

As a stronger statement of Condition 4, we have the following properties.

Lemma 17 For arbitrary tripartite state ρ^{ABC}, the state $\check{\rho}^{ABC}$ defined as

$$\check{\rho}^{ABC} := (\rho^{BC})^{\frac{1}{2}}(\rho^{B})^{-\frac{1}{2}}\rho^{AB}(\rho^{B})^{-\frac{1}{2}}(\rho^{BC})^{\frac{1}{2}} \tag{4.2}$$

is a Markov state conditioned by B.

Proof. Define a CPTP map $\mathcal{E}_\rho : B \to BC$ by

$$\mathcal{E}_\rho(\tau) = (\rho^{BC})^{\frac{1}{2}}(\rho^{B})^{-\frac{1}{2}}\tau^{B}(\rho^{B})^{-\frac{1}{2}}(\rho^{BC})^{\frac{1}{2}}. \tag{4.3}$$

This is indeed a CPTP map because

$$\mathrm{Tr}[\mathcal{E}_\rho(\tau)] = \mathrm{Tr}[(\rho^{B})^{-\frac{1}{2}}\tau^{B}(\rho^{B})^{-\frac{1}{2}}(\rho^{BC})] = \mathrm{Tr}[\tau^{B}(\rho^{B})^{-\frac{1}{2}}\rho^{B}(\rho^{B})^{-\frac{1}{2}})] = \mathrm{Tr}[\tau]. \tag{4.4}$$

We have

$$\check{\rho}^{ABC} = \mathcal{E}_\rho(\rho^{AB}) = \mathcal{E}_\rho(\check{\rho}^{AB}). \tag{4.5}$$

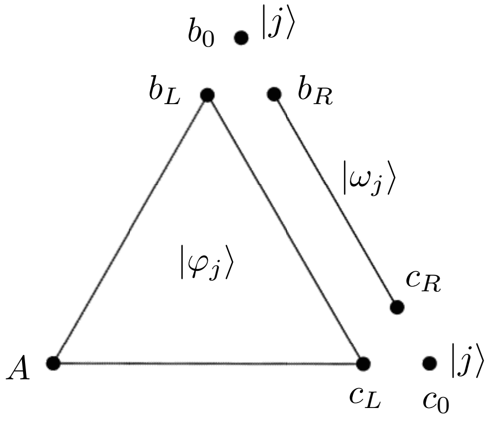


Figure 4.2: A graphical representation of the KI decomposition of tripartite pure states (4.9). Each vertex corresponds to a quantum system. The solid lines express pure states. The whole state is the superposition of the above states with the probability amplitude $\sqrt{p_j}$, namely, $\sum_{j\in J}\sqrt{p_j}|j\rangle^{b_0}|j\rangle^{c_0}|\varphi_j\rangle^{Ab_Lc_L}|\omega_j\rangle^{b_Rc_R}$.

Thus, by the data processing inequality, we have

$$0 \leq I(A:C|B)_{\check{\rho}} = I(A:BC)_{\check{\rho}} - I(A:B)_{\check{\rho}} \leq I(A:B)_{\check{\rho}} - I(A:B)_{\check{\rho}} = 0, \quad (4.6)$$

which implies $I(A:C|B)_{\check{\rho}} = 0$. ■

Corollary 18 Let Υ^{ABC} be a Markov state conditioned by B, and let $\{M_k\}_k$ be an arbitrary measurement on A. Let $p_k = \mathrm{Tr}[M_k\Upsilon M_k^\dagger]$ and $\Upsilon_k = p_k^{-1}M_k\Upsilon M_k^\dagger$. Then Υ_k are Markov states conditioned by B for all k.

Proof. Follows from (4.1). ■

Let us introduce the formal definition of Markovianization.

Definition 19 We say that a tripartite state Ψ^{ABC} is turned to a Markov state conditioned by B with the randomness cost R on A if, for any $\epsilon > 0$ and for sufficiently large n, there exist a random unitary operation $\mathcal{T}_n : \tau \mapsto 2^{-nR}\sum_{k=1}^{2^{nR}} V_k\tau V_k^\dagger$ on A^n and a Markov state $\Upsilon^{A^nB^nC^n}$ conditioned by B^n such that

$$\|\mathcal{T}_n(\Psi^{\otimes n}) - \Upsilon^{A^nB^nC^n}\|_1 \leq \epsilon. \quad (4.7)$$

The Markovianizing cost of Ψ^{ABC} is defined as $M_{A|B}(\Psi^{ABC}) := \inf\{R \mid \Psi^{ABC}$ is turned to a Markov state conditioned by B with the randomness cost R on $A\}$.

Let us introduce another adaptation of the KI decomposition for tripartite pure states.

Lemma 20 Let $|\Psi\rangle^{ABC}$ be a tripartite pure state and suppose that the KI decomposition of Ψ^{AB} on B is given by

$$\Gamma^B\Psi^{AB}\Gamma^{\dagger B} = \sum_{j\in J} p_j|j\rangle\langle j|^{b_0} \otimes \varphi_j^{Ab_L} \otimes \omega_j^{b_R}. \quad (4.8)$$

There exists a unitary isomorphism $\Gamma'^C : \mathcal{H}^C \to \mathcal{H}^{c_0} \otimes \mathcal{H}^{c_L} \otimes \mathcal{H}^{c_R}$ such that $|\Psi\rangle^{ABC}$ is decomposed as

$$(\Gamma^B \otimes \Gamma'^C)|\Psi\rangle^{ABC} = \sum_{j \in J} \sqrt{p_j} |j\rangle^{b_0} |j\rangle^{c_0} |\varphi_j\rangle^{Ab_Lc_L} |\omega_j\rangle^{b_Rc_R}, \tag{4.9}$$

where $\langle j|j'\rangle^{c_0} = \delta_{jj'}$, $|\varphi_j\rangle^{Ab_Lc_L}$ and $|\omega_j\rangle^{b_Rc_R}$ are purifications of $\varphi_j^{Ab_L}$ and $\omega_j^{b_R}$, respectively. Moreover, Γ'^C is the KI decomposition of C with respect to Ψ^{AC}.

Proof. Existence of Γ'^C follows from Uhlmann's theorem. The symmetric form of (4.9) in B and C implies that Γ'^C is the KI decomposition of $\mathcal{H}^C$ with respect to Ψ^{AC}. ■

We call (4.9) as the KI decomposition of $|\Psi\rangle^{ABC}$ on B and C (Figure 4.2). This decomposition gives the Markovianizing cost of $|\Psi\rangle$.

Theorem 21 Let $|\Psi\rangle^{ABC}$ be a pure state, and let

$$|\Psi_{KI}\rangle^{ABC} = \sum_{j \in J} \sqrt{p_j} |j\rangle^{a_0} |j\rangle^{b_0} |\omega_j\rangle^{a_Lb_L} |\varphi_j\rangle^{a_Rb_RC} \tag{4.10}$$

be the KI decomposition of $|\Psi\rangle^{ABC}$ on A and B. Then we have

$$M_{A|B}(\Psi^{ABC}) = H(\{p_j\}_{j \in J}) + 2\sum_{j \in J} p_j S(\varphi_j^{a_R}). \tag{4.11}$$

Proof. Follows from Lemma 22 and Lemma 23 shown below. ■

4.2 Upper Bound

We show the 'achievability' part of Theorem 21, namely, we show that the R.H.S. in (4.11) is the sufficient amount of randomness required for Markovianizing. The proof is based on random coding in terms of a random unitary ensemble designed according to the KI decomposition.

Lemma 22 Let $|\Psi\rangle^{ABC}$ be a pure state whose KI decomposition on A and B is given by (4.10). Then we have

$$M_{A|B}(\Psi^{ABC}) \leq H(\{p_j\}_{j \in J}) + 2\sum_{j \in J} p_j S(\varphi_j^{a_R}). \tag{4.12}$$

Proof. Fix arbitrary $\epsilon > 0$ and take sufficiently large n. Let $J_{n,\epsilon}$ be a set of all sequences $\boldsymbol{j} = j_1 \cdots j_n$ that are ϵ-strongly typical in terms of $\{p_j\}_{j \in J}$. The state $|\Psi_{KI}^{\otimes n}\rangle^{\bar{A}\bar{B}\bar{C}}$ is equal to the subnormalized state

$$|\Psi_{n,\epsilon}\rangle^{\bar{A}\bar{B}\bar{C}} = \sum_{\boldsymbol{j} \in J_{n,\epsilon}} \sqrt{p_{\boldsymbol{j}}} |\boldsymbol{j}\rangle^{\bar{a}_0} |\boldsymbol{j}\rangle^{\bar{b}_0} |\omega_{\boldsymbol{j}}\rangle^{\bar{a}_L \bar{b}_L} |\varphi_{\boldsymbol{j}}\rangle^{\bar{a}_R \bar{b}_R \bar{C}},$$

up to small error ϵ. Here, we introduce notations $\varphi_{\boldsymbol{j}} = \varphi_{j_1} \otimes \cdots \otimes \varphi_{j_n}$ and $\omega_{\boldsymbol{j}} = \omega_{j_1} \otimes \cdots \otimes \omega_{j_n}$.

For each $j \in J$ and $\boldsymbol{j} = j_1 \cdots j_n \in J_{n,\epsilon}$, define $\mathfrak{L}_{j,\boldsymbol{j}} := \{l | j_l = j, 1 \leq l \leq n\}$. The number of elements in the set is bounded as

$$n\left(p_j - \frac{\epsilon}{|J|}\right) \leq |\mathfrak{L}_{j,\boldsymbol{j}}| \leq n\left(p_j + \frac{\epsilon}{|J|}\right). \tag{4.13}$$

For each $\boldsymbol{j} \in J_{n,\epsilon}$, $(\mathcal{H}^{a_R})^{\otimes n} = \mathcal{H}^{a_{R_1}} \otimes \cdots \otimes \mathcal{H}^{a_{R_n}}$ is sorted as

$$(\mathcal{H}^{a_R})^{\otimes n} = \bigotimes_{j \in J} \bigotimes_{l \in \mathfrak{L}_{j,\boldsymbol{j}}} \mathcal{H}^{a_{R_l}}.$$

Let $\mathcal{H}_{j,\boldsymbol{j}}^{typ}$ be the ϵ-weakly typical subspace of $(\varphi_j^{a_R})^{\otimes |\mathfrak{L}_{j,\boldsymbol{j}}|}$ in $\bigotimes_{l \in \mathfrak{L}_{j,\boldsymbol{j}}} \mathcal{H}^{a_{R_l}}$, $\Pi_{j,\boldsymbol{j}}^{typ}$ be the projection onto $\mathcal{H}_{j,\boldsymbol{j}}^{typ}$, and let $\Pi_{\boldsymbol{j}}^{\bar{a}_R} := \bigotimes_{j \in J} \Pi_{j,\boldsymbol{j}}$. Define

$$\Pi^{\bar{A}} := \sum_{\boldsymbol{j} \in J_{n,\epsilon}} |\boldsymbol{j}\rangle\langle\boldsymbol{j}|^{\bar{a}_0} \otimes I^{\bar{a}_L} \otimes \Pi_{\boldsymbol{j}}^{\bar{a}_R}$$

and

$$|\Psi'_{n,\epsilon}\rangle^{\bar{A}\bar{B}\bar{C}} := \Pi^{\bar{A}} |\Psi_{n,\epsilon}\rangle^{\bar{A}\bar{B}\bar{C}} = \sum_{\boldsymbol{j} \in J_{n,\epsilon}} \sqrt{p_{\boldsymbol{j}}} |\boldsymbol{j}\rangle^{\bar{a}_0} |\boldsymbol{j}\rangle^{\bar{b}_0} |\omega_{\boldsymbol{j}}\rangle^{\bar{a}_L \bar{b}_L} \Pi_{\boldsymbol{j}}^{\bar{a}_R} |\varphi_{\boldsymbol{j}}\rangle^{\bar{a}_R \bar{b}_R \bar{C}}. \tag{4.14}$$

The subnormalized state $\Psi'_{n,\epsilon}$ is equal to $\Psi_{n,\epsilon}$ up to a small error ϵ.

Let $v_{j,\boldsymbol{j}}$ be any unitary acting on $\mathcal{H}_{j,\boldsymbol{j}}^{typ}$, and let $v_{\boldsymbol{j}} := \bigotimes_{j \in J} v_{j,\boldsymbol{j}}$. Define a unitary

$$V^{\bar{A}} := \sum_{\boldsymbol{j} \in J_{n,\epsilon}} |\boldsymbol{j}\rangle\langle\boldsymbol{j}|^{\bar{a}_0} \otimes I^{\bar{a}_L} \otimes v_{\boldsymbol{j}}^{\bar{a}_R} + P^{\bar{a}_0 \bar{a}_L \bar{a}_R}, \tag{4.15}$$

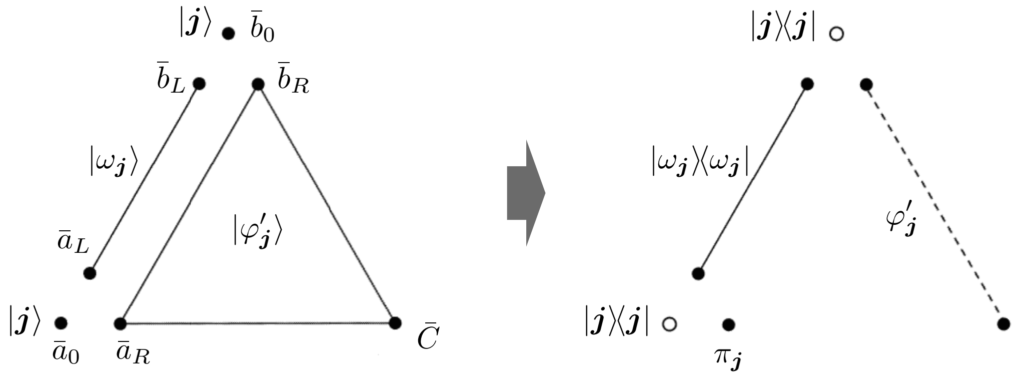
where $P^{\bar{a}_0 \bar{a}_L \bar{a}_R}$ is the projection onto the subspace that is not supported by the first term. We have

$$|\Psi'_{n,\epsilon}(V)\rangle^{\bar{A}\bar{B}\bar{C}} := V^{\bar{A}} |\Psi'_{n,\epsilon}\rangle^{\bar{A}\bar{B}\bar{C}} = \sum_{\boldsymbol{j} \in J_{n,\epsilon}} \sqrt{p_{\boldsymbol{j}}} |\boldsymbol{j}\rangle^{\bar{a}_0} |\boldsymbol{j}\rangle^{\bar{b}_0} |\omega_{\boldsymbol{j}}\rangle^{\bar{a}_L \bar{b}_L} v_{\boldsymbol{j}}^{\bar{a}_R} |\varphi'_{\boldsymbol{j}}\rangle^{\bar{a}_R \bar{b}_R \bar{C}},$$

where $|\varphi'_{\boldsymbol{j}}\rangle := \Pi_{\boldsymbol{j}}^{\bar{a}_R} |\varphi_{\boldsymbol{j}}\rangle$.

Let $\{p(dV), V\}$ be the ensemble of unitaries generated by choosing $v_{j,\boldsymbol{j}}$ randomly and independently according to the Haar measure for each j and $\boldsymbol{j}$ in (4.15). From (3.28) and (3.38), as an ensemble average, we have

$$\mathbb{E}\left[v_{\boldsymbol{j}}^{\bar{a}_R} |\varphi'_{\boldsymbol{j}}\rangle\langle\varphi'_{\boldsymbol{j}}| v_{\boldsymbol{j}}^{\dagger \bar{a}_R}\right] = \pi_{\boldsymbol{j}}^{\bar{a}_R} \otimes \varphi_{\boldsymbol{j}}'^{\bar{b}_R \bar{C}},$$



$|\boldsymbol{j}\rangle$
$\bar{b}_0$
$\bar{b}_L$
$\bar{b}_R$
$|\omega_{\boldsymbol{j}}\rangle$
$|\varphi'_{\boldsymbol{j}}\rangle$
$\bar{a}_L$
$|\boldsymbol{j}\rangle$
$\bar{a}_0$
$\bar{a}_R$
$\bar{C}$
$|\boldsymbol{j}\rangle\langle\boldsymbol{j}|$
$|\omega_{\boldsymbol{j}}\rangle\langle\omega_{\boldsymbol{j}}|$
$\varphi'_{\boldsymbol{j}}$
$|\boldsymbol{j}\rangle\langle\boldsymbol{j}|$
$\pi_{\boldsymbol{j}}$

Figure 4.3: A graphical representation of the state transformation from $\Psi'_{n,\epsilon}$ in (4.14) to $\bar{\Psi}_{n,\epsilon}$ in (4.16) by a random unitary operation $\{p(dV), V\}$.

where $\pi_{\boldsymbol{j}}^{\bar{a}_R} = \Pi_{\boldsymbol{j}}/\mathrm{Tr}\Pi_{\boldsymbol{j}}$, and

$$\mathbb{E}\left[v_{\boldsymbol{j}}^{\bar{a}_R} |\varphi'_{\boldsymbol{j}}\rangle\langle\varphi'_{\boldsymbol{j}'}| v_{\boldsymbol{j}'}^{\dagger\bar{a}_R} \right] = 0$$

for $\boldsymbol{j} \neq \boldsymbol{j}'$. Thus the average state is given by

$$\begin{aligned}
\bar{\Psi}_{n,\epsilon} &:= \mathbb{E}\left[\left|\Psi'_{n,\epsilon}(V)\right\rangle\!\left\langle\Psi'_{n,\epsilon}(V)\right|^{\bar{A}\bar{B}\bar{C}} \right] \\
&= \sum_{\boldsymbol{j}\in J_{n,\epsilon}} p_{\boldsymbol{j}} |\boldsymbol{j}\boldsymbol{j}\rangle\langle\boldsymbol{j}\boldsymbol{j}|^{\bar{a}_0\bar{b}_0} \otimes |\omega_{\boldsymbol{j}}\rangle\langle\omega_{\boldsymbol{j}}|^{\bar{a}_L\bar{b}_L} \otimes \pi_{\boldsymbol{j}}^{\bar{a}_R} \otimes \varphi_{\boldsymbol{j}}^{\prime\bar{b}_R\bar{C}}, \\
&= \sum_{\boldsymbol{j}\in J_{n,\epsilon}} p_{\boldsymbol{j}} |\boldsymbol{j}\rangle\langle\boldsymbol{j}|^{\bar{b}_0} \otimes \left(\pi_{\boldsymbol{j}}^{\bar{a}_R} \otimes |\boldsymbol{j},\omega_{\boldsymbol{j}}\rangle\langle\boldsymbol{j},\omega_{\boldsymbol{j}}|^{\bar{a}_0\bar{a}_L\bar{b}_L} \right) \otimes \varphi_{\boldsymbol{j}}^{\prime\bar{b}_R\bar{C}} \qquad (4.16)
\end{aligned}$$

which is a Markov state conditioned by $\bar{B}$ (Figure 4.3).

The minimum nonzero eigenvalue of $\bar{\Psi}_{n,\epsilon}$ is calculated as follows. First, due to the definition of $J_{n,\epsilon}$, we have

$$p_{\boldsymbol{j}} \geq \prod_{j\in J} p_j^{n(p_j+\epsilon/|J|)} = 2^{-n(H(\{p_j\}_j)+\eta(\epsilon))}.$$

Second, since the spectrums of $\varphi_{\boldsymbol{j}}^{\prime\bar{a}_R}$ and $\varphi_{\boldsymbol{j}}^{\prime\bar{b}_R\bar{C}}$ are the same, the minimum nonzero eigenvalue $\mu_{\boldsymbol{j}}$ of $\varphi_{\boldsymbol{j}}^{\prime\bar{b}_R\bar{C}}$ is bounded below as

$$\mu_{\boldsymbol{j}} \geq \prod_{j\in J} 2^{-n(p_j+\epsilon/|J|)(S(\varphi_j^{a_R})+\epsilon)} = 2^{-n\left(\sum_j p_j S(\varphi_j^{a_R})+\eta(\epsilon)\log d_A\right)}.$$

Third, we have

$$\mathrm{rank}\,\Pi_{j,\boldsymbol{j}} \leq 2^{|\mathfrak{L}_{j,\boldsymbol{j}}|(S(\varphi_j^{a_R})+\epsilon)} \leq 2^{n(p_j+\epsilon/|J|)(S(\varphi_j^{a_R})+\epsilon)}$$

and

$$\mathrm{rank}\,\Pi_{\boldsymbol{j}} \leq \prod_{j\in J} 2^{n(p_j+\epsilon/|J|)(S(\varphi_j^{a_R})+\epsilon)}.$$

Thus the nonzero eigenvalue $\nu_{\boldsymbol{j}}$ of $\pi_{\boldsymbol{j}}^{\bar{a}_R}$ is, in the same way as $\mu_{\boldsymbol{j}}$, bounded below as

$$\nu_{\boldsymbol{j}} \geq \prod_{j\in J} 2^{-n(p_j+\epsilon/|J|)(S(\varphi_j^{a_R})+\epsilon)}.$$

All in all, the minimum nonzero eigenvalue λ of $\bar{\Psi}_{n,\epsilon}$ is bounded as

$$\lambda = p_{\boldsymbol{j}}\mu_{\boldsymbol{j}}\nu_{\boldsymbol{j}} \geq 2^{-n\left[H(\{p_j\}_j)+2\sum_j p_j S(\varphi_j^{a_R})+\eta(\epsilon)\log d_A\right]}.$$

We also have

$$\mathrm{rank}\,\bar{\Psi}_{n,\epsilon} \leq |J_{n,\epsilon}| \times \mathrm{rank}\,\pi_{\boldsymbol{j}}^{\bar{a}_R} \times \mathrm{rank}\,\varphi_{\boldsymbol{j}}^{\prime\bar{a}_R} \leq d_A^{3n}, \tag{4.17}$$

where $d_A = \dim \mathcal{H}^A$.

