# Semidefinite Optimization for Quantum Information

**Xin Wang**

Centre for Quantum Software and Information

Faculty of Engineering and Information Technology

University of Technology Sydney

A dissertation submitted for the degree of

***Doctor of Philosophy***

Supervisors: Prof. Runyao Duan & Prof. Andreas Winter

July 2018

**Principle supervisor**:

Prof. Runyao Duan

**External supervisor**:

Prof. Andreas Winter

**Thesis examiners**:

Prof. Matthias Christandl

Prof. Barbara Kraus

## CERTIFICATE OF ORIGINAL AUTHORSHIP

I hereby declare that I am the sole author of this thesis. I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Xin Wang

This thesis is dedicated to my mother.

# Acknowledgements

During my PhD study, I enjoyed a lot doing research and always felt excited to learn new things and discover interesting results. However, none of this would have been possible without the support and help of many people.

First and foremost, I must express my sincere gratitude to my supervisor Runyao Duan for his inspiring supervision and support. When I was a third year undergraduate student, I learned many exciting things about quantum information from him via emails and got interested in research. Without him, I probably would not have entered the academic world. I particularly thank him for the energy and passion he always brings to our meetings, for giving me the freedom in exploring different research directions, and for his encouragement on studying important problems regardless of whether they are old fashioned or popular.

I would especially like to thank my external supervisor Andreas Winter for his enlightening supervision. I want to thank him for many discussions on entanglement theory, strong converse, channel capacities, state redistribution, and Lovász number. He really knows everything. I wish I had visited Barcelona for at least one semester. Moreover, I am grateful to my QIP mentors Fernando Brandão, Debbie Leung and Thomas Vidick for their valuable advice on both research and career.

I sincerely thank all my co-authors: Gerardo Adesso, Mario Berta, John Calsamiglia, Runyao Duan, María García Díaz, Kun Fang, Ji Guan, Ludovico Lami, Yinan Li, Shusen Liu, Rosanna Nichols, Youming Qiao, Bartosz Regula, Matteo Rosati, Michalis Skotiniotis, Marco Tomamichel, Andreas Winter, Wei Xie, and Mingsheng Ying. It has been a pleasure to collaborate with these wonderful people. I want to further extend my gratitude to Gerardo Adesso, Mario Berta, Matthias Christandl, Xiongfeng Ma, and Man-Hong Yung for their hospitality and helpful discussions. I profited a lot from these visits.

I am grateful to the Center for Quantum Software and Information at UTS for providing an excellent environment for research, especially the travel opportunities provided for graduate students. I thank all the past and present members of UTS:QSI for creating such a pleasant research atmosphere. Special thanks goes to my friends Kun Fang and Yinan Li, with whom I discussed and collaborated frequently. I particularly thank Yinan for driving me to school during the past two years. I am also grateful to Marco Tomamichel for his supervision on some of my research projects, to Cheng Guo, Ching-Yi Lai for helpful suggestions, and to Michael Bremner, Hao-Chung Cheng, Yuan Feng, Min-Hsiu Hsieh, Zhengfeng Ji, Ryan Mann, Youming Qiao, Peter Rohde, Kun Wang, and Nengkun Yu for interesting discussions.

I would furthermore like to thank Anurag Anshu, Charles H. Bennett, Eric Chitambar, Nilanjana Datta, Omar Fawzi, Li Gao, Mark W. Girard, Gilad Gour, Masahito Hayashi, Richard Josza, Felix Leditzky, Debbie Leung, Ke Li, Ziwen Liu, Laura Mančinska, Alexander Müller-Hermes, Martin Plenio, Peter W. Shor, John Watrous, Mark M. Wilde, Xiaodi Wu, Yunlong Xiao, Dong Yang, and Yuxiang Yang for interesting and stimulating discussions.

I am grateful in advance to my examiners Matthias Christandl and Barbara Kraus for reading my thesis and their feedback. The present version also benefited from the very helpful comments by Mark M. Wilde.

Finally, I thank all my great friends who have made my stay in Sydney full of fun and adventures. Many thanks to my family and friends for their encouragement and support, and I am most grateful to my dear Mom who has always been there to support me. Thank you all.

# Abstract

This thesis aims to improve our understanding of the structure of quantum entanglement and the limits of information processing with quantum systems. It presents new results relevant to three threads of quantum information: the theory of quantum entanglement, the communication capabilities of quantum channels, and the quantum zero-error information theory.

In the first part, we investigate the fundamental features of quantum entanglement and develop quantitative approaches to better exploit the power of entanglement. First, we introduce a computable and additive entanglement measure to quantify the amount of entanglement, which also plays an important role as the improved semidefinite programming (SDP) upper bound of distillable entanglement. Second, we show that the Rains bound is neither additive nor equal to the asymptotic relative entropy of entanglement. Third, we establish SDP lower bounds for the entanglement cost and demonstrate the irreversibility of asymptotic entanglement manipulation under positive-partial-transpose-preserving quantum operations, resolving a major open problem in quantum information theory.

In the second part, we develop a framework of semidefinite programs to evaluate the classical and quantum communication capabilities of quantum channels in both the non-asymptotic and asymptotic regimes. In particular, we establish the first general SDP strong converse bound on the classical capacity of an arbitrary quantum channel and give in particular the best known upper bound on the classical capacity of the amplitude damping channel. We further establish a finite resource analysis of classical communication over quantum erasure channels, including the first second-order expansion of classical capacity beyond entanglement-breaking channels. For quantum communication, we establish the best SDP-computable strong converse bound and refine it as the so-called max-Rains information.

In the third part, we investigate the quantum zero-error information theory. In contrast to the conventional Shannon theory, there is a very different-looking information theory when errors are required to be precisely zero, where the communication problem reduces to the analysis of the so-called confusability graph (non-commutative graph) of a classical channel (quantum channel). We develop an activated communication model and explore its novel properties. Notably, we separate the quantum Lovász number and the entanglement-assisted zero-error capacity, resolving an intriguing open problem in the area of zero-error information.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Information theory, the theory of information processing and transmission, is one of the cornerstones of the last century. In a single paper [Sha48], Shannon initiated the study of information theory as an abstract discipline and led a revolution in communication theory by proving two fundamental theorems, the noiseless and noisy coding theorems. Without information theory, one could not imagine today's highly information-based society, where information and communication have become central to our modern world.

Quantum information theory, a generalization of Shannon information theory, is the theory of the ultimate performance of information processing and transmission with *quantum systems*. On one hand, the information processing and transmission realized by physical systems are ultimately governed by the laws of quantum physics, another great theme of the 20th century. On the other hand, the miraculous features of quantum mechanics led to the revolution of the classical information technologies and further enabled various applications which are currently not feasible on conventional platforms. Quantum entanglement, one of the most fundamental concepts of quantum physics [HHHH09], plays a key role in the advantages gained by considering applications of quantum mechanics. For instance, quantum entanglement can be applied to boost the communication rate as well as to secure the tasks of computation and communication via quantum cryptography [GRTZ02, SBPC$^+$09].

The era of quantum computing also relies on faithful quantum information processors and stable quantum networks. With the aim to construct next generation of networks and computers, the study of quantum information focuses on the capabilities and limitations of computation and information processing in a quantum world. Its main goal is to resolve the following:

- How can quantum information be compressed and manipulated?

- How much classical/quantum/private information can be transmitted faithfully using quantum channels?
- How to detect, quantify, understand, distribute and use entanglement?

There are two ways to explore the above major topics. One is to consider these information processing tasks under the *asymptotic regime* with the simplifying assumption that available resources are unbounded, which reveals the ultimate nature of information processing. Another one is the *non-asymptotic regime*, which is also known as the *finite resource information theory*. This regime is motivated by the realistic thought that the resource is finite. Although the industry and academia have invested a lot to realize the *small-sized* quantum processors, we still have to meet the experimental and theoretical challenges that there are certain limitations to control the large-size quantum systems coherently and accurately. Hence, it is of great practical relevance and theoretical value to study quantum information processing in a scenario involving only a small and medium number of bits or qubits.

In order to investigate quantum information processing under both the asymptotic and non-asymptotic regime, we require new efficient technical tool-kits. Semidefinite optimization (also known as semidefinite programming) [VB96, Tod01, BV04], a relatively new field of optimization with both theoretical and practical interests, has become an ideal and powerful tool-kit for the theory of quantum information. It is concerned with choosing a symmetric matrix to optimize a linear function subject to linear constraints and a further crucial constraint that the matrix has to be positive semidefinite. Its elegant duality theory and its connections to various information measures lead us to a better exploration of quantum information with both analytical and numerical solutions.

This thesis aims to contribute to the development of quantum Shannon theory, entanglement theory, and zero-error information theory, with focuses on the structure of quantum entanglement and the limits of elementary information processing tasks in a quantum world. In the following, I will overview my research in the depicted areas.

## 1.1 Overview

The research during my PhD study explores the fundamental properties of quantum entanglement and establishes efficiently computable approximations for elementary tasks in quantum information theory by using semidefinite optimization [VB96, Tod01, BV04], matrix analysis [HJ12, Bha09], and information measures [OP04, Tom16].

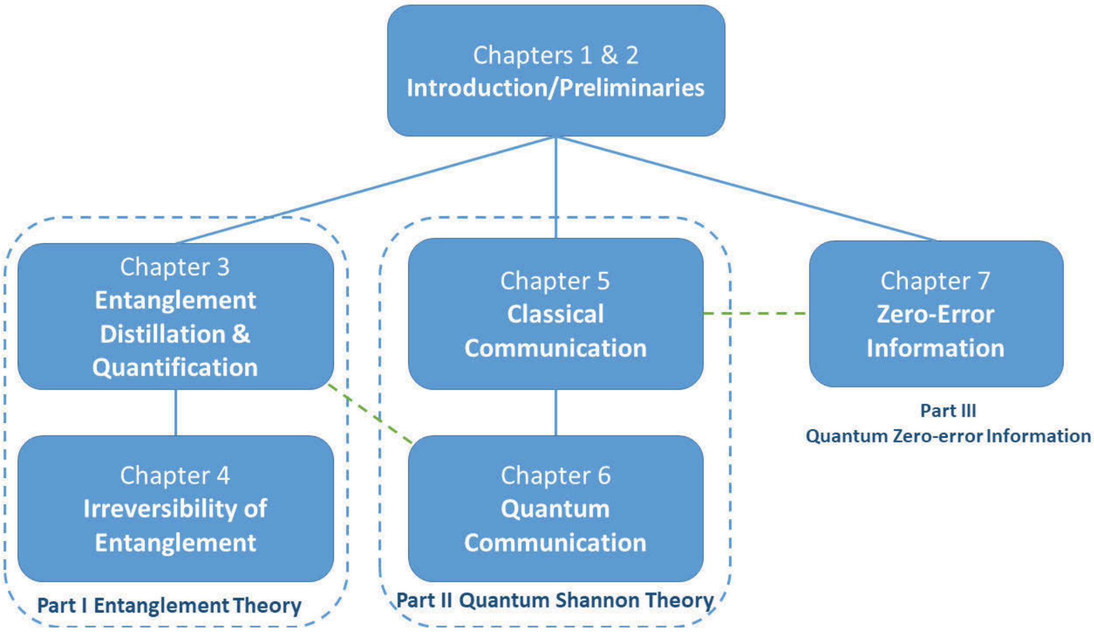


Figure 1.1: Structure of this thesis

After the introduction and preliminaries in the first two chapters, this thesis is divided into three halves: Chapters 3-4 discuss quantum entanglement, Chapters 5 and 6 focus on quantum Shannon theory, and Chapter 7 studies the quantum zero-error information. Here, we give a brief overview of the contents of the individual chapters and briefly discuss our contributions. (We refer to the corresponding chapters for a more extensive introduction and literature on the corresponding topic).

### Chapter 2 - Preliminaries

This chapter introduces the mathematical basics necessary for dealing with quantum information: state vectors, density operators, superoperators, distance measures, and so on. We then give an overview of quantum entanglement and introduce the framework of local and nonlocal bipartite quantum operations. After that, we introduce the basics of semidefinite optimization as well as some other useful toolkits for quantum information such as smoothed entropies.

### Chapter 3 - Entanglement distillation and quantification

Quantum entanglement plays a crucial role in quantum physics and is a key ingredient in many quantum information processing tasks. The concept of entanglement as a resource motivates us to develop a quantitative theory to explore the structure and the power of entanglement. This chapter focuses on the quantification and distillation of quantum entanglement. First, we introduce a computable and additive entanglement measure to quantify the amount of entanglement in quantum states. This entanglement measure also plays an important role as an improved SDP upper bound of the distillable entanglement—the rate at which Bell states can be distilled from the given states through local operations and classical communication (LOCC). Second, we study deterministic entanglement distillation and provide characterizations and estimates of the distillation rates in both the one-shot and asymptotic settings. Third, we show that the Rains bound (the best known upper bound on distillable entanglement) is neither additive nor equal to the asymptotic relative entropy of entanglement.

**Chapter 4 - Irreversibility of asymptotic entanglement manipulation**

The irreversibility is crucial to every resource theory and various approaches have been considered to enlarge the class of free operations to ensure the reversible interconversion of quantum entanglement. A natural candidate is the class of quantum operations that completely preserve positivity of partial transpose (PPT). In this chapter, we demonstrate that PPT operations do not lead to a reversible entanglement theory, resolving a longstanding open problem in quantum information theory [APE03, HOH02, Ple05b]. This means that even if we relax the free operations from LOCC operations to PPT operations, the asymptotic transformation between quantum states is still irreversible. Our key contribution is an efficiently computable lower bound for the entanglement cost, which quantifies the amount of Bell states required to reconstruct a specific state in the asymptotic regime.

**Chapter 5 - Classical communication with quantum systems**

This chapter studies the fundamental limits of classical communication over quantum channels in both the asymptotic and non-asymptotic regime. First, we contribute a framework of semidefinite programs (SDPs) to estimate the coding rate and success probability for classical communication over quantum channels, with or without entanglement assistance. Second, we establish the first general SDP upper bound on the classical capacity of a quantum channel and give the best known upper bound for the classical capacity of the amplitude damping channel. Third, we introduce the constant-bounded subchannels and use them to derive a meta-converse on the amount of information that can be transmitted over a single use of a quantum channel. In particular, we establish a finite resource analysis of quantum erasure channels, including the first second-order expansion of classical capacity beyond entanglement-breaking channels.

### Chapter 6 - Quantum communication with quantum systems

This chapter investigates the capabilities of a noisy quantum channel to transmit quantum information in both the non-asymptotic and asymptotic regime. First, we provide improved SDP converse bounds in the context of quantum communication with finite resources. Second, we establish an SDP strong converse bound on the quantum capacity, which means the fidelity of any sequence of codes with a rate exceeding this bound will vanish exponentially fast as the number of channel uses increases. Third, we refine our SDP strong converse bound as the so-called *max-Rains information* and show that it improves the *partial transposition bound* given by Holevo and Werner [HW01]. We further compare it with other well-known bounds on quantum capacity.

### Chapter 7 - Quantum zero-error information theory

This chapter studies the zero-error communication via quantum channels from the perspective of non-commutative graphs. The celebrated Lovász number [Lov79] and its quantum generalization [DSW13] were proved to be upper bounds on the zero-error capacity even assisted by entanglement. However, it remained unknown whether the quantum Lovász number is always achievable via the assistance of quantum entanglement [LMM$^+$12, DSW13, CLMW11]. The first main result of this chapter resolves this intriguing open problem by separating the quantum Lovász number and the entanglement-assisted zero-error capacity via an explicit construction of the non-commutative graph. After that, we further introduce an activated communication model and discuss its properties.

During the time of my PhD study at UTS, I had the pleasure to collaborate with many excellent researchers. Parts of this thesis are based on material contained in the following papers.

- **X. Wang** and R. Duan, *Improved semidefinite programming upper bound on distillable entanglement*, Physical Review A 94, 050301 (Rapid communication), 2016, [WD16b].
(Chapter 3)

- **X. Wang** and R. Duan, *Nonadditivity of Rains bound for distillable entanglement*, Physical Review A 95, 062322, 2017, [WD17b].
(Chapter 3)

- **X. Wang** and R. Duan, *Irreversibility of Asymptotic Entanglement Manipulation Under Quantum Operations Completely Preserving Positivity of Partial Transpose*, Physical Review Letters 119, 180506, 2017, [WD17a].
(Chapter 4)

- **X. Wang**, W. Xie, and R. Duan, *Semidefinite programming strong converse bounds for classical capacity*, IEEE Transactions on Information Theory 64(1), 640-653, 2018, [WXD18].
(Chapter 5)

- **X. Wang**, K. Fang, and M. Tomamichel, *On converse bounds for classical communication over quantum channels*, submitted, available at arXiv:1709.05258, 2017, [WFT17].
(Chapter 5)

- **X. Wang**, K. Fang, and R. Duan, *Semidefinite programming converse bounds for quantum communication*, submitted, available at arXiv:1709.00200, 2017, [WFD17].
(Chapter 6)

- **X. Wang** and R. Duan, *A semidefinite programming upper bound of quantum capacity*, In Proceedings of IEEE International Symposium on Information Theory (ISIT 2016), pages 1690-1694, 2016, [WD16a].
(Chapter 6)

- **X. Wang** and R. Duan, *Separation between quantum Lovász number and entanglement-assisted zero-error classical capacity*, IEEE Transactions on Information Theory 64(3), 1454-1460, 2016, [WD18].
(Chapter 7)

- R. Duan and **X. Wang**, *Activated zero-error classical capacity of quantum channels in the presence of quantum no-signalling correlations*, arXiv:1510.05437, 2015, [DW15].
(Chapter 7)

Other publications or preprints on which this manuscript does not focus:

- **X. Wang** and Runyao Duan, *On the quantum no-signalling assisted zero-error simulation cost of non-commutative bipartite graphs*, In Proceedings of the IEEE International Symposium on Information Theory, pages 2244-2248, 2016, [WD16c].

- K. Fang, **X. Wang**, M. Tomamichel, and R. Duan, *Non-asymptotic entanglement distillation*, arXiv:1706.06221, 2017, [FWTD17].

- K. Fang, **X. Wang**, L. Lami, B. Regula, and G. Adesso, *Probabilistic coherence distillation*, Physical Review Letters (in press), 2018, [FWL$^+$18].

- Y. Li, Y. Qiao, **X. Wang**, and R. Duan, *Tripartite-to-bipartite Entanglement Transformation by Stochastic Local Operations and Classical Communication and the Classification of Matrix Spaces*, Communications in Mathematical Physics 358(2), 791–814, 2018, [LQWD18].

- Y. Li, **X. Wang**, and R. Duan, *Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario*, Physical Review A 95, 052346, 2017, [LWD17].

- B. Regula, K. Fang, **X. Wang**, and G. Adesso, *One-shot coherence distillation*, Physical Review Letters 121, 010401, 2018, [RFWA18].

- W. Xie, K. Fang, **X. Wang**, and R. Duan, *Approximate broadcasting of quantum correlations*, Physical Review A 96, 022302, 2017, [XFWD17].
- L. Lami, B. Regula, **X. Wang**, R. Nichols, A. Winter, and G. Adesso, *Gaussian quantum resource theories*, arXiv:1801.05450, 2018, [LRW$^+$18].
- K. Fang, **X. Wang**, M. Tomamichel, and M. Berta, *Quantum Channel Simulation and the Channel's Max-Information*, In Proceedings of the IEEE International Symposium on Information Theory, 2018, [FWTB18].
- W. Xie, **X. Wang**, and R. Duan, *Converse bounds for classical communication over quantum networks*, arXiv:1712.05637, 2017, [XWD17].
- S. Liu, **X. Wang**, L. Zhou, J. Guan, Y. Li, Y. He, R. Duan, and M. Ying, *$Q|SI\rangle$: a quantum programming environment*, arXiv:1710.09500, 2017, [LWZ$^+$17].
- M. G. Díaz, K. Fang, **X. Wang**, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, *Using and reusing coherence to realize quantum processes*, arXiv:1805.04045, 2018, [DFW$^+$18].

# Chapter 2

# Preliminaries

## 2.1 Basics of linear algebra

A Hilbert space $\mathcal{H}$ is a complex vector space equipped with an inner product $(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$. We use symbols such as $\mathcal{H}_A$ (or $\mathcal{H}_{A'}$) and $\mathcal{H}_B$ (or $\mathcal{H}_{B'}$) to denote Hilbert spaces associated with Alice and Bob, respectively. In this thesis we restrict ourselves to finite-dimensional Hilbert spaces. We denote $\mathcal{L}(A)$ as the set of linear operators acting on Hilbert space $\mathcal{H}_A$. We denote $\mathcal{P}(A)$ as the subset of positive semidefinite operators acting on $\mathcal{H}_A$ and write $X \geq 0$ if $X \in \mathcal{P}(A)$.

Given two quantum systems $A$ and $B$, we consider them jointly by defining the composite quantum system $AB$. Its Hilbert space is the tensor product of the Hilbert spaces of its parts, i.e., $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Note that for a linear operator $M$, we define $|M| = \sqrt{M^\dagger M}$, and the trace norm of $M$ is given by $\|M\|_1 = \operatorname{Tr}|M|$, where $M^\dagger$ is the conjugate transpose of $M$. The operator norm $\|M\|_\infty$ is defined as the maximum eigenvalue of $|M|$. Trace norm and operator norm are dual to each other, in the sense that $\|M\|_\infty = \max_{\|C\|_1 \leq 1} \operatorname{Tr} MC$.

An overview of the basic notations in this thesis can be found in Table 2.1. The expert reader may directly proceed to Chapter 3.

## 2.2 The formalism of quantum information

Here we present the essential formalism of quantum information. We start by briefly recalling the necessary concepts from linear algebra and then introduce the basic elements of quantum information.

| **General** | |
|---|---|
| $\mathbb{C}$, $\mathbb{R}$, $\mathbb{N}$ | complex, real, and natural numbers |
| log | logarithm with base 2 |
| $\langle\cdot\|$, $\|\cdot\rangle$ | bra and ket |
| $d_A$ | dimension of the Hilbert space $A$ |
| $\mathbb{1}_A$, $\mathrm{id}_A$ | identity operator and identity map on $A$ |
| Tr, $\mathrm{Tr}_A$ | trace and partial trace |
| $\|S\|$ | cardinality of the set $S$ |

| **Operators** | |
|---|---|
| $\mathcal{L}(A)$ | set of bounded linear operators acting on $\mathcal{H}_A$ |
| $\mathcal{P}(A)$ | set of positive semidefinite operators acting on $\mathcal{H}_A$ |
| $\mathcal{S}(A)$ | set of density operators acting on $\mathcal{H}_A$ |
| $\mathcal{S}(A \otimes B)$ | set of density operators acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ |
| $\mathcal{S}_{\leq}(A)$ | set of subnormalized density operators acting on $\mathcal{H}_A$ |
| $\mathrm{supp}(X)$ | support of the operator $X$ |
| $\mathrm{rank}(X)$ | rank of the operator $X$ |
| $X \ll Y$ | support of $X$ is contained in the support of $Y$ |
| $X^T$ | transpose of the operator $X$ |
| $X^\dagger$ | conjugate transpose of the operator $X$ |
| $X_{AB}^{T_B}$ | Partial transpose on system $B$ of $X_{AB}$ |
| $X_A \otimes Y_B$ | tensor product of operators $A$ and $B$ |
| $X_A \oplus X_B$ | direct sum of operators $A$ and $B$ |
| $\lambda_{\max}(X)$ | largest eigenvalue of a Hermitian operator $X$ |

| **Norms** | |
|---|---|
| $\\|X\\|_1$ | trace norm of $X \in \mathrm{Herm}(A)$ |
| $\\|X\\|_\infty$ | spectral norm of $X \in \mathrm{Herm}(A)$ |
| $\\|\mathcal{E}\\|_\diamond$ | diamond norm of $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$ |

Table 2.1: Overview of notational conventions

### 2.2.1 Quantum states

A quantum state on $\mathcal{H}_A$ is an operator $\rho_A \in \mathcal{P}(A)$ with $\mathrm{Tr}\,\rho_A = 1$. The set of quantum states on $\mathcal{H}_A$ is denoted by $\mathcal{S}(A)$. The set of subnormalized states on $\mathcal{H}_A$ is denoted by $\mathcal{S}_{\leq}(A) := \{\rho_A \in \mathcal{P}(A) : \mathrm{Tr}\,\rho_A \leq 1\}$. A state is called pure if it is a projector, i.e. $\rho = |\psi\rangle\langle\psi|$ for a vector $|\psi\rangle$. If a state $\rho$ is not pure, we call it mixed.

### 2.2.2 Quantum channels and measurements

In this subsection, we briefly introduce the unitary evolution, quantum channels, and quantum measurements.

### Unitary evolution

The evolution of any closed quantum system is described by a unitary evolution U that maps

$$\rho \to U\rho U^{\dagger}, \tag{2.1}$$

where $U^{\dagger}U = \mathbb{1}$.

### Quantum channels

The dynamical evolution of an open quantum system with Hilbert space $\mathcal{H}$ is given by a quantum channel $\mathcal{N}$, which is defined to be a linear completely positive (CP) and trace-preserving (TP) map from $\mathcal{L}(A')$ to $\mathcal{L}(B)$. We also call $\mathcal{N}$ quantum channel. The class of physical mappings should at least always map positive operators to positive operators. The complete positivity of a map ensures that this remains true if the quantum system is regarded as a part of a larger system.

There are several equivalent representations of a quantum channel:

1. **Choi-Kraus representation** [Kra71, Cho75]: A linear map $\mathcal{N}$ from $\mathcal{L}(A')$ to $\mathcal{L}(B)$ is CP if and only if there exists a set of linear operators $\{E_k\}$ from $\mathcal{H}_{A'}$ to $\mathcal{H}_B$ such that

$$\mathcal{N}(\rho) = \sum_k E_k \rho E_k^{\dagger},\ \forall \rho \in \mathcal{S}\left(A'\right), \tag{2.2}$$

   where $E_k$ is also referred to as a Kraus operator. Furthermore, $\mathcal{N}$ is TP if and only if

$$\sum_k E_k^{\dagger} E_k = \mathbb{1}. \tag{2.3}$$

2. **Stinespring Representation** [Sti55]: A linear map $\mathcal{N}$ from $\mathcal{L}(A')$ to $\mathcal{L}(B)$ is CPTP if and only if there exists a Hilbert space $\mathcal{H}_E$ and an isometry $V$ such that

$$\mathcal{N}(\rho_{A'}) = \operatorname{Tr}_E V \rho_{A'} V^{\dagger},\ \forall \rho_{A'} \in \mathcal{S}\left(A'\right). \tag{2.4}$$

   Such $V$ is called a Stinespring dilation of $\mathcal{N}$.

3. **Choi-Jamiołkowski representation** [Cho75, Jam72]: For a linear map $\mathcal{N}$ from

$\mathcal{L}(A')$ to $\mathcal{L}(B)$, its Choi-Jamiołkowski matrix is given by

$$J_{\mathcal{N}} = \sum_{ij} |i_A\rangle\langle j_A| \otimes \mathcal{N}\left(|i_{A'}\rangle\langle j_{A'}|\right), \tag{2.5}$$

where $\{|i_A\rangle\}$ and $\{|i_{A'}\rangle\}$ are orthonormal bases for isomorphic Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_{A'}$, respectively. The map $\mathcal{N}$ is CP if and only if

$$J_{\mathcal{N}} \geq 0, \tag{2.6}$$

and $\mathcal{N}$ is TP if and only if

$$\operatorname{Tr}_B J_{\mathcal{N}} = \mathbb{1}_A. \tag{2.7}$$

This Choi representation allows us to represent a quantum channel by a positive semidefinite operator obeying certain linear constraints.

### Inverse Choi-Jamiołkowski transformation

For a given quantum channel $\mathcal{N}_{A\to B}$ and input state $\rho_A$, we have

$$\mathcal{N}_{A\to B}(\rho_A) = \operatorname{Tr}_A J_{\mathcal{N}}\left(\rho_A^T \otimes \mathbb{1}_B\right) \tag{2.8}$$

$$= \operatorname{Tr}_A J_{\mathcal{N}}^{T_A}\left(\rho_A \otimes \mathbb{1}_B\right), \tag{2.9}$$

where $T_A$ means the partial transpose on system $A$, i.e., $\left(|i_A j_B\rangle\langle k_A l_B|\right)^{T_A} = |k_A j_B\rangle\langle i_A l_B|$, and $\{|i_A\rangle\}$, $\{|j_B\rangle\}$ are orthonormal bases for Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

### Measurements

To realize the advantages quantum technology promises, we actually have to understand how to extract classical information from quantum states. Such a process is called *quantum measurement*.

A *quantum measurement* is a CPTP map from a quantum system to a classical register containing the measurement outcome and a system with the state after measurement. It can be described by a collection of Choi-Kraus operators $\{E_j\}_{j=1}^n$, where the indices $j \in \{1, ..., n\}$ indicate the outcomes of the states. If the system is initially prepared in the state $\rho \in \mathcal{S}(A)$, outcome $j$ will be observed with probability $p_j = \operatorname{Tr} E_j^\dagger E_j \rho$ and the resulting state is $\rho_j = \frac{1}{p_j} E_j \rho E_j^\dagger$. The concise CPTP map of the measurement here is given by

$$\mathcal{M}(\rho) = \sum_{j=1}^{n} p_j |j\rangle\langle j| \otimes \rho_j. \tag{2.10}$$

A *generalized quantum measurement* is defined in terms of a positive operator-valued measure (POVM). A POVM is a family of positive semidefinite matrices $\{M_j\}_{j=1}^{n}$ such that $\sum_{j=1}^{n} M_j = \mathbb{1}$. The probability of getting outcome $j$ is $\operatorname{Tr} M_j \rho$. A POVM fully characterizes the probability distribution the measurement induces on the classical register. The POVM is very useful when we are only interested in the classical outcomes.

### 2.2.3 Bipartite quantum states

#### Entangled states

The set of quantum states on $\mathcal{H}_A \otimes \mathcal{H}_B$ is denoted by $\mathcal{S}(A \otimes B)$. We call a bipartite quantum state *separable* if it can be written as convex combination of tensor product states. The set of separable states on system $A \otimes B$ is denoted as $\mathrm{SEP}(A : B)$. If $\rho \notin \mathrm{SEP}(A : B)$, $\rho$ is called *entangled*.

The most important entangled state is arguably the *Bell state*

$$|\Phi\rangle = \frac{1}{2}(|00\rangle + |11\rangle), \tag{2.11}$$

which is deemed to be the currency of quantum information processing. As its generalization, we denote

$$\Phi(d) = \frac{1}{d} \sum_{i,j=0}^{d-1} |i_A i_B\rangle\langle j_A j_B| \tag{2.12}$$

as the *maximally entangled state* on Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $d$ is the dimension of $\mathcal{H}_A$ and $\mathcal{H}_B$, $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ are the standard, orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. Moreover, the identity operator on Hilbert space $\mathcal{H}_A$ is denoted as $\mathbb{1}_A = \sum_{i=0}^{d-1} |i_A\rangle\langle i_A|$.

#### Positive partial transpose (PPT)

A positive semidefinite operator $E_{AB} \in \mathcal{P}(A \otimes B)$ is said to be PPT if $E_{AB}^{T_B} \geq 0$, where $T_B$ means the partial transpose on system $B$. The set of all PPT states on system $A \otimes B$ is denoted as

$$\mathrm{PPT}(A : B) := \left\{\rho \in \mathcal{S}(A \otimes B) : \rho^{T_B} \geq 0\right\}. \tag{2.13}$$

One of the most useful methods for detecting entanglement is the positive partial transpose, or Peres-Horodecki, criterion [HHH96, Per96]:

$$\mathrm{SEP}\,(A:B) \subsetneq \mathrm{PPT}\,(A:B)\,. \tag{2.14}$$

In addition, the Rains set [ADVW02], a superset of PPT $(A:B)$, is defined as

$$\mathrm{PPT}'\,(A:B) := \left\{ M \in \mathcal{P}\,(A\otimes B) : \left\| M^{T_B} \right\|_1 \leq 1 \right\}. \tag{2.15}$$

## 2.3 Bipartite quantum operations

In this section, we introduce the hierarchy of local and non-local quantum bipartite operations. The charaterizations of different classes of quantum bipartite operations are also summarized.

### 2.3.1 Local operations and classical communication

**Local operations (Unassisted code)**

For two distant quantum systems held by Alice and Bob, a bipartite operation is called a local operation (LO) if it corresponds to the product of separate operations implemented by Alice and Bob, i.e., $\Pi = \mathcal{D}_{B\to B'} \otimes \mathcal{E}_{A\to A'}$. We also call such bipartite operation the *unassisted code* (UA).

**Local operations and classical communication (LOCC)**

When a quantum system is distributed to spatially separated parties, it is natural to consider how the system evolves when the parties perform local quantum operations with classical communication.

If one-way classical communication is allowed from Alice to Bob (or Bob to Alice), the corresponding bipartite operation is called *1-LOCC*. A 1-LOCC operation ($A \to B$) $\Lambda$ can be mathematically described by

$$\Lambda\,(\rho_{AB}) = \sum_{i,j} \left(E_{A,i} \otimes F_{B,i,j}\right) \rho_{AB} \left(E_{A,i} \otimes F_{B,i,j}\right)^{\dagger}, \tag{2.16}$$

where $\sum_i E^{\dagger}_{A,i} E_{A,i} = \mathbb{1}$ and $\sum_j F^{\dagger}_{B,i,j} F_{B,i,j} = \mathbb{1}$ for all $i$.

Or, equivalently,

$$\Lambda = \sum_j \mathcal{E}^j_{A\to A'} \otimes \mathcal{F}^j_{B\to B'}, \tag{2.17}$$

where $\{\mathcal{E}^j_{A\to A'}\}_j$ is a set of CP maps such that $\sum_j \mathcal{E}^j_{A\to A'}$ is trace preserving, and $\{\mathcal{F}^j_{B\to B'}\}_j$ is a set of CPTP maps.

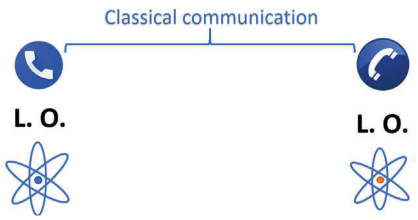


Figure 2.1: Local operations and classical communication

If both parties are allowed to communicate with each other with unlimited rounds, the corresponding bipartite operation is called *LOCC*. A LOCC operation can be decomposed into sequences of 1-LOCC operations and the round of communication can be finite or infinite. The mathematical structure of the LOCC operation is complicated and more details can be found in [CLM$^+$14].

### Separable operations

Considering that the structure of LOCC operations is exceedingly complex, leaving many important physical problems unsolved, the sets of separable and PPT operations were introduced to explore the fundamental limits of the resource theory of entanglement. A bipartite quantum operation $\Pi_{AB\to A'B'}$ is said to be a *SEP operation* if its Choi-Jamiołkowski matrix

$$J_\Pi = \sum_{i,j,m,k} |i_A j_B\rangle\langle m_A k_B| \otimes \Pi\left(|i_A j_B\rangle\langle m_A k_B|\right) \tag{2.18}$$

is separable under the partition of $AA' : BB'$, where $\{|i_A\rangle\}$ and $\{|j_B\rangle\}$ are orthonormal bases for Hilbert spaces $A$ and $B$, respectively. Separable operations were first studied in [Rai97, VP98] and the distillation of entanglement using separable operations was studied in [Rai97].

### PPT operations

A bipartite quantum operation $\Pi_{AB\to A'B'}$ is said to be a *PPT operation* if its Choi-Jamiołkowski matrix $J_\Pi$ is positive under partial transpose under the partition of $AA' : BB'$. The entanglement theory under PPT operations was first studied in [Rai99, Rai01]. A well-known fact is that the classes of PPT, separable (SEP) and

LOCC operations obey the following strict inclusions [HHHH09]:

$$\text{1-LOCC} \subsetneq \text{LOCC} \subsetneq \text{SEP} \subsetneq \text{PPT}. \tag{2.19}$$

The most intriguing is the non-equivalence LOCC $\neq$ SEP which follows from the non-locality without entanglement [BDF+99].

**Quantum supermap (or superchannel)**

A bipartite quantum channel $\Pi_{AB\to A'B'}$ is called a superchannel (or supermap) if it maps all quantum channels to quantum channels. Quantum superchannels describe all possible transformations between elementary quantum objects. Interesting, the mathematical structure of quantum superchannels is closely related to semi-causal quantum operations [CDP08]: a CPTP map $\Pi_{AB\to A'B'}$ is a superchannel if and only if $\Pi$ is no-signalling from $B$ to $A$ (see Section 2.3.3 for more details).

**PPT codes**

If a PPT operation $\Pi_{AB\to A'B'}$ is also a superchannel, such $\Pi_{AB\to A'B'}$ is called *PPT code* since it can seem as a general code to simulate a new physical channel from $\mathcal{N}_{A'\to B}$. A PPT operation $\Pi_{AB\to A'B'}$ is a PPT code if and only if it is also B to A no-signalling (cf. Eq. (2.20)). We note the PPT codes [LM15] could be applied to study the communication capability of a quantum channel (see e.g., Part. II).

### 2.3.2 Non-local operations

**Local operation with shared entanglement (Entanglement-assisted code)**

A *local operation with shared entanglement* corresponds to a bipartite operation of the form $\Pi = \mathcal{D}_{B\hat{B}\to B'}\mathcal{E}_{A\hat{A}\to A'}\Psi_{\hat{A}\hat{B}}$, where $\Psi_{\hat{A}\hat{B}}$ can be any entangled state shared between Alice and Bob. We also call a local operation with shared entanglement *an entanglement-assisted code*. See Figure 2.2 for details.