Suppose $V_1, \cdots, V_N$ are unitaries that are randomly and independently chosen from the ensemble $\{p(dV), V\}$. From Lemma 16, we have

$$\Pr\left\{\frac{1}{N}\sum_{i=1}^{N}\Psi'_{n,\epsilon}(V_i) \notin [(1-\epsilon)\bar{\Psi}_{n,\epsilon}, (1+\epsilon)\bar{\Psi}_{n,\epsilon}]\right\} \leq 2d_A^{3n}\exp\left(-\frac{N\lambda\epsilon^2}{2}\right).$$

Therefore, if $N = 2^{nR}$ and

$$R > H(\{p_j\}_j) + 2\sum_j p_j S(\varphi_j^{a_R}) + \eta(\epsilon)\log d_A$$

holds, there exists a set of unitaries $V_1, \cdots, V_N$ such that

$$\left\|\frac{1}{N}\sum_{i=1}^{N}\Psi'_{n,\epsilon}(V_i) - \bar{\Psi}_{n,\epsilon}\right\|_1 \leq \epsilon.$$

for sufficiently large n. Thus we obtain (4.12). ■

4.3 Lower Bound

In this section, we prove the converse part in Theorem 21, namely, we show that the R.H.S. in (4.11) is the necessary amount of randomness required for Markovianization. If we would assume that $\Upsilon^{A^nB^nC^n}$ in (4.7) is a Markov state with respect to $(\Gamma_\Psi^B)^{\otimes n}$, where Γ_Ψ is the KI decomposition of B with respect to Ψ^{BC}, it is not difficult to show that the amount of randomness required for turning $\Psi^{\otimes n}$ to $\Upsilon^{A^nB^nC^n}$ per copy is bounded below by the R.H.S. in (4.11). However, it might be possible in general that the amount of randomness can be further reduced by appropriately choosing $\Upsilon^{A^nB^nC^n}$ and the corresponding KI decomposition of $\bar{B}$. We show that this is impossible. At the core of the proof is Inequality (4.28), which will be discussed in detail in Section 4.4.

Lemma 23 Let $|\Psi\rangle^{ABC}$ be a state whose KI decomposition on A and B is given by (4.10). Then we have

$$M_{A|B}(\Psi^{ABC}) \geq H(\{p_j\}_{j\in J}) + 2\sum_{j\in J} p_j S(\varphi_j^{a_R}). \tag{4.18}$$

Proof. Take arbitrary $R > M_{A|B}(\Psi^{ABC})$, $\epsilon > 0$ and choose sufficiently large n. There exist a random unitary operation $\mathcal{T}_n : \tau \mapsto 2^{-nR}\sum_{k=1}^{2^{nR}} V_k\tau V_k^\dagger$ on A^n and a Markov state $\Upsilon^{A^nB^nC^n}$ conditioned by B^n such that

$$\left\|\mathcal{T}_n(\Psi^{\otimes n}) - \Upsilon^{A^nB^nC^n}\right\|_1 \leq \epsilon. \tag{4.19}$$

By tracing out A^n, we have

$$\left\|(\Psi^{\otimes n})^{\bar{B}\bar{C}} - \Upsilon^{\bar{B}\bar{C}}\right\|_1 \leq \epsilon. \tag{4.20}$$

Due to Uhlmann's theorem (see Section 2.2), there exists a state $|\chi\rangle^{\bar{A}\bar{B}\bar{C}}$ such that $\chi^{\bar{B}\bar{C}} = \Upsilon^{\bar{B}\bar{C}}$ and

$$\left\|(\Psi^{\otimes n})^{\bar{A}\bar{B}\bar{C}} - \chi^{\bar{A}\bar{B}\bar{C}}\right\|_1 \leq \eta(\epsilon). \tag{4.21}$$

Let $\Gamma_\Upsilon : \mathcal{H}^{\bar{B}} \to \mathcal{H}^{\hat{b}_0} \otimes \mathcal{H}^{\hat{b}_L} \otimes \mathcal{H}^{\hat{b}_R}$ be the Markov decomposition of $\mathcal{H}^{\bar{B}}$ with respect to $\Upsilon^{\bar{A}\bar{B}\bar{C}}$, and let

$$\Upsilon_{Mk}^{\bar{A}\bar{B}\bar{C}} := \Gamma_\Upsilon^{\bar{B}}\Upsilon^{\bar{A}\bar{B}\bar{C}}\Gamma_\Upsilon^{\dagger\bar{B}} = \sum_i q_i|i\rangle\langle i|^{\hat{b}_0} \otimes \sigma_i^{\bar{A}\hat{b}_L} \otimes \phi_i^{\hat{b}_R\bar{C}} \tag{4.22}$$

be the Markov decomposition of $\Upsilon^{\bar{A}\bar{B}\bar{C}}$. Due to $\chi^{\bar{B}\bar{C}} = \Upsilon^{\bar{B}\bar{C}}$, the KI decomposition of $\mathcal{H}^{\bar{B}}$ with respect to $|\chi\rangle$ is equal to $\Gamma_\Upsilon^{\bar{B}}$. Thus the KI decomposition of $|\chi\rangle$ on $\bar{A}$ and $\bar{B}$ is given by

$$|\chi_{KI}\rangle := (\Gamma_\chi^{\bar{A}} \otimes \Gamma_\Upsilon^{\bar{B}})|\chi\rangle = \sum_i \sqrt{q_i}|i\rangle^{\hat{a}_0}|i\rangle^{\hat{b}_0}|\xi_i\rangle^{\hat{a}_L\hat{b}_L}|\phi_i\rangle^{\hat{a}_R\hat{b}_R\bar{C}}, \tag{4.23}$$

where $|\xi_i\rangle^{\hat{a}_L\hat{b}_L}$ and $|\phi_i\rangle^{\hat{a}_R\hat{b}_R\bar{C}}$ are purifications of $\sigma_i^{\hat{b}_L}$ and $\phi_i^{\hat{b}_R\bar{C}}$, respectively, and $\Gamma_\chi^{\bar{A}} : \bar{A} \to \hat{a}_0\hat{a}_L\hat{a}_R$ is the KI decomposition of $\bar{A}$ with respect to $\chi^{\bar{A}\bar{C}}$.

Define a map $\mathcal{T}_n' := \mathcal{T}_n \circ \mathcal{E}_{\Gamma_\chi^\dagger}$ on $\bar{A}$, where $\mathcal{E}_{\Gamma_\chi^\dagger}$ is a unitary isomorphic operation corresponding to $\Gamma_\chi^\dagger$, and let $\mathcal{D}^{\hat{b}_0}$ be the complete dephasing operation on $\hat{b}_0$ with respect to the basis $\{|i\rangle^{\hat{b}_0}\}_i$. From (4.19), (4.21), (4.22) and (4.23), we have

$$\left\|(\mathcal{T}_n' \otimes \mathcal{D}^{\hat{b}_0})(|\chi_{KI}\rangle\langle\chi_{KI}|) - \Upsilon_{Mk}^{\bar{A}\bar{B}\bar{C}}\right\|_1 \leq \eta(\epsilon). \tag{4.24}$$

We also have

$$\mathcal{D}^{\hat{b}_0}(|\chi_{KI}\rangle\langle\chi_{KI}|) = \sum_i q_i|i\rangle\langle i|^{\hat{a}_0} \otimes |i\rangle\langle i|^{\hat{b}_0} \otimes |\xi_i\rangle\langle\xi_i|^{\hat{a}_L\hat{b}_L} \otimes |\phi_i\rangle\langle\phi_i|^{\hat{a}_R\hat{b}_R\bar{C}}. \tag{4.25}$$

From (4.22), (4.24) and (4.25), by tracing out $\hat{b}_R\bar{C}$, we obtain

$$\sum_i q_i \left\| \mathcal{T}_n'(|i,\xi_i\rangle\langle i,\xi_i|^{\hat{a}_0\hat{a}_L\hat{b}_L} \otimes \phi_i^{\hat{a}_R}) - \sigma_i^{\bar{A}\hat{b}_L} \right\|_1 \leq \eta(\epsilon).$$

Thus we have

$$\sum_i q_i S(\sigma_i^{\bar{A}\hat{b}_L}) \geq \sum_i q_i S(\phi_i^{\hat{a}_R}) - n\eta(\epsilon)\log(d_A d_B), \tag{4.26}$$

since the von Neumann entropy is nondecreasing under random unitary operations.

The von Neumann entropy of the state $\Upsilon^{\bar{A}\bar{B}\bar{C}}$ is then bounded below as

$$\begin{aligned} S(\bar{A}\bar{B}\bar{C})_{\Upsilon} &= S(\bar{A}\hat{b}_0\hat{b}_L\hat{b}_R\bar{C})_{\Upsilon_{Mk}} = S(\hat{b}_0)_{\Upsilon_{Mk}} + S(\bar{A}\hat{b}_L\hat{b}_R\bar{C}|\hat{b}_0)_{\Upsilon_{Mk}} \\ &= H(\{q_i\}_i) + \sum_i q_i(S(\sigma_i^{\bar{A}\hat{b}_L}) + S(\phi_i^{\hat{b}_R\bar{C}})) \\ &= H(\{q_i\}_i) + \sum_i q_i(S(\sigma_i^{\bar{A}\hat{b}_L}) + S(\phi_i^{\hat{a}_R})) \\ &\geq H(\{q_i\}_i) + 2\sum_i q_i S(\phi_i^{\hat{a}_R}) - n\eta(\epsilon)\log(d_A d_B). \end{aligned} \tag{4.27}$$

We prove in Lemma 24 that (4.21) implies

$$H(\{q_i\}_i) + 2\sum_i q_i S(\phi_i^{\hat{a}_R}) \geq n\left(H(\{p_j\}_{j\in J}) + 2\sum_{j\in J} p_j S(\varphi_j^{a_R}) - \eta_{\Psi}(\epsilon)\log d_C\right). \tag{4.28}$$

Here, $\eta_{\Psi}(\epsilon)$ is a proper function of $\epsilon > 0$ and Ψ satisfying $\lim_{\epsilon\to 0}\eta_{\Psi}(\epsilon) = 0$, which is continuous with respect to ϵ, and does not depend on n. Thus, from (4.19), (4.27) and (4.28), we finally obtain

$$\begin{aligned} R &\geq \frac{1}{n}S(\bar{A}\bar{B}\bar{C})_{\mathcal{T}_n(\Psi^{\otimes n})} \geq \frac{1}{n}S(\bar{A}\bar{B}\bar{C})_{\Upsilon} - \eta(\epsilon)\log(d_A d_B d_C) \\ &\geq H(\{p_j\}_{j\in J}) + 2\sum_{j\in J} p_j S(\varphi_j^{a_R}) - \eta_{\Psi}(\epsilon)\log(d_A d_B d_C), \end{aligned}$$

which implies (4.18). ■

4.4 Proof of Entropic Inequality for KI decomposition

In this section, we prove Inequality (4.28). Our proof is based on the data compression theorem for mixed state quantum information sources proposed in [24, 25].

Lemma 24 Suppose the KI decomposition of $|\Psi\rangle^{ABC}$ on A and B is given by (4.10). For any n and $\epsilon > 0$, let $\chi^{\bar{A}\bar{C}}$ be a state that satisfies

$$\left\| (\Psi^{\otimes n})^{\bar{A}\bar{C}} - \chi^{\bar{A}\bar{C}} \right\|_1 \leq \epsilon, \tag{4.29}$$

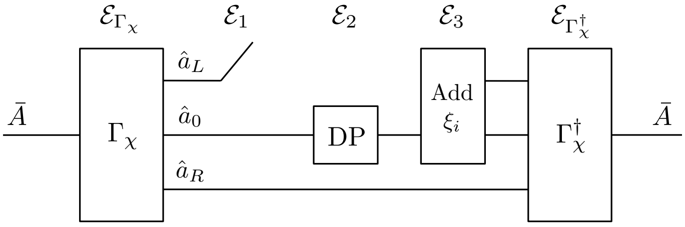


Figure 4.4: A graphical representation of the channel $\mathcal{E}^{\bar{A}}$. Due to the complete dephasing operation denoted as DP, the channel $\hat{a}_0$ has some capacity to transmit classical information, but has no capacity to transfer entanglement. The channel $\hat{a}_R$ has some capacity for both.

and let

$$\chi_{KI}^{\bar{A}\bar{C}} = \sum_i q_i |i\rangle\langle i|^{\hat{a}_0} \otimes \xi_i^{\hat{a}_L} \otimes \phi_i^{\hat{a}_R\bar{C}} \tag{4.30}$$

be the KI decomposition of $\chi^{\bar{A}\bar{C}}$ on $\bar{A}$. Then we have

$$H(\{q_i\}_i) + 2\sum_i q_i S(\phi_i^{\hat{a}_R}) \geq n\left(H(\{p_j\}_{j\in J}) + 2\sum_{j\in J} p_j S(\varphi_j^{a_R}) - \eta_\Psi(\epsilon)\log(d_A d_C)\right). \tag{4.31}$$

Proof. We prove two inequalities,

$$\sum_i q_i S(\phi_i^{\hat{a}_L}) \geq n\left(\sum_{j\in J} p_j S(\varphi_j^{a_R}) - \eta_\Psi(\epsilon)\log(d_A d_C)\right) \tag{4.32}$$

and

$$H(\{q_i\}_i) + \sum_i q_i S(\phi_i^{\hat{a}_R}) \geq n\left(H(\{p_j\}_{j\in J}) + \sum_{j\in J} p_j S(\varphi_j^{a_R}) - \eta_\Psi(\epsilon)\log(d_A d_C)\right), \tag{4.33}$$

which together imply (4.31).

Let Γ_χ be the KI decomposition of $\bar{A}$ with respect to $\chi^{\bar{A}\bar{C}}$. Consider a quantum channel $\mathcal{E}$ on $\bar{A}$ defined by

$$\mathcal{E}(\tau) = \Gamma_\chi^\dagger\left(\sum_i |i\rangle\langle i|^{\hat{a}_0}\mathrm{Tr}_{\hat{a}_L}[\Gamma_\chi \tau \Gamma_\chi^\dagger]|i\rangle\langle i|^{\hat{a}_0} \otimes \xi_i^{\hat{a}_L}\right)\Gamma_\chi.$$

This channel is decomposed as $\mathcal{E}_{\Gamma_\chi^\dagger} \circ \mathcal{E}_3 \circ \mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_{\Gamma_\chi}$, where $\mathcal{E}_{\Gamma_\chi}$ and $\mathcal{E}_{\Gamma_\chi^\dagger}$ are an isometry operation corresponding to Γ_χ and $\Gamma_\chi^\dagger$, $\mathcal{E}_1$ is discarding $\hat{a}_L$, $\mathcal{E}_2$ is the completely dephasing operation on $\hat{a}_0$ with respect to the basis $|i\rangle$, and $\mathcal{E}_3$ is addition

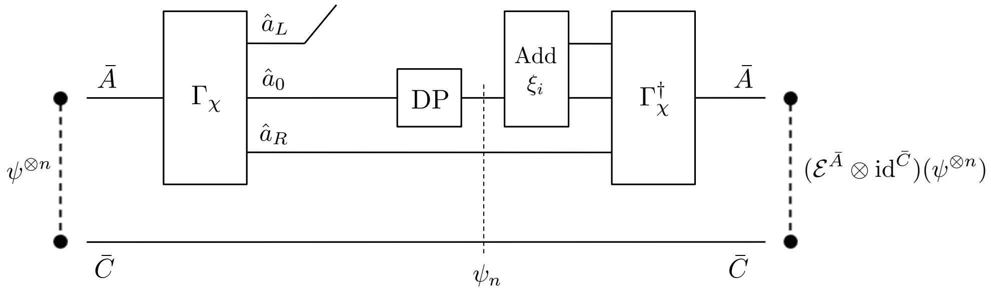


Figure 4.5: A graphical representation of state transformation of $\psi^{\otimes n}$ under $\mathcal{E}$. The channel $\mathcal{E}$ does not much change the state $\psi^{\otimes n}$ on average. In particular, it almost conserves entanglement that $\psi^{\otimes n}$ initially has. Thus the intermediate state ψ_n has the same amount of entanglement because of the monotonicity. Since $\hat{a}_0$ has no capacity for transferring entanglement, $\hat{a}_R$ holds all entanglement in ψ_n with $\bar{C}$. We evaluate entanglement by the conditional entropy.

of $\hat{a}_L$ in the state $\xi_i^{\hat{a}_L}$, conditioned by $\hat{a}_0$ (Figure 4.4). From (4.29) and (4.30), we have $\mathcal{E}(\chi^{\bar{A}\bar{C}}) = \chi^{\bar{A}\bar{C}}$ and thus

$$\left\| \mathcal{E}(\Psi^{\otimes n})^{\bar{A}\bar{C}} - (\Psi^{\otimes n})^{\bar{A}\bar{C}} \right\|_1 \leq \eta(\epsilon). \tag{4.34}$$

There exists a POVM $\{M_\mu\}_\mu$ on C such that the KI decomposition of ensemble $\{p_\mu, \Psi_\mu\}_\mu$, where

$$p_\mu := \mathrm{Tr}[(I^A \otimes M_\mu^C)\Psi^{AC}] \tag{4.35}$$

and

$$\Psi_\mu^A := p_\mu^{-1}\mathrm{Tr}_C[(I^A \otimes M_\mu^C)\Psi^{AC}], \tag{4.36}$$

is equivalent to the KI decomposition of Ψ^{AC}. Let $\boldsymbol{\mu} = \mu_1 \cdots \mu_n$, $p_{\boldsymbol{\mu}} = p_{\mu_1} \cdots p_{\mu_n}$ and $\Psi_{\boldsymbol{\mu}}^{\bar{A}} = \Psi_{\mu_1}^{A_1} \otimes \cdots \otimes \Psi_{\mu_n}^{A_n}$. From (4.34), we obtain

$$\eta(\epsilon) \geq \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} \left\| \mathcal{E}(\Psi_{\boldsymbol{\mu}}^{\bar{A}}) - \Psi_{\boldsymbol{\mu}}^{\bar{A}} \right\|_1 \geq \max_{1 \leq l \leq n} \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} \left\| \mathcal{E}(\Psi_{\boldsymbol{\mu}}^{\bar{A}})^{A_l} - \Psi_{\mu_l}^{A_l} \right\|_1. \tag{4.37}$$

To prove (4.32), let ψ^{AC} be the state such that the KI decomposition of A with respect to ψ^{AC} is the same as one with respect to Ψ^{AC}, and that it is decomposed as

$$\psi_{KI}^{AC} := \Gamma_\Psi^A \psi^{AC} \Gamma_\Psi^{\dagger A} = \sum_{j \in J} p_j |j\rangle\langle j|^{a_0} \otimes \omega_j^{a_L} \otimes |\tilde{\varphi}_j\rangle\langle\tilde{\varphi}_j|^{a_R C},$$

where $|\tilde{\varphi}_j\rangle^{a_R C}$ is a purification of $\varphi_j^{a_R}$ (Figure 4.5). The state satisfies $\psi^A = \Psi^A$. It is proved in [24, 25] that the condition (4.37) implies

$$\max_{1 \leq l \leq n} \left\| \mathcal{E}(\psi^{\otimes n})^{A_l C_l} - \psi^{A_l C_l} \right\|_1 \leq \eta_\Psi(\epsilon).$$

Here, $\mathcal{E}(\psi^{\otimes n})^{A_l C_l}$ denotes

$$\mathcal{E}(\psi^{\otimes n})^{A_l C_l} = \Big((\mathcal{E}^{\bar{A}} \otimes \mathrm{id}^{\bar{C}})(\psi^{\otimes n}) \Big)^{A_l C_l}. \tag{4.38}$$

Thus we have

$$\frac{1}{n} \sum_{l=1}^{n} \left\| \mathcal{E}(\psi^{\otimes n})^{A_l C_l} - \psi^{A_l C_l} \right\|_1 \leq \eta_{\Psi}(\epsilon), \tag{4.39}$$

and consequently,

$$\sum_{l=1}^{n} \left| S(C_l|A_l)_{\mathcal{E}(\psi^{\otimes n})} - S(C|A)_{\psi} \right| \leq n\eta_{\Psi}(\epsilon) \log d_C. \tag{4.40}$$

from (2.86). We also have

$$\begin{aligned} S(\bar{C}|\bar{A})_{\mathcal{E}(\psi^{\otimes n})} &= S(C_1 \cdots C_n | A_1 \cdots A_n)_{\mathcal{E}(\psi^{\otimes n})} \\ &= \sum_{l=1}^{n} S(C_l | A_1 \cdots A_n C_1 \cdots C_{l-1})_{\mathcal{E}(\psi^{\otimes n})} \\ &\leq \sum_{l=1}^{n} S(C_l|A_l)_{\mathcal{E}(\psi^{\otimes n})} \end{aligned} \tag{4.41}$$

Combining (4.40) and (4.41), we obtain

$$nS(C|A)_{\psi} \geq S(\bar{C}|\bar{A})_{\mathcal{E}(\psi^{\otimes n})} - n\eta_{\Psi}(\epsilon) \log d_C. \tag{4.42}$$

Let $\psi_n^{\hat{a}_0 \hat{a}_R \bar{C}} := (\mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_{\Gamma_\chi})((\psi^{\otimes n})^{\bar{A}\bar{C}})$. By the data processing inequality, we have

$$S(\bar{C}|\bar{A})_{\mathcal{E}(\psi^{\otimes n})} \geq S(\bar{C}|\hat{a}_0 \hat{a}_R)_{\psi_n}. \tag{4.43}$$

Since $\psi_n^{\hat{a}_0 \hat{a}_R \bar{C}}$ is a classical-quantum state between $\hat{a}_0$ and $\hat{a}_R \bar{C}$, we have

$$S(\bar{C}|\hat{a}_0 \hat{a}_R)_{\psi_n} = S(\hat{a}_R \bar{C}|\hat{a}_0)_{\psi_n} - S(\hat{a}_R|\hat{a}_0)_{\psi_n} \geq -S(\hat{a}_R|\hat{a}_0)_{\psi_n}. \tag{4.44}$$

On the other hand, from the condition (4.29) and the fact that $\psi^A = \Psi^A$, we have

$$|S(\hat{a}_R|\hat{a}_0)_{\chi_{KI}} - S(\hat{a}_R|\hat{a}_0)_{\psi_n}| \leq n\eta(\epsilon) \log d_A. \tag{4.45}$$

Combining (4.42), (4.43), (4.44) and (4.45), we obtain

$$nS(C|A)_{\psi} \geq -S(\hat{a}_R|\hat{a}_0)_{\chi_{KI}} - n\eta_{\Psi}(\epsilon) \log (d_A d_C).$$

Since we have, from (2.74), that

$$S(C|A)_{\psi} \;=\; S(C|a_0 a_L a_R)_{\psi_{KI}} = \sum_{j \in J} p_j S(C|a_R)_{\tilde{\varphi}_j} = -\sum_{j \in J} p_j S(\varphi_j^{a_R})$$

and

$$S(\hat{a}_R|\hat{a}_0)_{\chi_{KI}} = \sum_i q_i S(\phi_i^{\hat{a}_R}),$$

we obtain (4.32).

To prove (4.33), let

$$|\tilde{\varphi}_j\rangle^{a_R C} = \sum_k \sqrt{p_{k|j}} |\tilde{\varphi}_{k|j}\rangle^{a_R} |\tilde{\varphi}_{k|j}\rangle^C \tag{4.46}$$

be the Schmidt decomposition of $|\tilde{\varphi}_j\rangle^{a_R C}$, and define a classical-classical state

$$\tilde{\varphi}_{j,\mathrm{Cl}}^{a_R C} := \sum_k p_{k|j} |\tilde{\varphi}_{k|j}\rangle\langle\tilde{\varphi}_{k|j}|^{a_R} \otimes |\tilde{\varphi}_{k|j}\rangle\langle\tilde{\varphi}_{k|j}|^C. \tag{4.47}$$

Since $\tilde{\varphi}_{j,\mathrm{Cl}}^{a_R} = \tilde{\varphi}_{\mathrm{Cl}}^{a_R}$, we have $\psi_{\mathrm{Cl}}^A = \psi^A = \Psi^A$. Let $\psi_{\mathrm{Cl}}^{ACC'}$ be the state such that the KI decomposition of A with respect to $\psi_{\mathrm{Cl}}^{ACC'}$ is the same as one with respect to Ψ^{AC}, and that it is decomposed as

$$\begin{aligned} \psi_{\mathrm{Cl},KI}^{ACC'} &:= \Gamma_\Psi^A \psi_{\mathrm{Cl}}^{ACC'} \Gamma_\Psi^{\dagger A} \\ &= \sum_{j\in J} p_j |j\rangle\langle j|^{a_0} \otimes \omega_j^{a_L} \otimes \tilde{\varphi}_{j,\mathrm{Cl}}^{a_R C} \otimes |j\rangle\langle j|^{C'}. \end{aligned}$$

Similar to (4.39), the condition (4.37) implies

$$\frac{1}{n}\sum_{l=1}^n \left\| \mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})^{A_l C_l C_l'} - \psi_{\mathrm{Cl}}^{A_l C_l C_l'} \right\|_1 \leq \eta_\Psi(\epsilon).$$

Thus we have

$$\sum_{l=1}^n \left| I(A_l : C_l C_l')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} - I(A : CC')_{\psi_{\mathrm{Cl}}} \right| \leq n\eta_\Psi(\epsilon)\log(d_A d_C). \tag{4.48}$$

We also have

$$\begin{aligned} I(\bar{A} : \bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} &= S(\bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} - S(\bar{C}\bar{C}'|\bar{A})_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} \\ &= S(\bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} - \sum_{l=1}^n S(C_l C_l'|A_1 \cdots A_n C_1 C_1' \cdots C_{l-1} C_{l-1}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} \\ &\geq S(\bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} - \sum_{l=1}^n S(C_l C_l'|A_l)_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} \\ &= \sum_{l=1}^n I(A_l : C_l C_l')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})}. \end{aligned} \tag{4.49}$$

Here, we use the fact that $\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})^{\bar{C}\bar{C}'} = (\psi_{\mathrm{Cl}}^{\otimes n})^{\bar{C}\bar{C}'}$, and thus

$$S(\bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} = S(\bar{C}\bar{C}')_{\psi_{\mathrm{Cl}}^{\otimes n}} = nS(CC')_{\psi_{\mathrm{Cl}}} = \sum_{l=1}^n S(C_l C_l')_{\psi_{\mathrm{Cl}}^{\otimes n}} = \sum_{l=1}^n S(C_l C_l')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})}.$$

Let $\psi_{n,\mathrm{Cl}}^{\hat{a}_0\hat{a}_R\bar{C}\bar{C}'} := (\mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_{\Gamma_\chi})((\psi_{\mathrm{Cl}}^{\otimes n})^{\bar{A}\bar{C}\bar{C}'})$. By the data processing inequality, we have

$$I(\bar{A}:\bar{C}\bar{C}')_{\mathcal{E}(\psi_{\mathrm{Cl}}^{\otimes n})} \leq I(\hat{a}_0\hat{a}_R:\bar{C}\bar{C}')_{\psi_{n,\mathrm{Cl}}} \leq S(\hat{a}_0\hat{a}_R)_{\psi_{n,\mathrm{Cl}}}. \tag{4.50}$$

From (4.29) and the fact that $\psi_{\mathrm{Cl}}^A = \Psi^A$, we have

$$\left|S(\hat{a}_0\hat{a}_R)_{\chi_{KI}} - S(\hat{a}_0\hat{a}_R)_{\psi_{n,\mathrm{Cl}}}\right| \leq n\eta(\epsilon)\log d_A. \tag{4.51}$$

Combining (4.48), (4.49), (4.50), and (4.51), we obtain

$$S(\hat{a}_0\hat{a}_R)_{\chi_{KI}} \geq nI(A:CC')_{\psi_{\mathrm{Cl}}} - n\eta_{\Psi}(\epsilon)\log(d_A d_C).$$

Therefore we obtain (4.33), since we have

$$S(\hat{a}_0\hat{a}_R)_{\chi_{KI}} = S(\hat{a}_0)_{\chi_{KI}} + S(\hat{a}_R|\hat{a}_0)_{\chi_{KI}} = H(\{q_i\}_i) + \sum_i q_i S(\phi_i^{\hat{a}_R})$$

and

$$\begin{aligned} I(A:CC')_{\psi_{\mathrm{Cl}}} &= I(a_0a_R:CC')_{\psi_{\mathrm{Cl},KI}} = I(a_0:CC')_{\psi_{\mathrm{Cl},KI}} + I(a_R:CC'|a_0)_{\psi_{\mathrm{Cl},KI}} \\ &= H(\{p_j\}_{j\in J}) + \sum_{j\in J} p_j S(\varphi_j^{a_R}). \end{aligned}$$

■

4.5 Measurement-Induced Markovianization

In this section, we consider Markovianization by performing measurements instead of applying random unitary operations. We analyze changes of entropic quantities under a state transformation induced by a measurement. We show that a lower bound on entropy changes is given by the Markovianizing cost derived in the previous sections. Results obtained in this section are used in the next chapter.

Theorem 25 For an arbitrary pure state $|\Psi\rangle^{ABC}$ and any $n, \epsilon > 0$, let $\{M_k^{\bar{A}A_0\to A'}\}_k$ be a measurement on $\bar{A}A_0$, $|\phi_{res}\rangle^{A_0G}$ be an arbitrary pure state, $r_k := \| M_k|\Psi^{\otimes n}\rangle^{\bar{A}\bar{B}\bar{C}}$ $|\phi_{res}\rangle^{A_0G}\|_1^2$ and $|\Psi_k\rangle^{A'\bar{B}\bar{C}G} := r_k^{-1/2} M_k|\Psi\rangle^{\bar{A}\bar{B}\bar{C}}$ $|\phi_{res}\rangle^{A_0G}$. Suppose that the following conditions are satisfied.

1. The measurement does not much change the reduced state on $\bar{B}\bar{C}$ on average:

$$\sum_k r_k \left\|(\Psi^{\otimes n})^{\bar{B}\bar{C}} - \Psi_k^{\bar{B}\bar{C}}\right\|_1 \leq \epsilon.$$

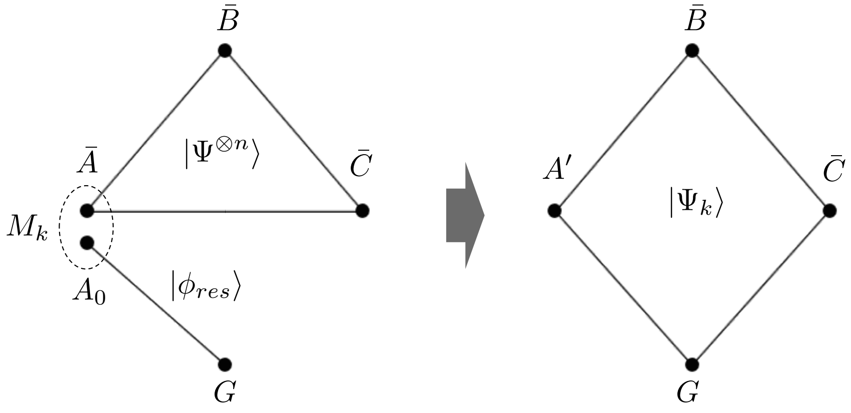


Figure 4.6: Markovianizing by measurements with an auxiliary entangled resource is considered. We want to obtain the state after the measurement that is close to a Markov state on $A'\bar{B}\bar{C}$ conditioned by $\bar{B}$.

2. There exist Markov states $\Upsilon_k^{A'\bar{B}\bar{C}}$ conditioned by $\bar{B}$ such that

$$\sum_k r_k \left\| {\Psi_k}^{A'\bar{B}\bar{C}} - \Upsilon_k^{A'\bar{B}\bar{C}} \right\|_1 \leq \epsilon.$$

3. The dimension of $\mathcal{H}^G$, denoted by d_G, is bounded above as

$$\log d_G \leq n\kappa \log\left(d_A d_B d_C\right) \tag{4.52}$$

with a constant κ. This condition is only for a technical reason.

Define the entropy decrease due to the measurement on AA_0 by $\Delta S(A')_k := nS(A)_\Psi + S(A_0)_{\phi_{res}} - S(A')_{\Psi_k}$ and the average entropy decrease by $\Delta S(A')_{av} := \sum_k r_k \Delta S(A')_k$. Define also $\Delta\hat{S}_k := nS(A)_\Psi - S(A')_{\Psi_k} + S(G)_{\Psi_k} = \Delta S(A')_k - S(G)_{\phi_{res}} + S(G)_{\Psi_k}$ and $\Delta\hat{S}_{av} := \sum_k r_k \Delta\hat{S}_k$. Then we have

$$\begin{aligned} \Delta S(A')_{av} &\geq nM_{A|B}(\Psi^{ABC}) - n\eta_\Psi(\epsilon)\log\left(d_A d_B d_C\right), & (4.53)\\ \Delta\hat{S}_{av} &\geq nM_{A|B}(\Psi^{ABC}) - n\eta_\Psi(\epsilon)\log\left(d_A d_B d_C\right) & (4.54)\end{aligned}$$

and

$$H(\{r_k\}_k) \geq nM_{A|B}(\Psi^{ABC}) - n\eta_\Psi(\epsilon)\log\left(d_A d_B d_C\right). \tag{4.55}$$

Proof. Let

$$\epsilon_k'' := \left\| (\Psi^{\otimes n})^{\bar{B}\bar{C}} - {\Psi_k}^{\bar{B}\bar{C}} \right\|_1, \quad \epsilon_k' := \left\| {\Psi_k}^{A'\bar{B}\bar{C}} - \Upsilon_k^{A'\bar{B}\bar{C}} \right\|_1, \quad \epsilon_k := \epsilon_k' + \epsilon_k''. \tag{4.56}$$

Fix one k for the moment. There exists a state $|\chi\rangle^{A'\bar{B}\bar{C}G}$ such that $\chi^{A'\bar{B}\bar{C}} = \Upsilon_k^{A'\bar{B}\bar{C}}$ and

$$\left\| {\Psi_k}^{A'\bar{B}\bar{C}G} - \chi^{A'\bar{B}\bar{C}G} \right\|_1 \leq \eta(\epsilon_k'). \tag{4.57}$$

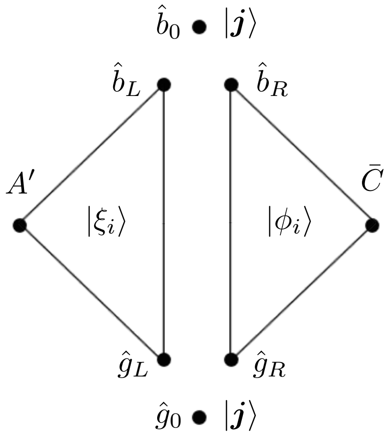


Figure 4.7: A graphical representation of $|\tilde{\chi}\rangle$. The state $|\chi\rangle$, which is a purification of $\Upsilon^{A'\bar{B}\bar{C}}$, is decomposed into $|\tilde{\chi}\rangle$, which is the superposition of the above state with the probability amplitude $\sqrt{q_i}$, namely, $|\tilde{\chi}\rangle = \sum_i \sqrt{q_i}|i\rangle^{\hat{b}_0}|i\rangle^{\hat{g}_0}|\xi_i\rangle^{A'\hat{b}_L\hat{g}_L}|\phi_i\rangle^{\hat{b}_R\hat{g}_R\bar{C}}$.

Let $\Gamma^{\bar{B}}_{\Upsilon} : \mathcal{H}^{\bar{B}} \to \mathcal{H}^{\hat{b}_0} \otimes \mathcal{H}^{\hat{b}_L} \otimes \mathcal{H}^{\hat{b}_R}$ be the Markov decomposition of $\mathcal{H}^{\bar{B}}$ with respect to $\Upsilon_k^{A'\bar{B}\bar{C}}$, and let

$$\Upsilon_{Mk}^{A'\bar{B}\bar{C}} := \Gamma^{\bar{B}}_{\Upsilon}\Upsilon_k^{A'\bar{B}\bar{C}}\Gamma^{\dagger\bar{B}}_{\Upsilon} = \sum_i q_i|i\rangle\langle i|^{\hat{b}_0} \otimes \sigma_i^{A'\hat{b}_L} \otimes \phi_i^{\hat{b}_R\bar{C}} \tag{4.58}$$

be the Markov decomposition of $\Upsilon_k^{A'\bar{B}\bar{C}}$. Since $\Gamma^{\bar{B}}_{\Upsilon}|\chi\rangle$ is a purification of $\Upsilon_{Mk}^{A'\bar{B}\bar{C}}$, there exists a unitary isomorphism $\Gamma^G_\chi : G \to \hat{g}_0\hat{g}_L\hat{g}_R$ such that

$$|\tilde{\chi}\rangle := (\Gamma^{\bar{B}}_{\Upsilon} \otimes \Gamma^G_\chi)|\chi\rangle = \sum_i \sqrt{q_i}|i\rangle^{\hat{b}_0}|i\rangle^{\hat{g}_0}|\xi_i\rangle^{A'\hat{b}_L\hat{g}_L}|\phi_i\rangle^{\hat{b}_R\hat{g}_R\bar{C}}, \tag{4.59}$$

where $|\xi_i\rangle^{A'\hat{b}_L\hat{g}_L}$ and $|\phi_i\rangle^{\hat{b}_R\hat{g}_R\bar{C}}$ are purifications of $\sigma_i^{A'\hat{b}_L}$ and $\phi_i^{\hat{b}_R\bar{C}}$, respectively. Therefore, as we prove in Appendix, we have

$$\Delta I_k := I(\bar{B}\bar{C} : G)_{\Psi_k} \geq nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon_k)\log(d_Ad_Bd_C). \tag{4.60}$$

The entropy decrease is then calculated as

$$\begin{aligned} \Delta S(A')_k &= nS(A)_\Psi + S(A_0)_{\phi_{res}} - S(A')_{\Psi_k} \\ &= S(\bar{B}\bar{C})_{\Psi^{\otimes n}} + S(A_0)_{\phi_{res}} - S(\bar{B}\bar{C}G)_{\Psi_k} \\ &\geq S(\bar{B}\bar{C})_{\Psi_k} + S(A_0)_{\phi_{res}} - S(\bar{B}\bar{C}G)_{\Psi_k} - n\eta(\epsilon_k)\log(d_Bd_C) \\ &= S(A_0)_{\phi_{res}} - S(G|\bar{B}\bar{C})_{\Psi_k} - n\eta(\epsilon_k)\log(d_Bd_C) \\ &= S(G)_{\phi_{res}} - S(G)_{\Psi_k} + I(\bar{B}\bar{C} : G)_{\Psi_k} - n\eta(\epsilon_k)\log(d_Bd_C) \\ &= S(G)_{\phi_{res}} - S(G)_{\Psi_k} + \Delta I_k - n\eta_{\Psi}(\epsilon_k)\log(d_Bd_C), \end{aligned} \tag{4.61}$$

where the second line follows from (4.56). Averaging over k, from the concavity of entropy and η, we obtain

$$\begin{aligned}\Delta S(A')_{ave} &= S(G)_{\phi_{res}} - \sum_k r_k S(G)_{\Psi_k} + \sum_k r_k \left(\Delta I_k - n\eta(\epsilon_k)\log(d)\right) \\ &\geq nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon)\log(d_A d_B d_C). \qquad (4.62)\end{aligned}$$

From (4.61), we also have

$$\Delta\hat{S}_k \geq nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon_k)\log(d_A d_B d_C) \qquad (4.63)$$

and thus

$$\Delta\hat{S}_{av} \geq nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon)\log(d_A d_B d_C). \qquad (4.64)$$

Let $V : \bar{A}A_0 \to A'E_0$ be an isometry such that the Naimark extension of $\{M_k\}_k$ is given by $M_k = \langle k|^{E_0} V$, and let

$$\Psi'^{A'E_0} := \sum_k |k\rangle\langle k|^{E_0} V((\Psi^{\otimes n})^{\bar{A}} \otimes \phi^{A_0})V^{\dagger}|k\rangle\langle k|^{E_0}. \qquad (4.65)$$

We have

$$\begin{aligned}H(\{r_k\}_k) &= S(E_0)_{\Psi'} = S(A'E_0)_{\Psi'} - S(A'|E_0)_{\Psi'} \\ &\geq S(\bar{A}A_0)_{\Psi^{\otimes n}\otimes\phi_{res}} - \sum_k r_k S(A')_{\Psi_k} = \Delta S(A')_{ave} \\ &\geq nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon)\log(d_A d_B d_C), \qquad (4.66)\end{aligned}$$

which concludes the proof. ■

Chapter 5

Distributed Quantum Computation

Distributed quantum computation is a task in which many distant parties perform a large quantum computation in collaboration with each other, by using classical communication, quantum communication or shared entanglement as resources. Among others, we consider implementation of bipartite unitaries on unknown input states by local operation and classical communication (LOCC) assisted by shared entanglement. We consider an asymptotic scenario in which the two parties perform the same bipartite unitary on infinitely many independent input pairs. We investigate to what extent the entanglement cost and the classical communication cost can be reduced by allowing nonzero but vanishing error in the asymptotic limit. The main result is that the optimal costs of entanglement, forward and backward classical communication in protocols consisting of three steps are given by the Markovianizing cost of a state associated with the unitary.

5.1 Previous Results

Suppose Alice and Bob have quantum systems A and B in unknown states $|\varphi\rangle^{AB}$, respectively, and try to apply a unitary U^{AB} by LOCC using some resource entanglement shared in advance. A trivial way is one using quantum teleportation, where Alice teleports her input state to Bob, Bob performs the unitary, and then he teleports back Alice's share of output. In the case of two-qubit unitaries, such a protocol consumes two Bell pairs and two bits of classical communication in both directions. However, in [26], a protocol is proposed for performing two-qubit controlled-unitaries deterministically and exactly by using one Bell pair and one bit of classical communication in both directions. This result indicates that there are cases where we can reduce the cost of resources depending on the unitary.

Protocols for this task are classified in terms of the success probability and the fidelity of the final state to the target state $U^{AB}|\varphi\rangle^{AB}$. A protocol is called

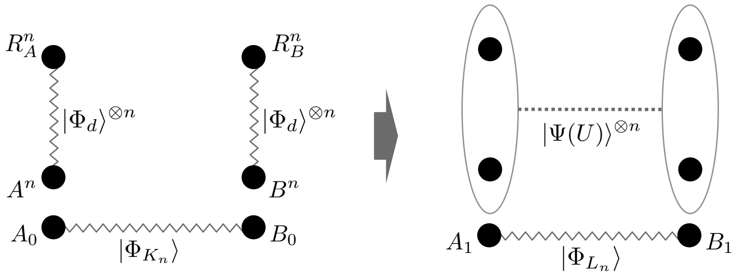

R_A^n
R_B^n
$|\Phi_d\rangle^{\otimes n}$
$|\Phi_d\rangle^{\otimes n}$
A^n
B^n
A_0
B_0
$|\Phi_{K_n}\rangle$
$|\Psi(U)\rangle^{\otimes n}$
A_1
B_1
$|\Phi_{L_n}\rangle$

Figure 5.1: The figure represents the task of applying $(U^{AB})^{\otimes n}$ on $(|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B})^{\otimes n}$ by using resource entanglement represented by $|\Phi_{K_n}^{A_0B_0}\rangle$. Black circles represents quantum systems, and wavy lines represent maximally entangled states. R_A and R_B are reference systems that Alice and Bob cannot access. The entanglement cost is defined as the difference between the amount of initial entanglement and that of final entanglement shared by Alice and Bob.

deterministic if it succeeds in implementing U^{AB} with probability one, otherwise it is called *probabilistic*. A protocol is called *exact* if the fidelity of the final state to the target state is one, otherwise called *approximate*.

A deterministic and exact protocol for two-qubit unitaries is proposed in [26], and those for general bipartite unitaries are studied in [27, 28]. In [29], the minimum entanglement cost of deterministic and exact protocols for two-qubit controlled-unitaries is investigated. It is shown that, in any protocol for that task, the resource state must be maximally entangled when it is a pure entangled state with Schmidt rank 2, despite the fact that the controlled-unitary can be almost equal to the identity operator. The result is generalized for arbitrary bipartite unitaries in [30]. It is also shown numerically in [30] that there exists a class of two-qubit controlled-unitaries which can be implemented exactly and deterministically by using entanglement resource with Schmidt rank 3, but with the entanglement entropy less than 1. There are several works to find more efficient protocols which consume less resources [26, 27, 28, 31, 32, 33, 34, 35, 36]. However, most studies so far have only focused on single-shot protocols in which *one* bipartite unitary is applied to *a pair of* inputs, and no result is known about an asymptotic scenario in which we consider infinitely many input pairs.

5.2 Formulation of the Problem

As an asymptotic version, we consider a task in which Alice and Bob apply $(U^{AB})^{\otimes n}$ on $(|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B})^{\otimes n}$ by LOCC using a resource state $\Phi_{K_n}^{A_0B_0}$ (Fig. 5.1). Here, R_A and R_B are imaginary reference systems that are inaccessible and invisible to Alice and Bob. Φ_d and Φ_{K_n} are maximally entangled states with Schmidt rank $d = \dim\mathcal{H}^A = \dim\mathcal{H}^B$ and K_n, respectively. We evaluate the efficiency of the protocol by the fidelity between the final state of the protocol and the desired state $(U^{AB}|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B})^{\otimes n}$. We do not require the fidelity to be unity for finite n. Instead, we require that the fidelity converge to one in the limit of $n \to \infty$. Our interest is to find the minimal cost of entanglement, forward classical communication and backward classical communication per copy for accomplishing the task. Rigorous definitions are given below.