**No-signalling operations (codes)**

Generally speaking, a bipartite quantum operation is no-signalling (NS) if it cannot be used by spatially separated parties to violate relativistic causality. In more specific language, a bipartite operation $\Pi_{AB\to A'B'}$ is non-signalling from Bob to Alice if the marginal state of Alice's output is given by some fixed operation applied to the

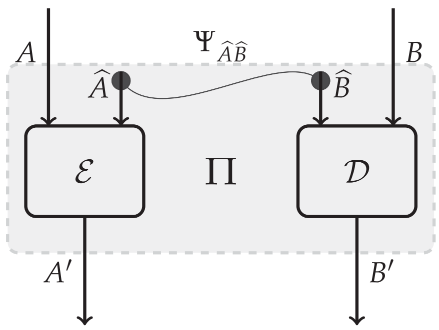


Figure 2.2: A bipartite operation $\Pi_{AB\to A'B'}$ is an entanglement-assisted code (or local operation with shared entanglement) if can be implemented by a shared entangled state $\Psi_{\hat{A}\hat{B}}$ and local operations $\mathcal{E}_{A\hat{A}\to A}$ and $\mathcal{D}_{B\hat{B}\to B}$.

marginal state of Alice's input. Its equivalent condition is

$$\mathrm{Tr}_{B'} J_\Pi = \mathrm{Tr}_{BB'} J_\Pi \otimes \frac{\mathbb{1}_B}{d_B}, \tag{2.20}$$

where $J_\Pi$ is the Choi-Jamiołkowski matrix of $J_\Pi$. Smilarily, $\Pi_{AB\to A'B'}$ is non-signalling from Alice to Bob if

$$\mathrm{Tr}_{A'} J_\Pi = \mathrm{Tr}_{AA'} J_\Pi \otimes \frac{\mathbb{1}_A}{d_A}, \tag{2.21}$$

Furthermore, $\Pi_{AB\to A'B'}$ is a no-signalling operation if it is no-signalling from Alice to Bob and vice versa. We also call a bipartite no-signalling operation *a non-signalling code*. It is worth noting that the set of NS-assisted codes includes all the operations that can be implemented via local operations and shared entanglement.

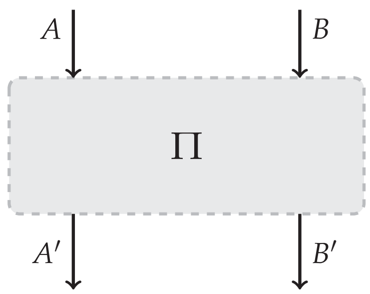


Figure 2.3: A bipartite operation $\Pi\,(A_iB_i \to A_oB_o)$ is a no-signalling operations (or NS-assisted) code if Alice and Bob cannot use $\Pi$ to communicate (or equivalently, the Choi-Jamiołkowski matrix of $\Pi$ satisfies the above Eqs. (2.20), (2.21)).

### Relationship between different classes of bipartite operations

In the following Figure 2.3.2, we briefly summarize the relationship between the different classes of bipartite operations we introduced in this section.

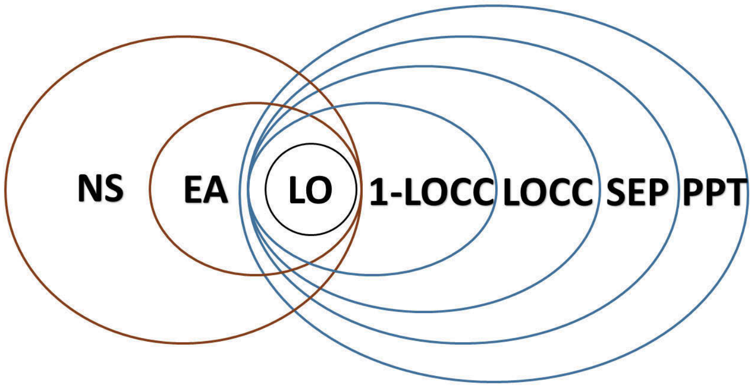


Figure 2.4: Hierarchy of quantum bipartite operations

### NS∩PPT operations (codes)

From the Figure 2.3.2, one can see that the set of local operations is a subset of the set of bipartite operations that are NS and PPT. We can use these NS∩PPT operations to simplify the behavior of local operations since both NS and PPT operations have mathematically tractable structures.

In the following two tables, we summarize the mathematical characterizations of the three main kinds of codes we will study in Part II of this thesis.

| Constraint | Mathematical charaterization |
|---|---|
| CP | $J_\Pi \geq 0$ |
| TP | $\operatorname{Tr}_{A_oB_o} J_\Pi = \mathbb{1}_{A_iB_i}$ |
| $A \not\to B$ | $\operatorname{Tr}_{A_o} J_\Pi = \mathbb{1}_{A_i}/d_{A_i} \otimes \operatorname{Tr}_{A_oA_i} J_\Pi$ |
| $B \not\to A$ | $\operatorname{Tr}_{B_o} J_\Pi = \mathbb{1}_{B_i}/d_{B_i} \otimes \operatorname{Tr}_{B_oB_i} J_\Pi$ |
| PPT | $J_\Pi^{T_{B_iB_o}} \geq 0$ |

Table 2.2: Mathematical charaterizations of the constraints of bipartite operations

### 2.3.3 Channel composition

**Definition 2.1.** A CPTP map $\Pi : \mathcal{L}(A_i \otimes B_i) \to \mathcal{L}(A_o \otimes B_o)$ is called a superchannel if it sends all CPTP map $\mathcal{N} : \mathcal{L}(A_o) \to \mathcal{L}(B_i)$ to another CPTP map $\mathcal{M} : \mathcal{L}(A_i) \to$

| Codes | Corresponding constraints |
|---|---|
| NS operations (codes) | CP, TP, $A \not\to B$, $B \not\to A$ |
| NS∩PPT operations (codes) | CP, TP, $A \not\to B$, $B \not\to A$, PPT |
| PPT operations | CP, TP, PPT |
| PPT codes | CP, TP, PPT, $B \not\to A$ |

Table 2.3: Different kinds of bipartite operations and codes

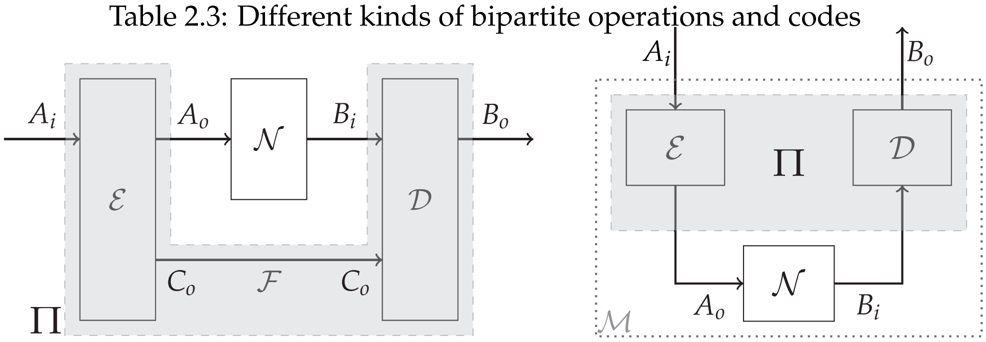


Figure 2.5: Simulation of a channel $\mathcal{M}\left(A_i \to B_o\right)$ from a channel $\mathcal{N}\left(A_o \to B_i\right)$ and a deterministic super-operator (general code) $\Pi\left(A_iB_i \to A_oB_o\right)$.

$\mathcal{L}\left(B_o\right)$. We also call such $\Pi$ a *general code* throughout this thesis.

The following proposition guarantees that if $\Pi$ is B to A no-signalling, then the composition of a bipartite quantum operation $\Pi : \mathcal{L}\left(A_i \otimes B_i\right) \to \mathcal{L}\left(A_o \otimes B_o\right)$ and any quantum channel $\mathcal{N} : \mathcal{L}\left(A_o\right) \to \mathcal{L}\left(B_i\right)$ is physical.

**Lemma 2.2** ([CDP08])**.** *A bipartite quantum operation $\Pi : \mathcal{L}\left(A_i \otimes B_i\right) \to \mathcal{L}\left(A_o \otimes B_o\right)$ is a deterministic supermap if and only if $\Pi$ is* B *to* A *no-signalling. (See an alternative proof and more related discussions in [DW16].)*

Now, let $\mathcal{M}\left(A_i \to B_o\right)$ denote the resulting composition channel of the deterministic bipartite quantum operation $\Pi_{A_iB_i\to A_oB_o}$ and the quantum channel $\mathcal{N}_{A_o\to B_i}$. We write $\mathcal{M} = \Pi \circ \mathcal{N}$ for simplicity. An interesting fact is that we can characterize the effective channel $\mathcal{M}$ via the Choi-Jamiołkowski matrices of $\mathcal{N}$ and $\Pi$, in the similar spirit of the above inverse Choi-Jamiołkowski transformation.

As $\Pi$ is a deterministic super-operator, there exist quantum channels $\mathcal{E}_{A_i\to A_oC_i}$ and $\mathcal{D}_{B_iC\to B_o}$ and $\mathcal{F}_{C_i\to C_o}$ such that [CDP08]

$$\mathcal{M}_{A_i\to B_o} = \mathcal{D}_{B_iC_o\to B_o} \circ \mathcal{F}_{C_i\to C_o} \circ \mathcal{N}_{A_o\to B_i} \circ \mathcal{E}_{A_i\to A_oC_i}. \tag{2.22}$$

And the bipartite operation is given by

$$\Pi_{A_iB_i\to A_oB_o} = \mathcal{D}_{B_iC_o\to B_o} \circ \mathcal{F}_{C_i\to C_o} \circ \mathcal{E}_{A_i\to A_oC_i}. \tag{2.23}$$

Based on this, we can apply the inverse Choi-Jamiołkowski transformation to get the following lemma.

**Lemma 2.3.** *[LM15] Given a deterministic super-operator $\Pi_{A_iB_i\to A_oB_o}$ and a quantum channel $\mathcal{N}_{A_o\to B_i}$, the effective channel $\mathcal{M}_{A_i\to B_o}$ composed via $\Pi\circ\mathcal{N}$ is characterized by*

$$J_{\mathcal{M}} = \operatorname{Tr}_{A_oB_i}\left(J_{\mathcal{N}}^T \otimes \mathbb{1}_{A_iB_o}\right) J_\Pi. \tag{2.24}$$

We give a proof sketch here. One could first use inverse Choi-Jamiołkowski transformation and Eq. (2.23) to show

$$J_\Pi = \operatorname{Tr}_{C_iC_o}\left(J_{\mathcal{D}}^{T_{C_o}} \otimes \mathbb{1}_{A_iA_oC_i}\right)\left(J_{\mathcal{F}}^{T_{C_i}} \otimes \mathbb{1}_{A_iA_oB_iB_o}\right)\left(J_{\mathcal{E}} \otimes \mathbb{1}_{B_iB_oC_o}\right). \tag{2.25}$$

Furthermore, one could use similar steps to get

$$\begin{aligned} \mathcal{M}\left(\rho_{A_i}\right) &= \mathcal{D}_{B_iC_o\to B_o} \circ \mathcal{F}_{C_i\to C_o} \circ \mathcal{N}_{A_o\to B_i} \circ \mathcal{E}_{A_i\to A_oC_i}\left(\rho_{A_i}\right) & (2.26)\\ &= \operatorname{Tr}_{A_i}\left(\operatorname{Tr}_{A_oB_i}\left(J_{\mathcal{N}}^T \otimes \mathbb{1}_{A_iB_o}\right) J_\Pi\right)\left(\rho_{A_i}^T \otimes \mathbb{1}_{B_o}\right), & (2.27) \end{aligned}$$

which means that the Choi-Jamiołkowski matrix of $\mathcal{M}$ is given by

$$J_{\mathcal{M}} = \operatorname{Tr}_{A_oB_i}\left(J_{\mathcal{N}}^T \otimes \mathbb{1}_{A_iB_o}\right) J_\Pi. \tag{2.28}$$

## 2.4 Semidefinite optimization

### 2.4.1 Basics of semidefinite programming

Semidefinite programming is a relatively new subfield of convex optimization concerned with the optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space (see, e.g., [WSV00, Tod01, LV16, BV04] for more details). Though the related research on semidefinite programming has been studied as far back as the 1940s [Boh48], the interest has grown vastly during the last twenty years. In the last decades, semidefinite programs (SDPs) have become an important tool for engineering, combinatorial optimization, complexity theory, and information theory (see e.g., [Lov79, GLS93, GW95]).

In the study of quantum information, the convexity and the semidefinite properties arise naturally. As a result, many useful tools from convex optimization can be used to deepen our understanding of quantum information. In the following, we briefly introduce the basics of semidefinite programming. This subsection is based on John Watrous' book [Wat18] and we restrict the definitions to positive operators.

**Definition 2.4.** A semidefinite program (SDP) is defined by a triplet $\{\Psi, C, D\}$, where $C \geq 0$ and $D \geq 0$ and $\Psi$ is a CP map.

| Primal problem | | Dual problem | |
|---|---|---|---|
| maximize: | $\mathrm{Tr}\, CX$ | minimize: | $\mathrm{Tr}\, DY$ |
| subject to: | $\Psi(X) \leq D,$ | subject to: | $\Psi^*(Y) \geq C,$ |
| | $X \geq 0.$ | | $Y \geq 0.$ |

where $\Psi^*$ is the dual map to $\Psi$ ($\mathrm{Tr}\, Y\Psi(X) = \mathrm{Tr}\, X\Psi^*(Y)$).

Either problem is called feasible if there exists a valid variable satisfying the corresponding constraint. If there exists a $X \geq 0$ such that $D - \Psi(X)$ is positive definite, then the primal problem is said to be strictly feasible. And the dual is strictly feasible if there is a $Y \geq 0$ such that $\Psi^*(Y) - C$ is positive definite.

For these two problems, we define their optimal attained values

$$\begin{aligned} \alpha &= \sup\{\mathrm{Tr}\,(CX) : \Psi(X) \leq D, X \geq 0\}, \\ \beta &= \inf\{\mathrm{Tr}\,(DY) : \Psi^*(Y) \geq C, Y \geq 0\}, \end{aligned} \tag{2.29}$$

where $\alpha = -\infty$ if the primal problem is not feasible and $\beta = +\infty$ if the dual problem is not feasible.

### 2.4.2 Duality of semidefinite programming

The duality between primal and dual problems is one of the most important properties of semidefinite programming.

#### Weak duality

For any semidefinite program, it holds that $\alpha \leq \beta$. This convenient relation allows us to immediately bound the optimal attained values of the primal problem by picking a valid variable of the dual problem, and vice versa.

#### Strong duality

For any semidefinite program that satisfies the following Slater's conditions, we have

$$\alpha = \beta. \tag{2.30}$$

This strong duality is remarkable as it allows us to determine the optimal attained values of many SDPs by picking valid variables of the prime and dual problems.

**Theorem 2.5** (Slater's conditions)**.** *For a semidefinite program $\{\Psi, C, D\}$ and $\alpha, \beta$ defined as in Eq.* (2.29)*, the following holds:*

- *if the primal problem is feasible and the dual is strictly feasible, then strong duality holds and there exists a valid choice $X$ for the dual problem with $\alpha = \operatorname{Tr} CX$;*
- *if the dual problem is feasible and the primal is strictly feasible, then strong duality holds and there exists a valid choice $Y$ for the dual problem with $\beta = \operatorname{Tr} DY$;*
- *if both problems are strictly feasible, then strong duality holds and there exist valid choices of $X, Y$ such that $\alpha = \beta = \operatorname{Tr} CX = \operatorname{Tr} DY$.*

Finally, there are many optimization problems that are not immediately represented by SDP but can be refined in that form. Examples include the fidelity between two states, the trace distance, the infinity norm, as well as most of the smooth entropies. We take the spectral norm as an example here. Let us take $\Psi(\cdot) = \operatorname{Tr}(\cdot), C = \rho, D = 1$, then

$$\begin{aligned} \|\rho\|_\infty &= \max\{\operatorname{Tr} \rho X : \operatorname{Tr} X \leq 1, X \geq 0\} \\ &= \min\{y : \rho \leq y\mathbb{1}\}. \end{aligned}$$

**Minimax theorem**

A minimax theorem is a theorem providing conditions that guarantee that the exchange between the minimization and maximization of a minimax problem will not change the optimal value. The first theorem in this sense is von Neumann's minimax theorem [vN28], which is considered the starting point of game theory.

The following Sion's minimax theorem [Sio58] is a generalization of John von Neumann's minimax theorem.

**Lemma 2.6** (Sion's minimax theorem [Sio58])**.** *Let $\mathcal{X}, \mathcal{Y}$ be convex compact sets and $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a continuous function that satisfies the following properties: $f(\cdot, y) : \mathcal{X} \to \mathbb{R}$ is convex for fixed $y$, and $f(x, \cdot) : \mathcal{Y} \to \mathbb{R}$ is concave for fixed $x$. Then it holds that [Sio58]*

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x, y) = \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} f(x, y). \tag{2.31}$$

### 2.4.3 Applications of semidefinite programming in quantum information

In the following, we briefly review the applications of semidefinite programming in quantum information and computation.

- Hierarchies for nonlocal correlations (see e.g., [NPA07, NPA08]).

- Quantum query complexity (see e.g., [Rei11, HLS07, LMR+11]).
- Quantum communication complexity (see e.g., [GKRdW09]).
- Quantum computational complexity (see e.g., [KW00, JJUW11])
- Quantum steering (see e.g., [CS17, KSC+15, PW15]).
- Quantum coin-flipping (see e.g., [ABDR03, NST15]).
- Quantum state discrimination (see e.g., [Eld03, YDY14]).
- Quantum program language (see e.g., [LY17, YYW17]).

We note that semidefinite programming also has applications in quantum error correction (e.g., [KSL08, FSW07]), nonlocal games (e.g., [CSUU08, Weh06, KRT10]) and many other topics in the area of quantum information.

## 2.5 Symmetries

In this subsection, we first briefly introduce the basics about complex representations of finite and compact groups, and then introduce the useful Schur's lemma. A group homomorphism from group $G$ to group $H$ is a map $\phi : G \to H$ such that $\phi(gg') = \phi(g)\phi(g')$ for all $g, g' \in G$. A representation of the group $G$ is a group homomorphism $\phi : G \to GL(V)$, where $V = \mathbb{C}^n$. Two representations $\phi_1 : G \to GL(V_1)$ and $\phi_2 : G \to GL(V_2)$ are said to be equivalent if there exists an isomorphism $M : V_1 \to V_2$ such that $\phi_1(g) M = M\phi_2(g)$ for all $g \in G$. Such $M$ is called an intertwiner (or intertwining operator). It turns out that for finite groups every representation is equivalent to a unitary representation. A representation $\phi : G \to GL(V)$ is called reducible, if there exists a decomposition $V = V_1 \oplus V_2$ such that $\phi(g) = \phi_1(g) \oplus \phi_2(g)$ for all $g \in G$, and otherwise it is irreducible. A useful fact is that every representation of a finite group can be expressed as a direct sum of irreducible representations. A detailed introduction of representation theory can be found in [FH04, Ste12].

### Schur's lemma

Schur's lemma [Sch05] is an elementary but useful statement in the representation theory, which has many applications in quantum information theory (see e.g., [Hay17b]). It shows that homomorphisms between irreducible representations of a group $G$ have a very simple structure.

**Lemma 2.7** (Schur's Lemma)**.** *Let $V_1$ and $V_2$ be two irreducible representations of a group $G$. If $M : V_1 \to V_2$ is an intertwiner operator, then the following hold.*

*i) Either $M = 0$ or $M$ is an isomorphism.*

*ii) If $V = W$ (as representations), then $M = \lambda \mathbb{1}$ for some $\lambda \in \mathbb{C}$.*

The introduction of other powerful tools such as Schur-Weyl duality from representation theory can be found in [Chr06, Har05, Hay17a].

#### Covariant channel

**Definition 2.8.** Let $G$ be a finite group, and for every $g \in G$, let $g \to U_A(g)$ and $g \to V_B(g)$ be unitary representations acting on the input and output spaces of the channel, respectively. Then a quantum channel $\mathcal{N}_{A\to B}$ is G-covariant if

$$\mathcal{N}_{A\to B}\left(U_A(g)\,\rho_A U_A^\dagger(g)\right) = V_B(g)\,\mathcal{N}_{A\to B}(\rho_A)\,V_B^\dagger(g)$$

for all $\rho_A \in \mathcal{S}(A)$. We also introduce the average state $\overline{\rho}_A = \frac{1}{|G|}\sum_g U_A(g)\,\rho_A U_A^\dagger(g)$.

## 2.6 Distance measures

On one hand, a fundamental question in quantum information theory is to distinguish different quantum states (or operations). A natural intuition is that if two states $\rho$ and $\sigma$ are too close, it will be difficult to distinguish them. Thus, we need distance measures to quantify the distinguishability. On the other hand, we are interested in optimizing the quantum information-processing protocols to simulate an ideal one. One way to quantify the efficiency is to show that the output state $\rho$ of the actual protocol is very close to the output state $\sigma$ of the ideal protocol. Therefore, we need distance measures to quantify how well the actual quantum protocol works.

### 2.6.1 Distance between states

In this subsection, we introduce two basic distance measures to quantify the closeness between two quantum states.

#### Trace distance

Given two states $\rho, \sigma \in \mathcal{S}(A)$, the trace distance between $\rho$ and $\sigma$ is given by

$$\|\rho - \sigma\|_1 = \max\{\operatorname{Tr} X(\rho - \sigma) : -\mathbb{1} \le X \le \mathbb{1}\}. \tag{2.32}$$

where $\|\cdot\|_1$ is the trace norm. This distance measure is operational in the sense that it quantifies the probability of distinguishing two states with an optimal measurement.

$$\frac{1}{2}\|\rho-\sigma\|_1 = \max_{0\leq M\leq \mathbb{1}} \operatorname{Tr} M\left(\rho-\sigma\right). \tag{2.33}$$

To see this, suppose that the spectral decomposition of $\rho-\sigma$ is as follows:

$$\rho-\sigma = \sum_i \lambda_i |i\rangle\langle i|, \tag{2.34}$$

where $\{|i\rangle\}$ is an orthonormal basis of eigenvectors and $\{\lambda_i\}$ is a set of real eigenvalues. Let us further choose $P_+ = \sum_{\lambda_i\geq 0}\lambda_i|i\rangle\langle i|$ and $P_- = \sum_{\lambda_i<0}\lambda_i|i\rangle\langle i|$. From Eq. (2.32), one can see that $\|\rho-\sigma\|_1 = \operatorname{Tr}\left(P_+ - P_-\right) = 2\operatorname{Tr} P_+$. Furthermore,

$$\frac{1}{2}\|\rho-\sigma\|_1 = \operatorname{Tr} P_+ = \max_{0\leq M\leq \mathbb{1}} \operatorname{Tr} M\left(\rho-\sigma\right). \tag{2.35}$$

A useful fact is that a measurement with one outcome that is likely causes a little disturbance (measured by trace distance) to the quantum state that we measure. Winter [Win99] originally proved the following "gentle measurement" lemma and later Ogawa and Nagaoka [ON07] subsequently improved this bound to $2\sqrt{\varepsilon}$.

**Lemma 2.9** (Gentle measurement)**.** *For a quantum state $\rho$ and an operator $0\leq X\leq \mathbb{1}$ satisfying that $1-\operatorname{Tr}\rho X\leq\varepsilon\leq 1$, it holds that*

$$\|\rho-\sqrt{X}\rho\sqrt{X}\|_1 \leq 2\sqrt{\varepsilon}. \tag{2.36}$$

### Fidelity

Another useful distance measure is the fidelity [Bur69, Uhl76]. For two states $\rho$ and $\sigma$, the fidelity between them is defined as

$$F\left(\rho,\sigma\right) = \|\sqrt{\rho}\sqrt{\sigma}\|_1. \tag{2.37}$$

A useful fact is that the fidelity between $\rho$ and $\sigma$ can be computed via semidefinite programming [Wat13]:

$$F\left(\rho,\sigma\right) = \sup\left\{\frac{1}{2}\operatorname{Tr}\left(X+X^\dagger\right) : \begin{bmatrix}\rho & X\\ X^\dagger & \sigma\end{bmatrix}\geq 0\right\} \tag{2.38}$$

$$= \inf\left\{\frac{1}{2}\operatorname{Tr}\left(\rho Y+\sigma Z\right) : \begin{bmatrix}Y & -\mathbb{1}\\ -\mathbb{1} & Z\end{bmatrix}\geq 0\right\}. \tag{2.39}$$

#### Purified distance and the $\varepsilon$-ball of a quantum state

The purified distance [TCR10] between two subnormalized states is defined as

$$P(\rho,\sigma) = C(\rho \oplus [1-\operatorname{Tr}\rho], \sigma \oplus [1-\operatorname{Tr}\sigma]), \tag{2.40}$$

where $C(\rho,\sigma) := \sqrt{1-F(\rho,\sigma)}$ [Ras06, GLN05, Ras02, Ras03]. The purified distance have nice properties and is very useful when it is applied to define the smooth min- and max-entropies.

**Definition 2.10.** The $\varepsilon$-ball of a state $\rho$ defined as

$$\mathcal{B}_\varepsilon(\rho) = \{\widetilde{\rho} \in \mathcal{S}_{\leq}(A) : P(\rho,\widetilde{\rho}) \leq \varepsilon\}. \tag{2.41}$$

#### Relations between Trace Distance and Fidelity

It is naturally to think that the trace distance should be small if the fidelity is high because the trace distance vanishes when the fidelity is one and vice versa. The next lemma explains the above intuition by establishing important relationships between the trace distance and fidelity.

**Lemma 2.11.** *([FvdG99]) Given two quantum states $\rho$ and $\sigma$, it holds that*

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

**Lemma 2.12** (Uhlmann's Theorem [Uhl76])**.** *Let $\rho_A, \sigma_A \in \mathcal{S}(A)$. Let $\rho_{AB} \in \mathcal{S}(A \otimes B)$ be a purification of $\rho_A$ and $\sigma_{AC} \in \mathcal{S}(A \otimes C)$ be a purification of $\sigma_A$. There exists an isometry $V : C \to B$ such that,*

$$F(|\tau\rangle\langle\tau|_{AB}, |\rho\rangle\langle\rho|_{AB}) = F(\rho_A, \sigma_A),$$

*where $|\tau\rangle_{AB} = (\mathbb{1}_A \otimes V)|\sigma\rangle_{AC}$.*

### 2.6.2 Distance between channels

For quantum channels, we use the completely bounded (cb) norm (or the diamond norm) to measure the bias in distinguishing two such mappings [Kit97, Pau02].

**Definition 2.13.** For a linear map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$, the diamond norm of $\varepsilon$ is defined as

$$\|\mathcal{E}\|_\diamond = \sup_{k \in N} \|\mathcal{E} \otimes id_k\|_1, \tag{2.43}$$

where $id_k$ denotes the identity map on states of a $k$-dimensional quantum system, and $\|\mathcal{N}\|_1 = \sup_\sigma \|\mathcal{N}(\sigma)\|_1$ with $\sigma \in S_{\leq}(A)$.

The supremum in Definition 2.13 is reached for $k = d_A$ [Kit97, Pau02]. We call two quantum channels $\varepsilon$-close if they are $\varepsilon$-close in the metric induced by the diamond norm.

The diamond norm is known to be efficiently computable by SDP in [Wat13]. To be specific, for a linear map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that

$$\begin{aligned} \|\mathcal{E}\|_\diamond &= \max \frac{1}{2} \operatorname{Tr}(J_\mathcal{E} X) + \frac{1}{2} \operatorname{Tr}\left(J_\mathcal{E} X^\dagger\right) \\ &\quad \text{s.t.} \begin{pmatrix} \rho_0 \otimes \mathbb{1}_B & X \\ X^\dagger & \rho_1 \otimes \mathbb{1}_B \end{pmatrix} \geq 0 \\ &= \min \frac{1}{2} \| \operatorname{Tr}_B Y_0 \|_\infty + \frac{1}{2} \| \operatorname{Tr}_B Y_1 \|_\infty \\ &\quad \text{s.t.} \begin{pmatrix} Y_0 & -J_\mathcal{E} \\ -J_\mathcal{E} & Y_1 \end{pmatrix} \geq 0,\ Y_0, Y_1 \geq 0. \end{aligned} \tag{2.44}$$

As a special case, for two given quantum channels $\mathcal{N}_1, \mathcal{N}_2 : \mathcal{L}(A) \to \mathcal{L}(B)$, the diamond norm of their difference is given by

$$\begin{aligned} \|\mathcal{N}_1 - \mathcal{N}_2\|_\diamond &= \max\{\operatorname{Tr}(J_{\mathcal{N}_1} - J_{\mathcal{N}_2}) X : \ \rho_A \in \mathcal{S}(A),\ 0 \leq X \leq \rho_A \otimes \mathbb{1}_B\} \\ &= \min\{t : \ \operatorname{Tr}_B Y \leq t\mathbb{1}_A,\ Y \geq J_{\mathcal{N}_1} - J_{\mathcal{N}_2},\ Y \geq 0\}. \end{aligned} \tag{2.45}$$

## 2.7 Entropies

### 2.7.1 Entropy of a single system

A fundamental concept in classical and quantum information theory is entropy. The Shannon entropy [Sha48] has played an important role in information theory in the *independent and identically distributed (i.i.d.) limit*: the asymptotic limit in which an average of the resource is counted over many independent repetitions. The *Shannon entropy* of a probability distribution $p(x)$ of a classical system $X$ is defined as

$$H(X) = -\log \sum_x p(x) \log p(x). \tag{2.46}$$

For the quantum information theory in the i.i.d. limit, the *von Neumann entropy* is the most important measure. It is defined as the Shannon entropy of the spectrum of a quantum state, or equivalently,

$$S(\rho) = -\operatorname{Tr} \rho \log \rho. \tag{2.47}$$

It is worth noting that the von Neumann entropy has the property of continuity, which is guaranteed by the following Fannes inequality [Fan73].

**Lemma 2.14** (Fannes inequality [Fan73]). *Given two quantum states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_A)$, such that $d_A = d$ and $\|\rho_1 - \rho_2\|_1 = \varepsilon \leq \mathrm{e}^{-1}$, it holds that*

$$|S(\rho_1) - S(\rho_2)| \leq \varepsilon \log(d) - \varepsilon \log \varepsilon. \tag{2.48}$$

*A sharp version of this Fannes inequality was introduced in [Aud07]:*

$$|S(\rho_1) - S(\rho_2)| \leq \frac{\varepsilon}{2} \log(d-1) + H_2\left(\frac{\varepsilon}{2}\right), \tag{2.49}$$

*where $H_2(x) = -x \log x - (1-x)\log(1-x)$ is the binary entropy.*

### 2.7.2 Relative entropies

In order to describe the relative amount of uncertainty a state contains with respect to another state, the relative entropy was introduced.

**Definition 2.15.** For $\rho \in \mathcal{S}(A)$ and $\sigma \in \mathcal{P}(A)$ the *relative entropy* between $\rho$ and $\sigma$ is defined as

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}\, \rho(\log \rho - \log \sigma) & \text{if } \rho \ll \sigma \\ +\infty & \text{otherwise}. \end{cases} \tag{2.50}$$

And the quantum information variance is defined by

$$V(\rho\|\sigma) := \mathrm{Tr}\, \rho(\log \rho - \log \sigma)^2 - D(\rho\|\sigma)^2. \tag{2.51}$$

The quantum relative entropy has a flavor of distance measure, as it is nonnegative and $D(\rho\|\sigma) = 0$ if and only if $\rho = \sigma$. It has the monotonicity under quantum channels [Lin75]: for quantum states $\rho$, $\sigma \in \mathcal{S}(A)$ and any quantum channel $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that

$$D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D(\rho\|\sigma). \tag{2.52}$$

This is also known as the *data processing inequality* of the relative entropy, which states that processing of information cannot increase the relative entropy. It is worthing noting that the quantum relative entropy is not a metric on the set of quantum states since it is not symmetric under the exchange of its arguments. But the quantum relative entropy can be related to the trace distance in the following way:

**Lemma 2.16** (Pinsker's inequality [OP04])**.** *. For quantum states $\rho$ and $\sigma$, it holds that*

$$D(\rho\|\sigma) \geq \frac{1}{2\ln 2}\|\rho-\sigma\|_1. \tag{2.53}$$

### Min- and max-relative entropies

Moreover, Rényi introduced a family of entropies as a generalization of the Shannon entropy [Rén60] and there are various generalizations of the Rényi entropies. In the Rényi entropy framework (see e.g., [Tom16]), two very useful information measures are the *min-relative entropy* and the *max-relative entropy* [Dat09]:

$$D_{\min}(\rho\|\sigma) = -\log\|\sqrt{\rho}\sqrt{\sigma}\|_1^2 \tag{2.54}$$

$$D_{\max}(\rho\|\sigma) = \min\{\lambda : \rho \leq 2^{\lambda}\sigma\}. \tag{2.55}$$

These two relative entropies both have interesting operational significances and obey the data processing inequality under quantum channels:

$$D_{\max}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_{\max}(\rho\|\sigma) \tag{2.56}$$

$$D_{\min}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \leq D_{\min}(\rho\|\sigma). \tag{2.57}$$

In particular, for bipartite states $\rho_{AB}, \sigma_{AB} \in \mathcal{S}(A\otimes B)$, it holds that

$$D_{\max}(\rho_A\|\sigma_A) \leq D_{\max}(\rho_{AB}\|\sigma_{AB})$$
$$D_{\min}(\rho_A\|\sigma_A) \leq D_{\min}(\rho_{AB}\|\sigma_{AB}).$$

### Sandwiched Rényi relative entropy

A more general type of relative entropy with important applications in quantum information theory is the sandwiched Rényi relative entropy.

**Definition 2.17.** For any $\rho \in \mathcal{S}(A)$, $\sigma \in \mathcal{P}(A)$ and $\alpha \in (0,1)\cup(1,\infty)$, the sandwiched Rényi relative entropy is defined as [MLDS+13, WWY14],

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\operatorname{Tr}\left(\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^{\alpha}\right), \tag{2.58}$$

if $\operatorname{supp}(\rho) \subset \operatorname{supp}(\sigma)$ and it is equal to $+\infty$ otherwise.

### Conditional entropies

In classical information theory, the conditional entropy $H(Y|X)$ quantifies the amount of information needed to describe the outcome of a random variable $Y$ given that the value of another random variable $X$ is known. The conditional quantum entropy is a generalization of the classical conditional entropy:

**Definition 2.18.** The *conditional entropy* of a state $\rho_{AB}$ is defined by

$$H(A|B)_\rho := S(\rho_{AB}) - S(\rho_B) = -D(\rho_{AB}\|\mathbb{1}_A \otimes \rho_B), \tag{2.59}$$

where $\rho_B$ is the reduced state $\rho_B = \mathrm{Tr}_A\, \rho_{AB}$.

Note that $H(A|B)_\rho$ can be negative for some bipartite state $\rho_{AB}$ [CA97]. In the operational task of state merging, the conditional entropy quantifies the optimal entanglement cost when there is free classical communication [HOW05].

**Definition 2.19.** The *coherent information* [SN96] of a bipartite state $\rho_{AB}$ is defined by

$$I(A\rangle B)_\rho := -H(A|B)_\rho = S(\rho_B) - S(\rho_{AB}). \tag{2.60}$$

In entanglement theory, the widely used quantum Rényi entropies are the conditional min- and max-entropies.

**Definition 2.20.** The conditional min-entropy [KRS09, Tom12] of a bipartite state $\rho_{AB} \in \mathcal{S}(A \otimes B)$ is defined by

$$H_{\min}(A|B)_\rho := -\inf_{\sigma_B \in \mathcal{D}(B)} D_{\max}(\rho_{AB}\|\mathbb{1}_A \otimes \sigma_B). \tag{2.61}$$

The conditional max-entropy is defined as the dual of the conditional min-entropy in the sense that

$$H_{\max}(A|B)_\rho = -H_{\min}(A|C)_\rho, \tag{2.62}$$

where $\rho_{ABC}$ is a purification of $\rho_{AB}$.

### 2.7.3 Smoothed entropies

The smooth entropy framework [Ren05, Tom12] has many applications in quantum information theory [Wil17, Hay17c, Wat18] and quantum resource theories [CG18]. For example, in the single instance regime, the smoothed max-relative entropy characterizes the resource costs of many information-theoretic tasks (see, e.g., [BD11b, ZLY$^+$18, FWTB18]) while the smoothed min-relative entropy characterizes the amount of resource that can be generated in many information-theoretic tasks (see, e.g., [Hay17d, WR12, TBR16, BD11a, FWTD17, RFWA18]).

To be specific, the smoothed the min-relative entropy and the max-relative entropy are introduced as follows:

$$D_{\min}^{\varepsilon}(\rho\|\sigma) = \max_{\hat{\rho}\approx_{\varepsilon}\rho} D_{\min}(\widehat{\rho}\|\sigma), \tag{2.63}$$

$$D_{\max}^{\varepsilon}(\rho\|\sigma) = \min_{\hat{\rho}\approx_{\varepsilon}\rho} D_{\max}(\hat{\rho}\|\sigma), \tag{2.64}$$

where $\hat{\rho} \approx_{\varepsilon} \rho$ is equivalent to $\hat{\rho} \in \mathcal{B}_{\varepsilon}(\rho)$.

Another important smoothed quantity is the hypothesis testing relative entropy [WR12, BD10]:

$$\begin{aligned} D_H^{\varepsilon}(\rho_0\|\rho_1) &:= -\log \beta_{\varepsilon}(\rho_0\|\rho_1) & (2.65)\\ &= -\log\min\{\operatorname{Tr} Q\rho_1 : 1 - \operatorname{Tr} Q\rho_0 \le \varepsilon, 0 \le Q \le \mathbb{1}\}, & (2.66)\end{aligned}$$

where $\beta_{\varepsilon}(\rho_0\|\rho_1)$ is the minimum type-II error for the test while the type-I error is no greater than $\varepsilon$. The dual SDP of $D_H^{\varepsilon}(\rho_0\|\rho_1)$ is given by

$$-\log\max\{-\operatorname{Tr} X + (1-\varepsilon)\,t : X + \rho_1 - t\rho_0 \ge 0, X, t \ge 0\}. \tag{2.67}$$

Note that $\beta_{\varepsilon}$ is a fundamental quantity in quantum theory [Hel76, HP91, NO00]. It is worth noting that $D_H^{\varepsilon}(\cdot\|\cdot)$ interpolates between smoothed min- and max-relative entropies [DKF$^+$12].