Definition 26 Consider a unitary $U : \mathcal{H}^A \otimes \mathcal{H}^B \to \mathcal{H}^A \otimes \mathcal{H}^B$ acting on two d-level systems A and B. Let $|\Psi(U)\rangle := U^{AB}|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$. Define Alice and Bob have registers A_0, A_1 and B_0, B_1, respectively. We refer to the following quantum operation $\mathcal{M}_n$ as an EALOCC (entanglement-assisted LOCC) implementation of $U^{\otimes n}$ with the error ϵ_n, the entanglement cost $\log K_n - \log L_n$, the forward classical communication cost $C_n^{\rightarrow}$, and the backward classical communication cost $C_n^{\leftarrow}$. Here, $\mathcal{M}_n : A^n A_0 \otimes B^n B_0 \to A^n A_1 \otimes B^n B_1$ is a LOCC and

$$F(\rho(\mathcal{M}_n), |\Psi(U)\rangle^{\otimes n}|\Phi_{L_n}\rangle^{A_1B_1}) \geq 1 - \epsilon_n \tag{5.1}$$

holds for $\rho(\mathcal{M}_n) = \mathcal{M}_n(|\Phi_d^{AR_A}\rangle^{\otimes n}|\Phi_d^{BR_B}\rangle^{\otimes n}|\Phi_{K_n}\rangle^{A_0B_0})$. $C_n^{\rightarrow}$ and $C_n^{\leftarrow}$ are the total amount of classical communication transmitted from Alice to Bob and Bob to Alice in $\mathcal{M}_n$ measured by bits, respectively.

Definition 27 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is said to be achievable if there exists a sequence of EALOCC implementations of $U^{\otimes n}$ such that $\epsilon_n \to 0$, $\frac{1}{n}(\log K_n - \log L_n) \to R_E$, $\frac{1}{n}C_n^{\rightarrow} \to C^{\rightarrow}$ and $\frac{1}{n}C_n^{\leftarrow} \to C^{\leftarrow}$ in the limit of $n \to \infty$.

Condition (5.1) implies that, for *almost all* input states $|\varphi\rangle \in \mathcal{H}^{A^n} \otimes \mathcal{H}^{B^n}$, the final state $\hat{\mathcal{M}}_n(\varphi^{A^nB^n}) := \mathrm{Tr}_{A_1B_1}[\mathcal{M}_n(\varphi^{A^nB^n} \otimes \Phi_{K_n}^{A_0B_0})]$ is sufficiently close to the desired state $U^{\otimes n}|\varphi\rangle$. Indeed, due to the relation between entanglement fidelity (2.48) and ensemble fidelity (2.49), Condition (5.1) implies

$$\int_{\varphi} p(d\varphi)\, F(\hat{\mathcal{M}}_n(\varphi), U^{\otimes n}|\varphi\rangle) \geq 1 - \epsilon_n, \tag{5.2}$$

where the average is taken with respect to the Haar measure on $\mathcal{H}^{A^n} \otimes \mathcal{H}^{B^n}$. On the other hand, we do not require the protocol to be *universal*, i.e., there may be some input states φ for which the output state $\hat{\mathcal{M}}_n(\varphi^{A^nB^n})$ is not close to the desired state $U^{\otimes n}|\varphi\rangle$.

If $\hat{\mathcal{M}}_n$ implements $(U^{AB})^{\otimes n}$ on $|\Phi_d^{AR_A}\rangle^{\otimes n}|\Phi_d^{BR_B}\rangle^{\otimes n}$ with a small error, it also implements $(U^{AB})^{\otimes n}$ on $U_0^{AB}|\Phi_d^{AR_A}\rangle^{\otimes n}|\Phi_d^{BR_B}\rangle^{\otimes n}$ for any U_0 with a small error. More precisely, we have the following lemma.

Lemma 28 For any U_0, Condition (5.1) is equivalent to

$$F(\rho(\mathcal{M}_n, U_0), |\Psi(UU_0)\rangle^{\otimes n}|\Phi_{L_n}\rangle^{A_1B_1}) \geq 1 - \epsilon_n, \tag{5.3}$$

where $\rho(\mathcal{M}_n, U_0) := \mathcal{M}_n(|\Psi(U_0)\rangle^{AR_ABR_B}|\Phi_{K_n}\rangle^{A_0B_0})$.

Proof. Follows from invariance of fidelity under unitary operations (2.41) and the relation $U_0^{AB}|\Phi_d\rangle^{AR_A}\ |\Phi_d\rangle^{BR_B} = (U_0^T)^{R_AR_B}|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$, where the superscript T denotes transposition with respect to the Schmidt basis of Φ_d. ■

In Section 5.3, 5.4 and 5.5, we consider protocols for implementing $U^{\otimes n}$, consisting of three steps, in which Alice first performs a measurement, communicates the measurement results to Bob, Bob performs a measurement, communicates the results to Alice, and Alice performs a quantum operation. We refer to this class of protocols as *one-round protocols.* We first extend the notion of Markovianizing cost introduced in Chapter 4 to a given bipartite unitary.

Definition 29 Consider a "tripartite" state $|\Psi(U)\rangle^{AR_A(BR_B)} := (U^{AB}\otimes I^{R_AR_B})|\Phi_d\rangle^{AR_A}$ $|\Phi_d\rangle^{BR_B}$ by regarding B and R_B as a single system. The Markovianizing cost of U is defined as $M(U) := M_{A|R_A}(\Psi(U)^{AR_A(BR_B)})$.

The main result is that the optimal rates of costs of classical communication and entanglement in one-round protocols are given by the Markovianizing cost of U.

Theorem 30 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is achievable by one-round protocols in EALOCC implementation of U if and only if $R_E, C^{\rightarrow}, C^{\leftarrow} \geq M(U^\dagger)$.

Proof. Follows from Theorem 43 in Section 5.4 and Theorem 45 in Section 5.5. ■

5.3 Single-shot One-round Protocols

In this section, we consider $n = 1$ (single-shot) case, and analyze a single-shot protocol $\mathcal{M}$ for implementing U by one-round EALOCC. We derive several conditions for the protocol to succeed in high fidelity. The result obtained here is then applied to the asymptotic situation in the next section.

Let $\mathcal{M} : AA_0 \otimes BB_0 \rightarrow AA_1 \otimes BB_1$ be a one-round LOCC protocol for implementing U. When $\mathcal{M}$ succeeds in implementing U in high fidelity, it satisfies

$$F(\rho(\mathcal{M})^{AR_ABR_B}, |\Psi(U)\rangle) \geq 1 - \epsilon, \tag{5.4}$$

where $\rho(\mathcal{M}) := \mathcal{M}(|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}|\phi_{\mathrm{res}}\rangle^{A_0B_0})$ and ϕ_{res} is a pure resource state shared in advance. Lemma 28 indicates that, taking $U^\dagger$ for U_0, Condition (5.4) is equivalent to

$$F(\rho(\mathcal{M}, U^\dagger)^{ABR_AR_B}, |\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}) \geq 1 - \epsilon, \tag{5.5}$$

where $\rho(\mathcal{M}, U^\dagger) := \mathcal{M}(|\Psi(U^\dagger)\rangle^{AR_ABR_B}|\phi_{\mathrm{res}}\rangle^{A_0B_0})$. Thus it is necessary and sufficient that $\mathcal{M}$ transforms $|\Psi(U^\dagger)\rangle$ to $|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$ (Fig.5.2). While $|\Phi_d\rangle|\Phi_d\rangle$ obviously

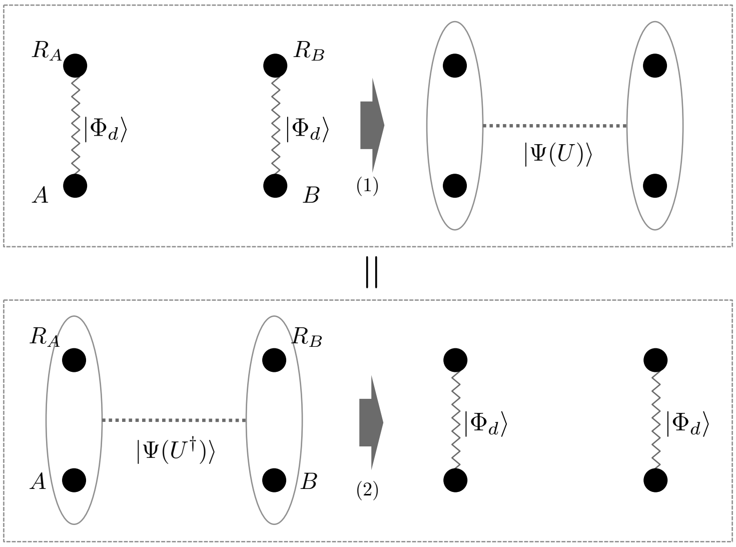


Figure 5.2: The task of performing U^{AB} on $|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$ represented by the upper process (1) is equivalent to transforming $|\Psi(U^\dagger)\rangle$ into $|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$ represented by the lower process (2). This equivalence is due to the fact that R_A and R_B are reference systems that are inaccessible and invisible to Alice and Bob.

has no correlation between AR_A and BR_B, $|\Psi(U^\dagger)\rangle$ has some amount of entanglement depending on $U^\dagger$. Thus, for a given initial state $|\Psi(U^\dagger)\rangle$, Alice and Bob need to decouple AR_A and BR_B while preserving the maximal entanglement between AB and R_AR_B. Observe that both $|\Psi(U^\dagger)\rangle$ and $|\Phi_d\rangle|\Phi_d\rangle$ are d^2-dimensional maximally entangled state between AB and R_AR_B.

Without loss of generality, we assume that the protocol $\mathcal{M}$ proceeds as follows.

I-1. Alice performs a measurement $\{M_k^{AA_0\to A'}\}_k$. The probability of obtaining measurement result k is given by $p_k = \|M_k|\Psi(U^\dagger)\rangle|\phi_{\mathrm{res}}\rangle\|_1^2$, and the state after the measurement is $|\Psi_k(U^\dagger)\rangle = p_k^{-1/2}M_k|\Psi(U^\dagger)\rangle|\phi_{\mathrm{res}}\rangle$.

I-2. Alice communicates the measurement result k to Bob.

I-3. Bob performs a measurement $\{N_{l|k}^{BB_0\to BB_1}\}_l$. The probability of obtaining measurement result l, conditioned by k, is given by $p_{l|k} = \|N_{l|k}|\Psi_k(U^\dagger)\rangle\|_1^2$ and the state after Bob's measurement is $|\Psi_{kl}(U^\dagger)\rangle = p_{l|k}^{-1/2}N_{l|k}|\Psi_k(U^\dagger)\rangle$.

I-4. Bob communicates the measurement result l to Alice.

I-5. Alice performs an operation which is described by a CPTP map $\mathcal{O}_{kl} : A' \to AA_1$. The final state is given by $\Psi_{kl}^{\mathrm{fin}}(U^\dagger) = \mathcal{O}_{kl}(\Psi_{kl}(U^\dagger))$.

In total, the final state is given by

$$\rho(\mathcal{M}, U^\dagger)^{ABR_AR_B} = \sum_{kl} p_{kl}(\Psi_{kl}^{\mathrm{fin}}(U^\dagger)). \tag{5.6}$$

Thus Condition (5.5) implies

$$\sum_{kl} p_{kl} F((\Psi^{\rm fin}_{kl}(U^\dagger)), |\Phi_d\rangle^{AR_A} |\Phi_d\rangle^{BR_B}) \geq 1 - \epsilon, \tag{5.7}$$

and consequently,

$$\sum_{kl} p_{kl} \left\| (\Psi^{\rm fin}_{kl}(U^\dagger)) - {\Phi_d}^{AR_A} \otimes {\Phi_d}^{BR_B} \right\|_1 \leq \epsilon. \tag{5.8}$$

The main goal of this section is to derive conditions on operations that comprise $\mathcal{M}$, for the protocol to succeed in high fidelity. As we show below, state transformations by Alice's first measurement play a central role for the success of the protocol. Thus we clarify conditions on Alice's measurement. In particular, we reveal that it is necessary that Alice's measurement Markovianizes the "tripartite" state $|\Psi(U^\dagger)\rangle^{AR_A(BR_B)}$. We also show that Markovianization by Alice's measurement is sufficient for the first half of the protocol to succeed in high fidelity.

Let us first discuss general conditions regarding state transformations caused by Alice's measurement. Let $\mathbb{M} = \{M_k^{AA_0 \to A'}\}_k$ be an arbitrary measurement on AA_0. For a fixed $\phi_{\rm res}$, each M_k induces the following state transformation.

$$\mathcal{E}_k(\tau^A) := p_k^{-1} M_k(\tau^A \otimes \phi_{res}^{A_0}) M_k^\dagger, \quad p_k := {\rm Tr}[M_k(\tau^A \otimes \phi_{res}^{A_0}) M_k^\dagger]. \tag{5.9}$$

Consequently, we have

$$\Psi_k(U^\dagger)^{A'R_ABR_B} = \mathcal{E}_k(\Psi(U^\dagger)). \tag{5.10}$$

For any linear operator $M : \mathcal{H}^A \otimes \mathcal{H}^{A_0} \to \mathcal{H}^{A'}$, we define a map

$$\mathcal{E}_M(\tau^A) := p_M^{-1} M(\tau^A \otimes \phi_{res}^{A_0}) M^\dagger, \quad p_M := {\rm Tr}[M(\tau^A \otimes \phi_{res}^{A_0}) M^\dagger]. \tag{5.11}$$

We call $\mathcal{E}_M$ as an *M-induced map.*

Definition 31 We say that an M-induced map is ϵ-linear if it satisfies

$$\left\| \Phi_M^{R_A} - \frac{1}{d} I^{R_A} \right\|_1 \leq \epsilon, \tag{5.12}$$

where $\Phi_M^{A'R_A} := \mathcal{E}_M^A(\Phi_d^{AR_A})$. We say that a measurement $\mathbb{M}$ is ϵ-linear on average if an M_k-induced map is ϵ_k-linear for each k and $\sum_k p_k \epsilon_k \leq \epsilon$.

Let us denote $\mathcal{E}_M(\Psi(U^\dagger))$ as $\Psi_M(U^\dagger)$. The following fact (Lemma 32) is used in proofs of most of the lemmas in this section.

Lemma 32 An M-induced map is ϵ-linear if and only if

$$\left\| \Psi_M(U^\dagger)^{R_ABR_B} - \Psi(U^\dagger)^{R_ABR_B} \right\|_1 \leq \epsilon. \tag{5.13}$$

Consequently, a measurement $\mathbb{M}$ is ϵ-linear on average if and only if

$$\sum_k p_k \left\| \Psi_k(U^\dagger)^{R_A B R_B} - \Psi(U^\dagger)^{R_A B R_B} \right\|_1 \leq \epsilon. \tag{5.14}$$

Proof. Follows from $U^{\dagger AB}|\Phi\rangle^{AR_A}|\Phi\rangle^{BR_B} = U^{tR_A R_B}|\Phi\rangle^{AR_A}|\Phi\rangle^{BR_B}$ and the invariance of trace distance under unitary operations. ■

Definition 33 We say that an M-induced map is ϵ-decoupling between AR_A and R_B if it satisfies

$$\left\| \Psi_M(U^\dagger)^{A'R_A R_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B} \right\|_1 \leq \epsilon, \tag{5.15}$$

We say that a measurement $\mathbb{M}$ is ϵ-decoupling between AR_A and R_B on average if an M_k-induced map is ϵ_k-decoupling between AR_A and R_B for each k and $\sum_k p_k \epsilon_k \leq \epsilon$.

Definition 34 We say that an M-induced map is ϵ-Markovianizing conditioned by R_A if there exists a Markov state $\Upsilon_M^{A'R_A(BR_B)}$ conditioned by R_A such that

$$\left\| \Psi_M(U^\dagger)^{A'R_A B R_B} - \Upsilon_M^{A'R_A B R_B} \right\|_1 \leq \epsilon. \tag{5.16}$$

We say that a measurement $\mathbb{M}$ is ϵ-Markovianizing conditioned by R_A on average if the map induced by M_k is ϵ_k-Markovianizing conditioned by R_A for each k and $\sum_k p_k \epsilon_k \leq \epsilon$.

Although two notions of ϵ-decoupling and ϵ-Markovianizing are introduced in two different contexts, they are closely related as shown in the following lemmas.

Lemma 35 An M-induced map is $\eta(\epsilon)$-Markovianizing conditioned by R_A if it is ϵ-linear and ϵ-decoupling between AR_A and R_B. Consequently, a measurement $\mathbb{M}$ is $\eta(\epsilon)$-Markovianizing conditioned by R_A on average if it is ϵ-linear and ϵ-decoupling between AR_A and R_B on average.

Proof. We describe a simplified version of the proof where we assume $\epsilon = 0$, to show why decoupling and Markovianizing are equivalent under the condition of linearity, as stated in this Lemma and Lemma 36. The details of the proof for $\epsilon > 0$ is given in Appendix. Let $\mathcal{E}$ be the map induced by M which is assumed to be an *exactly* linear map. Let $V : A \to A'E$ be an isometry such that the Naimark extension of $\mathcal{E}$ is given by $\mathcal{E}(\tau) = \mathrm{Tr}_E[V\tau V^\dagger]$, and let $|\Psi_V(U^\dagger)\rangle^{EA'R_A B B R_B} := V|\Psi(U^\dagger)\rangle$. For this state, we have

$$\begin{aligned} I(A' : BR_B|R_A) &= S(A'R_A) + S(R_A B R_B) - S(R_A) - S(A'R_A B R_B) \\ &= S(A'R_A) + S(EA') - S(R_A) - S(E) \\ &= S(A'R_A) + \log d - \log d - S(E) \\ &= S(A'R_A) - S(E). \end{aligned}$$

We also have

$$
\begin{aligned}
I(A'R_A : R_B) &= S(A'R_A) + S(R_B) - S(A'R_A R_B) \\
&= S(A'R_A) + S(R_B) - S(BE) \\
&= S(A'R_A) + S(R_B) - S(B) - S(E) \\
&= S(A'R_A) - S(E).
\end{aligned}
$$

Thus we have $I(A' : BR_B|R_A) = I(A'R_A : R_B)$, which implies that *exact* Markovianizing conditioned by R_A is equivalent to *exact* decoupling between $A'R_A$ and R_B. For rigorous proofs, we need to relax the "exact" condition ($\epsilon = 0$) to the "approximate" condition ($\epsilon > 0$). ■

Lemma 36 An M-induced map is $\eta(\epsilon)$-decoupling between AR_A and R_B if it is ϵ-linear and ϵ-Markovianizing conditioned by R_A. Consequently, a measurement $\mathbb{M}$ is $\eta(\epsilon)$-Markovianizing conditioned by R_A on average if it is ϵ-linear and ϵ-decoupling between AR_A and R_B on average.

Proof. See Appendix.

Let us now analyze conditions on Alice's measurement imposed by (5.5). Let $\mathbb{M} = \{M_k^{AA_0 \to A'}\}_k$ be Alice's measurement in protocol $\mathcal{M}$ that satisfies (5.5). Let A_E and B_E be ancillary systems on Alice and Bob, respectively. Let $W_k : BB_0 \to BB_1B_E$ be isometries such that the Naimark extension of Bob's measurement is given by $N_{l|k} = \langle l|^{B_E} W_k$, and let $V_{kl} : A' \to AA_1A_E$ be an isometry such that the Stinespring dilation of $\mathcal{O}_{kl}$ is given by $\mathcal{O}_{kl}(\tau) = \mathrm{Tr}_{A_E}[V_{kl}\tau^{A'}V_{kl}^\dagger]$. Consider the following protocol, which is equivalent to the procol described in the beginning of this section as a whole.

II-1. Alice performs a measurement $\mathbb{M}$ and obtains the measurement result k. The state after the measurement is $|\Psi_k(U^\dagger)\rangle^{A'R_ABR_BB_0}$.

II-2. Alice communicates k to Bob.

II-3. Bob performs W_k. The state becomes $\big|\Psi'_k(U^\dagger)\big\rangle^{A'R_ABR_BB_1B_E} := W_k\big|\Psi_k(U^\dagger)\big\rangle$.

II-4. Bob performs a projective measurement on B_E in the basis $\{|l\rangle\}_l$, and obtains the result l with probability $p_{l|k}$. The state after the measurement is

$$\big|\Psi_{kl}(U^\dagger)\big\rangle^{A'R_ABR_BB_1} := p_{l|k}^{-1/2}\langle l|^{B_E}\big|\Psi'_k(U^\dagger)\big\rangle.$$

II-5. Bob communicates l to Alice.

II-6. Alice performs V_{kl}. The state becomes $\big|\Psi_{kl}^{\mathrm{fin}}(U^\dagger)\big\rangle^{AA_1A_ER_ABR_BB_1} := V_{kl}\big|\Psi_{kl}(U^\dagger)\big\rangle$.

$\mathbb{I}$-7. Alice discards A_E.

The following lemma states that Alice's measurement $\mathbb{M}$ must be almost linear.

Lemma 37 The measurement $\mathbb{M}$ is $\eta(\epsilon)$-linear on average.

Proof. See Appendix.

Since the final state is close to $|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$, correlation between AR_A and R_B is destroyed by $\mathcal{M}$. The following lemma states that this part of decoupling must be accomplished by Alice's measurement alone.

Lemma 38 The measurement $\mathbb{M}$ is $\eta(\epsilon)$-decoupling between AR_A and R_B on average.

Proof. See Appendix.

Finally, we obtain the *necessary* condition that Alice's measurement must Markovianize $|\Psi(U^\dagger)\rangle$.

Lemma 39 The measurement $\mathbb{M}$ is $\eta(\epsilon)$-Markovianizing conditioned by R_A on average.

Proof. Follows from Lemma 35, 37 and Lemma 38. ■

Conversely, the following two lemmas state that Markovianization by Alice's measurement is also a sufficient condition for the success of the first half of $\mathcal{M}$, in which $\Phi_d^{BR_B}$ is obtained from $|\Psi(U^\dagger)\rangle$.

Lemma 40 If an M-induced map is ϵ-linear and ϵ-decoupling between AR_A and R_B, there exists an isometry $W_M : BB_0 \to BB_1$ and a pure state $|\Psi_M^p\rangle^{A'R_AB_1}$ such that

$$\left\|\Psi_M'(U^\dagger)^{A'R_ABB_1R_B} - \Psi_M^p(U^\dagger)^{A'R_AB_1} \otimes \Phi_d^{BR_B}\right\|_1 \leq \eta(\epsilon), \tag{5.17}$$

where $|\Psi_M'(U^\dagger)\rangle := W_M|\Psi_M(U^\dagger)\rangle$. Consequently, if a measurement $\mathbb{M}$ is ϵ-linear and ϵ-decoupling between AR_A and R_B on average, there exist isometries $\{W_k^{BB_0\to BB_1}\}_k$ and pure states $\{|\Psi_k^p\rangle^{A'R_AB_1}\}_k$ such that

$$\sum_k p_k \left\|\Psi_k'(U^\dagger)^{A'R_ABB_1R_B} - \Psi_k^p(U^\dagger)^{A'R_AB_1} \otimes \Phi_d^{BR_B}\right\|_1 \leq \eta(\epsilon), \tag{5.18}$$

where $|\Psi_k'(U^\dagger)\rangle := W_k|\Psi_k(U^\dagger)\rangle$.

Proof. Follows from Uhlmann's theorem (2.46). ■

Lemma 41 If a measurement $\mathbb{M}$ is ϵ-linear and ϵ-Markovianizing conditioned by R_A on average, there exist isometries $\{W_k^{BB_0\to BB_1}\}_k$ and pure states $\{|\Psi_k^p\rangle^{A'R_AB_1}\}_k$ such that (5.18) is satisfied.

Proof. Follows from Lemma 35 and Lemma 40. ■

The following lemma states that, after obtaining $\Phi_d^{BR_B}$, the remaining task is to obtain $\Phi_d^{AR_A}$ from a pure state shared among Alice, Bob and R_A, which implies that quantum state merging, introduced in Section 3.2, is applicable for this task.

Lemma 42 There exist pure states $\{|\Psi_k^p\rangle^{A'R_AB_1B_E}\}_k$ such that

$$\sum_k p_k \left\|\Psi'_k(U^{\dagger\otimes n}) - (\Psi_k^p)^{A'R_AB_1B_E} \otimes \Phi_d^{BR_B}\right\|_1 \leq \eta(\epsilon). \tag{5.19}$$

Proof. By tracing out AR_A in (5.8), we obtain

$$\begin{aligned}\eta(\epsilon) &\geq \sum_{kl} p_{kl} \left\|\Psi_{kl}^{\rm fin}(U^\dagger)^{BR_B} - \Phi_d^{BR_B}\right\|_1 = \sum_{kl} p_{kl} \left\|\Psi_{kl}(U^\dagger)^{BR_B} - \Phi_d^{BR_B}\right\|_1 \\ &\geq \sum_k p_k \left\|\sum_l p_{l|k}\Psi_{kl}(U^\dagger)^{BR_B} - \Phi_d^{BR_B}\right\|_1 = \sum_k p_k \left\|\Psi'_k(U^\dagger)^{BR_B} - \Phi_d^{BR_B}\right\|_1.\end{aligned}$$

Thus, due to (2.42) and (2.45), there exist $\{|\Psi_k^p\rangle^{A'R_AB_1B_E}\}_k$ satisfying (5.19). ■

Results obtained here are applied to an asymptotic scenario in the next two sections. In particular, Lemma 41 is important to show the lower bound on the three kinds of costs, and Lemma 39 and 42 are important to show the upper bound.

5.4 Achievable Rate with One-round Protocols

Let us return to the asymptotic scenario. We prove the forward implication ("achievability") of Theorem 30. Following Lemma 28, we consider a task of transforming a state $|\Psi(U^\dagger)^{\otimes n}\rangle^{\bar{A}\bar{B}\bar{R}_A\bar{R}_B}|\Phi_{K_n}\rangle^{A_0B_0}$ into $|\Phi_d^{\otimes n}\rangle^{\bar{A}\bar{R}_A}\,|\Phi_d^{\otimes n}\rangle^{\bar{B}\bar{R}_B}$. We consider a case where $L_n = 1$, that is, no entanglement is left after the protocol. Conditions obtained in Section 5.3 directly apply by the following correspondence.

$$\begin{aligned} A, B, R_A, R_B &\to \bar{A}, \bar{B}, \bar{R}_A, \bar{R}_B, \\ U &\to U^{\otimes n} \\ \Phi_d &\to \Phi_d^{\otimes n} \\ \phi_{\rm res} &\to \Phi_{K_n} \\ \Psi(U^\dagger) &\to \Psi(U^\dagger)^{\otimes n} \\ \frac{1}{d}I &\to \frac{1}{d^n}I^{\otimes n} \end{aligned} \tag{5.20}$$

Theorem 43 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is achievable by one-round protocols in EA-LOCC implementation of U if $R_E, C^{\rightarrow}, C^{\leftarrow} \geq M(U^{\dagger})$.

Proof. The proof is by construction. Take arbitrary $R > M(U^{\dagger})$, $\epsilon > 0$, $\delta > 0$, choose sufficiently large n and let $K_n = 2^{n(R+\delta)}$. Divide the resource state $\Phi_{K_n}^{A_0B_0}$ as $\Phi_{2^{nR}}^{A_0B_0} \otimes \Phi_{2^{n\delta}}^{A'_0B'_0}$. We consider a protocol consisting of the following steps.

1. **Alice's measurement**: By definition, there exists a random unitary operation $\mathcal{T}_n : \tau \mapsto 2^{-nR}\sum_{j=1}^{2^{nR}} V_j\tau V_j^{\dagger}$ on $\bar{A}$ and a Markov state $\Upsilon^{\bar{A}\bar{R}_A(\bar{B}\bar{R}_B)}$ such that
$$\left\|\mathcal{T}_n(\Psi(U^{\dagger})^{\otimes n}) - \Upsilon^{\bar{A}\bar{R}_A(\bar{B}\bar{R}_B)}\right\| \leq \epsilon. \tag{5.21}$$
Using V_j in $\mathcal{T}_n$, construct Alice's measurement $\mathbb{M} = \{M_k^{\bar{A}A_0\rightarrow\bar{A}}\}_k$ as
$$M_k^{\bar{A}A_0\rightarrow\bar{A}} = \frac{1}{\sqrt{2^{nR}}}\sum_{j=1}^{2^{nR}} \exp\left(i\frac{2\pi jk}{2^{nR}}\right)\langle j|^{A_0} \otimes V_j^{\bar{A}} \qquad (k = 1, \cdots, 2^{nR}). \tag{5.22}$$
$\mathbb{M}$ is 0-CPTP and ϵ-Markovianizing conditioned by $\bar{R}_A$. Indeed, we have $p_k = 2^{-nR}$ and
$$p_k^{-1}M_k\left(\tau^{\bar{A}} \otimes \Phi_{2^{nR}}^{A_0}\right)M_k^{\dagger} = \frac{1}{2^{nR}}\sum_{j=1}^{2^{nR}} V_j\tau V_j^{\dagger} \qquad (k = 1, \cdots, 2^{nR}) \tag{5.23}$$
for $\tau \in \mathcal{H}^{\bar{A}}$. Alice performs the measurement defined as above.

2. **Forward classical communication**: Alice sends the measurement result k to Bob. This requires nR bits of forward classical communication.

3. **Bob's isometry**: Due to Lemma 41, there exist isometries $\{W_k^{BB_0\rightarrow BB_1}\}_k$ and pure states $\{|\Psi_k^p\rangle^{A'R_AB_1}\}_k$ such that
$$\sum_k p_k \left\|\Psi'_k(U^{\dagger})^{A'R_ABB_1R_B} - \Psi_k^p(U^{\dagger})^{A'R_AB_1} \otimes \Phi_d^{BR_B}\right\| \leq \eta(\epsilon), \tag{5.24}$$
where $|\Psi'_k(U^{\dagger})\rangle := W_k|\Psi_k(U^{\dagger})\rangle$. Bob performs W_k.

4. **State merging**: Alice and Bob perform state merging of $|\tilde{\Psi}_n^p(U^{\dagger})\rangle^{\tilde{A}\bar{R}_A\tilde{B}} := |\Psi_n^p(U^{\dagger})\rangle^{\bar{A}\bar{R}_AB_1}\,|\Phi_{2^{n\delta}}\rangle^{A'_0B'_0}$, where $\tilde{A} = \bar{A}A'_0$ and $\tilde{B} = B_1B'_0$. From Theorem 12, by choosing $L = 1$, there exists state merging of $\tilde{\Psi}_n^p(U^{\dagger})$ with the error $\eta(\epsilon')$ where $\epsilon' = 2^{-n\delta/2} + 2^{-\log K_n - n\delta}$, the entanglement cost $-\log L = 0$ and the classical communication cost $C = \log K_n + n\delta = n(R+\delta)$. By performing an isometry after state merging, Alice can obtain $|\Phi_d^{\otimes n}\rangle^{\bar{A}\bar{R}_A}$.

In total, we have $E_n = n(R+\delta)$, $C_n^{\rightarrow} = nR$ and $C_n^{\leftarrow} = n(R+\delta)$. Thus $(R_E, C^{\rightarrow}, C^{\leftarrow}) = (R, R, R)$ is achievable. ■

5.5 Optimal Rate with One-round Protocols

We prove the backward implication ("optimality") of Theorem 30, which states that we cannot reduce the three kinds of costs below $M(U^\dagger)$. Let $\mathcal{M}_n$ be an EALOCC implementation of $U^{\otimes n}$ with the error ϵ, the entanglement cost $E_n = \log K_n - \log L_n$, the forward classical communication cost $C_n^{\rightarrow}$, and the backward classical communication cost $C_n^{\leftarrow}$. By definition, $\mathcal{M}_n$ satisfies Condition (5.1), which is equivalent to

$$F(\rho(\mathcal{M}_n, U^\dagger), |\Phi_d^{\otimes n}\rangle^{\bar{A}\bar{R}_A}|\Phi_d^{\otimes n}\rangle^{\bar{B}\bar{R}_B}|\Phi_{L_n}\rangle^{A_1B_1}) \geq 1-\epsilon \tag{5.25}$$

for $\rho(\mathcal{M}_n, U^\dagger) := \mathcal{M}_n(|\Psi(U^\dagger)^{\otimes n}\rangle^{\bar{A}\bar{R}_A\bar{B}\bar{R}_B}|\Phi_{K_n}\rangle^{A_0B_0})$. In particular, we have

$$F(\rho(\mathcal{M}_n, U^\dagger)^{\bar{A}\bar{R}_A\bar{B}\bar{R}_B}, |\Phi_d^{\otimes n}\rangle^{\bar{A}\bar{R}_A}|\Phi_d^{\otimes n}\rangle^{\bar{B}\bar{R}_B}) \geq 1-\epsilon.$$

Thus, from Lemma 38 and Lemma 39, the map induced by Alice's measurement in $\mathcal{M}_n$ is $\eta(\epsilon)$-CPTP, $\eta(\epsilon)$-decoupling between $\bar{A}\bar{R}_A$ and $\bar{R}_B$, and $\eta(\epsilon)$-Markovianizing conditioned by $\bar{R}_A$ on average. Lemma 32 implies that Alice's measurement does not change the reduced state on $\bar{R}_A\bar{B}\bar{R}_B$, up to an average error $\eta(\epsilon)$. We assume here for simplicity that K_n is bounded above as

$$\log K_n \leq n\kappa \log d \tag{5.26}$$

with a constant $\kappa > 0$. Then all the three conditions in Theorem 25 are satisfied.

Lemma 44

1. $C_n^{\rightarrow} \geq nM(U^\dagger) - n\eta_{U}(\epsilon)\log d$.
2. $n\log d + \log K_n - \sum_k p_k S(A')_{\Psi_k(U^\dagger)} \geq nM(U^\dagger) - n\eta_{U}(\epsilon)\log d$.
3. $\log K_n - \log L_n \geq nM(U^\dagger) - n\eta_{U}(\epsilon)\log d$.
4. $C_n^{\leftarrow} \geq nM(U^\dagger) - n\eta_{U}(\epsilon)\log d$.

Here, $\eta_{U}(\epsilon)$ is a proper function of $\epsilon > 0$ and U satisfying $\lim_{\epsilon\to 0}\eta_{U}(\epsilon) = 0$, which is continuous with respect to ϵ, and does not depend on n.

Proof. See Appendix.

Theorem 45 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is achievable by one-round protocols in EALOCC implementation of U only if $R_E, C^{\rightarrow}, C^{\leftarrow} \geq M(U^\dagger)$.

Proof. Follows from 3, 1 and 4 in Lemma 44, respectively. ■

Chapter 6

Computation of Markovianizing Cost of Unitaries

We discuss how to compute the Markovianizing cost of unitaries. It is in principle possible to compute $M(U)$ by finding the KI decomposition of $\Psi(U^\dagger)$. However, the algorithm for obtaining the KI decomposition proposed in [18] involves repeated application of decompositions of the Hilbert space into subspaces, and is difficult to execute in general. In this chapter, we introduce two methods to compute $M(U)$ without explicitly finding the KI decomposition. The first one is based on irreducibility of the KI decomposition, and is applicable for arbitrary bipartite unitaries . The second one is applicable for a class of unitaries called *generalized Clifford operators*, and is based on the commutation relation between generalized Clifford operators and generalized Pauli operators.

6.1 A Method Based on KI Decomposition

Similarly to the Schmidt decomposition of bipartite pure states, any bipartite unitary acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ can be decomposed as

$$U^{AB} = \sum_{s=1}^{S} c_s E_s^A \otimes F_s^B, \tag{6.1}$$

where $c_s > 0, E_s \in \mathcal{L}(\mathcal{H}^A), F_s \in \mathcal{L}(\mathcal{H}^B), \sum_{s=1}^S c_s^2 = 1$ and $d^{-1}\mathrm{Tr}[E_s^\dagger E_{s'}] = d^{-1}\mathrm{Tr}[F_s^\dagger F_{s'}] = \delta_{ss'}$. It is called the operator Schmidt decomposition and S is called the operator Schmidt number. The operator Schmidt number characterizes a nonlocal property of U, similarly to the Schmidt number for pure bipartite states. Another quantity to characterize a nonlocal property of unitaries is the Schmidt strength[37, 38], defined as

$$K(U) := -\sum_s c_s^2 \log c_s^2. \tag{6.2}$$

Let $|\phi_s\rangle := E_s^A|\Phi_d\rangle^{AR_A}$ and $|\psi_s\rangle := F_s^B|\Phi_d\rangle^{BR_B}$, where Φ_d is a maximally entangled state with Schmidt rank d. Then we have $\langle\psi_s|\psi_{s'}\rangle = \langle\phi_s|\phi_{s'}\rangle = \delta_{ss'}$ and $|\Psi(U)\rangle = \sum_{s=1}^S c_s|\psi_s\rangle^{AR_A}|\phi_s\rangle^{BR_B}$. Thus the entanglement entropy of $|\Psi(U)\rangle$ between AR_A and BR_B is equal to $K(U)$ [39], that is,

$$S(AR_A)_{\Psi(U)} = K(U). \tag{6.3}$$

Since U is a unitary, we have

$$I^A \otimes I^B = U^\dagger U = \sum_{s,s'} c_s^* c_{s'} (E_s^\dagger E_{s'})^A \otimes (F_s^\dagger F_{s'})^B \tag{6.4}$$

By tracing out B, we have

$$I = \sum_{s,t} \left(c_s^* c_{s'} \times d^{-1}\mathrm{Tr}[F_s^\dagger F_{s'}]\right) E_s^\dagger E_{s'} = \sum_s |c_s|^2 E_s^\dagger E_s. \tag{6.5}$$

In the same way, we also have

$$I = \sum_s |c_s|^2 E_s E_s^\dagger. \tag{6.6}$$

In the following part of this section, we investigate relations between the Markovianizing cost of U and the set of linear operators $\{E_s\}_s$. Consider two unitaries U and $\hat{U}$ such that their operator Schmidt decompositions are given by

$$U^{AB} = \sum_{s=1}^S c_s E_s^A \otimes F_s^B \tag{6.7}$$

and

$$\hat{U}^{AB} = \sum_{s=1}^S \hat{c}_s E_s^A \otimes F_s^B, \tag{6.8}$$

respectively. The following two Lemmas imply that the Markovianizing costs of U and $\hat{U}$ are the same.

Lemma 46 The KI decomposition of A with respect to $\Psi(U)^{A(BR_B)}$ is equivalent to one with respect to $\Psi(\hat{U})^{A(BR_B)}$.

Proof. By Definition 3, the KI decomposition of $\mathcal{H}^A$ with respect to $\Psi(U)^{A(BR_B)}$ is defined as the KI decomposition of a set of states given by

$$\begin{aligned}&\mathfrak{S}_{\Psi(U)^{BR_B\to A}} :=\\ &\quad\{\varphi \in \mathcal{S}(\mathcal{H}^A) \mid \exists M \in \mathcal{P}(\mathcal{H}^B \otimes \mathcal{H}^{R_B}) \text{ s.t. } \varphi = \mathrm{Tr}_{BR_B}[M^{BR_B}\Psi(U)^{A(BR_B)}]\}.\end{aligned}$$

Due to (6.1), $\Psi(U)^{A(BR_B)}$ is decomposed as

$$\Psi(U)^{A(BR_B)} = \frac{1}{d}\sum_{s,s'=1}^{S} c_s c_{s'}^* E_s E_{s'}^{\dagger A} \otimes |\psi_s\rangle\langle\psi_{s'}|^{BR_B}, \tag{6.9}$$

where $|\psi_s\rangle := F_s^B|\Phi_d\rangle^{BR_B}$. Thus we have, for any $M \in \mathcal{P}(\mathcal{H}^B \otimes \mathcal{H}^{R_B})$,

$$\mathrm{Tr}_{BR_B}[M^{BR_B}\Psi(U)^{A(BR_B)}] = \sum_{s,s'=1}^{S} \alpha_{ss'} c_s c_{s'}^* E_s E_{s'}^{\dagger A}, \tag{6.10}$$

where $\alpha_{ss'} := \langle\psi_{s'}|M|\psi_s\rangle = \alpha_{s's}^*$. Define

$$\hat{M}^{BR_B} := \sum_{s,s'=1}^{S} \frac{\alpha_{ss'} c_s c_{s'}^*}{\hat{c}_s \hat{c}_{s'}^*} |\psi_{s'}\rangle\langle\psi_s|^{BR_B}. \tag{6.11}$$

Then we have $\hat{M} \in \mathcal{P}(\mathcal{H}^B \otimes \mathcal{H}^{R_B})$ and

$$\mathrm{Tr}_{BR_B}[\hat{M}^{BR_B}\Psi(\hat{U})^{A(BR_B)}] = \sum_{s,s'=1}^{S} \alpha_{ss'} c_s c_{s'}^* E_s E_{s'}^{\dagger A}, \tag{6.12}$$

which implies

$$\mathfrak{S}_{\Psi(U)^{BR_B \to A}} = \mathfrak{S}_{\Psi(\hat{U})^{BR_B \to A}}. \tag{6.13}$$

Thus we obtain the proof. ■

Lemma 47 $M(U) = M(\hat{U})$.

Proof. Let Γ be the KI decomposition of A with respect to $\Psi(U)^{A(BR_B)}$, which is equivalent to the one with respect to $\Psi(\hat{U})^{A(BR_B)}$. Since $\Psi(U)^A = \Psi(\hat{U})^A = \frac{1}{d}I^A$, we have

$$\Gamma\Psi(U)^A\Gamma^\dagger = \Gamma\Psi(\hat{U})^A\Gamma^\dagger. \tag{6.14}$$

Hence, from Theorem 21 and Definition 29, we obtain the proof. ■

Two unitaries U and U' are called *local unitarily equivalent* if there exist unitary operators P, Q on $\mathcal{H}^A$ and R, S on $\mathcal{H}^B$ such that

$$(P^A \otimes R^B)U(Q^A \otimes S^B) = U'. \tag{6.15}$$

From (6.10), it is immediate to verify that $M(U) = M(U')$.

Let us continue to discuss relations between $M(U)$ and $\{E_s\}_s$. Define a CPTP map $\mathcal{E}$ on A by

$$\mathcal{E}(\tau) = \sum_s |c_s|^2 E_s \tau E_s^\dagger. \tag{6.16}$$

This is a CPTP map because

$$\mathrm{Tr}[\mathcal{E}(\tau)] = \sum_s |c_s|^2 \mathrm{Tr}[E_s \tau E_s^\dagger] = \sum_s |c_s|^2 \mathrm{Tr}[E_s^\dagger E_s \tau] = \mathrm{Tr}[\tau]. \tag{6.17}$$

holds. The adjoint map $\mathcal{E}^*$ of $\mathcal{E}$ is defined by

$$\mathcal{E}^*(\tau) = \sum_s |c_s|^2 E_s^\dagger \tau E_s, \tag{6.18}$$

which is also a CPTP map due to (6.6). Denote (s, s') by t, $E_s E_{s'}^\dagger$ by $\tilde{E}_t$ and $c_s c_{s'}^*$ by $\tilde{c}_t$. Define a CPTP map $\tilde{\mathcal{E}}$ by

$$\tilde{\mathcal{E}}(\tau) = (\mathcal{E} \circ \mathcal{E}^*)(\tau) = \sum_t |\tilde{c}_t|^2 \tilde{E}_t \tau \tilde{E}_t^\dagger, \tag{6.19}$$

which is equal to its adjoint map $\tilde{\mathcal{E}}^*$. We introduce another CPTP map in the following Lemma, which plays a central role in computing $M(U)$.

Lemma 48 Define a map $\mathcal{E}_\infty$ on $\mathcal{S}(\mathcal{H}^A)$ by

$$\mathcal{E}_\infty(\tau) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \tilde{\mathcal{E}}^n(\tau). \tag{6.20}$$

The limit exists and is a CPTP map. Moreover, $\mathcal{E}_\infty(X) = X$ if and only if

$$[X, \tilde{E}_t] = [X, \tilde{E}_t^\dagger] = 0 \tag{6.21}$$

for all t.

Proof. See Lemma 12 in Ref.[21]. ■

The following Theorem connects $M(U)$ and $\mathcal{E}_\infty$ with a simple formula.

Theorem 49 $M(U) = S(\Phi_\infty)$ where $\Phi_\infty := (\mathcal{E}_\infty^A \circ \mathrm{id}^R)(\Phi_d^{AR})$.

The proof of this theorem is based on the following lemma on irreducibility of the KI decomposition.