The three smoothed relative entropy measures presented above all satisfy the data processing inequality. Furthermore, they also obey the *asymptotic equipartition property* (AEP) in the i.i.d. limit:

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n} D_H^{\varepsilon}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = D(\rho\|\sigma), \tag{2.68}$$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n} D_{\max}^{\varepsilon}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = D(\rho\|\sigma). \tag{2.69}$$

Moreover, there are second-order expansion of quantum hypothesis testing relative entropy and max-relative entropy [TH13, Li14]:

$$D_H^{\varepsilon}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = nD(\rho\|\sigma) + \sqrt{nV(\rho\|\sigma)}\Phi^{-1}(\varepsilon) + O(\log n), \tag{2.70}$$

$$D_{\max}^{\varepsilon}\left(\rho^{\otimes n}\|\sigma^{\otimes n}\right) = nD(\rho\|\sigma) - \sqrt{nV(\rho\|\sigma)}\Phi^{-1}\left(\varepsilon^2\right) + O(\log n), \tag{2.71}$$

where $\Phi(x) = \int_{-\infty}^{x}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt$ is the cumulative distribution function of a standard normal random variable.

# Part I

# Entanglement Theory

# Chapter 3

# Entanglement distillation and quantification

## 3.1 Introduction

### 3.1.1 Background

In 1935, Einstein, Podolsky and Rosen (EPR) and Schrödinger first recognized a spooky feature of quantum mechanics [EPR35, Sch35]: the existence of global states of a composite system which cannot be written as a product of the states of individual subsystems. This phenomenon, known as *entanglement*, was originally called "Verschränkung" by Schrödinger [Sch35]. The EPR paradox argued that quantum mechanics as a physical theory is incomplete. In 1964, Bell dealt directly with the EPR thought experiment and showed that entanglement is incompatible with a certain local classical inequality which can be verified experimentally [Bel64].

With the development of quantum information science, quantum entanglement has been recognized as an essential resource for quantum computation and communication. The study of quantum entanglement is one of the most active and important areas in quantum information theory. A series of remarkable efforts have been devoted to this area (for reviews see, e.g., Refs. [PV07, HHHH09]).

#### Entanglement distillation

The maximally entangled state plays a role as the currency in quantum information since it has become a key ingredient in many quantum information processing tasks (e.g., teleportation [BBC$^+$93], superdense coding [BW92], and quantum cryptography [BB84, Eke91]). Then a natural question arises: *how many maximally entangled states can we obtain from a source of less entangled states using physically-motivated operations?*

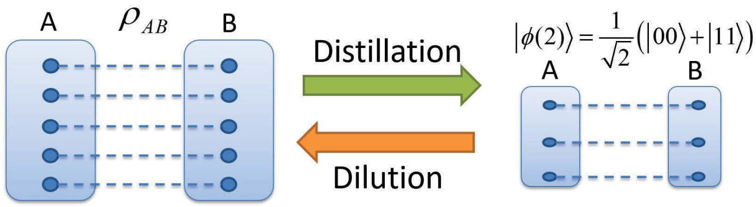


Figure 3.1: Entanglement distillation and formation

Imagine that Alice and Bob share a large supply of identically prepared states, and they want to convert these states to high fidelity Bell pairs. One ideal strategy is to use local operations and classical communication. We further define the *distillable entanglement* $E_D$ of $\rho$ to be the optimal rate $r$ of converting $\rho^{\otimes n}$ to $rn$ Bell pairs with an arbitrarily high fidelity in the limit of large $n$ by LOCC. The concise definition of entanglement of distillation by LOCC is given in as follows [PV07]:

$$E_D(\rho_{AB}) = \sup\{r : \lim_{n\to\infty}[\inf_{\Lambda\in LOCC} \|\Lambda(\rho_{AB}^{\otimes n}) - \Phi(2^{rn})\|_1] = 0\}, \tag{3.1}$$

where $\Lambda$ ranges over LOCC operations and $\Phi(d) = 1/d\sum_{i,j=1}^{d}|ii\rangle\langle jj|$ represents the standard $d\otimes d$ maximally entangled state. This can also be generalized to define the $\Omega$-assisted distillable entanglement $E_{D,\Omega}$ by replacing LOCC with $\Omega$ operations ($\Omega \in \{\text{1-LOCC}, \text{SEP}, \text{PPT}\}$):

$$E_{D,\Omega}(\rho_{AB}) = \sup\{r : \lim_{n\to\infty}[\inf_{\Lambda\in\Omega} \|\Lambda(\rho_{AB}^{\otimes n}) - \Phi(2^{rn})\|_1] = 0\}, \tag{3.2}$$

Entanglement distillation is also essential for quantum cryptography and quantum error correction. For given bipartite pure state $\psi_{AB}$ [BBPS96], it is known that

$$E_D(\psi_{AB}) = S(\mathrm{Tr}_A\,\psi_{AB}). \tag{3.3}$$

But for general quantum states, how to evaluate this fundamental quantity remains a formidable question.

#### Entanglement formation

The reverse task of entanglement distillation is called entanglement dilution. At this time, Alice and Bob share a large supply of Bell pairs and they are to convert $rn$ Bell pairs to $n$ high fidelity copies of the desired state $\rho^{\otimes n}$ using suitable operations. The *entanglement cost* $E_{C,\Omega}$ of a given bipartite state $\rho$ quantifies the optimal rate $r$ of converting $rn$ Bell pairs to $\rho^{\otimes n}$ with an arbitrarily high fidelity in the limit of large $n$.

The concise definition of entanglement cost using $\Omega$ operations is given as follows:

$$E_{C,\Omega}\left(\rho_{AB}\right) = \inf\{r : \lim_{n\to\infty} \inf_{\Lambda\in\Omega} \|\rho_{AB}^{\otimes n} - \Lambda\left(\Phi\left(2^{rn}\right)\right)\|_1 = 0\}, \tag{3.4}$$

where $\Omega \in \{\text{1-LOCC}, \text{LOCC}, \text{SEP}, \text{PPT}\}$ and we write $E_{C,\text{LOCC}} = E_C$ for simplification. For entanglement cost under LOCC operations, Hayden, Horodecki and Terhal [HHT01] proved that $E_C$ equals to the regularized entanglement of formation [BDSW96]:

$$E_C\left(\rho_{AB}\right) = \lim_{k\to\infty} \frac{E_F\left(\rho_{AB}^{\otimes k}\right)}{k}, \tag{3.5}$$

where

$$E_F\left(\rho_{AB}\right) := \inf\left\{\sum_i p_i S\left(\operatorname{Tr}_A |\psi_i\rangle\langle\psi_i|\right) : \rho_{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|\right\}. \tag{3.6}$$

In particular, for any bipartite pure state $\psi_{AB}$ [BBPS96], it is known that

$$E_C\left(\psi_{AB}\right) = E_D\left(\psi_{AB}\right) = S\left(\operatorname{Tr}_A \psi_{AB}\right), \tag{3.7}$$

from which we can see the reversibility between the asymptotic transformation between any pure states. However, little is known about the entanglement cost of general quantum states. More details and properties of the entanglement cost as well as the irreversibility of general quantum states will be discussed in the Chapter 4.

**Entanglement monotone**

As entanglement is a key resource, it is well motivated to develop quantifiers to measure it. In the past two decades, many entanglement measures have been proposed and studied [PV07, HHHH09]. To be a function to quantify entanglement, *entanglement monotone* is one of the most essential features. Motivated by the fact that it is not possible to create entanglement via LOCC, therefore the first property for an entanglement measure $E$ is that $E$ should be monotonically decreasing under LOCC operations.

There are different kinds of monotonicity considered in the literature. The simplest one states that $E$ should be monotonic under LOCC operations; i.e.,

$$E\left(\rho_{AB}\right) \geq E\left(\Lambda\left(\rho_{AB}\right)\right)$$

should hold for every state $\rho$ and every deterministic LOCC operation $\Lambda$. This is arguably the most important requirement for an entanglement measure. A direct consequence of this inequality is the invariance of $E$ under local unitaries.

There is another form of monotonicity which is known as *full monotonicity*. Any (non-negative) function $E(\cdot)$ over bipartite states is said to be a *full entanglement monotone* if it does not increase on average under general LOCC operations [Ple05a], i.e.,

$$E(\rho) \geq \sum_i p_i E(\rho_i), \tag{3.8}$$

where the state $\rho_i$ is obtained with probability $p_i$ in the LOCC protocol applied to $\rho$.

### 3.1.2 Outline

In this chapter, we focus on the study of different aspects of quantum entanglement and develop quantitative approaches to better exploit the power of entanglement. In section 3.2, we review and discuss the model of entanglement distillation under PPT operations introduced in [Rai01]. In section 3.3, we introduce a new computable and additive entanglement measure to quantify the amount of entanglement in the quantum states. Meanwhile, this entanglement measure also plays an important role as an improved semidefinite programming (SDP) upper bound of the distillable entanglement—the rate at which gold standard ebit states can be produced from the given states through local operations and classical communication. In section 3.4, we study deterministic entanglement distillation and provide characterizations and estimates of the distillation rates. In section 3.5, we show that the Rains bound (the best known upper bound on distillable entanglement) is neither additive nor equal to the asymptotic relative entropy of entanglement.

## 3.2 Distillation under PPT operations

Rains first studied entanglement distillation assisted with PPT operations and obtained an upper bound on the distillable entanglement [Rai99, Rai99, Rai01]. Considering entanglement manipulation under PPT operations provides us with a mathematically tractable framework to deepen our understanding of it.

#### Fidelity of PPT distillation

**Definition 3.1.** In deriving this bound, Rains introduced the "fidelity of $k$-state PPT distillation" by

$$F_{\mathrm{PPT}}(\rho_{AB}, k) := \max\{\operatorname{Tr} \Phi(k)\, \Pi(\rho_{AB}) : \Pi \in \mathrm{PPT}\} \tag{3.9}$$

which is the optimal entanglement fidelity of $k \otimes k$ maximally entangled states one can obtain from $\rho_{AB}$ by PPT operations (cf. Section 2.3.1).

In [Rai01], Rains simplified $F_{\text{PPT}}\left(\rho_{AB}, k\right)$ to

$$\begin{aligned} F_{\text{PPT}}\left(\rho_{AB}, k\right) = \max\ & \operatorname{Tr} \rho_{AB} Q_{AB}, \\ \text{s.t. } & 0 \leq Q_{AB} \leq \mathbb{1}, \\ & -\frac{1}{k}\mathbb{1} \leq Q_{AB}^{T_B} \leq \frac{1}{k}\mathbb{1}. \end{aligned} \tag{3.10}$$

**One-shot $\varepsilon$-infidelity PPT distillable entanglement**

**Definition 3.2.** For any bipartite quantum state $\rho_{AB}$, the one-shot $\varepsilon$-infidelity PPT distillable entanglement is defined as

$$E_{D,\text{PPT}}^{(1)}\left(\rho_{AB}, \varepsilon\right) := \log \max \left\{k : F_{\text{PPT}}\left(\rho_{AB}, k\right) \geq 1 - \varepsilon\right\}. \tag{3.11}$$

Using this SDP of fidelity of distillation in Eq. (3.10), it is easy to obtain

$$E_{D,PPT}^{(1)}\left(\rho_{AB}, \varepsilon\right) = -\log \min\{\eta :\ 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon, \|Q^{T_B}\|_\infty \leq \eta\}. \tag{3.12}$$

As mentioned in Section 2.7.2, the hypothesis testing relative entropy can be used to characterize the amount of standard entanglement that can be distilled from the quantum state: for any bipartite state $\rho_{AB}$ and infidelity tolerance $\varepsilon \in (0, 1)$,

$$E_{D,PPT}^{(1)}\left(\rho_{AB}, \varepsilon\right) = \min_{\left\|C^{T_B}\right\|_1 \leq 1} D_H^{\varepsilon}\left(\rho_{AB} \| C\right). \tag{3.13}$$

Note that $C$ need not be positive semidefinite.

Via the norm duality between the trace norm and the operator norm, it holds that

$$\begin{aligned} E_{D,PPT}^{(1)}\left(\rho_{AB}, \varepsilon\right) &= -\log \min\{\|Q^{T_B}\|_\infty : 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon\} \\ &= -\log \min_{Q} \max_{\|C\|_1 \leq 1} \{\operatorname{Tr} Q^{T_B} C : 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon\} \\ &= -\log \min_{Q} \max_{\|C\|_1 \leq 1} \{\operatorname{Tr} Q C^{T_B} : 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon\} \\ &= -\log \max_{\|C\|_1 \leq 1} \min_{Q} \{\operatorname{Tr} Q C^{T_B} : 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon\} \\ &= -\log \max_{\|C^{T_B}\|_1 \leq 1} \min_{Q} \{\operatorname{Tr} Q C : 0 \leq Q \leq \mathbb{1}, \operatorname{Tr} \rho_{AB} Q \geq 1 - \varepsilon\} \\ &= \min_{\|C^{T_B}\|_1 \leq 1} D_H^{\varepsilon}\left(\rho_{AB} \| C\right) \end{aligned} \tag{3.14}$$

In the fourth line, we apply the Sion minimax theorem [Sio58]. In the fifth line, we substitute $C$ with $C^{T_B}$.

We refer to [FWTD17] for more details about the non-asymptotic study of entanglement distillation. Moreover, the refinement of $E_{D,PPT}^{(1)}(\rho_{AB}, \varepsilon)$ also can be used to easily recover the Rains bound [Rai01] via the quantum Stein's lemma [HP91, ON99].

#### PPT distillable entanglement

**Definition 3.3.** For any bipartite quantum state $\rho_{AB}$, the asymptotic PPT distillable entanglement can be equivalently defined as

$$E_{D,\mathrm{PPT}}(\rho_{AB}) := \sup\{r : \lim_{n\to\infty} F_{\mathrm{PPT}}(\rho_{AB}^{\otimes n}, 2^{nr}) = 1\}. \tag{3.15}$$

The *logarithmic negativity* of a state $\rho_{AB}$ mentioned above is defined as [VW02, Ple05a]

$$E_N(\rho_{AB}) = \log \|\rho_{AB}^{T_B}\|_1. \tag{3.16}$$

As shown in Refs. [Rai01, VW02], the significance of $E_N$ is highlighted in the following

$$E_D(\rho_{AB}) \le E_{D,\mathrm{PPT}}(\rho_{AB}) \le E_N(\rho_{AB}).$$

## 3.3 Improved SDP upper bound on distillable entanglement

We are now ready to introduce an SDP upper bound $E_W$ on $E_{D,\mathrm{PPT}}$ and thus also on $E_D$, as follows:

$$E_W(\rho_{AB}) := \log W(\rho_{AB}),$$

where $W(\rho_{AB})$ is given by the following SDP:

$$\begin{aligned} W(\rho_{AB}) = \max\ & \mathrm{Tr}\,\rho_{AB}^{T_B} R_{AB}, \\ \text{s.t.}\ & -\mathbb{1} \le R_{AB} \le \mathbb{1}, R_{AB}^{T_B} \ge 0. \end{aligned} \tag{3.17}$$

Noticing that the constraint $-\mathbb{1} \le R_{AB} \le \mathbb{1}$ can be rewritten as $\|R_{AB}\|_\infty \le 1$, we can use Lagrange multiplier approach to obtain the dual SDP as follows:

$$\begin{aligned} W(\rho_{AB}) = \min\ & \mathrm{Tr}\,(U_{AB} + V_{AB}), \\ \text{s.t.}\ & U_{AB}, V_{AB} \ge 0, \\ & (U_{AB} - V_{AB})^{T_B} \ge \rho_{AB}. \end{aligned} \tag{3.18}$$

It is worth noting that the optimal values of the primal and the dual SDPs above coincide. This is a consequence of strong duality. By Slater's Theorem, one simply needs to show that there exists positive definite $U_{AB}$ and $V_{AB}$ such that $(U_{AB} - V_{AB})^{T_B} > \rho_{AB}$, which holds for $U_{AB} = 3V_{AB} = 3\mathbb{1}$. Introducing a new variable operator $X_{AB} = (U_{AB} - V_{AB})^{T_B}$, we can further simplify the dual SDP to

$$\begin{aligned} W(\rho_{AB}) = \min \|X_{AB}^{T_B}\|_1, \\ \text{s.t. } X_{AB} \geq \rho_{AB}. \end{aligned} \tag{3.19}$$

The function $E_W(\cdot)$ has the following remarkable properties which will be discussed in greater detail shortly:

i) Additivity (cf. Proposition 3.4):

$$E_W(\rho_{AB} \otimes \sigma_{A'B'}) = E_W(\rho_{AB}) + E_W(\sigma_{A'B'}). \tag{3.20}$$

ii) Upper bound on PPT distillable entanglement (cf. Theorem 3.5):

$$E_{D,\text{PPT}}(\rho_{AB}) \leq E_W(\rho_{AB}), \ \forall \rho_{AB}. \tag{3.21}$$

iii) Detecting genuine PPT distillable entanglement: $E_W(\rho_{AB}) > 0$ if and only if $\rho_{AB}$ is PPT distillable (cf. Proposition 3.6).

iv) Full entanglement monotone under general LOCC (or PPT) operations (cf. Theorem 3.12):

$$E_W(\rho_{AB}) \geq \sum_i p_i E_W(\rho_i). \tag{3.22}$$

v) Improved bound over logarithmic negativity (cf. Proposition 3.6):

$$E_W(\rho_{AB}) \leq E_N(\rho_{AB}), \ \forall \rho_{AB}, \tag{3.23}$$

and the inequality can be strict.

vi) An interpretation as the max-Rains relative entropy (cf. Proposition 3.10):

$$E_W(\rho) = \min_{\sigma \in \text{PPT}'} D_{\max}(\rho \| \sigma). \tag{3.24}$$

**Additivity of $E_W$**

Property i) is equivalent to the multiplicativity of the function $W(\cdot)$ under tensor products and can be proven directly by using the primal and dual SDPs of $W(\cdot)$.

**Proposition 3.4.** *For any two bipartite states $\rho_{AB}$ and $\sigma_{A'B'}$, we have*

$$W\left(\rho_{AB}\otimes\sigma_{A'B'}\right)=W\left(\rho_{AB}\right)W\left(\sigma_{A'B'}\right) \tag{3.25}$$

*Proof.* To see the super-multiplicativity, suppose that the optimal solutions to the primal SDP (3.17) of $W\left(\rho_{AB}\right)$ and $W\left(\sigma_{A'B'}\right)$ are $R_{AB}$ and $S_{A'B'}$, respectively.

We need to show that $R_{AB}\otimes S_{A'B'}$ is a feasible solution to the primal SDP (3.17) of $W\left(\rho_{AB}\otimes\sigma_{A'B'}\right)$. That will imply

$$W\left(\rho_{AB}\otimes\sigma_{A'B'}\right)\geq \operatorname{Tr}\left(\rho_{AB}^{T_B}\otimes\sigma_{A'B'}^{T_{B'}}\right)\left(R_{AB}\otimes S_{A'B'}\right)=W\left(\rho_{AB}\right)W\left(\sigma_{A'B'}\right). \tag{3.26}$$

The proof is quite straightforward. Indeed from $\|R_{AB}\|_\infty\leq 1$ and $\|S_{A'B'}\|_\infty\leq 1$, the inequality

$$\|R_{AB}\otimes S_{A'B'}\|_\infty\leq 1 \tag{3.27}$$

follows immediately. Also the positivity of $R_{AB}^{T_B}\otimes S_{A'B'}^{T_{B'}}$ is obvious. Hence we are done.

The sub-multiplicativity of $W\left(\cdot\right)$ can be proven similarly by employing dual SDP (3.19) of $W\left(\rho_{AB}\right)$. □

### Upper bound on distillable entanglement

Property ii) requires some effort and is presented in the following

**Theorem 3.5.** *For any state $\rho_{AB}$,*

$$E_{D,PPT}\left(\rho_{AB}\right)\leq E_W\left(\rho_{AB}\right). \tag{3.28}$$

*Proof.* Suppose $E_{D,\mathrm{PPT}}\left(\rho_{AB}\right)=r$. Then

$$\lim_{n\to\infty}F_{\mathrm{PPT}}\left(\rho_{AB}^{\otimes n},2^{nr}\right)=1.$$

For a given $k$, suppose that the optimal solution to the SDP (3.10) of $F_{\mathrm{PPT}}\left(\rho_{AB},k\right)$ is $Q_{AB}$. Let $R_{AB}=kQ_{AB}^{T_B}$. Then from the constraints of SDP (3.10), we have that $-\mathbb{1}\leq R_{AB}=kQ_{AB}^{T_B}\leq\mathbb{1}$. It is also clear that $R_{AB}^{T_B}\geq 0$. So $R_{AB}$ is a feasible solution to the primal SDP (3.17) of $W\left(\rho_{AB}\right)$. Therefore,

$$W\left(\rho_{AB}\right)\geq \operatorname{Tr}\rho_{AB}^{T_B}R_{AB}=k\operatorname{Tr}\rho_{AB}Q_{AB}=kF_{\mathrm{PPT}}\left(\rho_{AB},k\right).$$

Hence,

$$\lim_{n\to\infty} W\left(\rho_{AB}^{\otimes n}\right)/2^{nr} \geq \lim_{n\to\infty} F_{\text{PPT}}\left(\rho_{AB}^{\otimes n}, 2^{nr}\right) = 1.$$

Noticing that $W(\rho)$ is multiplicative, we have

$$\lim_{n\to\infty} W\left(\rho_{AB}^{\otimes n}\right)/2^{nr} = \lim_{n\to\infty} \left(W\left(\rho_{AB}\right)\right)^n/2^{nr} \geq 1.$$

Therefore, $W(\rho_{AB}) \geq 2^r$, and we are done. □

#### Detect entanglement

Property iii) suggests an interesting equivalent relation between $E_W$ and $E_\Gamma$ in the sense that $E_W$ can be used to detect whether a state is genuinely distillable under PPT operations.

**Proposition 3.6.** *For a state $\rho_{AB}$, $E_W(\rho_{AB}) > 0$ if and only if $E_{D,PPT}(\rho_{AB}) > 0$.*

*Proof.* We only need to show that $W(\rho_{AB}) > 1$ is equivalent to $\rho_{AB}$ is an non-positive partial transpose (NPPT) state. The rest proof then can be completed by combining this fact with an interesting result from [EVWW01]: any NPPT state is PPT distillable.

Firstly, if $\rho_{AB}$ is PPT, then $W(\rho_{AB}) \leq \|\rho_{AB}^{T_B}\|_1 = 1$. Assume now $\rho_{AB}$ is NPPT, we will show that $W(\rho_{AB}) > 1$. Let $P_-$ be the projection on the subspace spanned by the eigenvectors with negative eigenvalues of $\rho_{AB}^{T_B}$, and let $\lambda = \|P_-^{T_B}\|_\infty$. Introduce

$$R_{AB} = \mathbb{1}_{AB} - \frac{1}{\max\{\lambda, 0.5\}} P_-.$$

It is clear that $R_{AB}^{T_B} \geq 0$ by construction. Furthermore, we can easily verify that

$$-\mathbb{1} \leq \mathbb{1} - 2P_- \leq R_{AB} \leq \mathbb{1}. \tag{3.29}$$

So $R_{AB}$ is a feasible solution to the primal SDP (3.17) of $W(\rho_{AB})$. Noticing that $\rho_{AB}$ is NPPT, we have that

$$W(\rho_{AB}) \geq \operatorname{Tr} \rho_{AB}^{T_B} R_{AB} = 1 - \frac{\operatorname{Tr} P_- \rho_{AB}^{T_B}}{\max\{\lambda, 0.5\}} > 1,$$

where we have used the property that $\operatorname{Tr} P_- \rho_{AB}^{T_B} < 0$. □

**Comparison with logarithmic negativity:**

Now we discuss property iv). Before that, let us recall that $\|\rho_{AB}^{T_B}\|_1$ can be reformulated as

$$\begin{aligned} \|\rho_{AB}^{T_B}\|_1 = \max \ & \operatorname{Tr} \rho_{AB}^{T_B} R_{AB} \\ \text{s.t.} \ & \|R_{AB}\|_\infty \leq 1. \end{aligned} \tag{3.30}$$

**Proposition 3.7.** *For any state $\rho_{AB}$, $E_W(\rho_{AB}) \leq E_N(\rho_{AB})$, and the inequality can be strict. Moreover, $E_W(\rho_{AB}) = E_N(\rho_{AB})$ if and only if SDP (3.30) has an optimal solution with positive partial transpose.*

*Proof.* The definition of $E_N$ is given in Eq. (3.16). Noting that $\rho_{AB}$ is a feasible solution to the dual SDP (3.19) of $W(\rho_{AB})$, we have $E_W(\rho_{AB}) \leq \log \|\rho_{AB}^{T_B}\|_1 = E_N(\rho_{AB})$.

To see the above inequality can be strict, we focus on a class of two-qubit states $\sigma_{AB}^{(r)} = r|v_0\rangle\langle v_0| + (1-r)|v_1\rangle\langle v_1|$ $(0 < r < 1)$, where $|v_0\rangle = 1/\sqrt{2}(|10\rangle - |11\rangle)$ and $|v_1\rangle = 1/\sqrt{3}(|00\rangle + |10\rangle + |11\rangle)$. The fact that $E_W\left(\sigma^{(r)}\right)$ can be strictly smaller than $E_N\left(\sigma^{(r)}\right)$ is shown in Figure 3.2.

To prove the second part of the theorem, let us assume that the optimal solution to SDP (3.30) of $\|\rho_{AB}^{T_B}\|_1$ is $R_{AB}$. If $R_{AB}^{T_B} \geq 0$, then it is also a feasible solution to the primal SDP (3.17) of $W(\rho_{AB})$. That immediately implies $E_W(\rho_{AB}) = E_N(\rho_{AB})$. Conversely, assume that $E_W(\rho_{AB}) = E_N(\rho_{AB})$, then the optimal solution $R_{AB}$ to SDP (3.17) of $W(\rho_{AB})$ is also the optimal solution to the SDP (3.30) for $\|\rho_{AB}^{T_B}\|_1$ and it holds that $R_{AB}^{T_B} \geq 0$. Therefore, $E_W(\rho_{AB}) = E_N(\rho_{AB})$ if and only if SDP (3.30) for $\|\rho_{AB}^{T_B}\|_1$ has a PPT optimal solution. $\square$

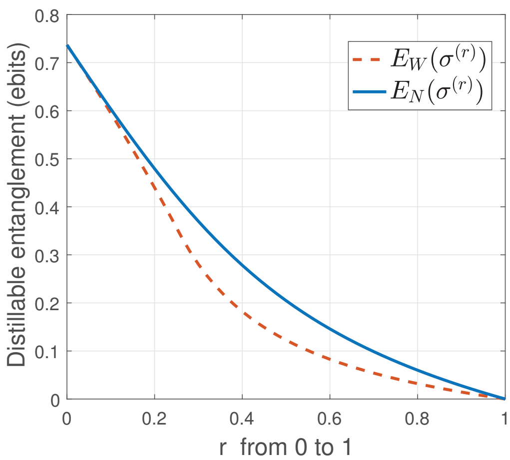


Figure 3.2: Comparsion between $E_W$ and $E_N$ for the class of states $\sigma^{(r)}$.

We further compare $E_W$ to $E_{D,PPT}$ and $E_N$ using a class of $3 \otimes 3$ states defined by

$$\rho^{(\alpha)} = \sum_{m=0}^{2} |\psi_m\rangle\langle\psi_m|/3 \;\; (0 < \alpha \leq 0.5)$$

with $|\psi_0\rangle = \sqrt{\alpha}|01\rangle + \sqrt{1-\alpha}|10\rangle$, $|\psi_1\rangle = \sqrt{\alpha}|02\rangle + \sqrt{1-\alpha}|20\rangle$ and $|\psi_2\rangle = \sqrt{\alpha}|12\rangle + \sqrt{1-\alpha}|21\rangle$.

**Proposition 3.8.** *For the class of states $\rho^{(\alpha)}$, we have that*

$$E_{D,PPT}\left(\rho^{(\alpha)}\right) \leq E_W\left(\rho^{(\alpha)}\right) < E_N\left(\rho^{(\alpha)}\right).$$

*In particular,*

$$E_{D,PPT}\left(\rho^{(0.5)}\right) = E_W\left(\rho^{(0.5)}\right) = \log\frac{3}{2} < \log\frac{5}{3} = E_N\left(\rho^{(0.5)}\right).$$

*Proof.* The first step is to show that

$$E_N\left(\rho^{(\alpha)}\right) = \log \| \left(\rho^{(\alpha)}\right)^{T_B} \|_1 = \log\left(1 + 4/3\sqrt{\alpha\,(1-\alpha)}\right). \tag{3.31}$$

Secondly, we can choose $X_{AB} = \rho^{(\alpha)} + \sqrt{\alpha\,(1-\alpha)}/3\,(|00\rangle\langle 00| + |11\rangle\langle 11| + |22\rangle\langle 22|)$ as a feasible solution to the dual SDP (3.19). By a routine calculation, we have

$$E_W\left(\rho^{(\alpha)}\right) = \log W\left(\rho^{(\alpha)}\right) \leq \log \|X_{AB}^{T_B}\|_1 \tag{3.32}$$

$$= \log\left(1 + \sqrt{\alpha\,(1-\alpha)}\right) < E_N\left(\rho^{(\alpha)}\right). \tag{3.33}$$

For $\alpha = 0.5$, choose $k_0 = 3/2$ and $Q = \sum_{m=0}^{2}\left(|\psi_m\rangle\langle\psi_m| + 1/3|\widehat{\psi}_m\rangle\langle\widehat{\psi}_m|\right)$ with

$$|\widehat{\psi}_0\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right), \tag{3.34}$$

$$|\widehat{\psi}_1\rangle = \frac{1}{\sqrt{2}}\left(|02\rangle - |20\rangle\right), \tag{3.35}$$

$$|\widehat{\psi}_2\rangle = \frac{1}{\sqrt{2}}\left(|12\rangle - |21\rangle\right). \tag{3.36}$$

Noticing that $\|Q^{T_B}\|_\infty = 2/3$, we have $-1/k_0\mathbb{1} \leq Q^{T_B} \leq 1/k_0\mathbb{1}$. Thus $Q$ is a feasible solution to the SDP (3.10) of $F_{\text{PPT}}\left(\rho^{(0.5)}, k_0\right)$, which has an optimal value 1 due to $1 \geq F_{\text{PPT}}\left(\rho^{(0.5)}, k_0\right) \geq \operatorname{Tr}\rho^{(0.5)}Q = 1$. Applying the definition of $E_{D,PPT}$, we have

$$E_{D,PPT}\left(\rho^{(0.5)}\right) \geq \log k_0 = \log 3/2. \tag{3.37}$$

Finally, combining Eqs. (3.31), (3.32), and (3.37), we obtain the desired chain of inequalities. □

**Remark 3.9.** It is worth noting that $\rho^{(0.5)}$ is in the subspace span$\{|01\rangle + |10\rangle, |02\rangle + |20\rangle, |12\rangle + |21\rangle\}$, which is not locally unitarily equivalent to the anti-symmetric subspace span$\{|01\rangle - |10\rangle, |02\rangle - |20\rangle, |12\rangle - |21\rangle\}$. For the corresponding $3 \otimes 3$ antisymmetric state $\sigma = \frac{1}{3}\left(|\widehat{\psi}_0\rangle\langle\widehat{\psi}_0| + |\widehat{\psi}_1\rangle\langle\widehat{\psi}_1| + |\widehat{\psi}_2\rangle\langle\widehat{\psi}_2|\right)$, it holds that $E_\Gamma(\sigma) = E_W(\sigma) = E_N(\sigma) = \log(5/3)$.

### 3.3.1 max-Rains relative entropy

**Proposition 3.10** (max-Rains relative entropy). *For any bipartite state $\rho \in \mathcal{S}(A \otimes B)$, it holds that*

$$E_W(\rho) = \min_{\sigma \in PPT'} D_{\max}(\rho \| \sigma). \tag{3.38}$$

*Consequently, we also call $E_W$ the max-Rains relative entropy.*

*Proof.* The following equality chain holds

$$\begin{aligned} E_W(\rho) &= \log \min \left\{ \left\| X^{T_B} \right\|_1 : \rho \leq X \right\} \\ &= \log \min \left\{ \mu : \rho \leq X, \left\| X^{T_B} \right\|_1 \leq \mu \right\} \\ &= \log \min \left\{ \mu : \rho \leq \mu\sigma, \left\| \mu\sigma^{T_B} \right\|_1 \leq \mu \right\} \\ &= \log \min \left\{ \mu : \rho \leq \mu\sigma, \left\| \sigma^{T_B} \right\|_1 \leq 1 \right\} \\ &= \min_{\sigma \in PPT'} D_{\max}(\rho \| \sigma). \end{aligned} \tag{3.39}$$

The first line follows from Eq. (3.19). In the second line, we introduce a new variable $\mu$. In the third line, we substitute $X$ with $\mu\sigma$. The last line follows from the definition of $D_{\max}$. □

**Remark 3.11.** This implies that $E_W$ can be considered as the max-Rains relative entropy, which also indicates that $E_W$ is always larger than the Rains bound, i.e.,

$$E_W(\rho) \geq R(\rho). \tag{3.40}$$

We note that the advantage of $E_W$ is that it can be represented as in both SDP problem and max-relative entropy form, which can lead to both theoretical and numerical insights for entanglement distillation as well as quantum communication (e.g., [WFD17, DBW17, BW18]). We will introduce the channel version of $E_W$ to study the quantum capacity of a general quantum channel in Chapter 6.

### $E_W$ is an entanglement monotone

We are going to prove that $E_W$ is entanglement monotone in the sense of Eq. (3.8) both under general LOCC operations as well as the more general PPT operations.

**Theorem 3.12.** *The function $E_W(\cdot)$ is an entanglement monotone both under general LOCC and PPT operations.*

*Proof.* Noting that PPT operations include LOCC as a subset, we only need to prove the case of PPT operations. Let us consider a general PPT operation $\mathcal{N} = \sum_i \mathcal{N}_i$ that maps bipartite state $\rho$ to $\mathcal{N}_i(\rho) / \operatorname{Tr}(\mathcal{N}_i(\rho))$ with probability $\operatorname{Tr} \mathcal{N}_i(\rho)$, where $\mathcal{N}_i$ is CP and PPT operation.

Refer to the dual SDP (3.19) of $W(\rho_{AB})$, we suppose that $X_{AB}$ is the optimal solution. It is easy to see that $\mathcal{N}_i(X_{AB}) \geq \mathcal{N}_i(\rho)$, then $\mathcal{N}_i(X_{AB})$ is feasible to the dual SDP (3.19) of $W(\mathcal{N}_i(\rho))$. Therefore,

$$W(\mathcal{N}_i(\rho)) \leq \| (\mathcal{N}_i(X_{AB}))^{T_B} \|_1 = \operatorname{Tr} |\mathcal{N}_i^{T_B}\left(X_{AB}^{T_B}\right)|,$$

where $\mathcal{N}_i^{T_B}(\sigma) = \left(\mathcal{N}\left(\sigma^{T_B}\right)\right)^{T_B}$. By the fact that $\mathcal{N}_i^{T_B}$ is CP [Rai99, Rai01], we have $W(\mathcal{N}_i(\rho)) \leq \operatorname{Tr} |\mathcal{N}_i^{T_B}\left(X_{AB}^{T_B}\right)| \leq \operatorname{Tr} \mathcal{N}_i^{T_B}\left(|X_{AB}^{T_B}|\right)$. Furthermore,

$$\begin{aligned}
\sum_i p_i E_W(\rho_i) &\leq \log \sum_i p_i W(\rho_i) = \log \sum_i W(\mathcal{N}_i(\rho)) \\
&\leq \log \sum_i \operatorname{Tr} \mathcal{N}_i^{T_B}\left(|X_{AB}^{T_B}|\right) \\
&= \log \sum_i \operatorname{Tr}[\mathcal{N}_i\left(|X_{AB}^{T_B}|^{T_B}\right)]^{T_B} \\
&= \log \operatorname{Tr} \mathcal{N}\left(|X_{AB}^{T_B}|^{T_B}\right) = E_W(\rho).
\end{aligned}$$

Thus, we obtain the monotonicity of $E_W$ under general PPT operations in the sense of Eq. (3.8). Similar to the logarithmic negativity, $E_W$ is also a full entanglement monotone that is not convex. □

## 3.4 Deterministic Distillable Entanglement

The deterministic entanglement distillation concerns about how to distill maximally entangled states exactly. The bipartite pure state case is completely solved in Refs. [MW08, DFJY05]. We will show that PPT deterministic distillable entanglement of a state $\rho$ depends only on the support $\operatorname{supp}(\rho)$, which is defined to be the space spanned by the eigenvectors with positive eigenvalues of $\rho$. We will study the deterministic distillable entanglement in both one-shot and asymptotic settings in this section.

### 3.4.1 One-copy deterministic distillable entanglement

The one-copy $\Omega$-assisted deterministic distillable entanglement of $\rho_{AB}$ is defined by

$$E^{(1)}_{0,\Omega}(\rho_{AB}) := \max\{\log k : F_\Omega(\rho_{AB}, k) = 1, k > 0\}, \tag{3.41}$$

where $\Omega \in \{\text{LOCC}, \text{SEP}, \text{PPT}\}$. Clearly $E^{(1)}_{0,\Omega}(\rho) \geq 0$ since $F_\Omega(\rho, 1) = 1$ trivially holds.

For LOCC operations, one-copy $\Omega$ deterministic distillable entanglement is still intractable. But for PPT operations, we could use the fidelity of PPT distillation to give a concrete characterization of the one-copy deterministic distillable entanglement. Replacing $k$ and $Q_{AB}$ in SDP (3.10) by $\operatorname{Tr}\rho_{AB}R_{AB}$ and $R_{AB}/\operatorname{Tr}\rho_{AB}R_{AB}$, respectively, we can further simplify $E^{(1)}_{0,D,PPT}(\rho_{AB})$ as follows:

$$\begin{aligned} E^{(1)}_{0,D,PPT}(\rho_{AB}) = \max\ & \log \operatorname{Tr}\rho_{AB}R_{AB}, \\ \text{s.t. } & 0 \leq R_{AB} \leq (\operatorname{Tr}\rho_{AB}R_{AB})\,\mathbb{1}_{AB}, \\ & |R^{T_B}_{AB}| \leq \mathbb{1}_{AB}. \end{aligned} \tag{3.42}$$

**Proposition 3.13.** *For bipartite state $\rho_{AB}$, it holds that $E^{(1)}_{0,D,PPT}(\rho_{AB}) = -\log W_0(\rho_{AB})$, where*

$$\begin{aligned} W_0(\rho_{AB}) = \min\ & \|R^{T_B}_{AB}\|_\infty, \\ \textit{s.t. } & P_{AB} \leq R_{AB} \leq \mathbb{1}_{AB}, \end{aligned} \tag{3.43}$$

*and $P_{AB}$ is the projection onto* $\operatorname{supp}(\rho_{AB})$.

*Proof.* The first constraint in SDP (3.42) implies that $\operatorname{Tr}\rho_{AB}R_{AB} \geq \|R_{AB}\|_\infty$. So any feasible $R_{AB}$ should be of the form $xP_{AB} + S_{AB}$, where $x \geq 0$, $P_{AB}$ is the projection onto $\operatorname{supp}(\rho_{AB})$, and $0 \leq S_{AB} \leq x(\mathbb{1} - P)_{AB}$. Replacing $S_{AB}/x + P_{AB}$ by $R_{AB}$ and noticing $E^{(1)}_{\Gamma,0}(\rho_{AB}) = \log W_0(\rho_{AB})$, we have

$$\begin{aligned} E^{(1)}_{0,D,PPT}(\rho_{AB}) = \max\ & -\log \|R^{T_B}_{AB}\|_\infty, \\ \text{s.t. } & P_{AB} \leq R_{AB} \leq \mathbb{1}_{AB}. \end{aligned} \tag{3.44}$$

□

In particular, $E^{(1)}_{0,D,PPT}(\rho_{AB}) \geq -\log \|P^{T_B}_{AB}\|_\infty$ when $R_{AB} = P_{AB}$. For bipartite pure entangled states this lower bound gives the exact value of the PPT deterministic distillable entanglement [MW08, DFJY05]. However, this is not the case for general mixed bipartite states.

### 3.4.2 Asymptotic deterministic distillable entanglement

The asymptotic deterministic distillable entanglement quantifies the rate of deterministic distillation in the asymptotic limit of large number of i.i.d. prepared states. Thus, it is in the form of regularization.

**Definition 3.14.** Given bipartite state $\rho_{AB}$, its asymptotic deterministic distillable entanglement under $\Omega$ operations is defined by

$$E_{0,\Omega}(\rho) := \sup_{n\geq 1} E_{0,\Omega}^{(1)}\left(\rho^{\otimes n}\right)/n = \lim_{n\geq 1} E_{0,\Omega}^{(1)}\left(\rho^{\otimes n}\right)/n, \tag{3.45}$$

where $\Omega \in \{\text{LOCC}, \text{SEP}, \text{PPT}\}$.

This deterministic distillable entanglement is computationally intractable due to regularization. However, using the technique of SDP, we will introduce an efficiently computable upper bound to evaluate this quantity.

For a bipartite quantum state $\rho$, we introduce

$$\begin{aligned} E_M(\rho) = -\log M(\rho) = -\log \max\ & \operatorname{Tr} P_{AB} V_{AB}, \\ \text{s.t.}\ & \operatorname{Tr}|V_{AB}^{T_B}| = 1, V_{AB} \geq 0, \end{aligned} \tag{3.46}$$

where $P_{AB}$ is the projection onto the support of $\rho$. And $M(\rho)$ is also given by the following SDP:

$$\begin{aligned} M(\rho) = \max \operatorname{Tr} & P_{AB} Z_{AB}, \\ \text{s.t.}\ & \operatorname{Tr}(X_{AB} + Y_{AB}) = 1, \\ & Z_{AB} \leq (X_{AB} - Y_{AB})^{T_B}, \\ & X_{AB}, Y_{AB}, Z_{AB} \geq 0, \end{aligned} \tag{3.47}$$

And its dual SDP is given by

$$M(\rho) = \min\{\|R_{AB}^{T_B}\|_\infty : R_{AB} \geq P_{AB}\}. \tag{3.48}$$

The optimal values of the primal and the dual SDPs above coincide by strong duality.

For any two bipartite states $\rho_{AB}$ and $\sigma_{A'B'}$, by utilizing semidefinite programming duality, it is not difficult to prove that

$$E_M(\rho_{AB} \otimes \sigma_{A'B'}) = E_M(\rho_{AB}) + E_M(\sigma_{A'B'}).$$

Furthermore, for any state bipartite $\rho$, $E_M(\rho) = 0$ if and only if $\operatorname{supp}(\rho)$ contains the support of a PPT state $\sigma$, i.e. $\operatorname{supp}(\sigma) \subseteq \operatorname{supp}(\rho)$. Too see this, if there exists PPT state $\sigma$ such that $\operatorname{supp}(\sigma) \subseteq \operatorname{supp}(\rho)$, then $E_M(\rho) = 0$. On the other hand, if any state $\sigma$ satisfies $\operatorname{supp}(\sigma) \subseteq \operatorname{supp}(\rho)$ is NPPT. Let the optimal solution to SDP (3.46) be $V$, where $V \geq 0$ and $\operatorname{Tr}|V^{T_B}| = 1$. It is clear that $\operatorname{Tr} V \leq 1$. Thus, we have $\operatorname{Tr} V = 1$ when $E_M(\rho) = 0$. Hence, $V$ is a PPT state and $\operatorname{supp}(V) \subseteq \operatorname{supp}(\rho)$. This leads to a contradiction.

We show that $E_M$ is the best upper bound on the deterministic distillable entanglement of bipartite states. The bipartite pure state case is completely solved in Refs. [MW08, DFJY05]. For a general bipartite state, the PPT-assisted deterministic distillation rates depend only on the support of this state.

**Theorem 3.15.** *For any bipartite state $\rho \in \mathcal{S}(A \otimes B)$,*

$$E_{0,D,PPT}(\rho) \leq E_M(\rho) \leq E_W(\rho).$$

*Proof.* To prove $E_{0,D,PPT}(\rho) \leq -\log M(\rho)$, let us first suppose that the optimal solution to SDP (3.43) of $W_0(\rho)$ is $R_0$. It is clear that $R_0$ is also a feasible solution to SDP (3.48) of $M(\rho)$. Thus, it holds that

$$W_0(\rho) = \|{R_0}^{T_B}\|_\infty \geq M(\rho), \tag{3.49}$$

Applying the additivity of $M(\rho)$, we have

$$W_0\left(\rho^{\otimes n}\right) \geq M\left(\rho^{\otimes n}\right) = M(\rho)^n. \tag{3.50}$$

Hence, we have

$$E_{0,D,PPT}(\rho) = \lim_{n\to\infty} -\frac{1}{n}\log W_0\left(\rho^{\otimes n}\right) \tag{3.51}$$

$$\leq \lim_{n\to\infty} -\frac{1}{n}\log M(\rho)^n = E_M(\rho). \tag{3.52}$$

Finally, to prove $E_M(\rho) \leq E_W(\rho)$, suppose that the optimal solution to SDP (3.48) is $R$, then we have $R \geq P \geq 0$. Let $R_1 = R/\|R^{T_B}\|_\infty$ and it is easy to see the positivity of $R_1$ and the fact that $|R_1^{T_B}| \leq \mathbb{1}$, which means that $R_1$ is a feasible solution to SDP (3.17). Therefore, $E_W(\rho) \geq \log \operatorname{Tr}\rho R_1 \geq \log \operatorname{Tr}\rho P/\|R^{T_B}\|_\infty = -\log\|R^{T_B}\|_\infty = E_M(\rho)$. □

**Remark 3.16.** For any bipartite state $\rho \in \mathcal{S}(A \otimes B)$, if the support of $\rho$ contains a PPT state $\sigma$, then $E_M(\rho) = 0$ and we have that $E_{0,D,PPT}(\rho) = 0$. Thus $\rho$ is bound entanglement for exact distillation under both LOCC or PPT operations.

We further show the estimation of Theorem 3.15 in the following figure by a class of $3 \otimes 3$ states in [WD16b] defined by

$$\rho^{(\alpha)} = \frac{1}{3} \sum_{m=0}^{2} \left(X^\dagger \otimes X\right)^m |\psi_0\rangle\langle\psi_0| \left(X \otimes X^\dagger\right)^m,$$

where $|\psi_0\rangle = \sqrt{\alpha}|00\rangle + \sqrt{1-\alpha}|11\rangle$ $(0 < \alpha \le 0.5)$ and $X = \sum_{j=0}^{2} |j \oplus 1\rangle\langle j|$. An interesting fact is that $E_M\left(\rho^{(\alpha)}\right)$ is tight for $E_{0,D,PPT}\left(\rho^{(\alpha)}\right)$ when $0 < \alpha \le 1/5$, which is proved in the following Proposition.

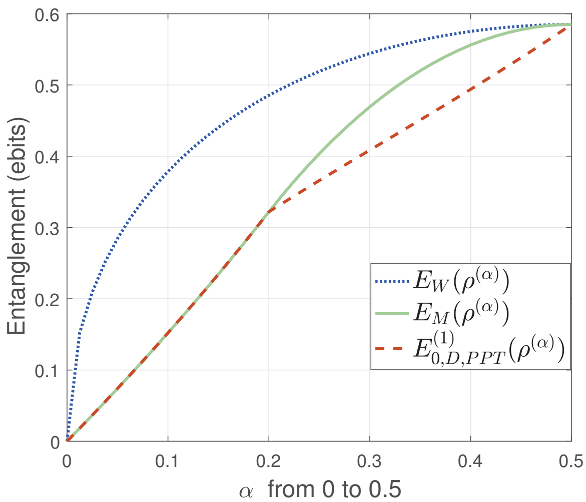


Figure 3.3: This plot presents the estimation of $E_{D,PPT}\left(\rho^{(\alpha)}\right)$ and $E_{0,D,PPT}\left(\rho^{(\alpha)}\right)$. The dot line depicts $E_W\left(\rho^{(\alpha)}\right)$, the dash line depicts $E^{(1)}_{0,D,PPT}\left(\rho^{(\alpha)}\right)$ and the solid line depicts $E_M\left(\rho^{(\alpha)}\right)$.

**Proposition 3.17.** *For any bipartite state $\rho \in \mathcal{S}\,(A \otimes B)$ with support projection $P$, suppose that the eigenvector $|\psi\rangle$ of $P^{T_B}$ with the eigenvalue $\|P^{T_B}\|_\infty$ is a product state, then*

$$E_{0,D,PPT}\,(\rho) = E_M\,(\rho) = -\log \|P^{T_B}\|_\infty \le E_{D,PPT}\,(\rho)\,. \tag{3.53}$$

*Proof.* From Eq. (3.44), it is easy to show that $E_{0,D,PPT}\,(\rho) \ge -\log \|P^{T_B}\|_\infty$. If $|\psi\rangle\langle\psi|$ is PPT, then we can choose $V = |\psi\rangle\langle\psi|$ and it is easy to see $V$ is a feasible solution to SDP (3.46) of $M\,(\rho)$. Thus,

$$E_M\,(\rho) \le -\log \operatorname{Tr} P^{T_B}|\psi\rangle\langle\psi| = -\log \|P^{T_B}\|_\infty. \tag{3.54}$$

□

For any pure state $|\phi\rangle\langle\phi|$, suppose that $|\phi\rangle$ has the Schmidt decomposition $|\phi\rangle = \sum_{i=1}^{m} \lambda_i |ii\rangle$ with $\lambda_1^2 \geq ... \geq \lambda_m^2$ and $\sum_{i=1}^{m} \lambda_i^2 = 1$. Then $|\phi\rangle\langle\phi|^{T_B} = \sum_{i=1}^{m} \lambda_i^2 |ii\rangle\langle ii| + \sum_{i\neq j} \lambda_i \lambda_j |ji\rangle\langle ij|$. Thus, $\|P^{T_B}\|_\infty = \lambda_1^2$ and the corresponding eigenvector is $|11\rangle\langle 11|$. Hence, by Proposition 3.17, $E_{0,D,PPT}\left(|\phi\rangle\langle\phi|\right) = E_M\left(|\phi\rangle\langle\phi|\right) = -\log \||\phi\rangle\langle\phi|^{T_B}\|_\infty$. This rate can be achieved by LOCC [DFJY05].

**Example 3.18.** For the $\rho^{(\alpha)}$, when $0 < \alpha \leq 1/5$, we have that

$$E_{0,D,PPT}\left(\rho^{(\alpha)}\right) = E_M\left(\rho^{(\alpha)}\right) = -\log\left(1-\alpha\right). \tag{3.55}$$

Let us choose $U = X^\dagger \otimes X$ with $X = \sum_{j=0}^{2} |j \oplus 1\rangle\langle j|$. The projection onto $\operatorname{supp}\left(\rho^{(\alpha)}\right)$ is

$$P_\alpha = \sum_{m=0}^{2} U^m |\psi_0\rangle\langle\psi_0| \left(U^\dagger\right)^m. \tag{3.56}$$

Therefore,

$$\begin{aligned} P_\alpha^{T_B} &= 2\sqrt{\alpha\left(1-\alpha\right)}|v_1\rangle\langle v_1| - \sqrt{\alpha\left(1-\alpha\right)}\left(|v_2\rangle\langle v_2| + |v_3\rangle\langle v_3|\right) \\ &+ \sum_{m=0}^{2} U^m [\left(1-\alpha\right)|11\rangle\langle 11| + \alpha|00\rangle\langle 00|] \left(U^\dagger\right)^m, \end{aligned}$$

where

$$|v_1\rangle = \frac{1}{\sqrt{3}}\left(|01\rangle + |10\rangle + |22\rangle\right), \tag{3.57}$$

$$|v_2\rangle = \frac{1}{\sqrt{6}}|01\rangle + \frac{1}{\sqrt{6}}|10\rangle - \sqrt{\frac{2}{3}}|22\rangle \tag{3.58}$$

$$|v_3\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right). \tag{3.59}$$

When $0 < \alpha \leq 1/5$, we always have $1-\alpha \geq 2\sqrt{\alpha\left(1-\alpha\right)}$. Therefore, $\|P_\alpha^{T_B}\|_\infty = 1-\alpha$ and the corresponding eigenvector is $|11\rangle\langle 11|$. Applying Proposition 3.17, the proof is done.

## 3.5 Nonadditivity of Rains bound

Rains bound is arguably the best known upper bound on the distillable entanglement and was conjectured to be additive and coincide with the asymptotic relative entropy of entanglement [ADVW02, PV07]. In this section, we disprove both conjectures by explicitly constructing a special class of mixed two-qubit states.

### 3.5.1 Rains bound on distillable entanglement

To evaluate $E_D$ efficiently, one possible way is to find computable upper bounds. A well-known upper bound of the distillable entanglement is the *relative entropy of entanglement* [VPRK97, VPJK97], i.e., for a given bipartite state $\rho$,

$$E_R(\rho) := \min_{\sigma \in SEP(A:B)} D(\rho \| \sigma). \tag{3.60}$$

The asymptotic relative entropy of entanglement is given by

$$E_R^\infty(\rho) := \inf_{n \geq 1} \frac{1}{n} E_R\left(\rho^{\otimes n}\right). \tag{3.61}$$

Similarly, for a given bipartite state $\rho$, the *PPT relative entropy of entanglement* is defined by

$$E_{R,PPT}(\rho) := \min_{\sigma \in PPT(A:B)} D(\rho \| \sigma) \tag{3.62}$$

$$= \min\{D(\rho \| \sigma) : \sigma, \sigma^{T_B} \geq 0, \operatorname{Tr} \sigma = 1\}, \tag{3.63}$$

the optimal solution $\sigma_{AB}$ is called the closest PPT state of $\rho$. The asymptotic PPT relative entropy of entanglement is given by

$$E_{R,PPT}^\infty(\rho) = \inf_{n \geq 1} \frac{1}{n} E_{R,PPT}\left(\rho^{\otimes n}\right). \tag{3.64}$$

An improved bound is the Rains bound [Rai01], which is arguably the best known upper bound of distillable entanglement and refined in [ADVW02] as a convex optimization problem as

$$R(\rho_{AB}) := \min_{\tau_{AB} \in PPT'(A:B)} D(\rho_{AB} \| \tau_{AB}) \tag{3.65}$$

$$= \min\{D(\rho_{AB} \| \tau_{AB}) : \ \tau_{AB} \geq 0, \operatorname{Tr} |\tau_{AB}^{T_B}| \leq 1\}. \tag{3.66}$$

In the following Table. 3.1, we introduce the entanglement measures that we will use in this thesis.

As Rains bound is proved to be equal to the asymptotic PPT relative entropy of entanglement for Werner states [AEJ+01] and orthogonally invariant states [ADVW02], one open problem is to determine whether these two quantities always coincide [PV07]. It is also significant to determine whether Rains bound is additive or not. In [ADVW02], it was conjectured that Rains bound might be additive for arbitrary quantum states.

| Measures | Acronym | Definition |
|---|---|---|
| Distillable entanglement | $E_D$ | $\sup\{r : \lim_{n\to\infty} \inf_{\Lambda\in\text{LOCC}} \|\Lambda(\rho^{\otimes n}), \Phi(2^{rn})\|_1 = 0\}$ |
| PPT distillable entanglement | $E_{D,\text{PPT}}$ | $\sup\{r : \lim_{n\to\infty} \inf_{\Lambda\in\text{PPT}} \|\Lambda(\rho^{\otimes n}), \Phi(2^{rn})\|_1 = 0\}$ |
| Entanglement cost | $E_C$ | $\inf\{r : \lim_{n\to\infty} \inf_{\Lambda\in\text{LOCC}} \|\rho^{\otimes n} - \Lambda(\Phi(2^{rn}))\|_1 = 0\}$ |
| PPT entanglement cost | $E_{C,\text{PPT}}$ | $\inf\{r : \lim_{n\to\infty} \inf_{\Lambda\in\text{PPT}} \|\rho^{\otimes n} - \Lambda(\Phi(2^{rn}))\|_1 = 0\}$ |
| Entanglement of formation | $E_F$ | $\inf_{\rho=\sum_i p_i\|\psi\rangle\langle\psi\|_i} \sum_i p_i S(\text{Tr}_A \|\psi\rangle\langle\psi\|_i)$ |
| REE | $E_R$ | $\min_{\sigma\in\text{SEP}(A:B)} D(\rho\\|\sigma)$ |
| AREE | $E_R^\infty$ | $\inf_{n\geq 1} \frac{1}{n} E_R(\rho^{\otimes n})$ |
| PPT REE | $E_{R,PPT}$ | $\min_{\sigma\in\text{PPT}(A:B)} D(\rho\\|\sigma)$ |
| PPT AREE | $E_{R,PPT}^\infty$ | $\inf_{n\geq 1} \frac{1}{n} E_{R,PPT}(\rho^{\otimes n})$ |
| Rains bound | $R$ | $\min_{\sigma\in\text{PPT}'(A:B)} D(\rho\\|\sigma)$ |
| Regularized Rains bound | $R^\infty$ | $\inf_{n\geq 1} \frac{1}{n} R(\rho^{\otimes n})$ |
| Max-Rains bound | $E_W$ $(R_{\max})$ | $\min_{\sigma\in\text{PPT}'(A:B)} D_{\max}(\rho\\|\sigma)$ |
| Logarithmic negativity | $E_N$ | $\log \\|\rho^{T_B}\\|_1$ |
| Squashed entanglement | $E_{sq}$ | $\inf\{\frac{1}{2} I(A;B\|E)_\rho : \rho_{AB} = \text{Tr}_E\, \rho_{ABE}\}$ |

Table 3.1: Partial zoo of entanglement measures

For a general bipartite state $\rho$, it holds that $E_{R,PPT}(\rho) \geq R(\rho)$. However, $E_{R,PPT}(\rho)$ equals to $R(\rho)$ for every two-qubit state $\rho$ [MI08] or the bipartite state with one qubit subsystem [GGF14]. In particular, a two-qubit full-rank state $\sigma_{AB}$ is the closest separable state of any state $\rho$ in the following form [MI08, FG11]:

$$\rho_{AB} = \sigma_{AB} - xG(\sigma_{AB}), \tag{3.67}$$

and

$$G(\sigma) = \sum_{i,j} G_{i,j} |v_i\rangle\langle v_i|_{AB} (|\phi\rangle\langle\phi|_{AB})^{T_B} |v_j\rangle\langle v_j|_{AB}, \tag{3.68}$$

with span $(|\phi\rangle_{AB})$ is the kernel (or null space) of $\sigma_{AB}^{T_B}$ and $G_{i,j} = \lambda_i$ when $\lambda_i = \lambda_j$ and $G_{i,j} = (\lambda_i - \lambda_j) / (\ln \lambda_i - \ln \lambda_j)$ when $\lambda_i \neq \lambda_j$, where $\lambda_i$ and $|v_i\rangle_{AB}$ are the eigenvalues and eigenvectors of $\sigma_{AB}$, respectively.

The numerical estimation of the PPT relative entropy of entanglement with respect to the PPT states was introduced in Refs. [ZFG10, GZFG15], i.e., can be estimated by a Matlab program. Suppose that the estimation of $E_{R,PPT}(\rho)$ in Refs. [ZFG10, GZFG15] is $E_R^+(\rho)$, and the inequality $E_R^+(\rho) = D(\rho\|\sigma) \geq E_{R,PPT}(\rho)$ holds since the algorithm indeed provides a feasible PPT state $\sigma$ which is almost optimal. This algorithm is implemented by CVX [GB08] (a Matlab software for disciplined convex programming) and QETLAB [Nat16]. In low dimensions, this algorithm provides an estimation $E_R^+(\rho)$ with an absolute error smaller than $10^{-3}$, i.e. $E_{R,PPT}(\rho) + 10^{-3} \geq E_R^+(\rho) \geq E_{R,PPT}(\rho)$.

### 3.5.2 Nonadditivity of Rains bound

We first introduce a class of two-qubit states $\rho_r$ whose closest separable states can be derived by the result in [MI08]. Thus, the Rains bound of $\rho_r$ is exactly given. Then we apply the algorithm in Refs. [ZFG10, GZFG15] to demonstrate the gap between $\frac{1}{2}R\left(\rho^{\otimes 2}\right)$ and $R(\rho)$.

**Theorem 3.19.** *There exists a two-qubit state ρ such that*

$$R\left(\rho^{\otimes 2}\right) < 2R(\rho). \tag{3.69}$$

*Meanwhile,*

$$E_{R,PPT}^{\infty}(\rho) < R(\rho). \tag{3.70}$$

*Proof.* Firstly, we construct two-qubit states $\rho_r$ and $\sigma_r$ satisfying Eq. (3.67). Then we have $R(\rho_r) = D(\rho_r \| \sigma_r)$. Suppose that

$$\begin{aligned}\sigma_r = &\frac{1}{4}|00\rangle\langle 00| + \frac{1}{8}|11\rangle\langle 11| + r|01\rangle\langle 01| \\ &+ \left(\frac{5}{8} - r\right)|10\rangle\langle 10| + \frac{1}{4\sqrt{2}}\left(|01\rangle\langle 10| + |10\rangle\langle 01|\right).\end{aligned} \tag{3.71}$$

The positivity of $\sigma_r$ requires that $\frac{5-\sqrt{17}}{16} \leq r \leq \frac{5+\sqrt{17}}{16}$. Assume that $r \geq 5/8 - r$ and we can further choose $0.3125 \leq r \leq 0.57$ for simplicity.

Meanwhile, let us choose

$$\begin{aligned}\rho_r = &\frac{1}{8}|00\rangle\langle 00| + x|01\rangle\langle 01| + \frac{7-8x}{8}|10\rangle\langle 10| \\ &+ \frac{32r^2 - (6+32x)r + 10x + 1}{4\sqrt{2}}\left(|01\rangle\langle 10| + |10\rangle\langle 01|\right)\end{aligned} \tag{3.72}$$

with

$$x = r + \frac{32r^2 - 10r + 1}{256r^2 - 160r + 33} + \frac{(16r-5)\,y^{-1}}{32\ln(5/8 - y) - 32\ln(5/8 + y)}, \tag{3.73}$$

$$y = \left(4r^2 - 5r/2 + 33/64\right)^{1/2}. \tag{3.74}$$

It is clear that $\operatorname{Tr}\rho_r = 1$ and we set $0.3125 \leq r \leq 0.5480$ to ensure the positivity of $\rho_r$.

One can readily verify that $\rho_r = \sigma_r - 3G(\sigma_r)/2$. Therefore, $\sigma_r$ is the closest separable state (CSS) for $\rho_r$ and we have that

$$R(\rho_r) = E_{R,PPT}(\rho_r) = D(\rho_r \| \sigma_r). \tag{3.75}$$

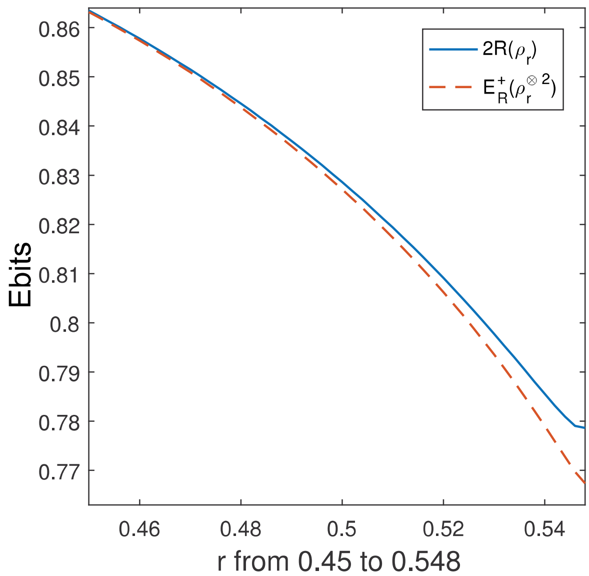


Figure 3.4: This plot demonstrates the difference between $2R(\rho_r)$ and $E_R^+(\rho_r^{\otimes 2})$ for $0.45 \le r \le 0.548$. The dashed line depicts $E_R^+(\rho_r^{\otimes 2})$ while the solid line depicts $2R(\rho_r)$.

In particular, let us first choose $r_0 = 0.547$, the Rains bound of $\rho_{r_0}$ is given by

$$R(\rho_{r_0}) = E_{R,PPT}(\rho_{r_0}) = D(\rho_{r_0} \| \sigma_{r_0}) \simeq 0.3891999. \tag{3.76}$$

Furthermore, applying the algorithm in Refs. [ZFG10, GZFG15], we can find a PPT state $\sigma_0$ such that

$$E_R^+(\rho_{r_0}^{\otimes 2}) = D(\rho_{r_0}^{\otimes 2} \| \sigma_0) \simeq 0.7683307. \tag{3.77}$$

The numerical value of relative entropy here is calculated based on the Matlab function "logm" [Mat13, AMHR13] and the function "Entropy" in QETLAB [Nat16]. In this case, the accuracy is guaranteed by the fact $\|e^{\text{logm}(\sigma_{r_0})} - \sigma_{r_0}\|_1 \le 10^{-16}$ and $\|e^{\text{logm}(\sigma_0)} - \sigma_0\|_1 \le 10^{-14}$. Noting that the difference between $2R(\rho_{r_0})$ and $E_R^+(\rho_{r_0}^{\otimes 2})$

is already $1.00691 \times 10^{-2}$, we have that

$$R\left(\rho_{r_0}^{\otimes 2}\right) \leq E_{R,PPT}\left(\rho_{r_0}^{\otimes 2}\right) \leq E_R^{+}\left(\rho_{r_0}^{\otimes 2}\right) < 2R\left(\rho_{r_0}\right). \tag{3.78}$$

It is also easy to observe that

$$E_{R,PPT}^{\infty}\left(\rho_{r_0}\right) \leq \frac{1}{2}E_{R,PPT}\left(\rho_{r_0}^{\otimes 2}\right) < R\left(\rho_{r_0}\right). \tag{3.79}$$

When $0.45 \leq r \leq 0.548$, we demonstrate the gap between $2R\left(\rho_r\right)$ and $E_R^{+}\left(\rho_r^{\otimes 2}\right)$ in Figure 3.4. □

Since Rains bound is not additive, the asymptotic Rains bound [Hay17c] can provide better upper bound on the distillable entanglement, i.e.,

$$E_{D,PPT}\left(\rho\right) \leq R^{\infty}\left(\rho\right) = \inf_{n\geq 1}\frac{1}{n}R\left(\rho^{\otimes n}\right) \leq R\left(\rho\right), \tag{3.80}$$

and the last inequality can be strict.

**Corollary 3.20.** *There exists bipartite quantum state ρ such that*

$$R^{\infty}\left(\rho\right) < R\left(\rho\right). \tag{3.81}$$

*As a consequence, the asymptotic Rains bound can provide a strictly better upper bound on $E_D$ than the Rains bound.*

## 3.6 Discussion

### 3.6.1 Summary

In this chapter, we have introduced an SDP-computable entanglement measure to evaluate the distillable entanglement and explored the deterministic entanglement distillation. This quantity enjoys additional good properties such as additivity and monotonicity under both general LOCC (or PPT) operations. We have also demonstrated the Rains bound is neither additive nor equal to the asymptotic relative entropy of entanglement by explicitly constructing a special class of mixed two-qubit states.

The main results in this chapter are summarized in the following box.

**Summary of Chapter 3**

(i) An additive SDP-computable entanglement measure:

$$E_{D,PPT}(\rho) \leq E_W(\rho) = R_{\max}(\rho) := \min_{\sigma \in PPT'} D_{\max}(\rho\|\sigma) = \min_{X \geq \rho} \log \|X^{T_B}\|_1.$$

(ii) Rains bound is not additive: $\exists \rho,\ R\left(\rho^{\otimes 2}\right) < 2R(\rho)$.

(iii) Rains bound and its regularization can be strictly smaller than the asymptotic PPT relative entropy of entanglement:

$$\exists\, \rho, \text{ such that } R^{\infty}(\rho) \leq R(\rho) < E_{R,PPT}^{\infty}(\rho). \tag{3.82}$$

(iv) Deterministic distillable entanglement:

$$E_{0,D,PPT}(\rho) \leq E_M(\rho) = \min_{X \geq P} \log \|X^{T_B}\|_{\infty}, \tag{3.83}$$

where $P$ is the projection onto supp $(\rho)$.

### 3.6.2 Outlook

In spite of a series of remarkable recent progress in the theory of entanglement (for reviews see, e.g., [PV07, Chr06, HHHH09, BŻ17]), many fundamental questions still remain open. It is of interest to determine whether the PPT distillable entanglement is given by $R^{\infty}$. Moreover, how to develop a resource theory of entanglement under one-way LOCC operations remains a challenging problem.

It is of great interest to explore the connections between non-local games [RV13, PV16] and fundamental entanglement measures [PV07] (e.g., distillable entanglement and entanglement cost). For example, a device-independent certification protocol of one-shot distillable entanglement was recently introduced in [AFB17].

A further direction is the distillation of secret key from quantum states [DW03]. It is important to develop both analytic and numerical methods to evaluate the rate of secret-key distillation [HHHO05, SBPC$^{+}$09] as well as the quantum key repeater rate [BCHW15] to extract private bits. Note that one-shot upper bounds for secret key were given in [WTB17]. However, due to the optimization over separable states, it is not clear whether the quantities are efficiently computable.

More generally, one may apply semidefinite optimization and the techniques in this chapter to investigate resource distillation and quantification in other quantum resource theories (e.g., [CG18, Reg18, SAP17, VHGE14, RBL18, GA15, GMN$^{+}$15]).

# Chapter 4

# Irreversibility of Asymptotic Entanglement Manipulation

## 4.1 Introduction

### 4.1.1 Background

In quantum information science, the resource theory of entanglement studies the transformation properties of entanglement under restricted classes of allowed operations. The irreversibility is crucial to this resource theory and it was sometimes argued to be the difference between entanglement and thermodynamics, as the Carnot cycle is reversible. When local operations and classical communication (LOCC) is available, the manipulation of entanglement is irreversible in the finite-copy regime. More precisely, the amount of pure entanglement that can be distilled from a finite number of copies of a state $\rho$ is usually strictly smaller than the amount of pure entanglement needed to prepare the same number of copies of $\rho$ [BDSW96]. Surprisingly, in the asymptotic limit of an arbitrarily large number of copies of the bipartite pure states, this process is known to be reversible [BBPS96]. But for mixed states, this asymptotic reversibility does not hold any more [VC01b, VC01a, VDC02, VWW04, CdOF11]. In particular, one requires a positive rate of pure state entanglement to generate the bound entanglement by LOCC [VC01b, YHHSR05], while it is well known that no pure state can be distilled from it [HHH98].

Various approaches have been considered to enlarge the class of operations to ensure reversible interconversion of entanglement in the asymptotic regime. A natural candidate is the class of quantum operations that completely preserve positivity of partial transpose (PPT) [Rai01], which include all quantum operations that can be implemented by LOCC. A remarkable result is that any state with a nonpositive

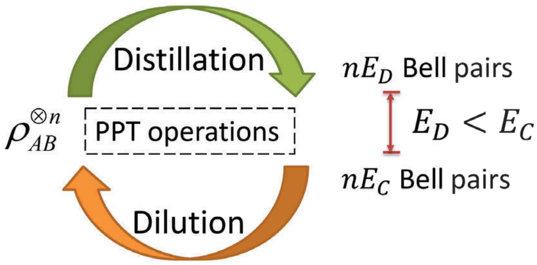


Figure 4.1: Illustration of entanglement irreversibility

partial transpose (NPT) is distillable under this class of operations [EVWW01]. This suggests the possibility of reversibility under PPT operations, and there are examples of mixed states which can be reversibly converted into pure states in the asymptotic setting, e.g., the class of antisymmetric states of arbitrary dimension [APE03]. However, the reversibility under PPT operations remained unsolved for over more than ten years [APE03, PV07, VWW04, BP10] and it was considered one of the major open problems in quantum information theory [Ple05b].

### 4.1.2 Outline

One approach to study the reversibility problem is to consider the transformations between the given state and Bell state, which naturally raise two fundamental entanglement measures: distillable entanglement and entanglement cost [BBPS96, BDSW96] (cf. 3.1.1). To be specific, the problem of reversibility under PPT operations is to determine whether distillable entanglement always coincides with entanglement cost under PPT operations, i.e.,

$$E_{C,\text{PPT}}(\rho) \overset{?}{=} E_{D,\text{PPT}}(\rho), \forall \rho \in \mathcal{S}(A \otimes B). \tag{4.1}$$

If $E_{C,\text{PPT}} = E_{D,\text{PPT}}$, then the transformation between any states under PPT operations is reversible. But this problem is still very difficult since for the general mixed states it is highly nontrivial to evaluate these two measures, both of which are given by a limiting procedure.

In this chapter, we resolve the open problem mentioned above by proving the irreversibility via the approach of semidefinite optimization.

Section 4.2 establishes an SDP lower bound for the entanglement cost under PPT operations. Using this new established lower bound, Section 4.3 further demonstrates the irreversibility of entanglement distillation under PPT operations via the standard rank-two state supported on the anti-symmetric subspace. As a byproduct, we also show that for this class of states, both the Rains bound and its regularization are strictly less than the asymptotic relative entropy of entanglement. So, in general, there is no unique entanglement measure for the manipulation of entanglement by PPT operations.

## 4.2 Lower bounds for entanglement cost

We note that the following entanglement measures we are going to use are summarized in Table 3.1 of entanglement measures.

### 4.2.1 Entanglement cost

Let us first recall the definition of entanglement cost using $\Omega$ operations:

$$E_{C,\Omega}\left(\rho_{AB}\right) = \inf\{r : \lim_{n\to\infty} \inf_{\Lambda\in\Omega} \|\rho_{AB}^{\otimes n} - \Lambda\left(\Phi\left(2^{rn}\right)\right)\|_1 = 0\}, \tag{4.2}$$

where $\Omega \in \{\text{1-LOCC}, \text{LOCC}, \text{SEP}, \text{PPT}\}$. The entanglement cost quantifies the amount of Bell states required to reconstruct the desired state using suitable operations. There are two known important lower bounds for entanglement cost, the squashed entanglement [CW04] and the asymptotic relative entropy of entanglement [AEJ+01].

#### Squashed entanglement

**Definition 4.1.** Given a bipartite state $\rho_{AB}$, Christandl and Winter [CW04] defined the squashed entanglement of $\rho_{AB}$ as

$$E_{\text{sq}}\left(A;B\right)_\rho =: \frac{1}{2}\inf_{\rho_{ABE}}\{I\left(A:B|E\right)_\rho : \rho_{AB} = \text{Tr}_E\,\rho_{ABE}\}, \tag{4.3}$$

where $I\left(A:B|E\right) := S\left(\rho_{AE}\right) + S\left(\rho_{BE}\right) - S\left(\rho_E\right) - S\left(\rho_{ABE}\right)$ is the conditional quantum mutual information of the extended state $\rho_{ABE}$. Alternatively, it can be represented as

$$E_{\text{sq}}\left(A;B\right)_\rho = \frac{1}{2}\inf_{\mathcal{M}_{E\to E'}} I\left(A:B|E'\right), \tag{4.4}$$

where the infimum is taken over all squashing channels $\mathcal{M}_{E\to E'}$ taking the $E$ system of the purification $\phi^{\rho}_{ABE}$ to a system $E'$ of arbitrary dimension.