Lemma 50 Let $\Gamma : \mathcal{H}^A \to \mathcal{H}^{a_0} \otimes \mathcal{H}^{a_L} \otimes \mathcal{H}^{a_R}$ be the KI decomposition of $\mathcal{H}^B$ with respect to a set of states $\mathfrak{S} := \{\rho_k\}_k$. By (2.87), state $\rho_k \in \mathfrak{S}$ is decomposed as

$$\Gamma \rho_k \Gamma^\dagger = \sum_{j \in J} p_{j|k} |j\rangle\langle j|^{a_0} \otimes \rho_{j|k}^{a_L} \otimes \sigma_j^{a_R}. \tag{6.22}$$

Let $\mathcal{H}_j^{a_L} := \mathrm{supp}(\sum_k \rho_{j|k})$. Then the following two properties hold.

1. If an operator Λ on $\mathcal{H}_j^{a_L}$ satisfies $\Lambda \rho_{j|k} = \beta \rho_{j|k} \Lambda$ for all k and for a complex number β, then $\Lambda = c I_j$, where c is a complex number and I_j is the identity operator on $\mathcal{H}_j^{a_L}$.

2. If an operator $\Lambda : \mathcal{H}_j^{a_L} \to \mathcal{H}_{j'}^{a_L}$ satisfies $\Lambda\rho_{j|k} = \alpha\rho_{j'|k}\Lambda$ for a positive number α, for some $j \neq j'$ and all k, then $\Lambda = 0$.

Proof. See Lemma 6 in Ref.[18]. ■

Proof of Theorem 49 The outline of the proof is as follows. Let Γ be the KI decomposition of A with respect to $\Psi(U)^{A(BR_B)}$. $\Psi(U)^A = \frac{1}{d}I$ is decomposed as

$$\Gamma\Psi(U)^A\Gamma^\dagger = \sum_{j\in J} p_j |j\rangle\langle j|^{a_0} \otimes \pi_j^{a_L} \otimes \pi_j^{a_R}, \tag{6.23}$$

where $\pi_j^{a_L}$ and $\pi_j^{a_R}$ are identity operators on subspaces $\mathcal{H}_j^{a_L} \subset \mathcal{H}^{a_L}$ and $\mathcal{H}_j^{a_R} \subset \mathcal{H}^{a_R}$, respectively. From (6.10), $\tilde{E}_t$ is decomposed as

$$\bar{E}_t := \Gamma\tilde{E}_t\Gamma^\dagger = \sum_{j\in J} p_j |j\rangle\langle j|^{a_0} \otimes \pi_j^{a_L} \otimes e_{j|t}^{a_R}, \tag{6.24}$$

where $e_{j|t}^{a_R}$ is a linear operator on $\mathcal{H}_j^{a_R}$. From Lemma 50, the set $\{e_{j|t}^{a_R}\}_t$ is irreducible on $\mathcal{H}_j^{a_R}$ for each j.

Let W be an arbitrary linear operator on $\mathcal{H}^A$. Since we have $(\tilde{\mathcal{E}} \circ \mathcal{E}_\infty) = \mathcal{E}_\infty$, from Lemma 48 we obtain

$$[\mathcal{E}_\infty(W), \tilde{E}_t] = 0 \tag{6.25}$$

for all t. Thus we have

$$[\bar{W}, \bar{E}_t] = 0 \tag{6.26}$$

for all t, where we defined $\bar{W} := \Gamma\mathcal{E}_\infty(W)\Gamma^\dagger$. Hence, from (6.24) and the irreducibility of $\{e_{j|t}^{a_R}\}_t$, $\bar{W}$ is decomposed as

$$\bar{W} = \sum_{j\in J} q_{j|W} |j\rangle\langle j|^{a_0} \otimes \eta_{j|W}^{a_L} \otimes \pi_j^{a_R}, \tag{6.27}$$

where $\eta_{j|W}^{a_L}$ is a linear operator on $\mathcal{H}_j^{a_L}$. In particular, when $W = \frac{1}{d}I$, we have $q_{j|W} = p_j$ and $\eta_{j|W}^{a_L} = \pi_j^{a_L}$. This implies that, with an isometry $\Gamma' : \mathcal{H}^R \to \mathcal{H}^{r_0} \otimes \mathcal{H}^{r_L} \otimes \mathcal{H}^{r_R}$, Φ_∞ is decomposed as

$$\Phi'_\infty := (\Gamma^A \otimes \Gamma'^R)\Phi_\infty(\Gamma^{\dagger A} \otimes \Gamma'^{\dagger R}) = \sum_{j\in J} p_j |j\rangle\langle j|^{a_0} \otimes |\Phi_j\rangle\langle\Phi_j|^{a_L r_L} \otimes \pi_j^{a_R} \otimes \pi_j^{r_R},$$

where $|\Phi_j\rangle\langle\Phi_j|^{a_L r_L}$ is a maximally entangled state with the Schmidt rank $\dim\mathcal{H}_j^{a_L}$, and $\pi_j^{r_R}$ is an identity operator on a subspace $\mathcal{H}_j^{r_R} \subset \mathcal{H}^{r_R}$ such that $\dim\mathcal{H}_j^{r_R} = \dim\mathcal{H}_j^{a_R}$. Therefore, we finally obtain

$$S(\Phi_\infty) = S(\Phi'_\infty) = H(\{p_j\}_j) + 2\sum_j p_j S(\pi_j^{a_R}) \tag{6.28}$$

which concludes the proof. ■

Using Theorem 49, we can compute $M(U)$ without explicitly obtaining the KI decomposition of $\Psi(U)$, if once Φ_∞ is provided. However, although $\tilde{\mathcal{E}}$ can be directly obtained from the operator Schmidt decomposition given by (6.1), the map $\mathcal{E}_\infty$ required to obtain Φ_∞ involves infinite series summation, and is not straightforwardly computable. To avoid this difficulty, we exploit a matrix representation of quantum operations [10]. Define a $d^2 - 1$ dimensional square matrix Ω by

$$[\Omega]_{pq,rs} = \mathrm{Tr}[\sigma_{pq}^\dagger \tilde{\mathcal{E}}(\sigma_{rs})] \ \ ((p,q),(r,s) \neq (0,0)), \tag{6.29}$$

which is a Hermitian matrix because of self-adjointness of $\tilde{\mathcal{E}}$. Moreover, due to (6.19), for an arbitrary $\omega \in \mathcal{L}(\mathcal{H})$ we have

$$\mathrm{Tr}[\omega^\dagger \tilde{\mathcal{E}}(\omega)] = \mathrm{Tr}[\mathcal{E}^*(\omega)^\dagger \mathcal{E}^*(\omega)] \leq \mathrm{Tr}[\mathcal{E}^*(\omega^\dagger \omega)] \tag{6.30}$$

by the Schwarz inequality (See Lemma 12 in [21]). Thus the absolute values of eigenvalues of Ω are not greater than 1. When the decomposition of $\tau \in \mathcal{S}(\mathcal{H})$ with respect to σ_{pq} is given by

$$\tau = \frac{1}{d} I + \sum_{(p,q)\neq(0,0)} f_{pq}\sigma_{pq}, \tag{6.31}$$

the decomposition of $\tilde{\mathcal{E}}(\tau)$ is given by

$$\tilde{\mathcal{E}}(\tau) = \frac{1}{d} I + \sum_{(p,q)\neq(0,0)} f'_{pq}\sigma_{pq}, \tag{6.32}$$

where

$$f'_{pq} = \sum_{rs} \Omega_{pq,rs} f_{rs}. \tag{6.33}$$

Thus the matrix representation of $\tilde{\mathcal{E}}^n$ and $\mathcal{E}_\infty$ are given by Ω^n and

$$\Omega_\infty := \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Omega^n, \tag{6.34}$$

respectively. It is straightforward to verify that Ω_∞ is the projection onto the eigensubspace of Ω corresponding to the eigenvalue 1. (If Ω does not have an eigenvalue 1, then $\Omega_\infty = 0$.) Thus Ω_∞ can be computed simply by diagonalizing Ω.

From Ω_∞ derived in this way, state Φ_∞ is obtained as follows. First, Φ is decomposed as

$$\begin{aligned}
|\Phi\rangle\langle\Phi|^{AR} &= \frac{1}{d}\sum_{p,q}\sigma_{pq}^{A}\mathrm{Tr}_A[(\sigma_{pq}^{\dagger A}\otimes I^R)|\Phi\rangle\langle\Phi|] \\
&= \frac{1}{d}\sum_{p,q}\sigma_{pq}^{A}\mathrm{Tr}_A[(I^A\otimes\sigma_{pq}^{*R})|\Phi\rangle\langle\Phi|] \\
&= \frac{1}{d^2}\sum_{p,q}\sigma_{pq}^{A}\otimes\sigma_{pq}^{*R} \\
&= \frac{1}{d^2}I^A\otimes I^R+\frac{1}{d^2}\sum_{(p,q)\neq(0,0)}\sigma_{pq}^{A}\otimes\sigma_{pq}^{*R}. \qquad (6.35)
\end{aligned}$$

Thus Φ_∞ is given by

$$\Phi_\infty^{AR}=\frac{1}{d^2}I^A\otimes I^R+\frac{1}{d^2}\sum_{(p,q)\neq(0,0)}(\Omega_\infty)_{pq,rs}\sigma_{pq}^{A}\otimes\sigma_{rs}^{*R}. \qquad (6.36)$$

The Markovianizing cost $M(U)=S(\Phi_\infty)$ is obtained by diagonalizing Φ_∞ and computing the Shannon entropy of its eigenvalues.

6.2 A Method Based on Decoupling Theorem

As indicated by Lemma 35 and 36, Markovianizing $\Psi(U)$ conditioned by R_A with an operation on A is equivalent to decoupling $\Psi(U)$ between AR_A and R_B with an operation on A. Therefore Markovianizing cost of U is equal to the cost of randomness required for decoupling $\Psi(U)$ between AR_A and R_B by a random unitary operation on A. To be more precise, we have the following statements.

Definition 51 We say that $\Psi(U)^{AR_AR_B}$ is decoupled between AR_A and R_B with the randomness cost R if, for any $\epsilon>0$ and for sufficiently large n, there exists a random unitary operation $\mathcal{T}_n^{\bar{A}}:\tau\mapsto 2^{-nR}\sum_{k=1}^{2^{nR}}V_k\tau V_k^\dagger$ on $\bar{A}$ such that

$$\left\|\mathcal{T}_n^{\bar{A}}(\Psi(U)^{\otimes n})^{\bar{A}\bar{R}_A\bar{R}_B}-\mathcal{T}_n^{\bar{A}}(\Psi(U)^{\otimes n})^{\bar{A}\bar{R}_A}\otimes(\Psi(U)^{\otimes n})^{\bar{R}_B}\right\|_1\leq\epsilon. \qquad (6.37)$$

The decoupling cost of U is defined as

$$\begin{aligned}D(U):=\inf\{R\mid \Psi(U)^{AR_AR_B}&\text{ is decoupled between}\\ &AR_A\text{ and }R_B\text{ with the randomness cost }R\}.\end{aligned}$$

Theorem 52 $M(U)=D(U)$.

Proof. Suppose $R>D(U)$. For any $\epsilon>0$ and for sufficiently large n, there exists a random unitary operation $\mathcal{T}_n^{\bar{A}}:\tau\mapsto 2^{-nR}\sum_{k=1}^{2^{nR}}V_k\tau V_k^\dagger$ such that (6.37) is satisfied.

The map $\mathcal{T}_n^{\bar{A}}$ is ϵ-decoupling between $\bar{A}\bar{R}_A$ and $\bar{R}_B$. Thus, from Lemma 35, it is $\eta(\epsilon)$-Markovianizing conditioned by $\bar{R}_A$. Hence we have $D(U) \geq M(U)$.

Conversely, suppose $R > M(U)$. For any $\epsilon > 0$ and for sufficiently large n, there exists a random unitary operation $\mathcal{T}_n^{\bar{A}} : \tau \mapsto 2^{-nR} \sum_{k=1}^{2^{nR}} V_k \tau V_k^\dagger$ that is ϵ-Markovianizing conditioned by $\bar{R}_A$. Thus, from Lemma 36, it is $\eta(\epsilon)$-decoupling between $\bar{A}\bar{R}_A$ and $\bar{R}_B$. Hence we have $D(U) \leq M(U)$. ■

The decoupling cost of U is an operationally defined function and is difficult to compute in general. But for a class of bipartite unitaries called *generalized Clifford operators*, it is possible to compute $D(U)$ by using the decoupling theorem (Lemma 15). A unitary U acting on two d-dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is called a generalized Clifford operator if, for any p, q, r, s and a phase θ_{pqrs}, there exist p', q', r' and s' such that $U(\sigma_{pq} \otimes \sigma_{rs})U^\dagger = e^{i\theta_{pqrs}} \sigma_{p'q'} \otimes \sigma_{r's'}$. Here σ_{pq} are generalized Pauli operators defined as (3.33).

The idea of the proof is as follows. Although R_B is a reference system that Alice and Bob cannot access, for the moment imagine that we can apply a random unitary operation on R_B. In particular, suppose that we can randomly apply generalized Pauli operators σ_{pq}^T with the uniform distribution. As it is presented in Section 3.3, the random Pauli operation decouples AR_A and R_B. In the asymptotic limit of infinite copies, the number of Pauli operators required to decouple AR_A and R_B is equal to $I(AR_A : R_B)_{\Psi(U^\dagger)} = K(U)$ per copy. But due to the commutation relation of generalized Pauli and Clifford operators, we can replace σ_{pq}^T on R_B by $\sigma_{p'q'} \otimes \sigma_{r's'}$ on AB. For the state $\Psi(U^\dagger)^{AR_AR_B}$, applying σ_{pq}^T on R_B is equivalent to applying $\sigma_{p'q'}$ on A. Thus the random Pauli operation on R_B is exactly substituted by a random Pauli operation on A. Hence the number of random Pauli operators on A that we need to decouple AR_A and R_B is at most $I(AR_A : R_B)_{\Psi(U^\dagger)} = K(U)$ per copy. We show the rigorous proof below.

Theorem 53 $M(U_{\mathrm{Cl}}) = K(U_{\mathrm{Cl}})$.

Proof. Let $|\Psi(U)\rangle = U^{AB}|\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}$. We have $I(AR_A : R_B)_{\Psi(U)} = S(AR_A)_{\Psi(U)} + S(R_B)_{\Psi(U)} - S(B)_{\Psi(U)} = S(AR_A)_{\Psi(U)} = K(U)$. From Theorem 15, we have $D(U) \geq K(U)$.

Fix arbitrary $\epsilon > 0$ and choose sufficiently large n. Let $\boldsymbol{\sigma}_{\vec{p}\vec{q}} := \sigma_{p_1q_1} \otimes \cdots \otimes \sigma_{p_nq_n}$ be tensor products of generalized Pauli operators on ${R_B}^{\otimes n} = \bar{R}_B$. Consider an ensemble of unitaries $\{\frac{1}{d^{2n}}, \boldsymbol{\sigma}_{\vec{p}\vec{q}}\}_{\vec{p}\vec{q}}$. Since $\Psi(U)^{R_B} = \frac{1}{d}I^{R_B}$, $\boldsymbol{\sigma}_{\vec{p}\vec{q}}$ is a unitary on the typical subspace of $(\Psi(U)^{\otimes n})^{\bar{R}_B}$. In the same way as (3.36), the ensemble satisfies

$$\frac{1}{d^{2n}} \sum_{\vec{p}\vec{q}} \boldsymbol{\sigma}_{\vec{p}\vec{q}}^{\bar{R}_B} |\phi\rangle\langle\phi|^{\bar{R}_B} \boldsymbol{\sigma}_{\vec{p}\vec{q}}^{\dagger\bar{R}_B} = \frac{1}{d^n} I^{\bar{R}_B}.$$

Thus, from (3.31) in Theorem 15, if $R \geq I(AR_A : R_B)_{\Psi(U)} + \epsilon = K(U) + \epsilon$, there exists a set of unitaries $\{\boldsymbol{\sigma}_{\vec{p}_k\vec{q}_k}\}_{k=1}^{2^{nR}}$ on $\bar{R}_B$ such that

$$\left\| \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} \boldsymbol{\sigma}_{\vec{p}_k\vec{q}_k}^{\bar{R}_B} (\Psi(U)^{\otimes n})^{\bar{A}\bar{R}_A\bar{R}_B} \boldsymbol{\sigma}_{\vec{p}_k\vec{q}_k}^{\dagger\bar{R}_B} - (\Psi(U)^{\otimes n})^{\bar{A}\bar{R}_A} \otimes \frac{1}{d^n} I^{\bar{R}_B} \right\|_1 \leq \eta(\epsilon).$$

When $U = U_{\mathrm{Cl}}$, we have

$$\begin{aligned} \sigma_{pq}^{R_B} |\Psi(U_{\mathrm{Cl}})\rangle^{ABR_AR_B} &= (U_{\mathrm{Cl}}^{AB} \otimes \sigma_{pq}^{R_B}) |\Phi_d\rangle^{AR_A} |\Phi_d\rangle^{BR_B} \\ &= e^{i\theta_{pq}} U_{\mathrm{Cl}}^{AB} (I^A \otimes \sigma_{pq}^B) |\Phi_d\rangle^{AR_A} |\Phi_d\rangle^{BR_B} \\ &= e^{i\theta_{pq}} (\sigma_{p'q'(pq)}^A \otimes \sigma_{r's'(pq)}^B) U_{\mathrm{Cl}}^{AB} |\Phi_d\rangle^{AR_A} |\Phi_d\rangle^{BR_B} \\ &= e^{i\theta_{pq}} (\sigma_{p'q'(pq)}^A \otimes \sigma_{r's'(pq)}^B) |\Psi(U_{\mathrm{Cl}})\rangle^{ABR_AR_B}. \end{aligned}$$

The third line follows from $\sigma_{pq}^{R_B} |\Phi_d\rangle^{BR_B} = (\sigma_{pq}^T)^B |\Phi_d\rangle^{BR_B} = e^{i\theta_{pq}} \sigma_{pq}^B |\Phi_d\rangle^{BR_B}$. In particular, we have

$$\sigma_{pq}^{R_B} \Psi(U_{\mathrm{Cl}})^{AR_AR_B} \sigma_{pq}^{\dagger R_B} = \sigma_{p'q'(pq)}^A \Psi(U_{\mathrm{Cl}})^{AR_AR_B} \sigma_{p'q'(pq)}^{\dagger A}.$$

Thus, for the state $\Psi(U_{\mathrm{Cl}})^{AR_AR_B}$, applying $\sigma_{pq}^{R_B}$ is equivalent to applying $\sigma_{p'q'(pq)}^A$. For the same reason, for $(\Psi(U_{\mathrm{Cl}})^{\otimes n})^{\bar{A}\bar{R}_A\bar{R}_B}$, applying $\boldsymbol{\sigma}_{\vec{p}_k\vec{q}_k}^{\bar{R}_B}$ is equivalent to applying $\boldsymbol{\sigma}_{\vec{p}'_k\vec{q}'_k}^{\bar{A}}$. Thus $\Psi(U_{\mathrm{Cl}})^{AR_AR_B}$ is decoupled between AR_A and R_B with the randomness cost $K(U_{\mathrm{Cl}})$. That is, we have $D(U_{\mathrm{Cl}}) \leq K(U_{\mathrm{Cl}})$.

From Theorem 52, we obtain $M(U_{\mathrm{Cl}}) = D(U_{\mathrm{Cl}}) = K(U_{\mathrm{Cl}})$. ■

Chapter 7

Implications of the Results

In this Chapter, we discuss implications of our results for distributed quantum computation obtained in Chapter 5 and Chapter 6. We consider two-qubit controlled-unitaries as one of the simplest examples of bipartite unitaries, and compute the Markovianizing costs of this class of unitaries. We show that it is *not* possible to reduce costs of entanglement and classical communication below the limit for single-shot ($n = 1$) protocols, even in an asymptotic scenario ($n \to \infty$), by one-round protocols. We also show that it *is* possible to reduce costs below the single-shot limit by considering a novel asymptotic scenario if we consider two-round protocols, that is, protocols consisting of concatenation of two one-round protocols. Thus we give an example of LOCC tasks for which there is a trade-off relation between entanglement cost and number of rounds. Such an example for asymptotic setting has not been reported before, and this result opens a new direction in the resource theory for distributed quantum computation.

7.1 Reviews on Single-Shot Protocols

Two qubit unitaries of the form

$$U^{AB} = |0\rangle\langle 0|^A \otimes I^B + |1\rangle\langle 1|^A \otimes u^B, \tag{7.1}$$

where u are a one-qubit unitary, are called *controlled-unitaries*. It is proved in [42] that any two-qubit controlled-unitaries have an operator Schmidt decompositions of the form

$$U_\theta'^{AB} = \cos\left(\frac{\theta}{2}\right) \cdot I^A \otimes I^B + i \sin\left(\frac{\theta}{2}\right) \cdot \sigma_z^A \otimes \sigma_z^B, \tag{7.2}$$

up to local unitary equivalence. Here, θ is a phase factor determined by u. Conversely, any two-qubit unitaries that has operator Schmidt number 2 is local unitarily equivalent to the following controlled-unitary:

$$U_\theta^{AB} = |0\rangle\langle 0|^A \otimes I^B + |1\rangle\langle 1|^A \otimes (e^{i\theta\sigma_z})^B \quad \left(0 < \theta \leq \frac{\pi}{2}\right). \tag{7.3}$$

In the following part of this chapter, we analyze and compare costs of entanglement and classical communication required for implementing U_θ^{AB} in a single-shot ($n = 1$), and those required in an asymptotic ($n \to \infty$) scenario. Note that U_θ is close to the identity when θ is sufficiently small.

As is discussed in Section 5.1, protocols for this task are classified in terms of the success probability and the fidelity of the final state to the target state $U^{AB}|\varphi\rangle^{AB}$. Deterministic and exact protocols are studied in [26, 29, 30]. Results are summarized as follows:

I-1. There exists a protocol for implementing U_θ^{AB} by using one Bell pair, one bit of forward and backward classical communication. The protocol is a one-round protocol, that is, composed of three steps [26].

I-2. If the resource entanglement is a pure state with Schmidt rank 2, then the state must be maximally entangled. In other words, if the resource entanglement is a state of two qubits, it must be a Bell pair, regardless of the number of steps in the protocol [29, 30].

I-3. There exists θ such that U_θ can be implemented by a protocol using entanglement resource with Schmidt rank 3, but with the entanglement entropy less than 1. The protocol consists of four steps. [30].