The *squashed entanglement* in Eq. (4.4) can be interpreted as the environment $E$ holding some purifying system of $\rho_{AB}$, and then squashing the correlations between $A$ and $B$ as much as possible by applying a channel $\mathcal{M}_{E\to E'}$ that minimizes the conditional mutual information $I\left(A;B|E'\right)$. It has various nice properties such as monotonicity under LOCC operations, additivity under tensor product, continuity, and normalization for the private state [CW04] (see also [Wil16] for approximate normalization of $E_{sq}$ for private states). Importantly, the squashed entanglement lies between the distillable entanglement and entanglement cost [CW04]: for any bipartite state $\rho \in \mathcal{S}(A \otimes B)$, it holds that

$$E_D(\rho) \leq K_D(\rho) \leq E_{\mathrm{sq}}(\rho) \leq E_C(\rho), \tag{4.5}$$

where $K_D(\rho)$ is the optimal number of private bits that can be generated from $\rho$ via LOCC operations in the i.i.d. limit.

**Asymptotic PPT relative entropy of entanglement**

Let us recall the definition of the asymptotic PPT relative entropy of entanglement (PPT AREE): given a biparitite state $\rho$, its PPT AREE is given by

$$E_{R,PPT}^{\infty}(\rho) = \inf_{n\geq 1} E_{R,PPT}\left(\rho^{\otimes n}\right)/n. \tag{4.6}$$

A useful fact is that $E_{R,PPT}^{\infty}$ lies between the PPT distillable entanglement and the PPT entanglement cost [Hay17c]: for any bipartite state $\rho \in \mathcal{S}(A \otimes B)$, it holds that

$$E_{D,PPT}(\rho) \leq E_{R,PPT}^{\infty}(\rho) \leq E_{C,PPT}(\rho). \tag{4.7}$$

### 4.2.2 Lower bounds for entanglement cost

The main difficulty of the problems above is that the regularized quantities are usually extremely difficult to determine or estimate. To figure out whether Rains bound always coincides with $E_R^{\infty}$, one necessarily has to evaluate $E_R^{\infty}(\rho)$ of an explicit state $\rho$. The problem of irreversibility under PPT operations is more intractable: one not only has to evaluate the PPT distillable entanglement, but also needs to determine the PPT entanglement cost.

Since computing the entanglement cost of a bipartite state is very difficult, we introduce an efficiently computable lower bound for evaluating entanglement cost.

Our key tool is an efficiently computable additive lower bound for the asymptotic REE. In the one-copy case, we need to do some relaxations of the minimization of $D(\rho\|\sigma)$ with respect to PPT states. By applying properties of the quantum relative entropy, we can relax the problem to minimizing $-\log \operatorname{Tr} P\sigma$ over all PPT state $\sigma$, where $P$ is the projection onto $\operatorname{supp}(\rho)$. Noting that this is SDP-computable, we can use SDP techniques to obtain the following bound

$$E_\eta(\rho) = \max\{-\log \|Y^{T_B}\|_\infty : -Y \leq P^{T_B} \leq Y\}. \tag{4.8}$$

Interestingly, $E_\eta(\cdot)$ is additive under tensor product, i.e.,

$$E_\eta(\rho_1 \otimes \rho_2) = E_\eta(\rho_1) + E_\eta(\rho_2),$$

so we can overcome the difficulty of estimating the regularized relative entropy of entanglement. The additivity of $E_\eta(\cdot)$ can be proved by utilizing the duality theory of semidefinite programming. A complete proof of the additivity of $E_\eta(\cdot)$ is presented in the following Lemma. 4.3.

**Proposition 4.2.** *For any bipartite state $\rho$,*

$$E^{\infty}_{R,PPT}(\rho) \geq E_\eta(\rho). \tag{4.9}$$

*Proof.* Firstly, let us introduce a CPTP map by $\mathcal{N}(\tau) = P\tau P + (\mathbb{1} - P)\tau(\mathbb{1} - P)$. Then we have that

$$\begin{aligned} D(\rho\|\sigma) &\geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \\ &= D(\rho\|P\sigma P/\operatorname{Tr} P\sigma P) - \log \operatorname{Tr} P\sigma \\ &\geq -\log \operatorname{Tr} P\sigma, \end{aligned} \tag{4.10}$$

where the first inequality is from the monotonicity of quantum relative entropy [Lin75, Uhl77] and the second inequality is due to the non-negativity of quantum relative entropy. (After we finished the proof, we found that this step already appeared in the Lemma 10 of [Dat09].)

Then, we can transform the original optimization problem to an SDP problem:

$$\min_{\sigma \in PPT(A:B)} D(\rho\|\sigma) \geq \min_{\sigma \in PPT(A:B)} -\log \operatorname{Tr} P\sigma. \tag{4.11}$$

Secondly, utilizing the weak duality of SDP, we can see that

$$\max_{\sigma \in PPT(A:B)} \operatorname{Tr} P\sigma \leq \min\{t : Y^{T_B} \leq t\mathbb{1}, P^{T_B} \leq Y\} \tag{4.12}$$

$$\leq \min\{t : -t\mathbb{1} \leq Y^{T_B} \leq t\mathbb{1}, -Y \leq P^{T_B} \leq Y\} \tag{4.13}$$

$$= \min\{\|Y^{T_B}\|_\infty : -Y \leq P^{T_B} \leq Y\}. \tag{4.14}$$

Thus,

$$E_{R,PPT}(\rho) \geq -\log \max_{\sigma \in PPT(A:B)} \operatorname{Tr} P\sigma \geq E_\eta(\rho). \tag{4.15}$$

Finally, noting that $E_\eta(\rho)$ is additive, we have that

$$\begin{aligned} E_{R,PPT}{}^{\infty}(\rho) &= \inf_{n\geq 1} E_R\left(\rho^{\otimes n}\right)/n \\ &\geq \inf_{n\geq 1} E_\eta\left(\rho^{\otimes n}\right)/n = E_\eta(\rho). \end{aligned}$$

□

**The additivity of $E_\eta$**

To see the additivity of $E_\eta(\rho)$, we reformulate it as $E_\eta(\rho) = -\log\eta(\rho)$, where

$$\begin{aligned} \eta(\rho) = \min\ & t \\ \text{s.t.}\ & -Y_{AB} \leq P_{AB}^{T_B} \leq Y_{AB}, \\ & -t\mathbb{1} \leq Y_{AB}^{T_B} \leq t\mathbb{1}, \end{aligned} \tag{4.16}$$

where $P_{AB}$ is the projection onto supp $(\rho)$.

The dual SDP of $\eta(\rho)$ is given by

$$\begin{aligned} \eta(\rho) = \max\ & \operatorname{Tr} P_{AB}\left(V_{AB} - F_{AB}\right)^{T_B}, \\ \text{s.t.}\ & V_{AB} + F_{AB} \leq \left(W_{AB} - X_{AB}\right)^{T_B}, \\ & \operatorname{Tr}\left(W_{AB} + X_{AB}\right) \leq 1, \\ & V_{AB}, F_{AB}, W_{AB}, X_{AB} \geq 0. \end{aligned} \tag{4.17}$$

The optimal values of the primal and the dual SDPs above coincide by strong duality.

**Lemma 4.3.** *For any two bipartite states $\rho_1$ and $\rho_2$, we have that*

$$E_\eta(\rho_1 \otimes \rho_2) = E_\eta(\rho_1) + E_\eta(\rho_2).$$

*Proof.* On one hand, suppose that the optimal solution to SDP (4.16) of $\eta(\rho_1)$ and $\eta(\rho_2)$ are $\{t_1, Y_1\}$ and $\{t_2, Y_2\}$, respectively. It is easy to see that

$$\begin{aligned} Y_1 \otimes Y_2 + P_1^{T_B} \otimes P_2^{T_{B'}} &= \frac{1}{2}[\left(Y_1 + P_1^{T_B}\right) \otimes \left(Y_2 + P_2^{T_{B'}}\right) + \left(Y_1 - P_1^{T_B}\right) \otimes \left(Y_2 - P_2^{T_{B'}}\right)] \geq 0, \\ Y_1 \otimes Y_2 - P_1^{T_B} \otimes P_2^{T_{B'}} &= \frac{1}{2}[\left(Y_1 + P_1^{T_B}\right) \otimes \left(Y_2 - P_2^{T_{B'}}\right) + \left(Y_1 - P_1^{T_B}\right) \otimes \left(Y_2 + P_2^{T_{B'}}\right)] \geq 0. \end{aligned}$$

Then, we have that $-Y_1 \otimes Y_2 \leq P_1^{T_B} \otimes P_2^{T_{B'}} \leq Y_1 \otimes Y_2$. Moreover,

$$\|Y_1^{T_B} \otimes Y_2^{T_{B'}}\|_\infty \leq \|Y_1^{T_B}\|_\infty \|Y_2^{T_{B'}}\|_\infty \leq t_1 t_2,$$

which means that

$$-t_1t_2\mathbb{1} \leq Y_1^{T_B} \otimes Y_2^{T_{B'}} \leq t_1t_2\mathbb{1}. \tag{4.18}$$

Therefore, $\{t_1t_2, Y_1 \otimes Y_2\}$ is a feasible solution to the SDP (4.16) of $\eta(\rho_1 \otimes \rho_2)$, which means that

$$\eta(\rho_1 \otimes \rho_2) \leq t_1t_2 = \eta(\rho_1)\eta(\rho_2). \tag{4.19}$$

On the other hand, suppose that the optimal solutions to SDP (4.17) of $\eta(\rho_1)$ and $\eta(\rho_2)$ are $\{V_1, F_1, W_1, X_1\}$ and $\{V_2, F_2, W_2, X_2\}$, respectively. Assume that

$$V = V_1 \otimes V_2 + F_1 \otimes F_2, \tag{4.20}$$

$$F = V_1 \otimes F_2 + F_1 \otimes V_2, \tag{4.21}$$

$$W = W_1 \otimes W_2 + X_1 \otimes X_2, \tag{4.22}$$

$$X = W_1 \otimes X_2 + X_1 \otimes W_2. \tag{4.23}$$

It is easy to see that

$$V + F = (V_1 + F_1) \otimes (V_2 + F_2) \tag{4.24}$$

$$\leq (W_1 - X_1)^{T_B} \otimes (W_2 - X_2)^{T_{B'}} = (W - X)^{T_{BB'}} \tag{4.25}$$

and $\mathrm{Tr}(W + X) = \mathrm{Tr}(W_1 + X_1) \otimes (W_2 + X_2) \leq 1$. Thus, $\{V, F, W, X\}$ is a feasible solution to the SDP (4.17) of $\eta(\rho_1 \otimes \rho_2)$. This means that

$$\eta(\rho_1 \otimes \rho_2) \geq \mathrm{Tr}(P_1 \otimes P_2)(V - F)^{T_{BB'}} \tag{4.26}$$

$$= \mathrm{Tr}(P_1 \otimes P_2)\left((V_1 - F_1)^{T_B} \otimes (V_2 - F_2)^{T_{B'}}\right) \tag{4.27}$$

$$= \eta(\rho_1)\eta(\rho_2). \tag{4.28}$$

Hence, combining Eq. (4.19) and Eq. (4.26), we have that

$$\eta(\rho_1 \otimes \rho_2) = \eta(\rho_1)\eta(\rho_2), \tag{4.29}$$

which directly leads to $E_\eta(\rho_1 \otimes \rho_2) = E_\eta(\rho_1) + E_\eta(\rho_2)$. □

**Lower bound for the regularized Rains bound**

Let us recall the upper bound $E_M$ (see Eq. (3.46)) on the deterministic distillable entanglement: for a bipartite quantum state $\rho$,

$$E_M(\rho) = -\log M(\rho) = -\log \max \ \operatorname{Tr} P_{AB} V_{AB}, \quad \text{s.t.} \ \operatorname{Tr} |V_{AB}^{T_B}| = 1, V_{AB} \geq 0, \tag{4.30}$$

**Proposition 4.4.** *For any bipartite state $\rho \in \mathcal{S}(A \otimes B)$,*

$$E_M(\rho) \leq R^\infty(\rho) \leq E_{C,PPT}(\rho).$$

*Proof.* Via similar techniques in Proposition 4.2, one can show that

$$E_M(\rho) \leq R(\rho). \tag{4.31}$$

Noting that $E_M(\cdot)$ is additive, we have that

$$E_M(\rho) \leq \inf_{n \geq 1} \frac{1}{n} R\left(\rho^{\otimes n}\right) = R^\infty(\rho).$$

Finally, it is clear that

$$E_M(\rho) \leq R^\infty(\rho) \leq E_{R,PPT}^\infty(\rho) \leq E_{C,PPT}(\rho),$$

where the last inequality is from [Hay17c]. □

**Remark 4.5.** As an application of this lower bound, one can also give an SDP lower bound for *the entanglement cost of quantum channels* [BBCW13], i.e. the rate of entanglement (ebits) needed to asymptotically simulate a quantum channel $\mathcal{N}$ with free classical communication.

## 4.3 Irreversibility of PPT entanglement manipulation

In this section, we focus on the following standard rank-two states supported on the three by three anti-symmetric subspace:

$$\rho_v = \frac{1}{2}\left(|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|\right), \tag{4.32}$$

where

$$|v_1\rangle = 1/\sqrt{2}\left(|01\rangle - |10\rangle\right), \tag{4.33}$$

$$|v_2\rangle = 1/\sqrt{2}\left(|02\rangle - |20\rangle\right). \tag{4.34}$$

The projection onto $\mathrm{supp}\left(\rho_v\right)$ is $P_v = |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2|$. The authors of [CD09] showed that this state can be transformed into some $2 \otimes 2$ pure entangled state by a suitable separable operation while no finite-round LOCC protocol can do that.

Our main result of this section is as follows.

**Theorem 4.6.** *For the state $\rho_v$, we have*

$$E_{D,PPT}\left(\rho_v\right) = R^{\infty}\left(\rho_v\right) < E^{\infty}_{R,PPT}\left(\rho_v\right) = E_{C,PPT}\left(\rho_v\right). \tag{4.35}$$

To see this, we first prove $E^{\infty}_{R,PPT}\left(\rho_v\right) = E_{C,PPT}\left(\rho_v\right) = 1$ in Proposition 4.7 and then show $E_{D,PPT}\left(\rho_v\right) = R^{\infty}\left(\rho_v\right) = \log\left(1 + 1/\sqrt{2}\right)$ in Proposition 4.8.

This result indicates that the asymptotic entanglement manipulation of $\rho_v$ under PPT operations is irreversible, thus resolving a long-standing open problem in quantum information theory [APE03, HOH02, Ple05b]. Furthermore, it also answers another open problem in [PV07] by showing a nonzero gap between the regularized Rains bound and the PPT AREE of $\rho_v$.

### 4.3.1 PPT entanglement cost of $\rho_v$

Applying the lower bound $E_{\eta}\left(\rho\right)$, we are now ready to show that the PPT entanglement cost of $\rho_v$ is still one ebit.

**Proposition 4.7.**

$$E_{C,PPT}\left(\rho_v\right) = E^{\infty}_{R,PPT}\left(\rho_v\right) = 1. \tag{4.36}$$

*Proof.* Firstly, let us choose a projector

$$Q = |01\rangle\langle 01| + |10\rangle\langle 10| + |02\rangle\langle 02| + |20\rangle\langle 20|. \tag{4.37}$$

Then we can easily prove that

$$E_{\eta}\left(\rho_v\right) \le E^{\infty}_{R,PPT}\left(\rho_v\right) \le 1 \tag{4.38}$$

by choosing a PPT state $\tau = Q/4$ such that $S\left(\rho_v \| \tau\right) = 1$.

Secondly, we are going to prove $E_\eta(\rho_v) \geq 1$. To see this, suppose that

$$Y = \frac{1}{2}\left(Q + |00\rangle\langle 00| + (|11\rangle + |22\rangle)(\langle 11| + \langle 22|)\right). \tag{4.39}$$

Noting that

$$Y - P_{AB}^{T_B} = \frac{1}{2}\left(|00\rangle + |11\rangle + |22\rangle\right)\left(\langle 00| + \langle 11| + \langle 22|\right),$$

it is clear that $P_{AB}^{T_B} \leq Y_{AB}$. Moreover,

$$Y + P_{AB}^{T_B} = Q + \frac{1}{2}\left(|00\rangle - |11\rangle - |22\rangle\right)\left(\langle 00| - \langle 11| - \langle 22|\right), \tag{4.40}$$

which means that $P_{AB}^{T_B} \geq -Y$.

Then $Y_{AB}$ is a feasible solution to the SDP (4.8) of $E_\eta(\rho_v)$. Thus,

$$E_\eta(\rho_v) \geq -\log \|Y^{T_B}\|_\infty = -\log 1/2 = 1, \tag{4.41}$$

and we can conclude that

$$E_\eta(\rho_v) = E_{R,PPT}^\infty(\rho_v) = 1. \tag{4.42}$$

Finally, it is clear that one Bell pair is sufficiently to prepare an exact copy of $\rho$ by LOCC. Combining with the above lower bounds, we have that

$$1 = E_\eta(\rho_v) \leq E_{R,PPT}^\infty(\rho_v) \leq E_{C,PPT}(\rho_v) \leq E_C(\rho) \leq 1. \tag{4.43}$$

□

It is worth pointing out that our approach to evaluating the PPT entanglement cost is to combine the lower bound $E_\eta$ and the upper bound $E_C$. This result provides a new proof of the entanglement cost of the rank-two $3 \otimes 3$ antisymmetric state in [Yur03]. Moreover, our result is stronger as it shows that the entanglement cost under PPT operations of this state is still one ebit.

### 4.3.2 PPT distillable entanglement of $\rho_v$

We are going to evaluate the PPT distillable entanglement of $\rho_v$ via the Rains bound and the SDP characterization of the one-copy PPT deterministic distillable entanglement in Eq. (3.42).

**Proposition 4.8.**

$$E_{D,PPT}\left(\rho_v\right) = R^{\infty}\left(\rho_v\right) = \log\left(1 + 1/\sqrt{2}\right). \tag{4.44}$$

*Proof.* Firstly, we need to introduce upper and lower SDP bounds to evaluate the entanglement of cost and the regularized Rains bound. The logarithmic negativity [VW02, Ple05a] is an upper bound on PPT distillable entanglement, i.e., $E_N\left(\rho\right) = \log \|\rho^{T_B}\|_1$.

Let us recall the one-copy PPT deterministic distillable entanglement:

$$\begin{aligned} E^{(1)}_{0,D,PPT}\left(\rho\right) = \max -\log \|R^{T_B}_{AB}\|_{\infty}, \\ \text{s.t. } P_{AB} \leq R_{AB} \leq \mathbb{1}_{AB}. \end{aligned} \tag{4.45}$$

where $P_{AB}$ is the projection onto supp $(\rho)$, the support of $\rho$. Note that supp $(\rho)$ is defined to be the subspace spanned by the eigenvectors of $\rho$ with positive eigenvalues. Clearly $E^{(1)}_{0,D,PPT}\left(\rho\right)$ is efficiently computable via SDP, and for a general bipartite state $\rho$ we have

$$E^{(1)}_{0,D,PPT}\left(\rho\right) \leq E_D\left(\rho\right) \leq R^{\infty}\left(\rho\right) \leq E_N\left(\rho\right),$$

which is very helpful to determine the exact values of PPT distillable entanglement for some states.

Now one can calculate that $\|\rho_v^{T_B}\|_1 = 1 + 1/\sqrt{2}$. Then we have

$$R^{\infty}\left(\rho_v\right) \leq E_N\left(\rho_v\right) \leq \log\left(1 + 1/\sqrt{2}\right). \tag{4.46}$$

On the other hand, let

$$R_{AB} = \left(3 - 2\sqrt{2}\right)\left(|r_1\rangle\langle r_1| + |r_2\rangle\langle r_2|\right) + P_{AB}$$

with $|r_1\rangle = \left(|01\rangle + |10\rangle\right)/\sqrt{2}$ and $|r_2\rangle = \left(|02\rangle + |20\rangle\right)/\sqrt{2}$. It is easy to check that $P_{AB} \leq R_{AB} \leq \mathbb{1}$, which means that $R_{AB}$ is a feasible solution to SDP (4.45) of $E^{(1)}_{0,D,PPT}\left(\rho_v\right)$. Therefore,

$$E^{(1)}_{0,D,PPT}\left(\rho_v\right) \geq -\log \|R^{T_B}_{AB}\|_{\infty} = \log\left(1 + 1/\sqrt{2}\right). \tag{4.47}$$

Finally, combining Eq. (4.46) and Eq. (4.47), we have that

$$E_{D,PPT}\left(\rho_v\right) = R^{\infty}\left(\rho_v\right) = \log\left(1 + 1/\sqrt{2}\right). \tag{4.48}$$

□

### 4.3.3 General irreversibility under PPT operations

We have shown the irreversibility of the asymptotic entanglement manipulation of $\rho_v$ under PPT operations. One can use similar technique to prove the irreversibility for any $\rho$ with spectral decomposition

$$\rho = p|u_1\rangle\langle u_1| + (1-p)\,|u_2\rangle\langle u_2| \;\; (0 < p < 1)\,,$$

where $|u_1\rangle = (|01\rangle - |10\rangle)\,/\sqrt{2}$, $|u_2\rangle = (|ab\rangle - |ba\rangle)\,/\sqrt{2}$. Interestingly, it holds that $E_{D,PPT}(\rho) < 1 = E_{C,PPT}(\rho)$. (See [WD17a] for a detailed proof). More generally, we can provide a sufficient condition for the irreversibility under PPT operations and construct a general class of such states.

It was shown in Chapter 3 that

$$E_{D,PPT}(\rho) \le E_W(\rho) \le E_N(\rho)\,,$$

and the second equality can be strict. It is straightforward to see that if $E_W(\rho) < E_\eta(\rho)$, then $E_{D,PPT}(\rho) < E_{C,PPT}(\rho)$.

Indeed, we can obtain a more specific condition if we use logarithmic negativity $E_N$ instead of $E_W$. That is, for a bipartite state $\rho$, if there is a Hermitian matrix $Y$ such that $P_{AB}^{T_B} \pm Y \ge 0$ and $\|\rho^{T_B}\|_1 < \|Y^{T_B}\|_\infty^{-1}$, we have $E_{D,PPT}(\rho) < E_{C,PPT}(\rho)$.

We further show the irreversibility in asymptotic manipulations of entanglement under PPT operations by a class of $3 \otimes 3$ states defined by

$$\rho^{(\alpha)} = (|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)\,/2, \tag{4.49}$$

where $|\psi_1\rangle = \sqrt{\alpha}|01\rangle - \sqrt{1-\alpha}|10\rangle$ and $|\psi_2\rangle = \sqrt{\alpha}|02\rangle - \sqrt{1-\alpha}|20\rangle$ with $0.42 \le \alpha \le 0.5$. Then the projection onto the range of $\rho^{(\alpha)}$ is $P_{AB} = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|$. One can easily calculate that

$$E_W\left(\rho^{(\alpha)}\right) \le \log \|\left(\rho^{(\alpha)}\right)^{T_B}\|_1 = \log\left(1 + \sqrt{2\alpha\,(1-\alpha)}\right).$$

We then construct a feasible solution to the dual SDP (4.8) of $E_\eta\left(\rho^{(\alpha)}\right)$, i.e.,

$$\begin{aligned} Y = &\ \alpha\,(|01\rangle\langle 01| + |02\rangle\langle 02|) + (1-\alpha)\,(|10\rangle\langle 10| + |20\rangle\langle 20|) \\ &+ \sqrt{\alpha\,(1-\alpha)}\,(|00\rangle\langle 00| + |11\rangle\langle 11| + |22\rangle\langle 22| + |11\rangle\langle 22| + |22\rangle\langle 11|)\,. \end{aligned} \tag{4.50}$$

It can be checked that $-Y \le P_{AB}^{T_B} \le Y$ and $\|Y^{T_B}\|_\infty \le 1-\alpha$. Thus, $E_\eta\left(\rho^{(\alpha)}\right) \ge -\log(1-\alpha)$.

When $0.42 \le \alpha \le 0.5$, it is easy to check that

$$-\log(1-\alpha) > \log\left(1+\sqrt{2\alpha(1-\alpha)}\right). \tag{4.51}$$

Therefore,

$$E_{D,PPT}\left(\rho^{(\alpha)}\right) \le E_W\left(\rho^{(\alpha)}\right) < E_\eta\left(\rho^{(\alpha)}\right) \le E_{C,PPT}\left(\rho^{(\alpha)}\right). \tag{4.52}$$

## 4.4 Discussion

### 4.4.1 Summary

In this chapter, we have explored semidefinite programs to evaluate the entanglement cost of bipartite entanglement, demonstrated irreversibility of entanglement theory under PPT operations, and established separations between fundamental entanglement measures.

The important results in this chapter are summarized in the following box.

**Summary of Chapter 4**

(i) Lower bound for the entanglement cost of a bipartite state $\rho \in \mathcal{S}(A \otimes B)$:

$$\begin{aligned} E_C(\rho) \ge E_{C,PPT}(\rho) \ge E_\eta(\rho) = \min\ & t \\ \text{s.t.}\ & -Y \le P^{T_B} \le Y, \\ & -t\mathbb{1} \le Y^{T_B} \le t\mathbb{1}, \end{aligned} \tag{4.53}$$

where $P$ is the projection onto $\operatorname{supp}(\rho)$.

(ii) Rains bound can be strictly larger than the asymptotic PPT relative entropy of entanglement:

$$\exists\, \rho \in \mathcal{S}(A \otimes B), \text{ such that } R(\rho) > E_{R,PPT}^{\infty}(\rho). \tag{4.54}$$

(iii) Asymptotic entanglement manipulation under PPT operations is irreversible:

$$\exists\, \rho \in \mathcal{S}(A \otimes B), \text{ such that } E_{C,PPT}(\rho) > E_{D,PPT}(\rho). \tag{4.55}$$

### 4.4.2 Outlook

The lower bound $E_\eta$ for entanglement cost is in general not tight and could be sometimes smaller than distillable entanglement. How to further refine the lower bound

$E_\eta$ remains an interesting problem. It will also be interesting to study the entanglement cost from the view of non-local games, see e.g., [AFY17].

By considering all asymptotically non-entangling transformations, a reversible theory of entanglement was obtained in Refs. [BP08, BP10]. Given the fact that the entanglement theory under PPT operations is not reversible, a very interesting question remains open: what is the smallest class of operations that permits a reversible entanglement theory?

Moreover, can we develop a non-asymptotic resource theory to efficiently evaluate the entanglement dilution with finite resources? For example, see [BD11b] for the study about the quantification of one-shot entanglement cost.

Finally, we end this part on entanglement theory with the following zoo of entanglement measures. The contributions of Chapters 3-4 are highlighted. The irreversibility of entanglement under PPT operations can be seen via the gap between $E^\infty_{R,PPT}$ and $R^\infty$.

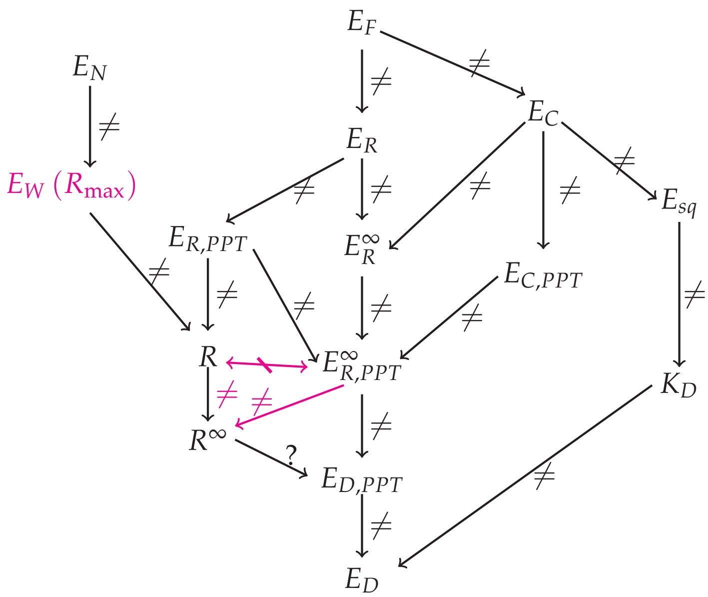


Figure 4.2: Zoo of entanglement measures. An arrow $E_A \longrightarrow E_B$ indicates that $E_A(\rho) \geq E_B(\rho)$ for any bipartite state $\rho$. $E_A \longleftrightarrow E_B$ indicates that $E_A$ and $E_B$ are not comparable. The detailed definitions of these entanglement measures can be found in Table 3.1.

# Part II

# Quantum Shannon Theory

# Chapter 5

# Classical communication via quantum channels

## 5.1 Introduction

The reliable transmission of classical information with quantum systems is central to the theory of quantum information. A natural question that arises is what are the maximum communication rates achievable over noisy communication channels? In 1948, Shannon stressed the nature of communication in his seminal work "A Mathematical Theory of Communication" [Sha48]:

> C. E. Shannon: *"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."*

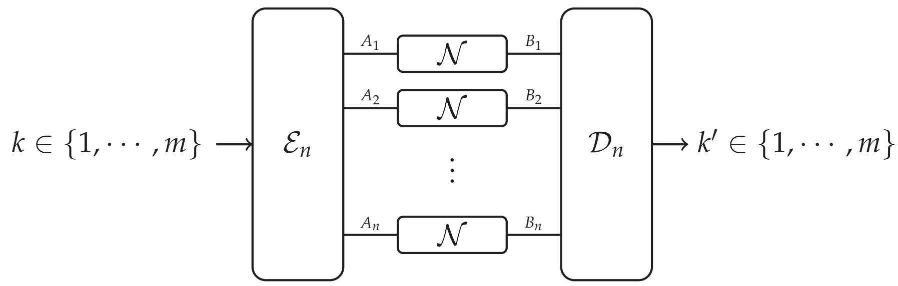


Figure 5.1: The sender (Alice) encodes the messages with an encoding operation $\mathcal{E}_n$ and then sends them trough the channel $\mathcal{N}^{\otimes n}$ to the receiver (Bob). Bob collects these registers and then applies a decoding operation $\mathcal{D}_n$ to extract the messages Alice sent.

The core of this view is the *channel* formalism, where any noisy communication line is depicted as a stochastic map connecting input signals selected by the sender of the message (Alice), to their corresponding output counterparts accessible to the receiver of the messages (Bob).

Besides the mathematical theory of communication, Rolf Landauer stressed the fact that information is physical [Lan96]: "Information is inevitably tied to a physical representation and therefore to restrictions and possibilities related to the laws of physics and the parts available in the universe." This view answers why we need an information theory based on quantum mechanics. It is worth noting that the quantum information theory not only extends but also completes the classical information theory (for reviews, see, e.g., [BS98, Wil17, Hay17c, CGLM14]).

### 5.1.1 Background

Any physical process can be represented as a quantum channel. The goal of building a classical communication system is to simulate a noiseless channel by using the actual noisy channel. In particular, the sender is able to apply any local physical operation to encode the messages to input to the channel. And the receiver may apply any local physical operation to decode the outputs of the channel.

The classical capacity of a noisy quantum channel is the highest rate at which it can convey classical information reliably over asymptotically many uses of the channel. (We refer to Eq. (5.7) for a formal definition.) The Holevo-Schumacher-Westmoreland (HSW) theorem [Hol73, Hol98b, SW97] gives a full characterization of the classical capacity of quantum channels.

**Theorem 5.1** (Classical capacaity (HSW theorem))**.** *Given a quantum channel $\mathcal{N}$, its classical capacity is given by the regularized Holevo capacity:*

$$C(\mathcal{N}) := \sup_{n\geq 1} \frac{\chi\left(\mathcal{N}^{\otimes n}\right)}{n}, \tag{5.1}$$

*where $\chi(\mathcal{N})$ is the Holevo capacity of the channel $\mathcal{N}$, defined as*

$$\chi(\mathcal{N}) := \max_{\{(p_i,\rho_i)\}} S\left(\sum_i p_i \mathcal{N}(\rho_i)\right) - \sum_i p_i S\left(\mathcal{N}(\rho_i)\right), \tag{5.2}$$

*and $\{(p_i,\rho_i)\}_i$ is an ensemble of quantum states on $A$.*

For certain classes of quantum channels (depolarizing channel [Kin03], erasure channel [BDS97], unital qubit channel [Kin02], etc. [AHW00, DHS04, Fuk05, KWW12]), the classical capacity of the channel is equal to the Holevo capacity, since their Holevo capacities are all additive. However, for a general quantum channel, our understanding of the classical capacity is still limited. The work of Hastings [Has09] shows that the Holevo capacity is generally not additive, thus the regularization in Eq. (5.1) is necessary in general. Since the complexity of computing the single-letter Holevo capacity of a channel is NP-complete [BS07], the regularized Holevo capacity of a general quantum channel is notoriously difficult to calculate and it is not even clear whether this regularized quantity is computable in the Turing sense. (See [SSMR16] for approaches to approximating Holevo information of a quantum channel.) Even for basic quantum channels such as the qubit amplitude damping channel, the classical capacity remains unknown.

#### Strong converse vs. weak converse

The converse part of the HSW theorem states that if the communication rate exceeds the capacity, then the error probability of any coding scheme cannot approach zero in the limit of many channel uses. This kind of "weak" converse suggests the possibility for one to increase communication rates by allowing an increased error probability. A *strong converse property* leaves no such room for the trade-off; i.e., the error probability necessarily converges to one in the limit of many channel uses whenever the rate exceeds the capacity of the channel. For classical channels, the strong converse property for the classical capacity was established by Wolfowitz [Wol78]. For quantum channels, the strong converse property for the classical capacity has been confirmed for several classes of channels [ON99, Win99, KW09, WW13, WWY14]. Winter [Win99] and Ogawa and Nagaoka [ON99] independently established the strong converse property for the classical capacity of classical-quantum channels. Koenig and Wehner [KW09] proved the strong converse property for particular covariant quantum channels. Recently, for the entanglement-breaking and Hadamard channels, the strong converse property was proved by Wilde, Winter and Yang [WWY14]. Moreover, the strong converse properties for the pure-loss bosonic channel and the quantum erasure channel were proved in [WW13] and [WW14], respectively. Unfortunately, for a general quantum channel, less is known about the strong converse property of the classical capacity, and it remains open whether this property holds for all quantum channels.

#### Strong converse bound

A *strong converse bound* for the classical capacity is a quantity such that the success probability of transmitting classical messages vanishes exponentially fast as the number of channel uses increases if the rate of communication exceeds this quantity, which forbids the trade-off between rate and error in the limit of many channel uses.

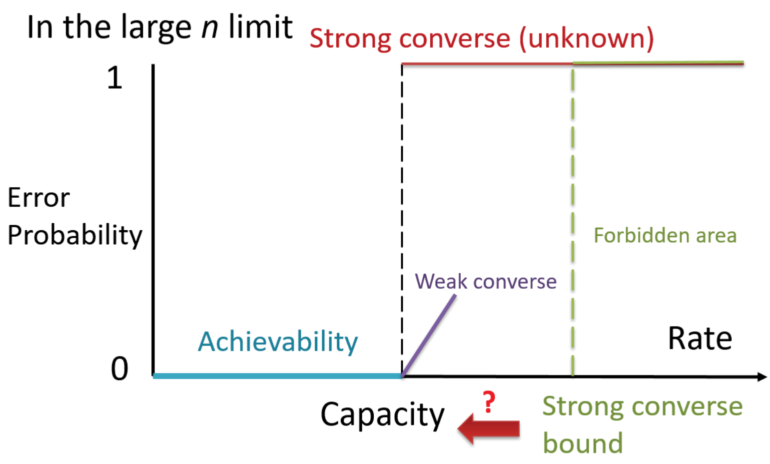


Figure 5.2: Strong vs. weak converse.

**Non-asymptotic classical communication**

Another fundamental problem, of both theoretical and practical interest, is the trade-off between the channel uses, communication rate and error probability in the non-asymptotic (or finite blocklength) regime. In a realistic setting, the number of channel uses is necessarily limited in quantum information processing. Therefore one has to make a trade-off between the transmission rate and error tolerance. Note that one only needs to study one-shot communication over the channel since it can correspond to a finite blocklength and one can also study the asymptotic capacity via the finite blocklength approach. The study of finite blocklength regime has recently garnered great interest in classical information theory (see e.g., [PPV10, Hay09, Mat12]) as well as in quantum information theory (see e.g., [MW14a, WR12, RR11, TH13, BCR11, LM15, TT15, BDL16, Tom16, TBR16]). For classical channels, Polyanskiy, Poor, and Verdú [PPV10] derived the finite blocklength converse bound via hypothesis testing and Matthews [Mat12] provided an alternative proof of this converse bound via classical no-signalling codes. For classical-quantum channels, the one-shot converse and achievability bounds were given in [MD09, WR12, RR11]. Recently, the one-shot converse bounds for entanglement-assisted and unassisted codes were given in [MW14a], which generalizes the hypothesis testing approach in [PPV10] to quantum channels.

### 5.1.2 Outline

To gain insights into the generally intractable problem of evaluating the capacities of quantum channels, a natural approach is to study the performance of extra free resources in the coding scheme. This scheme can be seemed as a deterministic super-operator performed jointly by the sender Alice and the receiver Bob to assist the communication, which we call *general code* (see Section 2.3 for details).

In this chapter, we derive a framework to evaluate the communication capabilities in both non-asymptotic and asymptotic regimes. In section 5.2, we show that the optimal coding success probability and one-shot $\varepsilon$-error classical capacity assisted with NS (and PPT) codes can be characterized by SDPs. We also show that the Matthews-Wehner meta-converse bound for entanglement-assisted classical communication can be achieved by activated, no-signalling assisted codes, suitably generalizing a result for classical channels. In section 5.3, we derive a new meta-converse for unassisted classical communication with application to a finite resource analysis of classical communication over quantum erasure channels. In section 5.4, we derive two SDP strong converse bounds for the classical capacity of a general quantum channel. We show an improved upper bound for the amplitude damping channel and discuss other potential upper bounds on classical capacity.

## 5.2 One-shot communication capability

### 5.2.1 Task of information processing

The aim of classical communication is to transmit classical messages from one side to another side via a noisy channel, which is equivalent to simulate a noiseless classical channel via suitable encoders and decoders.

Based on the previous results on channel composition [CDP08, DW16], one can simulate a channel $\mathcal{M}$ with the channel $\mathcal{N}$ and code $\Pi$, where $\Pi$ is a bipartite CPTP operation from $A_iB_i$ to $A_oB_o$ which is B to A no-signalling. We say such $\Pi$ is an $\Omega$-assisted code if it can be implemented by local operations with $\Omega$-assistance.

In the following, we eliminate $\Omega$ for the case of unassisted codes and write $\Omega =$ E and $\Omega =$ NS for entanglement-assisted and no-signalling-assisted (NS-assisted) codes, respectively. We refer to Section 2.3 for more details about the mathematical description of these codes.

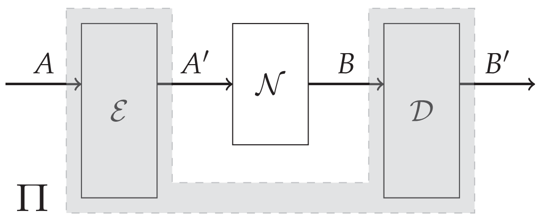


Figure 5.3: General code scheme

Suppose Alice wants to send the classical message labelled by $\{1, \ldots, m\}$ to Bob using the composite channel $\mathcal{M} = \Pi \circ \mathcal{N}$, where $\Pi$ is a deterministic super-operator that generalizes the usual encoding scheme $\mathcal{E}$ and decoding scheme $\mathcal{D}$. After the action of $\mathcal{E}$ and $\mathcal{N}$, the message results in quantum state at Bob's side. Bob then performs a POVM with $m$ outcomes on the resulting quantum state. The POVM is a component of the operation $\mathcal{D}$. Since the results of the POVM and the input messages are both classical, it is natural to assume that $\mathcal{M}$ is with classical registers throughout this chapter, that is, $\Delta \circ \mathcal{M} \circ \Delta = \mathcal{M}$ for some completely dephasing channel $\Delta$.

**Definition 5.2.** Given a quantum channel $\mathcal{N}_{A\to B}$ and a fixed $\Omega$-assisted code $\Pi$ with size $m$, the optimal average success probability of $\mathcal{N}$ to transmit $m$ messages is defined as

$$p_s\left(\mathcal{N}, \Pi, m\right) := \frac{1}{m}\sum_{k=1}^{m} \operatorname{Tr} \mathcal{M}\left(|k\rangle\langle k|\right)|k\rangle\langle k|, \tag{5.3}$$

where $\Omega \in \{\mathrm{UA}, \mathrm{E}, \mathrm{NS}, \mathrm{NS} \cap \mathrm{PPT}\}$.

Furthermore, the optimal average success probability of $\mathcal{N}$ to transmit $m$ messages assisted with $\Omega$-class code is defined as

$$\begin{aligned} p_{succ,\Omega}\left(\mathcal{N}, m\right) := \frac{1}{m} \sup & \sum_{k=1}^{m} \operatorname{Tr} \mathcal{M}\left(|k\rangle\langle k|\right)|k\rangle\langle k|, \\ & \text{s.t. } \mathcal{M} = \Pi \circ \mathcal{N} \text{ is the effective channel}, \\ & \Pi \in \Omega, \end{aligned} \tag{5.4}$$

where the maximum is over the codes in class $\Omega$.

With this in hand, we now say that a triplet $(r, n, \varepsilon)$ is achievable on the channel $\mathcal{N}$ with $\Omega$-assisted codes if

$$\frac{1}{n}\log m \geq r, \text{ and } p_{succ,\Omega}\left(\mathcal{N}^{\otimes n}, m\right) \geq 1-\varepsilon. \tag{5.5}$$

We are interested in the following boundary of the non-asymptotic achievable region:

$$C_\Omega^{(1)}\left(\mathcal{N}, \varepsilon\right) := \sup\{\log m : p_{succ,\Omega}\left(\mathcal{N}, m\right) \geq 1-\varepsilon\}. \tag{5.6}$$

We also define $p_{succ,\Omega}\left(\mathcal{N}, \rho_A, m\right)$ and $C_\Omega^{(1)}\left(\mathcal{N}, \rho_A, \varepsilon\right)$ as the same optimization but only using codes with a fixed average input $\rho_A$. The $\Omega$-assisted capacity of a quantum channel is

$$C_\Omega\left(\mathcal{N}\right) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}C_\Omega^{(1)}\left(\mathcal{N}^{\otimes n}, \varepsilon\right), \tag{5.7}$$

where $\Omega \in \{\mathrm{UA}, \mathrm{E}, \mathrm{NS}, \mathrm{NS} \cap \mathrm{PPT}\}$. Throughout the thesis, we omit $\Omega$ when $\Omega = \mathrm{UA}$.

The HSW theorem (cf. Theorem 5.1) tells us that for unassisted codes, it holds that

$$C(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n}\chi\left(\mathcal{N}^{\otimes n}\right). \tag{5.8}$$

Moreover, the entanglement-assisted classical capacity has a single-letter formula [BSST02]:

$$C_E(\mathcal{N}) = \max_{\rho_A} I(\rho_A;\mathcal{N}), \tag{5.9}$$

where $I(\rho_A;\mathcal{N}) := S(\rho_A) + S(\mathcal{N}(\rho_A)) - S\left((\mathrm{id}\otimes\mathcal{N})\,\phi_{\rho_A}\right)$, and $\phi_{\rho_A}$ is a purification of $\rho_A$.

### 5.2.2 Semidefinite programs for optimal success probability

We are now able to derive the one-shot characterization of classical communication assisted with NS (or NS∩PPT) codes. See Section 2.3 for details about general codes.

Let us recall that the NS and PPT codes can be characterized in a mathematically tractable way: a bipartite operation $\Pi(A_iB_i \to A_oB_o)$ is no-signalling and PPT-preserving if and only if its Choi-Jamiołkowski matrix $Z_{A_iB_iA_oB_o}$ satisfies:

$$\begin{aligned} Z_{A_iB_iA_oB_o} &\geq 0, && \text{(CP)}\\ \mathrm{Tr}_{A_oB_o} Z_{A_iB_iA_oB_o} &= \mathbb{1}_{A_iB_i}, && \text{(TP)}\\ Z^{T_{B_iB_o}}_{A_iB_iA_oB_o} &\geq 0, && \text{(PPT)}\\ \mathrm{Tr}_{A_o} Z_{A_iB_iA_oB_o} &= \frac{\mathbb{1}_{A_i}}{d_{A_i}} \otimes \mathrm{Tr}_{A_oA_i} Z_{A_iB_iA_oB_o}, && (A \not\to B)\\ \mathrm{Tr}_{B_o} Z_{A_iB_iA_oB_o} &= \frac{\mathbb{1}_{B_i}}{d_{B_i}} \otimes \mathrm{Tr}_{B_oB_i} Z_{A_iB_iA_oB_o}, && (B \not\to A) \end{aligned} \tag{5.10}$$

where the five lines correspond respectively to $\Pi$ being completely positive, trace-preserving, PPT-preserving, no-signalling from A to B, no-signalling from B to A, respectively.

**Theorem 5.3.** *For a given quantum channel $\mathcal{N}$, the optimal success probability of $\mathcal{N}$ to transmit $m$ messages assisted by NS∩PPT codes is given by*

$$\begin{aligned} p_{s,NS\cap PPT}(\mathcal{N}, m) = \max\ & \mathrm{Tr}\, J_{\mathcal{N}} F_{AB}\\ \text{s.t. } & 0 \leq F_{AB} \leq \rho_A \otimes \mathbb{1}_B, \mathrm{Tr}\,\rho_A = 1,\\ & \mathrm{Tr}_A F_{AB} = \mathbb{1}_B/m, 0 \leq F_{AB}^{T_B} \leq \rho_A \otimes \mathbb{1}_B\ \ (PPT). \end{aligned} \tag{5.11}$$

*Similarly, when assisted by NS codes, one can remove the PPT constraint to obtain the optimal* *success probability as follows:*

$$
\begin{aligned}
p_{s,NS}(\mathcal{N},m) = \max\ & \operatorname{Tr} J_{\mathcal{N}} F_{AB} \\
s.t.\ & 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1, \\
& \operatorname{Tr}_A F_{AB} = \mathbb{1}_B/m.
\end{aligned}
\tag{5.12}
$$

*Proof.* In this proof, we first use the Choi-Jamiołkowski representations of quantum channels to refine the average success probability and then exploit symmetry to simplify the optimization over all possible codes. Finally, we impose the no-signalling and PPT-preserving constraints to obtain the semidefinite program of the optimal average success probability.

Without loss of generality, we assume that $A_i$ and $B_o$ are classical registers with size $m$, i.e., the inputs and outputs are $\{|k\rangle_{A_i}\}_{k=1}^m$ and $\{|k'\rangle_{B_i}\}_{k'=1}^m$, respectively. For some NS∩PPT code $\Pi$, the Choi-Jamiołkowski matrix of $\mathcal{M} = \Pi \circ \mathcal{N}$ is given by $J_{\mathcal{M}} = \sum_{ij} |i\rangle\langle j|_{A_i} \otimes \mathcal{M}\left(|i\rangle\langle j|_{A_i'}\right)$, where $A_i'$ is isomorphic to $A_i$. Then, we can simplify $f(\mathcal{N},\Pi,m)$ to

$$
\begin{aligned}
p_s(\mathcal{N},\Pi,m) &= \frac{1}{m}\sum_{k=1}^{m} \operatorname{Tr}\left(\mathcal{M}\left(|k\rangle\langle k|_{A_i'}\right)|k\rangle\langle k|_{B_o}\right) \\
&= \frac{1}{m}\operatorname{Tr}\left(\sum_{i,j=1}^{m}\left(|i\rangle\langle j|_{A_i}\otimes\mathcal{M}\left(|i\rangle\langle j|_{A_i'}\right)\right)\sum_{k=1}^{m}|kk\rangle\langle kk|_{A_iB_o}\right) \\
&= \frac{1}{m}\operatorname{Tr} J_{\mathcal{M}}\sum_{k=1}^{m}|kk\rangle\langle kk|_{A_iB_o}.
\end{aligned}
\tag{5.13}
$$

Then, denoting $D_{A_iB_o} = \sum_{k=1}^m |kk\rangle\langle kk|_{A_iB_o}$, we have

$$
p_{s,NS\cap PPT}(\mathcal{N},m) = \max_{\mathcal{M}=\Pi\circ\mathcal{N}} \frac{1}{m}\operatorname{Tr}\left(J_{\mathcal{M}} D_{A_iB_o}\right),
$$

where $\mathcal{M} = \Pi \circ \mathcal{N}$ and $\Pi$ is any feasible NS∩PPT bipartite operation. (See Figure 6.2 for the implementation of $\mathcal{M}$.) Noting that $J_{\mathcal{M}} = \operatorname{Tr}_{A_oB_i}\left(J_{\mathcal{N}}^T \otimes \mathbb{1}_{A_iB_o}\right) Z_{A_iA_oB_iB_o}$, we can further simplify $f(\mathcal{N},m)$ as

$$
\begin{aligned}
p_{s,NS\cap PPT}(\mathcal{N},m) = \max \operatorname{Tr}&\left(J_{\mathcal{N}}^T\otimes\mathbb{1}_{A_iB_o}\right) Z_{A_iA_oB_iB_o}\left(\mathbb{1}_{A_oB_i}\otimes D_{A_iB_o}\right)/m, \\
&\text{s.t. } Z_{A_iA_oB_iB_o} \text{ satisfies Eq. (5.10)}
\end{aligned}
\tag{5.14}
$$

The next step is to simplify $f(\mathcal{N},m)$ by exploiting symmetry. For any permutation $\tau \in S_m$, where $S_m$ is the symmetric group of degree $m$, if $Z_{A_iA_oB_iB_o}$ is feasible (satisfying the constraints in Eq. (5.10)), then it is not difficult to check that

$$Z'_{A_iA_oB_iB_o} = (\tau_{A_i} \otimes \tau_{B_o} \otimes \mathbb{1}_{A_oB_i}) Z_{A_iA_oB_iB_o} (\tau_{A_i} \otimes \tau_{B_o} \otimes \mathbb{1}_{A_oB_i})^{\dagger} \tag{5.15}$$

is also feasible. And any convex combination $\lambda Z' + (1-\lambda) Z''$ $(0 \leq \lambda \leq 1)$ of two operators satisfying Eq. (5.10) can also be checked to be feasible. Therefore, if $Z_{A_iA_oB_iB_o}$ is feasible, so is

$$\begin{aligned} &\widetilde{Z}_{A_iA_oB_iB_o} \\ &= \mathcal{P}_{A_iB_o}(Z_{A_iA_oB_iB_o}) \\ &= \frac{1}{m!} \sum_{\tau_{A_i},\tau_{B_o} \in S_m} (\tau_{A_i} \otimes \tau_{B_o} \otimes \mathbb{1}_{A_oB_i}) Z_{A_iA_oB_iB_o} (\tau_{A_i} \otimes \tau_{B_o} \otimes \mathbb{1}_{A_oB_i})^{\dagger}, \end{aligned} \tag{5.16}$$

where $\mathcal{P}_{A_iB_o}$ is a twirling operation on $A_iB_o$.

Noticing that $\mathcal{P}_{A_iB_o}(D_{A_iB_o}) = D_{A_iB_o}$, we have

$$\operatorname{Tr}_{A_iB_o} Z_{A_iB_iA_oB_o} (\mathbb{1}_{A_oB_i} \otimes D_{A_iB_o}) \tag{5.17}$$

$$= \operatorname{Tr}_{A_iB_o} Z_{A_iB_iA_oB_o} (\mathbb{1}_{A_oB_i} \otimes \mathcal{P}_{A_iB_o}(D_{A_iB_o})) \tag{5.18}$$

$$= \operatorname{Tr}_{A_iB_o} \widetilde{Z}_{A_iA_oB_iB_o} (\mathbb{1}_{A_oB_i} \otimes D_{A_iB_o}). \tag{5.19}$$

Thus, it is easy to see that the optimal success probability equals to

$$\begin{aligned} p_{s,NS\cap PPT}(\mathcal{N}, m) = \max \operatorname{Tr} \left(J_{\mathcal{N}}^T \otimes \mathbb{1}_{A_iB_o}\right) \widetilde{Z}_{A_iA_oB_iB_o} (\mathbb{1}_{A_oB_i} \otimes D_{A_iB_o}) / m \\ \text{s.t. } \widetilde{Z}_{A_iA_oB_iB_o} \text{ satisfies Eq. } (5.10). \end{aligned}$$

By Schur's lemma, $\widetilde{Z}_{A_iA_oB_iB_o}$ can be rewritten as

$$\widetilde{Z}_{A_iA_oB_iB_o} = F_{A_oB_i} \otimes D_{A_iB_o} + E_{A_oB_i} \otimes (\mathbb{1} - D_{A_iB_o}),$$

for some operators $E_{A_oB_i}$ and $F_{A_oB_i}$. Thus, the objective function can be simplified to $\operatorname{Tr} J_{\mathcal{N}}^T F$. Also, the CP and PPT constraints are equivalent to

$$E_{A_oB_i} \geq 0, F_{A_oB_i} \geq 0, E_{A_oB_i}^{T_{B_i}} \geq 0, F_{A_oB_i}^{T_{B_i}} \geq 0. \tag{5.20}$$

Furthermore, the $B \not\to A$ constraint is equivalent to $\operatorname{Tr}_{B_o} \widetilde{Z}_{A_iA_oB_iB_o} = \operatorname{Tr}_{B_oB_i} \widetilde{Z}_{A_iA_oB_iB_o} \otimes \mathbb{1}_{B_i}/d_{B_i}$, i.e.

$$F_{A_oB_i} + (m-1) E_{A_oB_i} = \operatorname{Tr}_{B_i} (F_{A_oB_i} + (m-1) E_{A_oB_i}) \otimes \frac{\mathbb{1}_{B_i}}{d_{B_i}} =: \rho_{A_o} \otimes \mathbb{1}_{B_i}. \tag{5.21}$$

and the TP constraint holds if and only if $\operatorname{Tr}_{A_oB_o} Z_{A_iA_oB_iB_o} = \mathbb{1}_{A_iB_i}$, i.e.,

$$\operatorname{Tr}_{A_o}\left(F_{A_oB_i} + (m-1)\, E_{A_oB_i}\right) = \mathbb{1}_{B_i}, \tag{5.22}$$

which is equivalent to

$$\operatorname{Tr}\rho_{A_o} = \operatorname{Tr}\left(F_{A_oB_i} + (m-1)\, E_{A_oB_i}\right)/d_{B_i} = \operatorname{Tr}\mathbb{1}_{B_i}/d_{B_i} = 1. \tag{5.23}$$

As $\Pi$ is no-signalling from A to B, we have $\operatorname{Tr}_{A_o} \widetilde{Z}_{A_iA_oB_iB_o} = \operatorname{Tr}_{A_oA_i} \widetilde{Z}_{A_iA_oB_iB_o} \otimes \frac{\mathbb{1}_{A_i}}{m}$, i.e.,

$$\begin{aligned} &\operatorname{Tr}_{A_o} F_{A_oB_i} \otimes D_{A_iB_o} + \operatorname{Tr}_{A_o} E_{A_oB_i} \otimes (\mathbb{1} - D_{A_iB_o}) \\ =&\operatorname{Tr}_{A_o}\left(F_{A_oB_i} + (m-1)\, E_{A_oB_i}\right) \otimes \frac{\mathbb{1}_{A_iB_o}}{m} = \mathbb{1}_{A_iB_iB_o}/m. \end{aligned} \tag{5.24}$$

Since $D_{A_iB_o}$ and $\mathbb{1} - D_{A_iB_o}$ are orthogonal positive operators, we have

$$\operatorname{Tr}_{A_o} F_{A_oB_i} = \operatorname{Tr}_{A_o} E_{A_oB_i} = \mathbb{1}_{B_i}/m. \tag{5.25}$$

Finally, combining Eq. (5.20), (5.21), (5.23), (5.25), we have that

$$\begin{aligned} p_{s,NS\cap PPT}(\mathcal{N}, m) = \max\ & \operatorname{Tr} J_{\mathcal{N}} F_{A_oB_i} \\ \text{s.t. } & 0 \le F_{A_oB_i} \le \rho_{A_o} \otimes \mathbb{1}_{B_i}, \operatorname{Tr}\rho_{A_o} = 1, \\ & \operatorname{Tr}_{A_o} F_{A_oB_i} = \mathbb{1}_{B_i}/m, \\ & 0 \le F_{A_oB_i}^{T_{B_i}} \le \rho_{A_o} \otimes \mathbb{1}_{B_i}\ \ (\text{PPT}). \end{aligned} \tag{5.26}$$

This gives the SDP in Theorem 5.3, where we assume that $A_o = A$ and $B_i = B$ for simplification. □

**Remark:** The dual SDP for $p_{s,NS\cap PPT}(\mathcal{N}, m)$ is given by

$$\begin{aligned} p_{s,NS\cap PPT}(\mathcal{N}, m) = \min\ & t + \operatorname{Tr} S_B/m \\ \text{s.t. } & J_{\mathcal{N}} \le X_{AB} + \mathbb{1}_A \otimes S_B + (W_{AB} - Y_{AB})^{T_B}, \\ & \operatorname{Tr}_B (X_{AB} + W_{AB}) \le t\mathbb{1}_A, \\ & X_{AB}, Y_{AB}, W_{AB} \ge 0. \end{aligned} \tag{5.27}$$

To remove the PPT constraint, set $Y_{AB} = W_{AB} = 0$. It is worth noting that the strong duality holds here since the Slater's condition can be easily checked. Indeed, choosing $X_{AB} = Y_{AB} = W_{AB} = \|J_{\mathcal{N}}\|_\infty \mathbb{1}_{AB}$, $S_B = \mathbb{1}_B$ and $t = 3d_B\|J_{\mathcal{N}}\|_\infty$ in SDP (5.27), we have $(X_{AB}, Y_{AB}, W_{AB}, S_B, t)$ is in the relative interior of the feasible region.

It is worthing noting that $f_{\text{NS}}(\mathcal{N}, m)$ can be obtained by removing the PPT constraint and it corresponds with the optimal NS-assisted channel fidelity in [LM15].

### 5.2.3 Semidefinite programs for coding rates

For given $0 \le \varepsilon < 1$, the *one-shot $\varepsilon$-error classical capacity assisted by $\Omega$-class codes* is defined as

$$C_{\Omega}^{(1)}(\mathcal{N}, \varepsilon) := \sup\{\log \lambda : 1 - p_{succ,\Omega}(\mathcal{N}, \lambda) \le \varepsilon\}. \tag{5.28}$$

We now derive the one-shot $\varepsilon$-error classical capacity assisted by NS or NS∩PPT codes as follows.

**Theorem 5.4.** *For given channel $\mathcal{N}$ and error threshold $\varepsilon$, the one-shot $\varepsilon$-error* NS∩PPT-assisted and NS-assisted capacities are given by

$$\begin{aligned} C_{\text{NS}\cap\text{PPT}}^{(1)}(\mathcal{N}, \varepsilon) = & -\log \min \eta \\ & \text{s.t. } 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr} \rho_A = 1, \operatorname{Tr}_A F_{AB} = \eta \mathbb{1}_B, \\ & \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1 - \varepsilon, 0 \le F_{AB}^{T_B} \le \rho_A \otimes \mathbb{1}_B \ (\text{PPT}), \end{aligned} \tag{5.29}$$

and

$$\begin{aligned} C_{NS}^{(1)}(\mathcal{N}, \varepsilon) = & -\log \min \eta \\ & \text{s.t. } 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr} \rho_A = 1, \\ & \operatorname{Tr}_A F_{AB} = \eta \mathbb{1}_B, \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1 - \varepsilon, \end{aligned} \tag{5.30}$$

respectively.

*Proof.* When assisted by NS∩PPT codes, by Eq. (5.28), we have that

$$C_{\text{NS}\cap\text{PPT}}^{(1)}(\mathcal{N}, \varepsilon) = \log \max \lambda \ \text{ s.t. } p_{s,NS\cap PPT}(\mathcal{N}, \lambda) \ge 1 - \varepsilon. \tag{5.31}$$

To simplify Eq. (5.31), we suppose that

$$\begin{aligned} \Upsilon(\mathcal{N}, \varepsilon) = & -\log \min \eta \\ & \text{s.t. } 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr} \rho_A = 1, \operatorname{Tr}_A F_{AB} = \eta \mathbb{1}_B, \\ & \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1 - \varepsilon, 0 \le F_{AB}^{T_B} \le \rho_A \otimes \mathbb{1}_B \ (\text{PPT}). \end{aligned} \tag{5.32}$$

On one hand, for given $\varepsilon$, suppose that the optimal solution to the SDP (5.32) of $\Upsilon(\mathcal{N}, \varepsilon)$ is $\{\rho, F, \eta\}$. Then, it is clear that $\{\rho, F\}$ is a feasible solution of the SDP (5.11) of $p_{s,NS\cap PPT}\left(\mathcal{N},\eta^{-1}\right)$, which means that

$$p_{s,NS\cap PPT}\left(\mathcal{N},\eta^{-1}\right) \geq \operatorname{Tr} J_{\mathcal{N}} F \geq 1-\varepsilon. \tag{5.33}$$

Therefore,

$$C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}\left(\mathcal{N},\varepsilon\right) \geq \log \eta^{-1} = \Upsilon\left(\mathcal{N},\varepsilon\right). \tag{5.34}$$

On the other hand, for given $\varepsilon$, suppose that the value of $C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}\left(\mathcal{N},\varepsilon\right)$ is $\log\lambda$ and the optimal solution of $p_{s,NS\cap PPT}\left(\mathcal{N},\lambda\right)$ is $\{\rho, F\}$. It is easy to check that $\{\rho, F, \lambda^{-1}\}$ satisfies the constraints in SDP (5.32) of $\Upsilon\left(\mathcal{N},\varepsilon\right)$. Therefore,

$$\Upsilon\left(\mathcal{N},\varepsilon\right) \geq -\log\lambda^{-1} = C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}\left(\mathcal{N},\varepsilon\right). \tag{5.35}$$

Hence, combining Eqs. (5.32), (5.34) and (5.35), it is clear that

$$\begin{aligned} C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}\left(\mathcal{N},\varepsilon\right) = \Upsilon\left(\mathcal{N},\varepsilon\right) = -\log\min\ & \eta \\ \text{s.t.}\ & 0 \leq F_{AB} \leq \rho_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr}\rho_A = 1, \operatorname{Tr}_A F_{AB} = \eta\mathbb{1}_B, \\ & \operatorname{Tr} J_{\mathcal{N}} F_{AB} \geq 1-\varepsilon, \\ & 0 \leq F_{AB}^{T_B} \leq \rho_A \otimes \mathbb{1}_B\ \ (\mathrm{PPT}). \end{aligned} \tag{5.36}$$

And one can obtain $C^{(1)}_{\mathrm{NS}}\left(\mathcal{N},\varepsilon\right)$ by removing the PPT constraint. □

Considering the hierarchy of quantum codes in Figure 2.3.2, we know that NS codes are potentially stronger than the entanglement-assisted codes, which means $C^{(1)}_{\mathrm{NS}}\left(\mathcal{N},\varepsilon\right)$ can provide converse bounds of classical communication with entanglement assistance. Moreover, NS∩PPT codes are more powerful than the unassisted codes, and this implies that $C^{(1)}\left(\mathcal{N},\varepsilon\right) \leq C^{(1)}_{NS\cap PPT}\left(\mathcal{N},\varepsilon\right)$. Therefore, we have the following corollary:

**Corollary 5.5.** *For a given channel $\mathcal{N}$ and error threshold $\varepsilon$,*

$$\begin{aligned} C^{(1)}_{E}\left(\mathcal{N},\varepsilon\right) &\leq C^{(1)}_{NS}\left(\mathcal{N},\varepsilon\right), \\ C^{(1)}\left(\mathcal{N},\varepsilon\right) &\leq C^{(1)}_{NS\cap PPT}\left(\mathcal{N},\varepsilon\right). \end{aligned}$$

In the asymptotic regime, it is worth noting that the entanglement-assisted classical capacity of a quantum channel is equal to the NS-assisted classical capacity [LY16, WXD18]: for any quantum channel $\mathcal{N}$,

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}C^{(1)}_{NS}\left(\mathcal{N}^{\otimes n},\varepsilon\right) = C_E\left(\mathcal{N}\right).$$

### Classical-quantum channel

For classical-quantum channels, the one-shot $\varepsilon$-error NS-assisted (or NS$\cap$PPT-assisted) capacity can be further simplified based on the structure of the channel.

**Proposition 5.6.** *For the classical-quantum channel that acts as $\mathcal{N} : x \to \rho_x$, the Choi matrix of $\mathcal{N}$ is given by $J_{\mathcal{N}} = \sum_x |x\rangle\langle x| \otimes \rho_x$. Then, the SDP (5.30) of $C^{(1)}_{\mathrm{NS}}(\mathcal{N}, \varepsilon)$ and the SDP (5.29) of $C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}(\mathcal{N}, \varepsilon)$ can be simplified to*

$$
\begin{aligned}
C^{(1)}_{\mathrm{NS}}(\mathcal{N}, \varepsilon) = C^{(1)}_{\mathrm{NS}\cap\mathrm{PPT}}(\mathcal{N}, \varepsilon) = \log\max \sum s_x& \\
\text{s.t. } & 0 \le Q_x \le s_x \mathbb{1}_B, \forall x, \\
& \sum_x Q_x = \mathbb{1}_B, \\
& \sum_x \operatorname{Tr} Q_x \rho_x \ge \sum_x (1-\varepsilon) s_x.
\end{aligned} \tag{5.37}
$$

*Proof.* When $J_{\mathcal{N}} = \sum_x |x\rangle\langle x| \otimes \rho_x$, the SDP (5.30) easily simplifies to

$$
\begin{aligned}
C^{(1)}_{NS}(\mathcal{N}, \varepsilon) = -\log\min \eta& \\
\text{s.t. } & 0 \le F_x \le p_x \mathbb{1}_B, \forall x, \\
& \sum_x p_x = 1, \\
& \sum_x F_x/\eta = \mathbb{1}_B, \\
& \sum_x \operatorname{Tr} F_x \rho_x \ge (1-\varepsilon).
\end{aligned} \tag{5.38}
$$

By assuming that $Q_x = F_x/\eta$ and $s_x = p_x/\eta$, the above SDP simplifies to

$$
\begin{aligned}
C^{(1)}_{NS}(\mathcal{N}, \varepsilon) = \log\max \sum s_x& \\
\text{s.t. } & 0 \le Q_x \le s_x \mathbb{1}_B, \forall x, \\
& \sum_x Q_x = \mathbb{1}_B, \\
& \sum_x \operatorname{Tr} Q_x \rho_x \ge (1-\varepsilon) \sum s_x,
\end{aligned} \tag{5.39}
$$

where we use the fact $\sum s_x = \sum p_x/\eta = 1/\eta$. One can use similar to method to simplify $C^{(1)}_{NS\cap PPT}(\mathcal{N}, \varepsilon)$ as well. □

### Reduction to Polyanskiy-Poor-Verdú converse bound

For classical channels, Polyanskiy, Poor, and Verdú [PPV10] derived the finite blocklength converse via hypothesis testing. In [Mat12], an alternative proof of PPV converse was provided by considering the assistance of the classical no-signalling correlations. Here, we are going to show that both $C_{NS}^{(1)}(\mathcal{N},\varepsilon)$ and $C_{NS\cap PPT}^{(1)}(\mathcal{N},\varepsilon)$ will reduce to the PPV converse.

Let us first recall the linear program for the PPV converse bound of a classical channel $\mathcal{N}(y|x)$ [PPV10, Mat12]:

$$\begin{aligned} R^{PPV}(\mathcal{N},\varepsilon) = \max \sum_x s_x & \\ \text{s.t. } & Q_{xy} \le s_x, \forall x,y, \\ & \sum_x Q_{xy} \le 1, \forall y, \\ & \sum_{x,y} \mathcal{N}(y|x)\, Q_{xy} \ge (1-\varepsilon)\sum_x s_x. \end{aligned} \tag{5.40}$$

For classical channels, we can further simplify the SDP (5.37) to a linear program which coincides with the Polyanskiy-Poor-Verdú converse bound.

**Proposition 5.7.** *For a classical channel $\mathcal{N}(y|x)$,*

$$C_{NS}^{(1)}(\mathcal{N},\varepsilon) = C_{NS\cap PPT}^{(1)}(\mathcal{N},\varepsilon) = R^{PPV}(\mathcal{N},\varepsilon). \tag{5.41}$$

*Proof.* The idea is to further simplify the SDP (5.37) via the structure of classical channels. For input $x$, the corresponding outputs can be seemed as $\rho_x = \sum_y \mathcal{N}(y|x)\,|y\rangle\langle y|$. Then, $Q_x$ should be diagonal for any $x$, i.e., $Q_x = \sum_y Q_{xy}$. Thus, SDP (5.37) can be easily simplified to

$$\begin{aligned} C_{NS}^{(1)}(\mathcal{N},\varepsilon) = C_{NS\cap PPT}^{(1)}(\mathcal{N},\varepsilon) = \log\max \sum_x s_x & \\ \text{s.t. } & Q_{xy} \le s_x, \forall x,y, \\ & \sum_x Q_{xy} = 1, \forall y, \\ & \sum_x\sum_y \mathcal{N}(y|x)\, Q_{xy} \ge (1-\varepsilon)\sum_x s_x. \end{aligned} \tag{5.42}$$

Using the similar technique in [Mat12], the constraint $\sum_x Q_{xy} = 1$ can be relaxed to $\sum_x Q_{xy} \le 1$ in this case, which means that the linear program (5.42) is equal to the linear program (5.40). □

### 5.2.4 Matthews-Wehner converse via activated NS codes

For classical communication over quantum channels with entanglement assistance, Matthews and Wehner [MW14a] proved a meta-converse bound in terms of the hypothesis testing relative entropy which generalizes Polyanskiy, Poor and Verdú's approach [PPV10] to quantum channels. Given a quantum channel $\mathcal{N}_{A'\to B}$, they proved [MW14a] that

$$C_E^{(1)}(\mathcal{N},\varepsilon) \le R(\mathcal{N},\varepsilon) := \max_{\rho_{A'}} \min_{\sigma_B} D_H^{\varepsilon}\left(\mathcal{N}\left(\phi_{A'A}\right) \| \rho_{A'} \otimes \sigma_B\right), \tag{5.43}$$

where $\phi_{AA'} = \left(\mathbb{1}_A \otimes \rho_{A'}^{1/2}\right) \widetilde{\Phi}_{AA'} \left(\mathbb{1}_A \otimes \rho_{A'}^{1/2}\right)$ is a purification of $\rho_{A'}$ and $\widetilde{\Phi}_{AA'} = \sum_{ij} |i_A i_{A'}\rangle\langle j_A j_{A'}|$ denotes the unnormalized maximally entangled state. In the above expression, $D_H^{\varepsilon}(\cdot\|\cdot)$ is the quantum hypothesis testing relative entropy [WR12, BD10]. We refer to Section 2.7.3 for details.

The hypothesis testing relative entropy bound in Eq. (5.43) is an SDP and it holds that

$$\begin{aligned} R(\mathcal{N},\varepsilon) = -\log\min \lambda& \\ \text{s.t. } 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1,& \\ \operatorname{Tr}_A F_{AB} \le \lambda \mathbb{1}_B, \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1-\varepsilon.& \end{aligned} \tag{5.44}$$

Here $J_{\mathcal{N}}$ is the Choi-Jamiołkowski matrix of $\mathcal{N}$.

For classical channels, the hypothesis testing relative entropy bound is exactly equal to the one-shot classical capacity assisted by no-signalling (NS) codes [Mat12]. For quantum channels the one-shot $\varepsilon$-error capacity assisted by NS codes is given by [WXD18]

$$\begin{aligned} C_{NS}^{(1)}(\mathcal{N},\varepsilon) = -\log\min \eta& \\ \text{s.t. } 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1,& \\ \operatorname{Tr}_A F_{AB} = \eta \mathbb{1}_B, \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1-\varepsilon.& \end{aligned} \tag{5.45}$$

Note that the only difference between the SDPs (5.44) and (5.45) is the partial trace constraint of $F_{AB}$. However, unlike in the classical special case, the SDPs in (5.44) and (5.45) are not equal in general [WXD18].

In this section, we are going to show that this gap can be closed by considering activated, NS-assisted codes. The concept of activated capacity follows the idea of potential capacities of quantum channels introduced by Winter and Yang [WY15]. The model is described as follows. For a quantum channel $\mathcal{N}$ assisted by NS codes, we can first borrow a noiseless classical channel $\mathcal{I}_m$ whose capacity is $\log m$, then we can use $\mathcal{N} \otimes \mathcal{I}_m$ coherently to transmit classical messages. After the communication finishes, we just pay back the capacity of $\mathcal{I}_m$. This kind of communication method was also studied in zero-error information theory [ADR+17, DW15].

**Definition 5.8.** For any quantum channel $\mathcal{N}$, we define

$$C_{NS,a}^{(1)}(\mathcal{N},\varepsilon) := \sup_{m\geq 1}\left[C_{NS}^{(1)}(\mathcal{N}\otimes\mathcal{I}_m,\varepsilon) - \log m\right], \tag{5.46}$$

where $\mathcal{I}_m(\rho) = \sum_{i=1}^m \operatorname{Tr}(\rho|i\rangle\langle i|)\,|i\rangle\langle i|$ the classical noiseless channel with capacity $\log m$.

The following is the main result of this section:

**Theorem 5.9.** *For any quantum channel $\mathcal{N}_{A\to B}$ and error tolerance $\varepsilon\in(0,1)$, we have*

$$C_{NS,a}^{(1)}(\mathcal{N},\varepsilon) = R(\mathcal{N},\varepsilon) = \max_{\rho_{A'}}\min_{\sigma_B} D_H^{\varepsilon}(\mathcal{N}_{A\to B}(\phi_{A'A})\,\|\rho_{A'}\otimes\sigma_B). \tag{5.47}$$

The proof outline is as follows. We first show that the $\mathcal{I}_2$ is enough to activate the channel to achieve the bound $R(\mathcal{N},\varepsilon)$ in the following Lemma 5.10, i.e.,

$$C_{NS,a}^{(1)}(\mathcal{N},\varepsilon) \geq C_{\mathrm{NS}}^{(1)}(\mathcal{N}\otimes\mathcal{I}_2,\varepsilon) - 1 \geq R(\mathcal{N},\varepsilon). \tag{5.48}$$

We then show that $R(\mathcal{N},\varepsilon)$ is additive for noiseless channel in the following Lemma 5.11, i.e., $R(\mathcal{N}\otimes\mathcal{I}_m,\varepsilon) = R(\mathcal{N},\varepsilon)+\log m$. This implies that $R(\mathcal{N},\varepsilon)$ is also a converse bound for the activated capacity, i.e.,

$$C_{NS,a}^{(1)}(\mathcal{N},\varepsilon) = \sup_{m\geq 1}\left[C_{NS}^{(1)}(\mathcal{N}\otimes\mathcal{I}_m,\varepsilon) - \log m\right] \leq \sup_{m\geq 1}\left[R(\mathcal{N}\otimes\mathcal{I}_m,\varepsilon) - \log m\right] = R(\mathcal{N},\varepsilon). \tag{5.49}$$

The theorem thus directly follows from Lemmas 5.10 and 5.11.