Although a restriction on the Schmidt rank is imposed in Result 2, the result is counterintuitive in the sense that consuming one Bell pair is necessary regardless of θ, whereas the power of the unitary to generate entanglement depends on θ

Probabilistic protocols are investigated, e.g., in [31, 32, 33, 34, 35, 36]. In particular, we describe the protocol introduced in [34]. The protocol consists of one-round LOCC, and implements U_θ by using the following state as resource:

$$|\phi_\alpha\rangle^{A_0B_0} = \cos\left(\frac{\alpha}{2}\right)|0\rangle|0\rangle + i\sin\left(\frac{\alpha}{2}\right)|1\rangle|1\rangle. \tag{7.4}$$

We only consider cases where $\cos\alpha(\sin\theta + \cos\theta) \geq 1$ for simplicity. Then the success probability is given by

$$p(\alpha, \theta) = \frac{\sin^2\alpha}{2(1 - \cos\theta\cos\alpha)}. \tag{7.5}$$

If it fails, then another controlled-unitary $U_{\theta'}$ is applied to the input pair. This protocol, although probabilistic at this stage, can be transformed to a deterministic one by adding another round [34]. If the protocol in the first round succeeds, then Alice and Bob do nothing in the second round. If it fails, Alice and Bob perform the protocol described in I-1 to apply $U_{\theta-\theta'}$ by consuming one Bell pair. Note that $U_{\theta-\theta'}U_{\theta'} = U_\theta$. Thus the protocol succeeds in implementing U_θ in total, regardless of the failure in the intermediate step. The average entanglement cost, measured by entanglement entropy, is given by

$$\bar{E}(\alpha,\theta) = p(\alpha,\theta)h\left(\cos\left(\alpha/2\right)\right) + (1-p(\alpha,\theta))(1+h\left(\cos\left(\alpha/2\right)\right)). \tag{7.6}$$

Here, $h(x)$ is the *binary entropy* defined as $h(x) = -x\log x - (1-x)\log\left(1-x\right)$. It is proved in [34] that the minimum average entanglement cost for θ, defined as

$$\bar{E}^*(\theta) := \min_{\alpha} \bar{E}(\alpha,\theta), \tag{7.7}$$

is strictly smaller than 1 when θ is smaller than about 0.75. Moreover, $\bar{E}^*(\theta)$ is a continuous function of θ that satisfies $\lim_{\theta\to 0}\bar{E}^*(\theta) = 0$.

7.2 Trade-off Relation between Entanglement and Number of Rounds

Let us now consider implementation of two-qubit controlled-unitaries U_θ in the asymptotic scenario as we analyzed in Chapter 5. Since U_θ^{AB} is almost equal to the identity if θ is sufficiently small, it would be natural to expect that the costs of entanglement and classical communication can be reduced by considering the asymptotic scenario. That is, it might be possible that the costs of entanglement and classical communication per input pair can be reduced below the single-shot limit (one Bell pair for deterministic and exact protocols). In this section, we investigate if such a reduction of resources is possible.

First, we compute the Markovianizing cost of U_θ. Since U_θ given by (7.3) is local unitarily equivalent to U'_θ given by (7.2), from Lemma 47, $M(U_\theta) = M(U'_\theta)$ does not depend on θ. Moreover, U'_θ is a Clifford operator when $\theta = \pi/2$, and thus we have

$$M(U_{\pi/2}) = M(U'_{\pi/2}) = K(U'_{\pi/2}) = 1. \tag{7.8}$$

Thus we have $M(U_\theta) = 1$ regardless of θ. Consequently, we have the following corollary stating that it is *not* possible to reduce resource costs in the asymptotic scenario by one-round protocols.

Corollary 54 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is achievable in one-round EALOCC implementation of a one-qubit controlled-unitary if and only if $R_E, C^{\rightarrow}, C^{\leftarrow} \geq 1$.

Second, we show that it *is* possible to reduce the entanglement cost in the asymptotic scenario by EALOCC protocols consisting of two rounds. Our proof is by construction based on the probabilistic protocol introduced in Section 7.1. Take arbitrary $\epsilon, \delta > 0$ and choose sufficiently large n. The protocol $\mathcal{M}_n$ for n input pairs proceeds as follows:

1. Alice and Bob initially share, as resource entanglement, a maximally entangled state with Schmidt rank $2^{n(\bar{E}(\alpha,\theta)+\epsilon)}$.

2. By a local operation, they transform the resource entanglement to n copies of $|\phi_\alpha\rangle$ and $n(1-p(\alpha,\theta)+\epsilon)$ Bell pairs.

3. By using n copies of $|\phi_\alpha\rangle$ as resources, they perform U_θ on each input pair by a probabilistic protocol described in Section 7.1. With probability greater than $1-\delta$, the number of pairs for which the protocol succeeds in implementing U_θ is at least $n(p(\alpha,\theta)-\epsilon)$, and the number of pairs for which $U_{\theta'}$ has been applied is at most $n(1-p(\alpha,\theta)+\epsilon)$.

4. By using $n(1-p(\alpha,\theta)+\epsilon)$ Bell pairs, they perform $U_{\theta-\theta'}$ on pairs for which $U_{\theta'}$ has been applied.

Since ϵ and δ can be arbitrarily small, we obtain the following corollary.

Corollary 55 A rate triplet $(R_E, C^{\rightarrow}, C^{\leftarrow})$ is achievable in two-round EALOCC implementation of U_θ if $R_E, C^{\rightarrow}, C^{\leftarrow} \geq \bar{E}^*(\theta)$.

We note that $\bar{E}^*(\theta)$ is strictly smaller than 1 when θ is smaller than about 0.75. Combined with Corollary 54, we conclude that EALOCC implementation of U_θ is, for a proper parameter region, an example of tasks for which there is a trade-off relation between the required cost of entanglement and number of rounds. In particular, the ratio between the minimum costs of entanglement in one-round protocols and two-round protocols can be arbitrarily large, since $\lim_{\theta\to 0}\bar{E}^*(\theta)=0$ whereas $M(U_\theta)=1$ regardless of θ.

To our knowledge, this is the first reported case where the required amount of entanglement resource in an EALOCC task depends on the number of rounds in an asymptotic scenario. This result indicates a new interesting property of a trade-off relation between entanglement requirement and the number of rounds in EALOCC tasks, and suggests that, in addition to entanglement cost and forward/backward classical communication cost, the number of rounds plays an important role in the resource theory of distributed quantum computation.

Chapter 8

Summary and Discussion

We have studied distributed quantum computation by developing concepts and techniques in quantum Shannon theory, and derived the minimum costs of entanglement and classical communication required for implementing bipartite unitaries by entanglement-assisted LOCC consisting of three steps. In the following, we summarize the results and discuss some issues.

In Chapter 4, we introduce and analyze a task of Markovianization, in which a tripartite states is transformed to a quantum Markov chain by a randomizing operation on one of the three subsystems. By extending the Koashi-Imoto (KI) decomposition for tripartite pure states, we derive the Markovianizing cost, that is, the minimum cost of randomness required for Markovianization. For the proof of the upper bound, we develop random coding method based on the Haar distributed random unitary ensemble, which has originally been extensively used in quantum Shannon theory, by taking the structure of the KI decompsition into account. For the proof of the lower bound, we prove an entropic inequality regarding the KI decomposition of two different states, by applying a proof technique of the previously known data compression theorem for quantum mixed state signals. We also consider Markovianization induced by a measurement. We derive lower bounds on various entropic quantities regarding the state transformation by Markovianizing measurements.

In Chapter 5, we analyze implementation of bipartite unitaries by LOCC (local operations and classical communication) assisted by shared entanglement. We consider an asymptotic scenario in which the two parties perform the same bipartite unitaries on infinitely many independent input pairs. As the first nontrivial case, we consider protocols consisting of three steps. Our main result is that the minimum costs of entanglement and classical communication are given by the Markovianizing cost of a tripartite state associated with the unitary.

The result indicates that the KI decomposition divides information of input states into three parts: one of which only information of amplitude in a basis is required for computation (information of phase is not), another of which both amplitude and phase information are required, and the other of which no information is involved in computation. This corresponds to the fact that the KI decomposition of an ensemble divides information of each element into three pieces: the classical part, the quantum part and the redundant part. Our construction of a protocol that requires the smallest amount of resources is based on this decomposition.

A desirable refinement of our result is to lift the condition that the average input state is maximally mixed. This refinement would be possible by exploiting a representation theoretical methods used, e.g., in [40].

In Chapter 6, we propose two methods to compute the Markovianizing cost of a unitary. One is based on a group theoretical property of the KI decomposition, and the other is based on the commutation relation of generalized Clifford operators. As a simplest example, we show that the Markovianizing cost of any two-qubit controlled-phase gate is equal to 1. This result indicates that, in this example, it is not possible to reduce resource costs by three-step protocols by considering the asymptotic scenario, below the single-shot limit.

In Chapter 7, we consider two-qubit controlled-unitaries as a particular example. By applying results obtained in Chapter 5 and Chapter 6, we compare the minimum entanglement cost in one-round protocols and two-round protocols. We show that, for a particular class of two-qubit controlled unitaries, the minimum entanglement cost in two-round protocols is strictly smaller than that in one-round protocols. Thus we find a new interesting property of trade-off relation between entanglement cost and number of rounds required for implementing a class of controlled-unitaries, and suggests that the number of rounds plays an important role in the resource theory of EALOCC tasks.

Chapter 9

Appendix

Proof of Theorem 13 Without loss of generality, we assume that the protocol $\mathcal{M}$ proceeds as follows.

I-1. Bob performs a measurement $\{M_k^{BB_0\to B_1}\}_k$. The probability of obtaining the measurement result k is given by $p_k = \|M_k|\Psi\rangle|\Phi_K\rangle\|_1^2$, and the state after the measurement is $|\Psi_k\rangle = p_k^{-1/2}M_k|\Psi\rangle|\Phi_K\rangle$.

I-2. Bob communicates the measurement result k to Alice.

I-3. Alice performs an operation which is described by a CPTP map $\mathcal{O}_k : AA_0 \to A\hat{A}A_1$. The final state is given by $\Psi_k^{\rm fin} = \mathcal{O}_k(\Psi_k)$.

In total, the final state is given by

$$\rho(\mathcal{M})^{A\hat{A}RA_1B_1} = \sum_k p_k \Psi_k^{\rm fin}. \tag{9.1}$$

Thus Condition (3.20) implies

$$\sum_k p_k F(\Psi_k^{\rm fin}, |\Psi\rangle^{AA'R}|\Phi_L\rangle^{A_1B_1}) \geq 1-\epsilon, \tag{9.2}$$

and consequently,

$$\sum_k p_k \left\| \Psi_k^{\rm fin} - \Psi^{AA'R}\otimes {\Phi_L}^{A_1B_1} \right\|_1 \leq \eta(\epsilon). \tag{9.3}$$

Consider the following protocol, which is as a whole equivalent to the protocol described above.

II-1. Bob performs a CPTP map $\mathcal{E}_1 : BB_0 \to B_1C$ defined by

$$\mathcal{E}_1(\tau) = \sum_k |k\rangle\langle k|^C \otimes M_k\tau M_k^\dagger. \tag{9.4}$$

The state after the operation is

$$\Psi' = \sum_k p_k |k\rangle\langle k|^C \otimes |\Psi_k\rangle\langle\Psi_k|^{AA_0B_1R}. \tag{9.5}$$

Ⅱ-2. Bob transmits system C to Alice.

Ⅱ-3. Bob performs a CPTP map $\mathcal{E}_2 : CAA_0 \to A\hat{A}A_1$ defined as

$$\mathcal{E}_2(\tau) = \sum_k \mathcal{O}_k(\tau_k^{AA_0}), \tag{9.6}$$

where $\tau_k^{AA_0} := \langle k|^C \tau^{CAA_0} |k\rangle^C$. The state after the operation is

$$\mathcal{E}_2(\Psi') = \sum_k p_k \mathcal{O}_k(\Psi_k) = \rho(\mathcal{M}). \tag{9.7}$$

By the data processing inequality, we have

$$\begin{aligned}
2S(A)_\Psi + 2\log K &= I(AA_0 : BB_0R)_{\Psi\otimes\Phi_K} \geq I(AA_0 : B_1CR)_{\Psi'} \\
&= I(AA_0 : C)_{\Psi'} + I(AA_0 : B_1R|C)_{\Psi'} \\
&= I(AA_0 : C)_{\Psi'} + I(AA_0C : B_1R)_{\Psi'} - I(C : B_1R)_{\Psi'} \\
&= I(AA_0C : B_1R)_{\Psi'} + S(AA_0|C)_{\Psi'} - S(B_1R|C)_{\Psi'} \\
&= I(AA_0C : B_1R)_{\Psi'} + \sum_k p_k \left(S(AA_0)_{\Psi_k} - S(B_1R)_{\Psi_k}\right) \\
&= I(AA_0C : B_1R)_{\Psi'} \geq I(A\hat{A}A_1 : B_1R)_{\rho(\mathcal{M})} \\
&\geq I(A\hat{A}A_1 : B_1R)_{\Psi\otimes\Phi_L} - \eta(\epsilon)\log(d_Ad_Bd_RL) \\
&= I(A\hat{A} : R)_\Psi + I(A_1 : B_1)_{\Phi_L} - \eta(\epsilon)\log(d_Ad_Bd_RL^2) \\
&= 2S(R)_\Psi + 2(1-\eta(\epsilon))\log L - \eta(\epsilon)\log(d_Ad_Bd_R).
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\log K - (1-\eta(\epsilon))\log L &\geq S(R)_\Psi - S(A)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= S(AB)_\Psi - S(A)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= S(B|A)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R).
\end{aligned}$$

We also have

$$\begin{aligned}
2S(R)_\Psi &= I(A\hat{A} : R)_\Psi \leq I(A\hat{A} : R)_{\rho(\mathcal{M})} + \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= I(AA_0C : R)_{\Psi'} + \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= I(AA_0 : R)_{\Psi'} + I(C : R|AA_0)_{\Psi'} + \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= I(AA_0 : R)_{\Psi\otimes\Phi_K} + I(C : AA_0R)_{\Psi'} - I(C : AA_0)_{\Psi'} + \eta(\epsilon)\log(d_Ad_Bd_R) \\
&\leq I(A : R)_\Psi + S(C)_{\Psi'} + \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= I(A : R)_\Psi + H(\{p_k\}_k) + \eta(\epsilon)\log(d_Ad_Bd_R).
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
C &\geq H(\{p_k\}_k) \geq 2S(R)_\Psi - I(A : R)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= S(R)_\Psi + S(AR)_\Psi - S(A)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= S(R)_\Psi + S(B)_\Psi - S(BR)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R) \\
&= I(B : R)_\Psi - \eta(\epsilon)\log(d_Ad_Bd_R).
\end{aligned}$$

■

Proof of Inequality (4.60)

$$
\begin{aligned}
\Delta I_k &:= I(\bar{B}\bar{C} : G)_{\Psi_k} \\
\text{①} \quad &\geq I(\bar{B}\bar{C} : G)_{\chi} - n\eta(\epsilon'_k)\log(d_B d_C) - \eta(\epsilon'_k)\log(d_G) \\
\text{②} \quad &= I(\hat{b}_0\hat{b}_L\hat{b}_R\bar{C} : \hat{g}_0\hat{g}_L\hat{g}_R)_{\tilde{\chi}} - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{③} \quad &= I(\hat{b}_0 : \hat{g}_0)_{\tilde{\chi}} + I(\hat{b}_L\hat{b}_R\bar{C} : \hat{g}_0\hat{g}_L\hat{g}_R|\hat{b}_0)_{\tilde{\chi}} - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{④} \quad &\geq I(\hat{b}_0 : \hat{g}_0)_{\tilde{\chi}} + I(\hat{b}_R\bar{C} : \hat{g}_R|\hat{b}_0)_{\tilde{\chi}} - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{⑤} \quad &\geq S(\hat{b}_0)_{\tilde{\chi}} + 2S(\hat{b}_R\bar{C}|\hat{b}_0)_{\tilde{\chi}} - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{⑥} \quad &\geq 2S(\bar{C}|\hat{b}_0\hat{b}_L\hat{b}_R)_{\tilde{\chi}} + S(\hat{b}_0)_{\tilde{\chi}} + 2S(\hat{b}_R|\hat{b}_0)_{\tilde{\chi}} - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{⑦} \quad &= 2S(\bar{C}|\bar{B})_{\chi} + H(\{q_i\}_i) + 2\sum_i q_i S(\phi_i^{\hat{b}_R}) - n\eta(\epsilon'_k)\log(d_A d_B d_C) \\
\text{⑧} \quad &\geq n\left(2S(C|B)_{\Psi} + H(\{p_j\}_j) + 2\sum_j p_j S(\varphi_j^{b_R}) - \eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\right) \\
\text{⑨} \quad &= n\Big(2S(A)_{\Psi} - 2S(b_0 b_L b_R)_{\Psi_{KI}} + S(b_0)_{\Psi_{KI}} + 2S(b_R|b_0)_{\Psi_{KI}} \\
&\qquad -\eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\Big) \\
\text{⑩} \quad &= n\left(2S(A)_{\Psi} - S(b_0)_{\Psi_{KI}} - 2S(b_L|b_0)_{\Psi_{KI}} - \eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\right) \\
\text{⑪} \quad &= n\left(2S(a_0 a_L a_R)_{\Psi_{KI}} - S(a_0)_{\Psi_{KI}} - 2S(a_L|a_0)_{\Psi_{KI}} - \eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\right) \\
\text{⑫} \quad &= n\left(S(a_0)_{\Psi_{KI}} + 2S(a_R|a_0)_{\Psi_{KI}} - \eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\right) \\
\text{⑬} \quad &= n\left(H(\{p_j\}_j) + 2\sum_j p_j S(\varphi_j^{a_R}) - \eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right)\right) \\
\text{⑭} \quad &= nM_{A|B}(\Psi^{ABC}) - n\eta_{\Psi}(\epsilon_k)\log\left(d_A d_B d_C\right).
\end{aligned}
\tag{9.8}
$$

Here, the reason for each line is as follows:

① (4.57) and the continuity of the mutual information.

② (4.59) and Condition 3.

③ The chain rule of the mutual information.

④ The data processing inequality of the conditional mutual information.

⑤ $I(\hat{b}_0 : \hat{g}_0)_{\tilde{\chi}} \geq S(\hat{b}_0)_{\tilde{\chi}}$ follows from $\mathcal{D}^{\hat{b}_0}(\tilde{\chi}^{\hat{b}_0\hat{g}_0}) = \sum_i q_i |i\rangle\langle i|^{\hat{b}_0} \otimes |i\rangle\langle i|^{\hat{g}_0}$ and the data processing inequality. $I(\hat{b}_R\bar{C} : \hat{g}_R|\hat{b}_0)_{\tilde{\chi}} = 2S(\hat{b}_R\bar{C}|\hat{b}_0)_{\tilde{\chi}}$ follows from $\tilde{\chi}^{\hat{b}_0\hat{b}_R\bar{C}\hat{g}_R} = \sum_i q_i |i\rangle\langle i|^{\hat{b}_0} \otimes |\phi_i\rangle\langle\phi_i|^{\hat{b}_R\bar{C}\hat{g}_R}$.

⑥ $S(\hat{b}_R\bar{C}|\hat{b}_0) = S(\hat{b}_R|\hat{b}_0) + S(\bar{C}|\hat{b}_0\hat{b}_R) \geq S(\hat{b}_R|\hat{b}_0) + S(\bar{C}|\hat{b}_0\hat{b}_L\hat{b}_R)$.

⑦ From (4.59).

⑧ From (4.56) and $\chi^{\bar{B}\bar{C}} = \Upsilon_k^{\bar{B}\bar{C}}$, we have $\|(\Psi^{\otimes n})^{\bar{B}\bar{C}} - \chi^{\bar{B}\bar{C}}\|_1 \leq \epsilon'_k + \epsilon''_k = \epsilon_k$, thus $|S(\bar{C}|\bar{B})_\chi - nS(C|B)_\Psi| \leq n\eta(\epsilon_k)\log d_C$. Lemma 24 is also applied by replacing A with B.

⑨ (4.10) and $S(C|B)_\Psi = S(BC)_\Psi - S(B)_\Psi = S(A)_\Psi - S(B)_\Psi$.

⑩ $S(b_0b_Lb_R)_{\Psi_{KI}} = S(b_0)_{\Psi_{KI}} + S(b_Lb_R|b_0)_{\Psi_{KI}} = S(b_0)_{\Psi_{KI}} + S(b_L|b_0)_{\Psi_{KI}} + S(b_R|b_0)_{\Psi_{KI}}$.

⑪ From (4.10), we have $S(b_0)_{\Psi_{KI}} = H(\{p_j\}_j) = S(a_0)_{\Psi_{KI}}$ and $S(b_L|b_0)_{\Psi_{KI}} = \sum_j p_j S(\omega_j^{b_L}) = \sum_j p_j S(\omega_j^{a_L}) = S(a_L|a_0)_{\Psi_{KI}}$.

⑫ $S(a_0a_La_R)_{\Psi_{KI}} = S(a_0)_{\Psi_{KI}} + S(a_La_R|a_0)_{\Psi_{KI}} = S(a_0)_{\Psi_{KI}} + S(a_L|a_0)_{\Psi_{KI}} + S(a_R|a_0)_{\Psi_{KI}}$.

⑬ From (4.10).

⑭ Theorem 21.