**Lemma 5.10.** *We have $C_{NS}^{(1)}(\mathcal{N}\otimes\mathcal{I}_2,\varepsilon) - 1 \geq R(\mathcal{N},\varepsilon)$.*

*Proof.* This proof is based on a key observation that the additional one-bit noiseless channel can provide a larger solution space to help the activated capacity achieve the quantum hypothesis testing converse. Suppose that the optimal solution to SDP (5.44) of $R(\mathcal{N},\varepsilon)$ is $\{\lambda,\rho_{A_1},F_{A_1B_1}\}$. We are going to use this optimal solution to construct a feasible solution of the SDP (5.45) of $C_{\mathrm{NS}}^{(1)}(\mathcal{N}\otimes I_2,\varepsilon)$.

Let us choose $\rho_{A_1A_2} = \rho_{A_1}\otimes\frac{1}{2}(|0\rangle\langle 0|+|1\rangle\langle 1|)_{A_2}$ and

$$F_{A_1A_2B_1B_2} = \frac{F_{A_1B_1}}{2}\otimes(|00\rangle\langle 00|+|11\rangle\langle 11|)_{A_2B_2} + \frac{\widetilde{F}_{A_1B_1}}{2}\otimes(|01\rangle\langle 01|+|10\rangle\langle 10|)_{A_2B_2}, \tag{5.50}$$

where $\widetilde{F}_{A_1B_1} = \rho_{A_1}\otimes(\lambda\mathbb{1}_{B_1} - \operatorname{Tr}_{A_1}F_{A_1B_1})$. We see that $F_{A_1A_2B_1B_2}\geq 0$, $\rho_{A_1A_2}\geq 0$ and

$\operatorname{Tr} \rho_{A_1 A_2} = 1$. Moreover, this construction ensures that

$$\operatorname{Tr}_{A_1 A_2} F_{A_1 A_2 B_1 B_2} = \operatorname{Tr}_{A_1} \left( \left( \frac{F_{A_1 B_2}}{2} + \frac{\widetilde{F}_{A_1 B_1}}{2} \right) \otimes \mathbb{1}_{B_2} \right) = \frac{\lambda}{2} \mathbb{1}_{B_1 B_2}, \tag{5.51}$$

and

$$\begin{aligned} \operatorname{Tr} \left( J_{\mathcal{N}} \otimes D_{A_2 B_2} \right) F_{A_1 A_2 B_1 B_2} &= \operatorname{Tr} J_{\mathcal{N}} F_{A_1 B_1} \otimes \frac{1}{2} \operatorname{Tr} D_{A_2 B_2} \left( |00\rangle\langle 00| + |11\rangle\langle 11| \right) && (5.52) \\ &= \operatorname{Tr} J_{\mathcal{N}} F_{A_1 B_1} \geq 1 - \varepsilon, && (5.53) \end{aligned}$$

where $D_{A_2 B_2} = \sum_{i=0,1} |ii\rangle\langle ii|$ is the Choi-Jamiołkowski matrix of $\mathcal{I}_2$. Furthermore, $\rho_{A_1} \otimes \mathbb{1}_{B_1} - \widetilde{F}_{A_1 B_1} \geq 0$ and consequently we find that $\rho_{A_1 A_2} \otimes \mathbb{1}_{B_1 B_2} - F_{A_1 A_2 B_1 B_2} \geq 0$. Hence, $\left\{ \frac{1}{2} \lambda, \rho_{A_1 A_2}, F_{A_1 A_2 B_1 B_2} \right\}$ is a feasible solution, ensuring that $C_{NS}^{(1)} \left( \mathcal{N} \otimes \mathcal{I}_2, \varepsilon \right) - 1 \geq R \left( \mathcal{N}, \varepsilon \right)$. □

**Lemma 5.11.** *We have* $R \left( \mathcal{N} \otimes \mathcal{I}_m, \varepsilon \right) = R \left( \mathcal{N}, \varepsilon \right) + \log m$.

*Proof.* On one hand, it is easy to prove that $R \left( \mathcal{N} \otimes \mathcal{I}_m, \varepsilon \right) \geq R \left( \mathcal{N}, \varepsilon \right) + \log m$. To see the other direction, we are going to use the dual SDP of $R \left( \mathcal{N}, \varepsilon \right)$:

$$\begin{aligned} R \left( \mathcal{N}, \varepsilon \right) = - \log \max & \left[ s \left( 1 - \varepsilon \right) - t \right] \\ \text{s.t. } & X_{AB} + \mathbb{1}_A \otimes Y_B \geq s J_{\mathcal{N}}, \\ & \operatorname{Tr}_B X_{AB} \leq t \mathbb{1}_A, \operatorname{Tr} Y_B \leq 1, \\ & X_{AB}, Y_B, s \geq 0. \end{aligned} \tag{5.54}$$

We note that the strong duality holds here.

Suppose that the optimal solution to the dual SDP (5.54) of $R \left( \mathcal{N}, \varepsilon \right)$ is $\left\{ \widehat{X}_{AB}, \widehat{Y}_B, \widehat{s}, \widehat{t} \right\}$. Let us choose $X_{AA'BB'} = \frac{1}{m} \widehat{X}_{AB} \otimes D_m$, $Y_{BB'} = \frac{1}{m} \widehat{Y}_B \otimes \mathbb{1}_m$, $s = \frac{1}{m} \widehat{s}$, $t = \frac{1}{m} \widehat{t}$, with $D_m = \sum_{i=0}^{m-1} |ii\rangle\langle ii|$. Then it can be easily checked that

$$X_{AA'BB'} + \mathbb{1}_{AA'} \otimes Y_{BB'} \geq \left( \widehat{X}_{AB} + \mathbb{1}_A \otimes \widehat{Y}_B \right) \otimes \frac{D_m}{m} \geq s J_{\mathcal{N}} \otimes D_m. \tag{5.55}$$

The other constraints can be verified similarly. Thus, $\{X_{AA'BB'}, Y_{BB'}, s, t\}$ is a feasible solution to the SDP (5.54) of $R \left( \mathcal{N} \otimes \mathcal{I}_m, \varepsilon \right)$, which means that

$$R \left( \mathcal{N} \otimes \mathcal{I}_m, \varepsilon \right) \leq - \log [s \left( 1 - \varepsilon \right) - t] = R \left( \mathcal{N}, \varepsilon \right) + \log m. \tag{5.56}$$

□

## 5.3 Non-asymptotic communication capability

### 5.3.1 New meta-converse for classical communication

This subsection provides a new meta-converse that upper bounds the amount of information that can be transmitted with a single use of the channel by unassisted codes. This meta-converse, in the spirit of the classical meta-converse by Polyanskiy, Poor and Verdú [PPV10] as well as Nagaoka and Hayashi (see, e.g., [Nag01], [Hay06, Section 4.6]), relates the channel coding problem to a binary composite hypothesis test between the actual channel and a class of subchannels that are generalizations of the useless channels for classical communication.

Recall that the only useless quantum channel for classical communication is the *constant channel* $\mathcal{N}(\cdot) = \sigma$, which maps all states $\rho$ on $A$ to a constant state $\sigma$ on $B$. As a natural extension, we say a subchannel $\mathcal{N}$ is *constant-bounded* if it maps all states $\rho$ to positive semidefinite operators that are smaller than or equal to a constant state $\sigma$, i.e.,

$$\mathcal{N}(\rho) \leq \sigma, \forall \rho \in \mathcal{S}(A). \tag{5.57}$$

Here we denote $\mathcal{S}(A) := \{\rho_A \geq 0 : \operatorname{Tr} \rho_A = 1\}$ as the set of quantum states on $A$, and a subchannel $\mathcal{N}$ is a linear completely positive (CP) map that is trace non-increasing, i.e., $\operatorname{Tr} \mathcal{N}(\rho) \leq 1$ for all states $\rho$.

We also define the set of *constant-bounded subchannels*:

$$\mathcal{V} := \{\mathcal{M} \in \mathrm{CP}(A : B) : \exists\, \sigma \in \mathcal{S}(B) \text{ s.t. } \mathcal{M}(\rho) \leq \sigma, \forall \rho \in \mathcal{S}(A)\}, \tag{5.58}$$

where $\mathrm{CP}(A : B)$ denotes the set of all CP maps from $A$ to $B$. Clearly, $\mathcal{V}$ is convex and closed.

This inspires the following new one-shot converse bound:

**Theorem 5.12.** *For any quantum channel $\mathcal{N}_{A' \to B}$ and error tolerance $\varepsilon \in (0, 1)$, we have*

$$C^{(1)}(\mathcal{N}, \varepsilon) \leq \max_{\rho_{A'}} \min_{\mathcal{M} \in \mathcal{V}} D_H^{\varepsilon}\left(\mathcal{N}_{A' \to B}(\phi_{A'A}) \,\middle\|\, \mathcal{M}_{A' \to B}(\phi_{A'A})\right) \tag{5.59}$$

$$= \min_{\mathcal{M} \in \mathcal{V}} \max_{\rho_{A'}} D_H^{\varepsilon}\left(\mathcal{N}_{A' \to B}(\phi_{A'A}) \,\middle\|\, \mathcal{M}_{A' \to B}(\phi_{A'A})\right), \tag{5.60}$$

*where $\phi_{A'A}$ is a purification of $\rho_{A'}$.*

*Proof.* Consider an unassisted code with inputs $\{\rho_k\}_{k=1}^m$ and POVM $\{M_k\}_{k=1}^m$ whose average input state is $\rho_{A'} = \sum_{k=1}^m \frac{1}{m} \rho_k$, the success probability to transmit $m$ messages is given by

$$p_{succ} = \frac{1}{m}\sum_{k=1}^{m} \operatorname{Tr} \mathcal{N}\left(\rho_k\right) M_k = \operatorname{Tr} J_{\mathcal{N}} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes M_k\right)$$
$$= \operatorname{Tr} \mathcal{N}_{A'\to B}\left(\phi_{AA'}\right) \left(\rho_A^T\right)^{-1/2} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes M_k\right) \left(\rho_A^T\right)^{-1/2}. \tag{5.61}$$

Denote $E = \left(\rho_A^T\right)^{-1/2} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes M_k\right) \left(\rho_A^T\right)^{-1/2}$. Then

$$0 \le E \le \left(\rho_A^T\right)^{-1/2} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes \mathbb{1}_B\right) \left(\rho_A^T\right)^{-1/2} = \mathbb{1}_{AB}. \tag{5.62}$$

For any $\mathcal{M} \in \mathcal{V}$, we assume that the output states of $\mathcal{M}$ are bounded by the state $\sigma_B$, then

$$\operatorname{Tr} \mathcal{M}_{A'\to B}\left(\phi_{AA'}\right) E = \operatorname{Tr} \mathcal{M}_{A'\to B}\left(\phi_{AA'}\right) \left(\rho_A^T\right)^{-1/2} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes M_k\right) \left(\rho_A^T\right)^{-1/2} \tag{5.63}$$
$$= \operatorname{Tr} J_{\mathcal{M}} \left(\sum_{k=1}^{m} \frac{1}{m}\rho_k^T \otimes M_k\right) \tag{5.64}$$
$$= \frac{1}{m}\sum_{k=1}^{m} \operatorname{Tr} \mathcal{M}\left(\rho_k\right) M_k \tag{5.65}$$
$$\le \frac{1}{m}\sum_{k=1}^{m} \operatorname{Tr} \sigma_B M_k = \frac{1}{m}. \tag{5.66}$$

The second line follows from the fact that $J_{\mathcal{M}} = \left(\rho_A^T\right)^{-1/2} \mathcal{M}_{A'\to B}\left(\phi_{AA'}\right) \left(\rho_A^T\right)^{-1/2}$. In the third line, we use the inverse Choi-Jamiołkowski transformation $\mathcal{M}_{A'\to B}\left(\rho_{A'}\right) = \operatorname{Tr}_A J_{\mathcal{M}} \left(\rho_A^T \otimes \mathbb{1}_B\right)$. The forth line follows since any output state of $\mathcal{M}$ is bounded by the state $\sigma_B$.

Therefore, combining Eqs. (5.61) and (5.66), we know that

$$\operatorname{Tr} \mathcal{N}_{A'\to B}\left(\phi_{AA'}\right) E \ge 1 - \varepsilon, \tag{5.67}$$
$$\operatorname{Tr} \mathcal{M}_{A'\to B}\left(\phi_{AA'}\right) E \le \frac{1}{m}. \tag{5.68}$$

Thus,

$$C^{(1)}\left(\mathcal{N}, \rho_{A'}, \varepsilon\right) \le \min_{\mathcal{M}\in\mathcal{V}} D_H^{\varepsilon}\left(\mathcal{N}_{A'\to B}\left(\phi_{AA'}\right) \| \mathcal{M}_{A'\to B}\left(\phi_{AA'}\right)\right). \tag{5.69}$$

Maximizing over all average input $\rho_{A'}$, we can obtain the desired result of (5.59).

Since the function $\beta_\varepsilon\left(\mathcal{N}_{A'\to B}\left(\phi_{A'A}\right)\|\mathcal{M}_{A'\to B}\left(\phi_{A'A}\right)\right)$ is convex in $\rho_{A'}$ and concave in $\mathcal{M}$ [MW14a], we can exchange the maximization and minimization by applying Sion's minimax theorem [Sio58] and obtain the result of (5.60). □

**Remark 5.13.** Noting that the operator $E$ above also satisfies $0 \le E^{T_B} \le \mathbb{1}$, we can further obtain

$$C^{(1)}\left(\mathcal{N},\varepsilon\right) \le \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}} D^{\varepsilon}_{H,PPT}\left(\mathcal{N}_{A'\to B}\left(\phi_{A'A}\right)\middle\|\mathcal{M}_{A'\to B}\left(\phi_{A'A}\right)\right), \tag{5.70}$$

where $D^{\varepsilon}_{H,PPT}\left(\rho_0\|\rho_1\right) := -\log\min\{\operatorname{Tr} E\rho_1 : 1-\operatorname{Tr} E\rho_0 \le \varepsilon, 0\le E, E^{T_B}\le \mathbb{1}\}$.

If we consider $\max_{\rho_{A'}} D^{\varepsilon}_{H}\left(\mathcal{N}_{A'\to B}\left(\phi_{A'A}\right)\middle\|\mathcal{M}_{A'\to B}\left(\phi_{A'A}\right)\right)$ as the "distance" between the channel $\mathcal{N}$ and CP map $\mathcal{M}$. Then our new meta-converse can be treated as the "distance" between the given channel $\mathcal{N}$ with the class of useless constant-bounded subchannels.

To make this meta-converse bound efficiently computable, we then restrict the set of constant-bounded subchannels $\mathcal{V}$ to an SDP-tractable set of CP maps. Let us define

$$\mathcal{V}_\beta := \{\mathcal{M}\in \text{CP}\left(A:B\right) : \beta\left(J_{\mathcal{M}}\right)\le 1\}, \quad \text{where} \tag{5.71}$$

$$\beta\left(J_{\mathcal{M}}\right) := \min\left\{\operatorname{Tr} S_B : -R_{AB}\le J_{\mathcal{M}}^{T_B}\le R_{AB}, -\mathbb{1}_A\otimes S_B \le R_{AB}^{T_B}\le \mathbb{1}_A\otimes S_B\right\}. \tag{5.72}$$

Here $J_{\mathcal{M}}$ is the Choi-Jamiołkowski matrix of $\mathcal{M}$ and $T_B$ means the partial transpose on system $B$. The set $\mathcal{V}_\beta$ satisfies some basic properties such as convexity and invariance under composition with unitary maps.

**Lemma 5.14.** *The set $\mathcal{V}_\beta$ is a subset of $\mathcal{V}$, i.e., $\mathcal{V}_\beta \subset \mathcal{V}$.*

*Proof.* Given a CP map $\mathcal{M}$ in $\mathcal{V}_\beta$, suppose that the optimal solution of $\beta\left(J_{\mathcal{M}}\right)$ is $\{R, S_B\}$, we write $S_B = \sigma_B$ since $\beta\left(J_{\mathcal{M}}\right) = \operatorname{Tr} S_B \le 1$. For any input $\rho$, the output $\mathcal{M}\left(\rho\right)$ satisfies that

$$\mathcal{M}_{A\to B}\left(\rho_A\right) = \operatorname{Tr}_A \sqrt{\rho_A^T} J_{\mathcal{M}} \sqrt{\rho_A^T} = \left(\operatorname{Tr}_A \sqrt{\rho_A^T} J_{\mathcal{M}}^{T_B} \sqrt{\rho_A^T}\right)^T \tag{5.73}$$

$$\le \left(\operatorname{Tr}_A \sqrt{\rho_A^T} R \sqrt{\rho_A^T}\right)^T = \operatorname{Tr}_A \sqrt{\rho_A^T} R^{T_B} \sqrt{\rho_A^T} \tag{5.74}$$

$$\le \operatorname{Tr}_A \sqrt{\rho_A^T}\left(\mathbb{1}_A\otimes \sigma_B\right)\sqrt{\rho_A^T} = \sigma_B. \tag{5.75}$$

□

As a consequence, we have the following meta-converse.

**Corollary 5.15.** *For any quantum channel $\mathcal{N}_{A'\to B}$ and error tolerance $\varepsilon \in (0,1)$, we have*

$$C^{(1)}(\mathcal{N},\varepsilon) \le \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}_\beta} D_H^\varepsilon\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right) \tag{5.76}$$

$$= \min_{\mathcal{M}\in\mathcal{V}_\beta} \max_{\rho_{A'}} D_H^\varepsilon\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right), \tag{5.77}$$

*where $\phi_{A'A}$ is a purification of $\rho_{A'}$.*

There are several other converses for the one-shot $\varepsilon$-error capacity of a general quantum channel, e.g., the Matthews-Wehner converse [MW14a], the Datta-Hsieh converse [DH13], and the SDP converse via no-signaling (NS) and positive-partial-transpose-preserving (PPT) codes in Theorem 5.4. Note that the Datta-Hsieh converse is not known to be efficiently computable. Also, our meta-converse is tighter than the Matthews-Wehner converse in Eq. (5.43). As we will show later, our meta-converse will lead to new results in both finite blocklength and asymptotic regime.

### 5.3.2 Second-order analysis for quantum erasure channel

The quantum erasure channel is denoted by $\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$, where $|e\rangle$ is orthogonal to the input Hilbert space. The classical capacity of a quantum erasure channel is given by [BDS97]

$$C(\mathcal{E}_p) = (1-p)\log d, \tag{5.78}$$

where $d$ is the dimension of input space. In [WW14], the strong converse property for the classical capacity of $\mathcal{E}_p$ is established.

In this section, applying our new meta-converse, we derive the second-order expansion of quantum erasure channel in the following Theorem 5.16. This is the first second-order expansion of classical capacity beyond classical-quantum channels (more generally, the image-additive channels introduced in [TT15]).

**Theorem 5.16.** *For any quantum erasure channel $\mathcal{E}_p$ with parameter $p$ and input dimension $d$, we have*

$$C^{(1)}\left(\mathcal{E}_p^{\otimes n},\varepsilon\right) = n(1-p)\log d + \sqrt{np(1-p)(\log d)^2}\,\Phi^{-1}(\varepsilon) + O(\log n), \tag{5.79}$$

*where $\Phi$ is the cumulative distribution function of a standard normal random variable.*

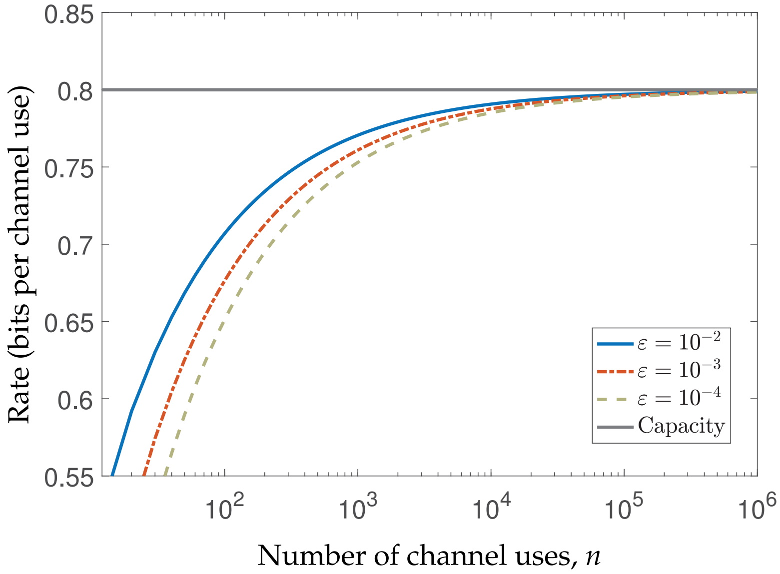


Figure 5.4: Approximation of the non-asymptotic achievable rate region for the quantum erasure channel $\mathcal{E}_p$ with noise parameter $p = 0.2$.

*Proof.* For the converse part, we have

$$C^{(1)}\left(\mathcal{E}_p^{\otimes n}, \varepsilon\right) \leq \min_{\mathcal{M}\in\mathcal{V}} D_H^{\varepsilon}\left(\mathcal{E}_p^{\otimes n}\left(\Phi_{A'^n A^n}\right) \| \mathcal{M}_{A'^n \to B^n}\left(\Phi_{A'^n A^n}\right)\right). \tag{5.80}$$

Note that quantum erasure channels are covariant under the discrete Heisenberg-Weyl unitary group acting on $A'$, and this covariance allows us to restrict the form of the optimal input states to the maximally entangled states. See Lemma 5.31 for details. (One can also refer to Proposition 2 of [TWW17] and find more discussions of the generalized channel divergence in [LKDW18].)

Let us consider the subchannel $\mathcal{M}(\rho) = \frac{1-p}{d}\rho + p|e\rangle\langle e|$ whose Choi-Jamiołkowski matrix is given by

$$J_{\mathcal{M}} = \frac{1-p}{d}\sum_{i,j=0}^{d-1} |ii\rangle\langle jj| + p\sum_{i=0}^{d-1} |i\rangle\langle i| \otimes |d\rangle\langle d|. \tag{5.81}$$

It is easy to see that $\mathcal{M}$ is a constant-bounded subchannel since

$$\mathcal{M}(\rho) \leq \frac{1-p}{d}\mathbb{1}_d + p|e\rangle\langle e|, \ \forall \rho. \tag{5.82}$$

When the number of channel uses is $n$, let us choose $\mathcal{M}_{A'^n \to B^n} = \mathcal{M}_{A' \to B}^{\otimes n}$. Then we can obtain apply Theorem 5.15 and obtain

$$\begin{aligned}
&C^{(1)}\left(\mathcal{E}_p^{\otimes n},\varepsilon\right)\\
&\leq D_H^{\varepsilon}\left(\mathcal{E}_p^{\otimes n}\left(\Phi_{A'^n A^n}\right)\|\mathcal{M}_{A'\to B}^{\otimes n}\left(\Phi_{A'^n A^n}\right)\right)\\
&= nD\left(\mathcal{E}_p\left(\Phi_{A'A}\right)\|\mathcal{M}\left(\Phi_{A'A}\right)\right)+\sqrt{nV\left(\mathcal{E}_p\left(\Phi_{A'A}\right)\|\mathcal{M}\left(\Phi_{A'A}\right)\right)}\,\Phi^{-1}\left(\varepsilon\right)+O\left(\log n\right)\\
&= n\left(1-p\right)\log d+\sqrt{np\left(1-p\right)\left(\log d\right)^2}\,\Phi^{-1}\left(\varepsilon\right)+O\left(\log n\right). \qquad (5.83)
\end{aligned}$$

In the third line, we use second-order expansion of quantum hypothesis testing relative entropy [TH13, Li14] (cf. Eq. (2.70)). The fourth line follows by direct calculation.

For the direct part, denote $\mathcal{F}_1\left(\rho\right)=\sum_{i=0}^{d-1}\langle i|\rho|i\rangle|i\rangle\langle i|$, and $\mathcal{F}_2\left(\rho\right)=\sum_{i=0}^{d}\langle i|\rho|i\rangle|i\rangle\langle i|$, which are both classical channels. Then $\mathcal{N}_p=\mathcal{F}_2\circ\mathcal{E}_p\circ\mathcal{F}_1$ is a classical erasure channel. We have

$$C^{(1)}\left(\mathcal{E}_p^{\otimes n},\varepsilon\right)\geq C^{(1)}\left(\mathcal{N}_p^{\otimes n},\varepsilon\right) \qquad (5.84)$$

$$= n\left(1-p\right)\log d+\sqrt{np\left(1-p\right)\left(\log d\right)^2}\,\Phi^{-1}\left(\varepsilon\right)+O\left(\log n\right), \qquad (5.85)$$

where the equality comes from the result in [PPV10]. □

## 5.4 Asymptotic communication via quantum channels

### 5.4.1 SDP strong converse bounds for the classical capacity

It is well known that evaluating the classical capacity of a general channel is extremely difficult. To the best of our knowledge, the only known nontrivial strong converse bound for the classical capacity is the entanglement-assisted capacity [BSST99] and there is also computable single-shot upper bound derived from entanglement measures [BEHY11]. In this section, we will derive two SDP strong converse bounds for the classical capacity of a general quantum channel. Our bounds are efficiently computable and do not depend on any special properties of the channel. We also show that for some classes of quantum channels, our bound can be strictly smaller than the entanglement-assisted capacity and the previous bound in [BEHY11].

Before introducing the strong converse bounds, we first introduce an SDP to estimate the optimal success probability of classical communication via multiple uses of the channel.

**Proposition 5.17.** *For any quantum channel $\mathcal{N}$ and given $m$,*

$$p_{s,NS\cap PPT}\left(\mathcal{N},m\right)\leq f^{+}\left(\mathcal{N},m\right),$$

*where*

$$\begin{aligned} f^+(\mathcal{N},m) = \min\ & \operatorname{Tr} Z_B \\ s.t.\ & -R_{AB} \le J_{\mathcal{N}}^{T_B} \le R_{AB}, \\ & -m\mathbb{1}_A \otimes Z_B \le R_{AB}^{T_B} \le m\mathbb{1}_A \otimes Z_B. \end{aligned} \tag{5.86}$$

*Furthermore, it holds that* $p_{s,NS\cap PPT}(\mathcal{N}_1\otimes\mathcal{N}_2, m_1m_2) \le f^+(\mathcal{N}_1,m_1)\, f^+(\mathcal{N}_2,m_2)$. *Consequently,*

$$p_{s,NS\cap PPT}\left(\mathcal{N}^{\otimes n}, m^n\right) \le f^+(\mathcal{N},m)^n. \tag{5.87}$$

*Proof.* We utilize the duality theory of semidefinite programming in the proof. To be specific, the dual SDP of $f^+(\mathcal{N},m)$ is given by

$$\begin{aligned} f^+(\mathcal{N},m) = \max\ & \operatorname{Tr} J_{\mathcal{N}}(V_{AB}-X_{AB})^{T_B} \\ \text{s.t. } & V_{AB}+X_{AB} \le (W_{AB}-Y_{AB})^{T_B}, \\ & \operatorname{Tr}_A(W_{AB}+Y_{AB}) \le \mathbb{1}_B/m, \\ & V_{AB}, X_{AB}, W_{AB}, Y_{AB} \ge 0. \end{aligned} \tag{5.88}$$

It is worth noting that the optimal values of the primal and the dual SDPs above coincide. This is a consequence of strong duality. By Slater's condition, one simply needs to show that there exists positive definite $V_{AB}$, $X_{AB}$, $W_{AB}$ and $Y_{AB}$ such that $V_{AB}+X_{AB} < (W_{AB}-Y_{AB})^{T_B}$ and $\operatorname{Tr}_A(W_{AB}+Y_{AB}) < \mathbb{1}_B/m$, which holds for $W_{AB} = 2Y_{AB} = 5V_{AB} = X_{AB} = \mathbb{1}_{AB}/2md_A$.

In SDP (5.88), let us choose $X_{AB} = Y_{AB} = 0$ and $V_{AB}^{T_B} = W_{AB}$, then we have that

$$\begin{aligned} f^+(\mathcal{N},m) &\ge \max\{\operatorname{Tr} J_{\mathcal{N}} W_{AB} : W_{AB}, W_{AB}^{T_B} \ge 0, \operatorname{Tr}_A W_{AB} \le \mathbb{1}_B/m\} \\ &\ge p_{s,NS\cap PPT}(\mathcal{N},m), \end{aligned} \tag{5.89}$$

which means that the SDP (5.88) of $f^+(\mathcal{N},m)$ is a relaxation of the SDP (5.11) of $p_{s,NS\cap PPT}(\mathcal{N},m)$.

To see $p_{s,NS\cap PPT}(\mathcal{N}_1\otimes\mathcal{N}_2, m_1m_2) \le f^+(\mathcal{N}_1,m_1)\, f^+(\mathcal{N}_2,m_2)$, we first suppose that the optimal solution to SDP (5.86) of $f^+(\mathcal{N}_1,m_1)$ is $\{Z_1,R_1\}$ and the optimal solution to SDP (5.86) of $f^+(\mathcal{N}_2,m_2)$ is $\{Z_2,R_2\}$. Let us denote the Choi-Jamiołkowski matrix of $\mathcal{N}_1$ and $\mathcal{N}_2$ by $J_1$ and $J_2$, respectively. It is easy to see that

$$\begin{aligned} & R_1\otimes R_2 + J_1^{T_B}\otimes J_2^{T_{B'}} \\ = & \frac{1}{2}[\left(R_1+J_1^{T_B}\right)\otimes\left(R_2+J_2^{T_{B'}}\right) + \left(R_1-J_1^{T_B}\right)\otimes\left(R_2-J_2^{T_{B'}}\right)] \ge 0, \end{aligned} \tag{5.90}$$

and

$$\begin{aligned} & R_1 \otimes R_2 - J_1^{T_B} \otimes J_2^{T_{B'}} \\ = & \frac{1}{2}[\left(R_1 + J_1^{T_B}\right) \otimes \left(R_2 - J_2^{T_{B'}}\right) + \left(R_1 - J_1^{T_B}\right) \otimes \left(R_2 + J_2^{T_{B'}}\right)] \geq 0. \end{aligned} \tag{5.91}$$

Therefore, we have that

$$-R_1 \otimes R_2 \leq J_1^{T_B} \otimes J_2^{T_{B'}} \leq R_1 \otimes R_2.$$

Applying similar techniques, it is easy to prove that

$$-m_1 m_2 \mathbb{1}_{AA'} \otimes Z_1 \otimes Z_2 \leq R_1^{T_B} \otimes R_2^{T_{B'}} \leq m_1 m_2 \mathbb{1}_{AA'} \otimes Z_1 \otimes Z_2.$$

Hence, $\{Z_1 \otimes Z_2, R_1 \otimes R_2\}$ is a feasible solution to the SDP (5.86) of $f^+ (\mathcal{N}_1 \otimes \mathcal{N}_2, m_1 m_2)$, which means that

$$p_{s,NS\cap PPT} (\mathcal{N}_1 \otimes \mathcal{N}_2, m_1 m_2) \leq f^+ (\mathcal{N}_1 \otimes \mathcal{N}_2, m_1 m_2) \tag{5.92}$$

$$\leq \operatorname{Tr} Z_1 \otimes Z_2 = f^+ (\mathcal{N}_1, m_1) f^+ (\mathcal{N}_2, m_2). \tag{5.93}$$

□

Now, we are able to derive the strong converse bounds of the classical capacity.

**Theorem 5.18.** *For any quantum channel* $\mathcal{N}$*,*

$$C(\mathcal{N}) \leq C_{NS\cap PPT}(\mathcal{N}) \leq C_\beta(\mathcal{N}) = \log \beta(\mathcal{N}) \leq \log \left(d_B \|J_{\mathcal{N}}^{T_B}\|_\infty\right),$$

*where*

$$\begin{aligned} \beta(\mathcal{N}) = \min \ & \operatorname{Tr} S_B \\ \text{s.t.} \ & -R_{AB} \leq J_{\mathcal{N}}^{T_B} \leq R_{AB}, -\mathbb{1}_A \otimes S_B \leq R_{AB}^{T_B} \leq \mathbb{1}_A \otimes S_B. \end{aligned} \tag{5.94}$$

*In particular, when the communication rate exceeds* $C_\beta(\mathcal{N})$*, the error probability goes to one exponentially fast as the number of channel uses increases.*

*Proof.* For $n$ uses of the channel, we suppose that the rate of the communication is $r$. By Proposition 5.17, we have that

$$p_{s,NS\cap PPT} \left(\mathcal{N}^{\otimes n}, 2^{rn}\right) \leq f^+ (\mathcal{N}, 2^r)^n. \tag{5.95}$$

Therefore, the $n$-shot error probability satisfies that

$$\varepsilon_n = 1 - p_{s,NS\cap PPT}\left(\mathcal{N}^{\otimes n}, 2^{rn}\right) \geq 1 - f^{+}\left(\mathcal{N}, 2^{r}\right)^{n}. \tag{5.96}$$

Suppose that the optimal solution to the SDP (5.94) of $\beta(\mathcal{N})$ is $\{S_0, R_0\}$. It is easy to verify that $\{S_0/\operatorname{Tr} S_0, R_0\}$ is a feasible solution to the SDP (5.86) of $f^{+}(\mathcal{N}, \operatorname{Tr} S_0)$. Therefore,

$$f^{+}(\mathcal{N}, \beta(\mathcal{N})) \leq \operatorname{Tr}(S_0/\operatorname{Tr} S_0) = 1.$$

It is not difficult to see that $f^{+}(\mathcal{N}, m)$ monotonically decreases when $m$ increases. Thus, for any $2^r > \beta(\mathcal{N})$, we have $f^{+}(\mathcal{N}, 2^r) < 1$. Then, by Eq. (5.96), it is clear that the corresponding $n$-shot error probability $\varepsilon_n$ will go to one exponentially fast as $n$ increases. Hence, $C_\beta(\mathcal{N})$ is a strong converse bound for the NS∩PPT-assisted classical capacity of $\mathcal{N}$.

Furthermore, let us choose $R_{AB} = \|J_{\mathcal{N}}^{T_B}\|_\infty \mathbb{1}_{AB}$ and $S_B = \|J_{\mathcal{N}}^{T_B}\|_\infty \mathbb{1}_B$. It is clear that $\{R_{AB}, S_B\}$ is a feasible solution to the SDP (5.94) of $\beta(\mathcal{N})$, which means that

$$\beta(\mathcal{N}) \leq d_B \|J_{\mathcal{N}}^{T_B}\|_\infty. \tag{5.97}$$

□

**Remark** $C_\beta$ has some remarkable properties. For example, it is additive for different quantum channels $\mathcal{N}_1$ and $\mathcal{N}_2$:

$$C_\beta(\mathcal{N}_1 \otimes \mathcal{N}_2) = C_\beta(\mathcal{N}_1) + C_\beta(\mathcal{N}_2). \tag{5.98}$$

This can be proved by utilizing semidefinite programming duality.

With similar techniques, we are going to show another SDP strong converse bound for the classical capacity of a general quantum channel.

**Theorem 5.19.** *For a quantum channel $\mathcal{N}$, we derive the following strong converse bound for the NS∩PPT assisted classical capacity, i.e.,*

$$C(\mathcal{N}) \leq C_{NS\cap PPT}(\mathcal{N}) \leq C_\zeta(\mathcal{N}) = \log \zeta(\mathcal{N})$$

*with*

$$\begin{aligned} \zeta(\mathcal{N}) = \min\ & \operatorname{Tr} S_B \\ \text{s.t. } & V_{AB} \geq J_{\mathcal{N}}, -\mathbb{1}_A \otimes S_B \leq V_{AB}^{T_B} \leq \mathbb{1}_A \otimes S_B \end{aligned} \tag{5.99}$$

*And if the communication rate exceeds $C_\zeta(\mathcal{N})$, the error probability will go to one exponentially fast as the number of channel uses increase.*

*Proof.* We first introduce the following SDP to estimate the optimal success probability:

$$\begin{aligned} \widetilde{f}^{+}(\mathcal{N},m) = \min\ & \operatorname{Tr} S_B \\ \text{s.t. } & V_{AB} \geq J_{\mathcal{N}}, \\ & -m\mathbb{1}_A \otimes S_B \leq V_{AB}^{T_B} \leq m\mathbb{1}_A \otimes S_B. \end{aligned} \tag{5.100}$$

Similar to Proposition 5.17, we can prove that

$$p_{s,NS\cap PPT}\left(\mathcal{N}^{\otimes n}, m^n\right) \leq \widetilde{f}^{+}(\mathcal{N},m)^n. \tag{5.101}$$

Then, when the communication rate exceeds $C_\zeta(\mathcal{N})$, we can use the technique in Theorem 5.18 to prove that the error probability will go to one exponentially fast as the number of channel uses increase. □

As an example, we first apply our bounds to the qudit noiseless channel. In this case, the bounds are tight and strictly smaller than the entanglement-assisted classical capacity.

**Proposition 5.20.** *For the qudit noiseless channel $I_d(\rho) = \rho$, it holds that*

$$C(I_d) = C_\beta(I_d) = C_\zeta(I_d) = \log d < 2\log d = C_E(I_d). \tag{5.102}$$

*Proof.* It is clear that $C(I_d) \geq \log d$. By the fact that $\|J_{I_d}^{T_B}\|_\infty = 1$, it is easy to see that $C_\beta(I_d) \leq \log d \|J_{I_d}^{T_B}\|_\infty = \log d$. Similarly, we also have $C_\zeta(I_d) \leq \log d$. And $C_E(I_d) = 2\log d$ is due to the superdense coding [BW92]. □

### 5.4.2 Amplitude damping channel

Amplitude damping is the process of asymmetric relaxation in a quantum system, such as spontaneous emission observed in trapped ions [BLMW04]. It has been considered as a basic noise process in quantum information processing [NC10].