Proof of Lemma 35 Let $\Xi : R_AR_B \to R_ABR_B$ be a CPTP map defined by

$$\Xi(\tau^{R_AR_B}) = d^2 \cdot (\Psi(U^\dagger)^{R_ABR_B})^{\frac{1}{2}}(\tau^{R_AR_B} \otimes I^B)(\Psi(U^\dagger)^{R_ABR_B})^{\frac{1}{2}}.$$

This is indeed CPTP since we have

$$\begin{aligned}\mathrm{Tr}[\Xi(\tau^{R_AR_B})] &= d^2 \cdot \mathrm{Tr}[\tau^{R_AR_B}(\Psi(U^\dagger)^{R_ABR_B})] \\ &= \mathrm{Tr}[\tau^{R_AR_B}(I^{R_A} \otimes I^{R_B})] \\ &= \mathrm{Tr}[\tau^{R_AR_B}].\end{aligned}$$

Consider two states

$$\check{\Psi}(U^\dagger)^{AR_A(BR_B)} := \Xi^{R_AR_B}(\Psi(U^\dagger)^{AR_A} \otimes \Psi(U^\dagger)^{R_B}) \tag{9.9}$$

and

$$\dot{\Psi}(U^\dagger)^{AR_ABR_B} := \Xi^{R_AR_B}(\Psi(U^\dagger)^{AR_AR_B}).$$

The state (9.9) is a Markov state conditioned by R_A because of Theorem 17. We also have

$$\dot{\Psi}(U^\dagger)^{AR_ABR_B} = \Psi(U^\dagger)^{AR_ABR_B}, \tag{9.10}$$

since we have

$$\begin{aligned}(\Psi(U^\dagger)^{R_ABR_B})^{\frac{1}{2}} &= \left(U^{*R_AR_B}\left(\frac{1}{d}I^{R_A} \otimes |\Phi\rangle\langle\Phi|^{BR_B}\right)U^{tR_AR_B}\right)^{\frac{1}{2}} \\ &= U^{*R_AR_B}\left(\frac{1}{\sqrt{d}}I^{R_A} \otimes |\Phi\rangle\langle\Phi|^{BR_B}\right)U^{tR_AR_B}\end{aligned} \tag{9.11}$$

and

$$\Psi(U^\dagger)^{AR_AR_B} = U^{*R_AR_B}\left(|\Phi\rangle\langle\Phi|^{AR_A} \otimes \frac{1}{d}I^{R_B}\right)U^{tR_AR_B},$$

which together imply

$$\begin{aligned}\Xi^{R_AR_B}(\Psi(U^\dagger)^{AR_AR_B}) &= U^{*R_AR_B}\left(|\Phi\rangle\langle\Phi|^{AR_A} \otimes |\Phi\rangle\langle\Phi|^{BR_B}\right)U^{tR_AR_B}\\ &= U^{\dagger AB}\left(|\Phi\rangle\langle\Phi|^{AR_A} \otimes |\Phi\rangle\langle\Phi|^{BR_B}\right)U^{AB}.\end{aligned}$$

Define

$$\check{\Psi}_M(U^\dagger)^{A'R_A(BR_B)} := \Xi^{R_AR_B}(\Psi_M(U^\dagger)^{A'R_A} \otimes \Psi(U^\dagger)^{R_B}) \tag{9.12}$$

and

$$\dot{\Psi}_M(U^\dagger)^{A'R_ABR_B} := \Xi^{R_AR_B}(\Psi_M(U^\dagger)^{A'R_AR_B}). \tag{9.13}$$

Due to Theorem 18, $\check{\Psi}_M(U^\dagger)^{A'R_A(BR_B)}$ is a Markov state conditioned by R_A. From (9.10), we have $\dot{\Psi}_M(U^\dagger)^{A'R_ABR_B} = \Psi_M(U^\dagger)^{A'R_ABR_B}$. Therefore, by the monotonicity of trace distance, we have

$$\begin{aligned}&\left\|\Psi_M(U^\dagger)^{A'R_A(BR_B)} - \check{\Psi}_M(U^\dagger)^{A'R_A(BR_B)}\right\|_1\\ =\ &\left\|\dot{\Psi}_M(U^\dagger)^{A'R_A(BR_B)} - \check{\Psi}_M(U^\dagger)^{A'R_A(BR_B)}\right\|_1\\ \leq\ &\left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi(U^\dagger)^{R_B}\right\|_1\\ \leq\ &\left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B}\right\|_1\\ &+\left\|\Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi(U^\dagger)^{R_B}\right\|_1\\ \leq\ &\left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B}\right\|_1 + \left\|\Psi_M(U^\dagger)^{R_B} - \Psi(U^\dagger)^{R_B}\right\|_1\\ \leq\ &\eta(\epsilon).\end{aligned} \tag{9.14}$$

The last inequality follows from the assumption and Lemma 32. ■

Proof of Lemma 36 Let $\Upsilon^{A'R_A(BR_B)}$ be a Markov state conditioned by R_A such that

$$\left\|\Psi_M(U^\dagger)^{A'R_A(BR_B)} - \Upsilon^{A'R_A(BR_B)}\right\|_1 \leq \epsilon. \tag{9.15}$$

By tracing out B, we obtain

$$\begin{aligned}
& \left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B}\right\|_1 \\
= & \left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B}\right\|_1 \\
\leq & \left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Upsilon^{A'R_AR_B}\right\|_1 + \left\|\Upsilon^{A'R_AR_B} - \Upsilon^{A'R_A} \otimes \Upsilon^{R_B}\right\|_1 \\
& + \left\|\Upsilon^{A'R_A} \otimes \Upsilon^{R_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B}\right\|_1 \\
\leq & \left\|\Psi_M(U^\dagger)^{A'R_AR_B} - \Upsilon^{A'R_AR_B}\right\|_1 + \left\|\Upsilon^{A'R_AR_B} - \Upsilon^{A'R_A} \otimes \Upsilon^{R_B}\right\|_1 \\
& + \left\|\Upsilon^{A'R_A} - \Psi_M(U^\dagger)^{A'R_A}\right\|_1 + \left\|\Upsilon^{R_B} - \Psi_M(U^\dagger)^{R_B}\right\|_1 \\
\leq & \left\|\Upsilon^{A'R_AR_B} - \Upsilon^{A'R_A} \otimes \Upsilon^{R_B}\right\|_1 + \eta(\epsilon). \qquad (9.16)
\end{aligned}$$

From Lemma 32, we also have

$$\begin{aligned}
& \left\|\Upsilon^{R_AR_B} - \Upsilon^{R_A} \otimes \Upsilon^{R_B}\right\|_1 \\
\leq & \left\|\Upsilon^{R_AR_B} - \Psi_M(U^\dagger)^{R_AR_B}\right\|_1 + \left\|\Psi_M(U^\dagger)^{R_AR_B} - \Upsilon^{R_A} \otimes \Upsilon^{R_B}\right\|_1 \\
= & \left\|\Upsilon^{R_AR_B} - \Psi_M(U^\dagger)^{R_AR_B}\right\|_1 + \left\|\Psi_M(U^\dagger)^{R_A} \otimes \Psi_M(U^\dagger)^{R_B} - \Upsilon^{R_A} \otimes \Upsilon^{R_B}\right\|_1 \\
\leq & \left\|\Upsilon^{R_AR_B} - \Psi_M(U^\dagger)^{R_AR_B}\right\|_1 + \left\|\Psi_M(U^\dagger)^{R_A} - \Upsilon^{R_A}\right\|_1 + \left\|\Psi_M(U^\dagger)^{R_B} - \Upsilon^{R_B}\right\|_1 \\
\leq & \ \eta(\epsilon). \qquad (9.17)
\end{aligned}$$

Let $\Gamma_\Upsilon : \mathcal{H}^{R_A} \to \mathcal{H}^{r_0} \otimes \mathcal{H}^{r_L} \otimes \mathcal{H}^{r_R}$ be the Markov decomposition of $\mathcal{H}^{R_A}$ with respect to $\Upsilon^{AR_A(BR_B)}$. The Markov decomposition of $\Upsilon^{A'R_A(BR_B)}$ on R_A is given by

$$\Upsilon_{Mk}^{AR_A(BR_B)} = \sum_i q_i |i\rangle\langle i|^{b_0} \otimes \sigma_i^{A'r_L} \otimes \phi_i^{r_R(BR_B)}$$

Let

$$\tilde{\Upsilon}^{A'R_AR_B} := \mathrm{Tr}_B[\Upsilon_{Mk}^{A'R_A(BR_B)}] = \sum_i q_i |i\rangle\langle i|^{b_0} \otimes \sigma_i^{A'r_L} \otimes \phi_i^{r_RR_B}.$$

We have

$$\tilde{\Upsilon}^{A'R_A} \otimes \tilde{\Upsilon}^{R_B} = \sum_i q_i |i\rangle\langle i|^{b_0} \otimes \sigma_i^{A'r_L} \otimes \phi_i^{r_R} \otimes \phi^{R_B},$$

where $\phi^{R_B} := \sum_i q_i \phi_i^{R_B}$, and thus

$$\begin{aligned}
\left\|\tilde{\Upsilon}^{A'R_AR_B} - \tilde{\Upsilon}^{A'R_A} \otimes \tilde{\Upsilon}^{R_B}\right\|_1 &= \sum_i q_i \left\|\phi_i^{r_RR_B} - \phi_i^{r_R} \otimes \phi^{R_B}\right\|_1 \\
&= \left\|\tilde{\Upsilon}^{R_AR_B} - \tilde{\Upsilon}^{R_A} \otimes \tilde{\Upsilon}^{R_B}\right\|_1,
\end{aligned}$$

which implies

$$\left\|\Upsilon^{A'R_AR_B} - \Upsilon^{A'R_A} \otimes \Upsilon^{R_B}\right\|_1 = \left\|\Upsilon^{R_AR_B} - \Upsilon^{R_A} \otimes \Upsilon^{R_B}\right\|_1. \qquad (9.18)$$

Combining (9.16), (9.17) and (9.18), we obtain

$$\left\| \Psi_M(U^\dagger)^{A'R_AR_B} - \Psi_M(U^\dagger)^{A'R_A} \otimes \Psi_M(U^\dagger)^{R_B} \right\|_1 \leq \eta(\epsilon),$$

which concludes the proof. ■

Proof of Lemma 37 and 38 From (5.4), we have

$$\begin{aligned} 1-\epsilon &\leq \sum_{kl} p_{kl} F(\Psi_{kl}^{\rm fin}(U^\dagger)^{ABR_AR_B}, |\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}) \\ &= \sum_{kl} p_{kl} F\left(\left|\Psi_{kl}^{\rm fin}(U^\dagger)\right\rangle, |\Phi_d\rangle^{AR_A}|\Phi_d\rangle^{BR_B}|\phi_{kl}\rangle^{A_1B_1E_3} \right) \end{aligned}$$

for some states ϕ_{kl}. By using the relation between fidelity and the trace distance, we obtain

$$\begin{aligned} \eta(\epsilon) &\geq \sum_{kl} p_{kl} \left\| \Psi_{kl}^{\rm fin}(U^\dagger) - {\Phi_d}^{AR_A} \otimes \frac{1}{d} I^{R_B} \otimes {\phi_{kl}}^{A_1} \right\|_1 \\ &= \sum_{kl} p_{kl} \left\| \Psi_{kl}(U^\dagger)^{A'R_AR_B} - V_{kl}^\dagger(\Phi_d^{AR_A} \otimes \phi_{kl}^{A_1})V_{kl} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &\geq \sum_k p_k \left\| \sum_l p_{l|k} \Psi_{kl}(U^\dagger)^{A'R_AR_B} - \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &= \sum_k p_k \left\| \Psi_k'(U^\dagger)^{A'R_AR_B} - \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &= \sum_k p_k \left\| \Psi_k(U^\dagger)^{A'R_AR_B} - \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1, \end{aligned} \tag{9.19}$$

where we defined

$$\psi_k^{A'R_A} := \sum_l p_{l|k} V_{kl}^\dagger (\Phi_d^{AR_A} \otimes \phi_{kl}^{A_1}) V_{kl}. \tag{9.20}$$

Hence we obtain

$$\begin{aligned} &\sum_k p_k \left\| \Psi_k(U^\dagger)^{A'R_AR_B} - \Psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ \leq\ & \sum_k p_k \left\| \Psi_k(U^\dagger)^{A'R_AR_B} - \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &+ \sum_k p_k \left\| \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} - \Psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ \leq\ & 2 \sum_k p_k \left\| \Psi_k(U^\dagger)^{A'R_AR_B} - \psi_k^{A'R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ \leq\ & \eta(\epsilon), \end{aligned} \tag{9.21}$$

which implies that Alice's measurement is ϵ-decoupling between AR_A and R_B. From (9.19), we also have

$$\begin{aligned}\eta(\epsilon) &\geq \sum_k p_k \left\| \Psi_k(U^\dagger)^{R_A R_B} - \psi_k^{R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &= \sum_k p_k \left\| \Psi_k(U^\dagger)^{R_A R_B} - \frac{1}{d} I^{R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &= \sum_k p_k \left\| \Phi_k^{R_A} \otimes \frac{1}{d} I^{R_B} - \frac{1}{d} I^{R_A} \otimes \frac{1}{d} I^{R_B} \right\|_1 \\ &= \sum_k p_k \left\| \Phi_k^{R_A} - \frac{1}{d} I^{R_A} \right\|_1 , \end{aligned} \tag{9.22}$$

which implies that Alice's measurement is ϵ-linear. ■

Proof of Theorem 43

1. Follow from (4.55) in Theorem 25.

2. Follows from (4.53) in Theorem 25.

3. From Lemma 42, there exist pure states $|\Psi_k^p\rangle^{A'\bar{R}_A B_1 B_E}$ such that

$$\sum_k p_k \left\| \Psi_k'(U^{\dagger \otimes n}) - (\Psi_k^p)^{A'\bar{R}_A B_1 B_E} \otimes (\Phi_d^{\otimes n})^{\bar{B}\bar{R}_B} \right\| \leq \eta(\epsilon). \tag{9.23}$$

Let $\tilde{\mathcal{M}}_{n,k} : A'B_1B_E \to \bar{A}A_1B_1$ be the CPTP map that describes the procedure Ⅱ-4∼7, presented in Section 5.3. We have

$$\sum_k p_k \left\| \tilde{\mathcal{M}}_{n,k}(\Psi_k'(U^{\dagger \otimes n})) - \tilde{\mathcal{M}}_{n,k}(\Psi_k^p)^{\bar{A}\bar{R}_A A_1 B_1} \otimes (\Phi_d^{\otimes n})^{\bar{B}\bar{R}_B} \right\| \leq \eta(\epsilon). \tag{9.24}$$

From (5.25), we also have

$$\sum_k p_k \left\| \tilde{\mathcal{M}}_{n,k}(\Psi_k'(U^{\dagger \otimes n})) - (\Phi_d^{\otimes n})^{\bar{A}\bar{R}_A} \otimes (\Phi_d^{\otimes n})^{\bar{B}\bar{R}_B} \otimes \Phi_{L_n}^{A_1 B_1} \right\| \leq \eta(\epsilon). \tag{9.25}$$

Thus we obtain

$$\sum_k p_k \left\| \tilde{\mathcal{M}}_{n,k}(\Psi_k^p)^{\bar{A}\bar{R}_A A_1 B_1} - (\Phi_d^{\otimes n})^{\bar{A}\bar{R}_A} \otimes \Phi_{L_n}^{A_1 B_1} \right\| \leq \eta(\epsilon), \tag{9.26}$$

which implies that $\tilde{\mathcal{M}}_{n,k}$ is state merging of $|\Psi_k^p\rangle^{A'\bar{R}_A(B_1B_E)}$ with the average error $\eta(\epsilon)$ and the entanglement cost $-\log L_n$. Hence, from Theorem 13 we have

$$\begin{aligned} -\log L_n &\geq \sum_k p_k S(B_1B_E|A')_{\Psi_k^p} - n\eta(\epsilon)\log d \\ &= \sum_k p_k S(\bar{R}_A)_{\Psi_k^p} - \sum_k p_k S(A')_{\Psi_k^p} - n\eta(\epsilon)\log d \\ &\geq n\log d - \sum_k p_k S(A')_{\Psi_k(U^\dagger)} - n\eta(\epsilon)\log d, \end{aligned}$$

which concludes the proof combined with 2.

4. From Theorem 13, we have

$$\begin{aligned} C_n^{\leftarrow} &\geq \sum_k p_k I(B_1 B_E : \bar{R}_A)_{\Psi_k^p} - n\eta(\epsilon)\log d \\ &= \sum_k p_k \left(S(B_1B_E)_{\Psi_k^p} + S(\bar{R}_A)_{\Psi_k^p} - S(B_1B_E\bar{R}_A)_{\Psi_k^p} \right) - n\eta(\epsilon)\log d \\ &\geq \sum_k p_k \left(S(B_1B_E)_{\Psi_k^p} + n\log d - S(A')_{\Psi_k^p} \right) - n\eta(\epsilon)\log d. \qquad (9.27) \end{aligned}$$

From (9.23) and $\Psi_k(U^{\dagger\otimes n})^{\bar{B}B_0} = \Psi_k(U^{\dagger\otimes n})^{\bar{B}} \otimes \Psi_k(U^{\dagger\otimes n})^{B_0}$, we also have

$$\begin{aligned} S(B_1B_E)_{\Psi_k^p} &= S(\bar{B}B_1B_E)_{\Psi_k^p \otimes \Phi_d^{\otimes n}} - S(\bar{B})_{\Phi_d^{\otimes n}} \\ &\geq S(\bar{B}B_1B_E)_{\Psi'_k(U^{\dagger\otimes n})} - n\log d - n\eta(\epsilon_k)\log d \\ &= S(\bar{B}B_0)_{\Psi_k(U^{\dagger\otimes n})} - n\log d - n\eta(\epsilon_k)\log d \\ &= S(\bar{B})_{\Psi_k(U^{\dagger\otimes n})} + S(B_0)_{\Psi_k(U^{\dagger\otimes n})} - n\log d - n\eta(\epsilon_k)\log d \\ &\geq S(B_0)_{\Psi_k(U^{\dagger\otimes n})} - n\eta(\epsilon_k)\log d \end{aligned}$$

and

$$S(A')_{\Psi_k^p} \leq S(A')_{\Psi'_k(U^{\dagger\otimes n})} - n\eta(\epsilon_k)\log d = S(A')_{\Psi_k(U^{\dagger\otimes n})} - n\eta(\epsilon_k)\log d. \qquad (9.28)$$

Thus, from (4.54) in Theorem 25, we obtain

$$\begin{aligned} C_n^{\leftarrow} &\geq \sum_k p_k \left(n\log d - S(A')_{\Psi_k(U^{\dagger\otimes n})} + S(B_0)_{\Psi_k(U^{\dagger\otimes n})} \right) - n\eta(\epsilon)\log d \\ &\geq nM(U^{\dagger}) - n\eta_U(\epsilon)\log d, \qquad (9.29) \end{aligned}$$

which concludes the proof. ■

Acknowledgement

I would like to thank my supervisor, Prof. Mio Murao for helpful and stimulating discussions and correcting early drafts of this thesis. I would like to thank Dr. Akihito Soeda for having insightful discussions and correcting my English. I would like to thank Prof. Hosho Katsura, Prof. Yasuhiko Arakawa, Prof. Akira Shimizu, Prof. Masato Koashi and Prof. Tomohiro Ogawa for refereeing my thesis. I would like to thank all current and former members of Murao group for providing me nice environment for my Ph.D research.

Bibliography

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

[2] C. Bennett and G. Brassard, in Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India, 1984 (IEEE, New York,1984), pp. 175-179

[3] P. Shor, Proc. 35th Ann. Symp. Found. Comp. Sci. (IEEE Comp. Soc. Press, Los Alamitos, California, 1994), p.124.

[4] J. Bell, Physics 1, 195 (1964)

[5] C. Shannon, Bell Syst. Tech. J. 27, 379 (1948)

[6] A. Holevo, Probl. Peredachi Inf. 9, 3 (1973)

[7] B. Schumacher, Phys. Rev. A 51, 2738 (1995)

[8] M. Wilde, *Quantum Information Theory*, Cambridge University Press (2013)

[9] I. Devetak, A. Harrow and A. Winter, IEEE Trans. on Inf. Theory 54, 4587 (2008)

[10] M. Hayashi, *Quantum Information An Introduction*, Springer, (2006)

[11] B. Schumacher, M. Westmoreland, Quant. Inf. Proc. 1, 5 (2002)

[12] A. Uhlmann, Rep. Math. Phys. 9, 273 (1976)

[13] B. Schumacher, Phys. Rev. A 54, 2614 (1996)

[14] E. Lieb and M. Ruskai, J. Math. Phys. 14, 1938 (1973)

[15] B. Groismann, S. Popescu and A. Winter, Phys. Rev. A 72, 032317 (2005)

[16] M. Horodecki, J. Oppenheim and A. Winter, Nature 436, 673 (2005)

[17] M. Horodecki, J. Oppenheim and A. Winter, Comm. Math. Phys. 269, 107 (2007)

[18] M. Koashi and N. Imoto, Phys. Rev. A 66, 022318 (2001)

[19] A. Winter, IEEE Trans. Inf. Theory 45, 2481 (1999)

[20] A. Abeyesinghe, I. Devetak, P. Hayden and A. Winter, Proc. R. Soc. A 465, 2537 (2009)

[21] P. Hayden, R. Jozsa, D. Petz and A. Winter, Comm. in Math. Phys. 246, 359 (2004)

[22] T. Cover and J. Thomas, *Elements of Information Theory*, A Wiley-Interscience publication (2006)

[23] L. Zhang and J. Wu, J. Phys. A: Math. Theor. 47, 415303 (2014)

[24] M. Koashi and N. Imoto, Phys. Rev. Lett. 87, 017902 (2001)

[25] M. Koashi, J. Phys.: Conf. Ser. 95, 012022 (2008)

[26] J. Eisert, K. Jacobs, P. Papadopoulos and M. Plenio, Phys. Rev. A 62, 052317 (2000)

[27] L. Yu, R. Griffiths and S. Cohen, Phys. Rev. A 81, 062315 (2010)

[28] S. Cohen, Phys. Rev. A 81, 062316 (2010)

[29] A. Soeda, P. Turner and M. Murao, Phys. Rev. Lett. 107, 180501 (2011)

[30] S. Stahlke and R. Griffiths, Phys. Rev. A 84, 032316 (2011)

[31] J. Cirac, W. Dur, B. Kraus, M. Lewenstein Phys. Rev. Lett. 86, 544 (2001)

[32] B. Groisman and B. Reznik, Phys. Rev. A 71, 032322 (2005)

[33] L. Chen and Y.-X. Chen, Phys. Rev. A 71, 054302 (2005)

[34] M.-Y. Ye, Y.-S. Zhang and G.-C. Guo, Phys. Rev. A 73, 032337 (2006)

[35] D. Berry, Phys. Rev. A 75, 032349 (2007)

[36] N. Zhao and A. Wang, Phys. Rev. A 78, 014305 (2008)

[37] M. Nielsen, C. Dawson, J. Dodd, A. Gilchrist, D. Mortimer, T. Osborne, M. Harrow and A. Hines, Phys. Rev. A 67, 052301 (2003)

[38] J. Oppenheim and B. Reznik, Phys. Rev. A 70, 022312 (2004)

[39] C. Bennett, A. Harrow, D. Leung and J. Smolin, IEEE Trans. on Inf. Theory 49, 1895 (2003)

[40] C. Bennett, I. Devetak, A. Harrow, P. Shor and A. Winter, IEEE Trans. on Inf. Theory 60, 2926 (2014)

[41] D. Avis, P. Hayden and I. Savov, J. Phys. A: Math. Theor. 41, 115301 (2008)

[42] B. Kraus and I. Cirac, Phys. Rev. A 63, 062309 (2001)