The amplitude damping channel is given as

$$\mathcal{N}_\gamma^{AD} = \sum_{i=0}^{1} E_i \cdot E_i^\dagger, \tag{5.103}$$

where the Choi-Kraus operators $E_i$ for the channel are

$$E_0 = |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|, \tag{5.104}$$

$$E_1 = \sqrt{\gamma}|0\rangle\langle 1| \tag{5.105}$$

and we call $\gamma \in (0,1)$ the amplitude damping parameter.

The Holevo capacity of this channel is given by [GF05]

$$C\left(\mathcal{N}_\gamma^{AD}\right) \geq \max_{0\leq p\leq 1}\left\{H_2[(1-\gamma)\,p] - H_2\left(1+\frac{\sqrt{1-4\,(1-\gamma)\,\gamma p^2}}{2}\right)\right\}, \tag{5.106}$$

where $H_2$ is the binary entropy. However, its classical capacity remains unknown so far. The only known nontrivial and meaningful upper bound for the classical capacity of the amplitude damping channel was established in [BEHY11]. As an application of theorems 5.18 and 5.19, we show a strong converse bound for the classical capacity of the qubit amplitude damping channel. Remarkably, our bound improves the best previously known upper bound [BEHY11].

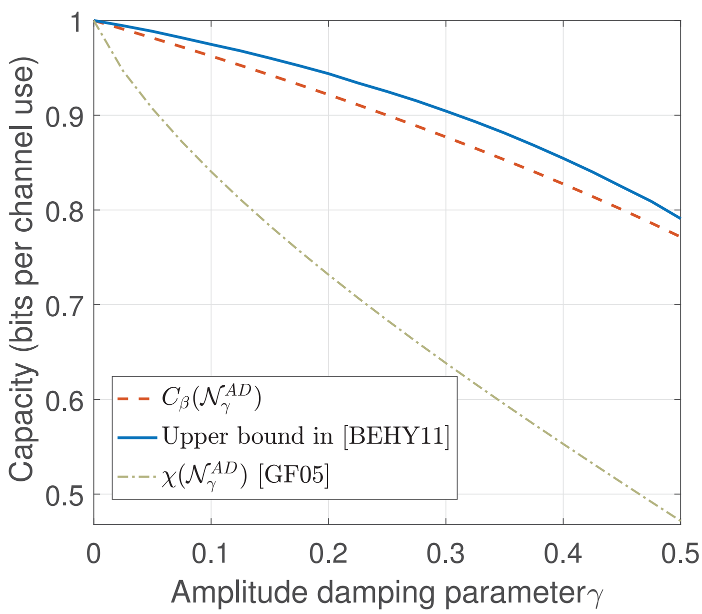


Figure 5.5: The solid line depicts $C_\beta\left(\mathcal{N}_\gamma^{AD}\right)$, the dashed line depicts the previous bound of $C\left(\mathcal{N}_\gamma^{AD}\right)$ [BEHY11], and the dotted line depicts the lower bound [GF05]. Our bound is tighter than the previous bound in [BEHY11].

**Theorem 5.21.** *For amplitude damping channel $\mathcal{N}_\gamma^{AD}$,*

$$C_{NS\cap PPT}\left(\mathcal{N}_\gamma^{AD}\right) \leq C_\zeta\left(\mathcal{N}_\gamma^{AD}\right) = C_\beta\left(\mathcal{N}_\gamma^{AD}\right) = \log\left(1+\sqrt{1-\gamma}\right).$$

*As a consequence,*

$$C\left(\mathcal{N}_\gamma^{AD}\right) \leq \log\left(1+\sqrt{1-\gamma}\right).$$

*Proof.* Suppose that

$$S_B = \frac{\sqrt{1-\gamma}+1+\gamma}{2}|0\rangle\langle 0| + \frac{\sqrt{1-\gamma}+1-\gamma}{2}|1\rangle\langle 1|$$

and

$$V_{AB} = J_\gamma^{AD} + \left(\sqrt{1-\gamma}-1+\gamma\right)|v\rangle\langle v|$$

with $|v\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

It is clear that $V_{AB} \geq J_\gamma^{AD}$. Moreover, it is easy to see that

$$\mathbb{1}_A \otimes S_B - V_{AB}^{T_B} = \frac{\sqrt{1-\gamma}+1-\gamma}{2}\left(|01\rangle - |10\rangle\right)\left(\langle 01| - \langle 10|\right) \geq 0$$

and

$$\begin{aligned}\mathbb{1}_A \otimes S_B + V_{AB}^{T_B} = &\left(\sqrt{1-\gamma}+1+\gamma\right)|00\rangle\langle 00| \\ &+ \left(\sqrt{1-\gamma}+1-\gamma\right)|11\rangle\langle 11| \\ &+ \frac{\sqrt{1-\gamma}+1-\gamma}{2}\left(|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01|\right) \\ &+ \frac{\sqrt{1-\gamma}+1+3\gamma}{2}|10\rangle\langle 10| \geq 0.\end{aligned}$$

Therefore, $\{S_B, V_{AB}\}$ is a feasible solution to SDP (5.99), which means that

$$C_\zeta\left(\mathcal{N}_\gamma^{AD}\right) \leq \log \operatorname{Tr} S_B = \log\left(1+\sqrt{1-\gamma}\right). \tag{5.107}$$

One can also use the dual SDP of $C_\beta$ to show that $C_\beta\left(\mathcal{N}_\gamma^{AD}\right) \geq \log\left(1+\sqrt{1-\gamma}\right)$. Hence, we have that

$$C_\zeta\left(\mathcal{N}_\gamma^{AD}\right) = \log\left(1+\sqrt{1-\gamma}\right). \tag{5.108}$$

Similarly, it can also be calculated that

$$C_\beta\left(\mathcal{N}_\gamma^{AD}\right) = \log\left(1+\sqrt{1-\gamma}\right). \tag{5.109}$$

□

**Remark:** We compare our bound with the previous upper bound [BEHY11] and lower bound [GF05] in Figure 5.5. It is also worth noting that our bound is strictly smaller than the entanglement-assisted capacity when $\gamma \leq 0.75$ as shown in the following Figure 5.6. It is clear that our bound provides a tighter bound to the classical

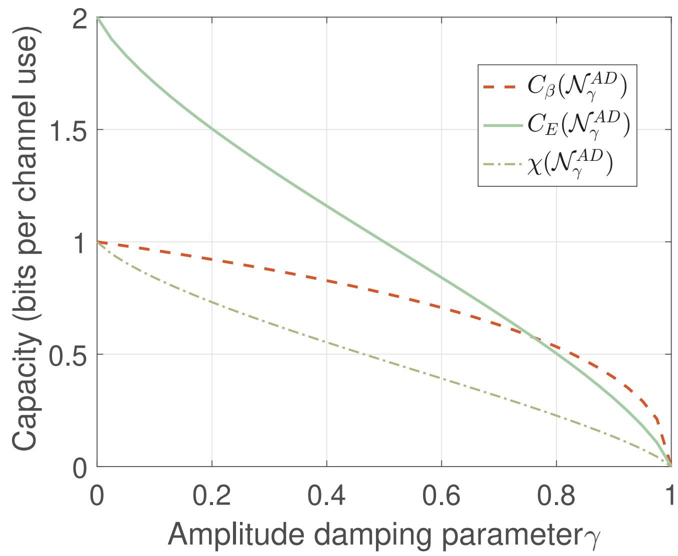


Figure 5.6: The solid line depicts $C_\beta\left(\mathcal{N}_\gamma^{AD}\right)$ while the dashed line depicts $C_E\left(\mathcal{N}_\gamma^{AD}\right)$. It is worth noting that $C_\beta\left(\mathcal{N}_\gamma^{AD}\right)$ is strictly smaller than $C_E\left(\mathcal{N}_\gamma^{AD}\right)$ for any $\gamma \leq 0.75$.

capacity than the previous bound [BEHY11].

### 5.4.3 A special class of quantum channels

In this chapter, we focus on a class of qutrit-to-qutrit channels which will be used to show the separation between quantum Lovász number and entanglement-assisted zero-error classical capacity in Chapter 7. It turns out that this class of channels also has strong converse property for classical and private communication. To be specific, the channel is given by $\mathcal{N}_\alpha(\rho) = E_\alpha \rho E_\alpha^\dagger + D_\alpha \rho D_\alpha^\dagger$ with

$$E_\alpha = \sin\alpha|0\rangle\langle 1| + |1\rangle\langle 2|, D_\alpha = \cos\alpha|2\rangle\langle 1| + |1\rangle\langle 0|.$$

This qutrit-qutrit channel $\mathcal{N}_\alpha$ is motivated in the similar spirit of the amplitude damping channel, which exhibits a significant difference from the classical channels.

The first Choi-Kraus operator $E_\alpha$ annihilates the ground state $|0\rangle\langle 0|$:

$$E_\alpha|0\rangle\langle 0|E_\alpha^\dagger = 0,$$

and it decays the state $|1\rangle\langle 1|$ to the ground state $|0\rangle\langle 0|$:

$$E_\alpha|1\rangle\langle 1|E_\alpha^\dagger = \sin^2\alpha|0\rangle\langle 0|.$$

Meanwhile, $E_\alpha$ also transfer the state $|2\rangle\langle 2|$ to $|1\rangle\langle 1|$, i.e., $E_\alpha|2\rangle\langle 2|E_\alpha^\dagger = |1\rangle\langle 1|$. On the other hand, the choice of $D_\alpha$ above ensures that

$$E_\alpha^\dagger E_\alpha + D_\alpha^\dagger D_\alpha = \mathbb{1},$$

which means that the operators $E_\alpha$ and $D_\alpha$ are valid Kraus operators for a quantum channel.

It follows that the complementary channel of $\mathcal{N}_\alpha$ is $\mathcal{N}_\alpha^c(\rho) = \sum_{i=0}^2 F_{i,\alpha}\rho F_{i,\alpha}^\dagger$ with

$$F_{0,\alpha} = \sin\alpha|0\rangle\langle 1|,\ F_{1,\alpha} = |0\rangle\langle 2| + |1\rangle\langle 0|,\ F_{2,\alpha} = \cos\alpha|1\rangle\langle 1|.$$

**Proposition 5.22.** *For $\mathcal{N}_\alpha$ $(0 < \alpha \le \pi/4)$, we have that*

$$C(\mathcal{N}_\alpha) = C_{NS\cap PPT}(\mathcal{N}_\alpha) = C_\beta(\mathcal{N}_\alpha) = 1.$$

*Proof.* Suppose the $Z_B = \sin^2\alpha|0\rangle\langle 0| + \cos^2\alpha|2\rangle\langle 2| + |1\rangle\langle 1|$ and

$$\begin{aligned} R_{AB} =& |01\rangle\langle 01| + |11\rangle\langle 11| + |21\rangle\langle 21| + \sin^2\alpha\left(|10\rangle\langle 10| + |20\rangle\langle 20|\right) \\ &+ \cos^2\alpha\left(|02\rangle\langle 02| + |12\rangle\langle 12|\right) + \sin\alpha\cos\alpha\left(|02\rangle\langle 20| + |20\rangle\langle 02|\right). \end{aligned}$$

It is easy to check that

$$-R_{AB} \le J_{\mathcal{N}_\alpha}^{T_B} \le R_{AB} \text{ and } -\mathbb{1}_A\otimes Z_B \le R_{AB}^{T_B} \le \mathbb{1}_A\otimes Z_B,$$

where $J_{\mathcal{N}_\alpha}$ is the Choi-Jamiołkowski matrix of $\mathcal{N}_\alpha$.

Therefore, $\{Z_B, R_{AB}\}$ is a feasible solution of SDP (5.94) of $\beta(\mathcal{N}_\alpha)$, which means that

$$\beta(\mathcal{N}_\alpha) \le \operatorname{Tr} Z_B = 2.$$

Noticing that we can use input $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ to transmit two messages via $\mathcal{N}$, we can conclude that

$$C(\mathcal{N}_\alpha) = C_{NS\cap PPT}(\mathcal{N}_\alpha) = 1.$$

□

**Remark 5.23.** We note that in Chapter 7, we show that the entanglement-assisted capacity of $\mathcal{N}_\alpha$ is given by

$$C_E(\mathcal{N}_\alpha) = 2.$$

Therefore, for $\mathcal{N}_\alpha$ $(0 < \alpha \le \pi/4)$, our bound $C_\beta$ is strictly smaller than the entanglement-assisted capacity. In this case, we also note that $C_\beta(\mathcal{N}_\alpha) < C_\zeta(\mathcal{N}_\alpha)$. However, it remains unknown whether $C_\beta$ is always smaller than or equal to $C_\zeta$.

Furthermore, it is easy to see that $\mathcal{N}_\alpha$ is neither an entanglement-breaking channel nor a Hadamard channel. Note also that $\mathcal{N}_\alpha$ does not belong to the three classes in [KW09], for which the strong converse for classical capacity has been established. Thus, our results show a new class of quantum channels which satisfy the strong converse property for classical capacity.

Moreover, we find that the strong converse property also holds for the private classical capacity [Dev05, CWY04] of $\mathcal{N}_\alpha$. Note that private capacity requires that no information leaked to the environment and is usually called $P(\mathcal{N})$. Recently, several converse bounds for private communication were established in [TGW14, PLOB17, CMH17, WTB17, Wil16].

**Proposition 5.24.** *The private capacity of $\mathcal{N}_\alpha$ is exactly one bit, i.e., $P(\mathcal{N}_\alpha) = 1$. In particular,*

$$Q(\mathcal{N}_\alpha) \le \log(1+\cos\alpha) < 1 = P(\mathcal{N}_\alpha) = C(\mathcal{N}_\alpha) = \frac{1}{2} C_E(\mathcal{N}_\alpha).$$

*Proof.* On one hand, it is easy to see that $P(\mathcal{N}_\alpha) \le C(\mathcal{N}_\alpha) = C_\beta(\mathcal{N}_\alpha) = 1$.

On the other hand, Alice can choose two input states $|\psi_0\rangle = |1\rangle$ and $|\psi_1\rangle = \cos\alpha|0\rangle + \sin\alpha|2\rangle$, then the corresponding output states Bob received are

$$\begin{aligned}
\mathcal{N}_\alpha(|\psi_0\rangle\langle\psi_0|) &= \sin\alpha^2|0\rangle\langle 0| + \cos\alpha^2|2\rangle\langle 2|,\\
\mathcal{N}_\alpha(|\psi_1\rangle\langle\psi_1|) &= |1\rangle\langle 1|.
\end{aligned}$$

It is clear that Bob can perfectly distinguish these two output states. Meanwhile, the corresponding outputs of the complementary channel $\mathcal{N}_\alpha^c$ are same, i.e.,

$$\mathcal{N}_\alpha^c(|\psi_0\rangle\langle\psi_0|) = \mathcal{N}_\alpha^c(|\psi_1\rangle\langle\psi_1|) = \sin\alpha^2|0\rangle\langle 0| + \cos\alpha^2|1\rangle\langle 1|,$$

which means that the environment obtain zero information during the communication.

Applying the SDP bound of the quantum capacity in [WD16a], the quantum capacity of $\mathcal{N}_\alpha$ is strictly smaller than $\log(1+\cos\alpha)$. □

Our result establishes the strong converse property for both the classical and private capacities of $\mathcal{N}_\alpha$. For the classical capacity, such a property was previously only known for classical channels, identity channel, entanglement-breaking channels, Hadamard channels and particular covariant quantum channels [WWY14, KW09]. For the private capacity, such a property was previously only known for generalized dephasing channels and quantum erasure channels [WTB17]. Moreover, our result also shows a simple example of the distinction between the private and the quantum capacities, which were discussed in [HHHO05, LLSS14].

### 5.4.4 New converse via channel divergence

Before introducing the new converse, we first recall the divergence radius representation of the Holevo capacity introduced in [SW01]:

$$\chi(\mathcal{N}) := \min_{\sigma_B} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\rho_{A'}) \,\|\sigma_B\right). \tag{5.110}$$

In the same spirit of the divergence radius, we are going to introduce a channel divergence to bound the capability of classical communication. By substituting the relative entropy for the hypothesis testing relative entropy in our meta-converse we define the following quantity, which we call the $\gamma$-information of the channel $\mathcal{N}$.

**Definition 5.25.** For a quantum channel $\mathcal{N} : \mathcal{L}(A') \to \mathcal{L}(B)$, we define

$$\gamma(\mathcal{N}) := \min_{\mathcal{M}\in\mathcal{V}} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\phi_{A'A}) \,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right), \tag{5.111}$$

where $\phi_{A'A}$ is a purification of $\rho_{A'}$.

We also introduce its regularization,

$$\gamma^{\infty}(\mathcal{N}) := \limsup_{n\to\infty} \frac{1}{n}\gamma\left(\mathcal{N}^{\otimes n}\right). \tag{5.112}$$

It is worth noting that one could exchange the min and max due to the fact that the function

$$D\left(\mathcal{N}_{A'\to B}(\phi_{A'A}) \,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right) \tag{5.113}$$

is concave in $\rho_{A'}$. (The detailed proof can be found in [WFT17].) This means

$$\gamma(\mathcal{N}) = \min_{\mathcal{M}\in\mathcal{V}} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\phi_{A'A}) \,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right) \tag{5.114}$$

$$= \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}} D\left(\mathcal{N}_{A'\to B}(\phi_{A'A}) \,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right). \tag{5.115}$$

**Proposition 5.26.** *For any channel $\mathcal{N}$, we have $\chi(\mathcal{N}) \le \gamma(\mathcal{N})$ and $C(\mathcal{N}) \le \Upsilon^{\infty}(\mathcal{N})$.*

*Proof.* We have the following chain of inequalities:

$$\gamma(\mathcal{N}) = \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}} D\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right) \tag{5.116}$$

$$= \min_{\mathcal{M}\in\mathcal{V}} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\|\mathcal{M}_{A'\to B}(\phi_{A'A})\right) \tag{5.117}$$

$$\geq \min_{\mathcal{M}\in\mathcal{V}} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\rho_{A'})\,\|\mathcal{M}_{A'\to B}(\rho_{A'})\right) \tag{5.118}$$

$$\geq \min_{\mathcal{M}\in\mathcal{V}} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\rho_{A'})\,\|\sigma_{\mathcal{M}}\right) \tag{5.119}$$

$$\geq \min_{\sigma_B} \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\rho_{A'})\,\|\sigma_B\right) \tag{5.120}$$

$$= \chi(\mathcal{N}). \tag{5.121}$$

The third line follows since we trace out $A$ system. The fourth line follows since for any $\mathcal{M}\in\mathcal{V}$ and $\rho_{A'}$, there exists a state $\sigma_{\mathcal{M}}$ independent of $\rho_{A'}$ such that $\mathcal{M}_{A'\to B}(\rho_{A'}) \leq \sigma_{\mathcal{M}}$. Due to the dominance of relative entropy, we have the inequality. The fifth line follows since we relax the feasible set of the minimization to a larger set.

Finally, according to the HSW theorem, we have

$$C(\mathcal{N}) = \limsup_{n\to\infty} \frac{1}{n}\chi\left(\mathcal{N}^{\otimes n}\right) \leq \limsup_{n\to\infty} \frac{1}{n}\gamma\left(\mathcal{N}^{\otimes n}\right) = \gamma^{\infty}(\mathcal{N}). \tag{5.122}$$

□

**Proposition 5.27.** *For any quantum channel $\mathcal{N}$, we have that*

$$\gamma(\mathcal{N}) \leq C_E(\mathcal{N}),\ \gamma^{\infty}(\mathcal{N}) \leq C_E(\mathcal{N}). \tag{5.123}$$

*Proof.* For any state $\sigma_B$ we introduce a trivial channel $\mathcal{M}$ that always outputs $\sigma_B$ via its Choi-Jamiołkowski matrix $J_{\mathcal{M}} = \mathbb{1}_A \otimes \sigma_B$. Then $\mathcal{M}\in\mathcal{V}$ and

$$\min_{\sigma_B} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'})\,\|\rho_A\otimes\sigma_B\right) \tag{5.124}$$

$$= \min_{\sigma_B} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'})\,\|\rho_A^{1/2}(\mathbb{1}_A\otimes\sigma_B)\,\rho_A^{1/2}\right) \tag{5.125}$$

$$\geq \min_{\mathcal{M}\in\mathcal{V}} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'})\,\|\mathcal{M}_{A'\to B}(\phi_{AA'})\right). \tag{5.126}$$

Take maximization over all input state $\rho_{A'}$ on both sides, we have

$$C_E(\mathcal{N}) \geq \gamma(\mathcal{N}). \tag{5.127}$$

Furthermore, since $C_E(\mathcal{N})$ is additive, we have

$$C_E(\mathcal{N}) = \limsup_{n\to\infty} \frac{1}{n} C_E\left(\mathcal{N}^{\otimes n}\right) \geq \limsup_{n\to\infty} \frac{1}{n} \gamma\left(\mathcal{N}^{\otimes n}\right) = \gamma^{\infty}(\mathcal{N}). \tag{5.128}$$

□

**Proposition 5.28.** *For any channel $\mathcal{N}$, we have*

$$\gamma(\mathcal{N}) \leq C_\beta(\mathcal{N}), \; \gamma^{\infty}(\mathcal{N}) \leq C_\beta(\mathcal{N}). \tag{5.129}$$

*Proof.* Take $\mathcal{M} = \frac{1}{\beta(J_{\mathcal{N}})}\mathcal{N}$, then $\mathcal{M} \in \mathcal{V}_\beta \subset \mathcal{V}$ and

$$\gamma(\mathcal{N}) = \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'}) \| \mathcal{M}_{A'\to B}(\phi_{AA'})\right) \tag{5.130}$$

$$\leq \max_{\rho_{A'}} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'}) \| \mathcal{N}_{A'\to B}(\phi_{AA'})\right) + \log\beta(J_{\mathcal{N}}) \tag{5.131}$$

$$= \log\beta(J_{\mathcal{N}}) = C_\beta(\mathcal{N}). \tag{5.132}$$

Furthermore, since $C_\beta(\mathcal{N})$ is additive, we have

$$\gamma^{\infty}(\mathcal{N}) = \limsup_{n\to\infty} \frac{1}{n} \gamma\left(\mathcal{N}^{\otimes n}\right) \leq \limsup_{n\to\infty} \frac{1}{n} C_\beta\left(\mathcal{N}^{\otimes n}\right) = C_\beta(\mathcal{N}). \tag{5.133}$$

□

One could focus on covariant channels which allow us to simplify the set of input states. We call a channel covariant if for any unitary $U_A$, there is a unitary $V_B$ such that $\mathcal{N}_{A\to B}\left(U_A\rho_A U_A^\dagger\right) = V_B \mathcal{N}_{A\to B}(\rho_A) V_B^\dagger$, for all $\rho_A \in \mathcal{S}(A)$.

For covariant channels, one could further show that the $\gamma$-information is a strong converse bound by using symmetry and sandwiched Rényi relative entropy. In the following, we are trying to establish the strong converse of $\gamma$-information and obtain some partial results. Specifically, we show that $\gamma$-information is a strong converse for covariant channels.

Let us introduce

$$\widetilde{\gamma}_\alpha(\mathcal{N}, \rho_{A'}) := \min_{\mathcal{M}\in\mathcal{V}} \widetilde{D}_\alpha\left(\mathcal{N}_{A'\to B}(\phi_{A'A}) \middle\| \mathcal{M}_{A'\to B}(\phi_{A'A})\right), \tag{5.134}$$

where $\phi_{AA'}$ is a purification of $\rho_{A'}$ as usual and $\widetilde{D}_\alpha(\cdot\|\cdot)$ is the sandwiched Rényi relative entropy [MLDS+13, WWY14] (see Eq. (2.58) for the formal definition).

First, we can establish the following estimation of the error probability via $\widetilde{\gamma}_\alpha$.

**Proposition 5.29.** *For any quantum channel $\mathcal{N}_{A'\to B}$ and unassisted code with achievable $(r,n,\varepsilon)$,*

$$\varepsilon \geq 1 - 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(r-\frac{1}{n}\widetilde{\gamma}_\alpha(\mathcal{N}^{\otimes n})\right)}, \tag{5.135}$$

*where* $\widetilde{\gamma}_\alpha(\mathcal{N}) := \max_{\rho_{A'}} \min_{\mathcal{M}\in\mathcal{V}} \widetilde{D}_\alpha\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\middle\|\,\mathcal{M}_{A'\to B}(\phi_{A'A})\right)$.

*Proof.* Suppose $(r,n,\varepsilon)$ is achieved by the average input state $\rho_{A'^n}$. From the proof of Theorem 5.15, we have $C^{(1)}(\mathcal{N}^{\otimes n},\rho_{A'^n},\varepsilon) \leq D_H^{\varepsilon}\left(\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A'^nA^n})\,\middle\|\,\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n})\right)$. Suppose now that the optimal test of $D_H^{\varepsilon}\left(\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A'^nA^n})\,\middle\|\,\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n})\right)$ is

$$\{F_{A^nB^n}, \mathbb{1} - F_{A^nB^n}\}. \tag{5.136}$$

Then, we have

$$nr \leq -\log \operatorname{Tr} F_{A^nB^n}\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n}), \tag{5.137}$$

$$1-\varepsilon \leq \operatorname{Tr} F_{A^nB^n}\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A'^nA^n}). \tag{5.138}$$

Due to the monotonicity of the sandwiched Rényi relative entropy under the test $\{F_{A^nB^n}, \mathbb{1} - F_{A^nB^n}\}$, we have

$$\begin{aligned}&\widetilde{D}_\alpha\left(\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A^nA^n})\,\middle\|\,\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n})\right)\\ \geq&\,\delta_\alpha\left(\operatorname{Tr} F_{A^nB^n}\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A'^nA^n})\,\middle\|\,\operatorname{Tr} F_{A^nB^n}\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n})\right),\end{aligned} \tag{5.139}$$

where $\delta_\alpha(p\|q) = \frac{1}{\alpha-1}\log\left(p^\alpha q^{1-\alpha} + (1-p)^\alpha(1-q)^{1-\alpha}\right)$. Using Eqs. (5.137) and (5.138), we thus find

$$\min_{\mathcal{M}\in\mathcal{V}} \widetilde{D}_\alpha\left(\mathcal{N}^{\otimes n}_{A'\to B}(\phi_{A'^nA^n})\,\middle\|\,\mathcal{M}_{A'^n\to B^n}(\phi_{A'^nA^n})\right) \geq \delta_\alpha\left(\varepsilon\,\middle\|\,1-2^{-nr}\right) \tag{5.140}$$

Maximizing over all average input state $\rho_{A'^n}$, we conclude that

$$\begin{aligned}\widetilde{\gamma}_\alpha(\mathcal{N}^{\otimes n}) &\geq \frac{1}{\alpha-1}\log\left(\varepsilon^\alpha\left(1-2^{-nr}\right)^{1-\alpha} + (1-\varepsilon)^\alpha\left(2^{-nr}\right)^{1-\alpha}\right) &(5.141)\\ &\geq \frac{1}{\alpha-1}\log(1-\varepsilon)^\alpha\left(2^{-nr}\right)^{1-\alpha} &(5.142)\\ &= \frac{\alpha}{\alpha-1}\log(1-\varepsilon) + nr, &(5.143)\end{aligned}$$

which implies that $\varepsilon \geq 1 - 2^{-n\left(\frac{\alpha-1}{\alpha}\right)\left(r-\frac{1}{n}\widetilde{\Upsilon}_\alpha(\mathcal{N}^{\otimes n})\right)}$. □

Then, for covariant channels, we could further establish the following result.

**Proposition 5.30.** *For any covariant channel $\mathcal{N}$,*

$$C(\mathcal{N}) \leq \gamma(\mathcal{N}). \tag{5.144}$$

*Moreover, $\gamma(\mathcal{N})$ is a strong converse bound.*

*Proof.* Exploring the symmetry, we can fix the average input state of $\widetilde{\gamma}_\alpha(\mathcal{N})$ to be the maximally mixed state. (See the following Lemma 5.31.)

Then $\widetilde{\gamma}_\alpha$ is subadditive, i.e., $\widetilde{\gamma}_\alpha(\mathcal{N}^{\otimes n}) \leq n\widetilde{\gamma}_\alpha(\mathcal{N})$. Thus from Eq. (5.135), we have

$$\varepsilon \geq 1 - 2^{-n\left(\frac{\alpha-1}{\alpha}\right)(r-\widetilde{\gamma}_\alpha(\mathcal{N}))}. \tag{5.145}$$

The quantity $\widetilde{\gamma}_\alpha(\mathcal{N})$ is monotonically increasing in $\alpha$. Following the proof of Lemma 3 in [TWW17], we can also show that

$$\lim_{\alpha\to 1^+} \widetilde{\gamma}_\alpha(\mathcal{N}) = \gamma(\mathcal{N}). \tag{5.146}$$

Hence, for $r > \gamma(\mathcal{N})$, there always exists an $\alpha > 1$ such that $r > \widetilde{\gamma}_\alpha(\mathcal{N})$. Therefore, $\varepsilon$ will to to 1 as $n$ goes to infinity. □

Let us recall the definition of G-covariant channel in Definition 2.8. Let $G$ be a finite group, and for every $g \in G$, let $g \to U_A(g)$ and $g \to V_B(g)$ be unitary representation acting on the input and output spaces of the channel, respectively. Then a quantum channel $\mathcal{N}_{A\to B}$ is G-covariant if $\mathcal{N}_{A\to B}\left(U_A(g)\rho_A U_A^\dagger(g)\right) = V_B(g)\mathcal{N}_{A\to B}(\rho_A)V_B^\dagger(g)$ for all $\rho_A \in \mathcal{S}(A)$. The average state is $\overline{\rho}_A = \frac{1}{|G|}\sum_g U_A(g)\rho_{A'}U_A^\dagger(g)$.

**Lemma 5.31.** *For any G-covariant channel $\mathcal{N}_{A'\to B}$, it holds that*

$$\widetilde{\gamma}_\alpha(\mathcal{N}, \rho_{A'}) \leq \widetilde{\gamma}_\alpha(\mathcal{N}, \overline{\rho}_{A'}). \tag{5.147}$$

*Proof.* The following proof is a direct adaption of Proposition 2 in [TWW17]. Consider the state $|\psi\rangle_{PAA'} = \sum_g \frac{1}{\sqrt{|G|}}|g\rangle \otimes (\mathbb{1}_A \otimes U_{A'}(g))|\phi^\rho_{AA'}\rangle$ which purifies $\overline{\rho}_{A'}$. Then for any fixed CP map $\mathcal{M}_{A'\to B} \in \mathcal{V}$, we have the following chain of inequalities:

$$\begin{aligned}
&\widetilde{D}_\alpha \left(\mathcal{N}_{A'\to B}\left(\psi_{PAA'}\right) \middle\| \mathcal{M}_{A'\to B}\left(\psi_{PAA'}\right)\right) \\
\geq &\widetilde{D}_\alpha \left(\sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{N}_{A'\to B} \circ \mathcal{U}_{A'}(g)\left(\phi_{A'A}\right) \middle\| \sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{M}_{A'\to B} \circ \mathcal{U}_{A'}(g)\left(\phi_{A'A}\right)\right) \\
= &\widetilde{D}_\alpha \left(\sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{V}_B(g) \circ \mathcal{N}_{A'\to B}\left(\phi_{A'A}\right) \middle\| \sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{M}_{A'\to B} \circ \mathcal{U}_{A'}(g)\left(\phi_{A'A}\right)\right) \\
= &\widetilde{D}_\alpha \left(\sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{N}_{A'\to B}\left(\phi_{A'A}\right) \middle\| \sum_g \frac{1}{|G|} |g\rangle\langle g|_P \otimes \mathcal{V}_B^\dagger(g) \circ \mathcal{M}_{A'\to B} \circ \mathcal{U}_{A'}(g)\left(\phi_{A'A}\right)\right) \\
\geq &\widetilde{D}_\alpha \left(\mathcal{N}_{A'\to B}\left(\phi_{A'A}\right) \middle\| \sum_g \frac{1}{|G|} \mathcal{V}_B^\dagger(g) \circ \mathcal{M}_{A'\to B} \circ \mathcal{U}_{A'}(g)\left(\phi_{A'A}\right)\right) \\
\geq &\min_{\mathcal{M}\in\mathcal{V}} \widetilde{D}_\alpha \left(\mathcal{N}_{A'\to B}\left(\phi_{A'A}\right) \middle\| \mathcal{M}_{A'\to B}\left(\phi_{A'A}\right)\right)
\end{aligned}$$

The second line follows from monotonicity of the sandwiched Rényi relative entropy under the CPTP map $\sum_g |g\rangle\langle g| \cdot |g\rangle\langle g|$. The third line follows from the $G$-invariance of $\mathcal{N}_{A'\to B}$. The fourth line follows from unitary invariance of the sandwiched Rényi relative entropy under $\sum_g |g\rangle\langle g| \otimes V_B^\dagger(g)$. The fifth line follows from monotonicity of the sandwiched Rényi relative entropy under the partial trace over $P$. The last line follows from the fact that $\sum_g \frac{1}{|G|}\mathcal{V}_B^\dagger(g) \circ \mathcal{M}_{A'\to B} \circ \mathcal{U}_{A'}(g)$ is still an element in $\mathcal{V}$.

Finally, we minimize over all maps $\mathcal{M} \in \mathcal{V}$. The conclusion then follows because all purifications are related by an isometry acting on the purifying system and the quantity $\widetilde{\Upsilon}_\alpha(\mathcal{N}, \rho_{A'})$ is invariant under isometries acting on the purifying system.

□

**Remark**: Note that in the proof we only use the monotonicity of the sandwiched Rényi relative entropy. The result can thus be easily generalized to other divergences and distance measures, including the hypothesis testing divergence.

### Operator radius and max-Holevo information

**Definition 5.32.** For a quantum channel $\mathcal{N}_{A'\to B}$, its *operator radius* is defined by

$$\eta(\mathcal{N}) := \{\min \operatorname{Tr} S : \mathcal{N}(\rho) \leq S, \forall \rho \in \mathcal{S}(A')\}. \tag{5.148}$$

The logarithmic operator radius is

$$\log \eta(\mathcal{N}) = \log\{\min \operatorname{Tr} S : \mathcal{N}(\rho) \leq S, \forall \rho \in \mathcal{S}(A')\}. \tag{5.149}$$

**Definition 5.33.** The max-Holevo information is defined by

$$\chi_{\max}(\mathcal{N}) := \min_{\sigma} \max_{\rho} D_{\max}(\mathcal{N}(\rho) \| \sigma). \tag{5.150}$$

**Lemma 5.34.** *The logarithmic operator radius of $\mathcal{N}$ can be refined as the max-Holevo information of $\mathcal{N}$, i.e.,*

$$\log \eta(\mathcal{N}) = \chi_{\max}(\mathcal{N}) \geq \chi(\mathcal{N}). \tag{5.151}$$

*Proof.*

$$\begin{aligned}
\log \eta(\mathcal{N}) &= \min \log\{t : \mathcal{N}(\rho) \leq t\sigma, \sigma \geq 0, \operatorname{Tr}\sigma = 1, \forall \rho \in \mathcal{S}(A')\} && (5.152)\\
&= \min_{\sigma} \min \log\{t : \mathcal{N}(\rho) \leq t\sigma, \forall \rho \in \mathcal{S}(A')\} && (5.153)\\
&= \min_{\sigma} \max_{\rho} \min \log\{t : \mathcal{N}(\rho) \leq t\sigma\} && (5.154)\\
&= \min_{\sigma} \max_{\rho} D_{\max}(\mathcal{N}(\rho) \| \sigma). && (5.155)
\end{aligned}$$

□

One could further use standard SDP techniques to show that the SDP strong converse bound $C_\beta$ is actually an additive upper bound on the max-Holevo information.

**Proposition 5.35.** *For any given channel $\mathcal{N}$, we have*

$$\eta(\mathcal{N}) \leq \beta(\mathcal{N}), \tag{5.156}$$

*where*

$$\beta(\mathcal{N}) = \min\left\{\operatorname{Tr} S_B : -R_{AB} \leq J_{\mathcal{N}}^{T_B} \leq R_{AB}, -\mathbb{1}_A \otimes S_B \leq R_{AB}^{T_B} \leq \mathbb{1}_A \otimes S_B\right\}. \tag{5.157}$$

*Consequently,*

$$C(\mathcal{N}) \leq \lim_{n\to\infty} \frac{1}{n} \chi_{\max}\left(\mathcal{N}^{\otimes n}\right) \leq C_\beta(\mathcal{N}). \tag{5.158}$$

In particular, for the amplitude damping channel with parameter $\gamma$, it holds that

$$C_\beta\left(\mathcal{N}_\gamma^{AD}\right) = \log \beta\left(\mathcal{N}_\gamma^{AD}\right) = \log \eta\left(\mathcal{N}_\gamma^{AD}\right) = \log\left(1 + \sqrt{1-\gamma}\right). \tag{5.159}$$

## 5.5 Discussion

### 5.5.1 Summary

We summarize the important results of this chapter in the following box.

**Summary of Chapter 5**

(i) Classical communication assisted by NS (and PPT) codes

$$\begin{aligned} p_{s,NS\cap PPT}(\mathcal{N},m) = \max \ & \operatorname{Tr} J_{\mathcal{N}} F_{AB} \\ \text{s.t. } & 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1, \\ & \operatorname{Tr}_A F_{AB} = \mathbb{1}_B / m, \\ & 0 \le F_{AB}^{T_B} \le \rho_A \otimes \mathbb{1}_B \ \text{(PPT)}. \end{aligned} \tag{5.160}$$

$$\begin{aligned} C^{(1)}_{NS\cap PPT}(\mathcal{N},\varepsilon) = -\log \min \ & \eta \\ \text{s.t. } & 0 \le F_{AB} \le \rho_A \otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1, \\ & \operatorname{Tr}_A F_{AB} = \eta \mathbb{1}_B, \operatorname{Tr} J_{\mathcal{N}} F_{AB} \ge 1-\varepsilon, \\ & 0 \le F_{AB}^{T_B} \le \rho_A \otimes \mathbb{1}_B \ \text{(PPT)}. \end{aligned} \tag{5.161}$$

(ii) An SDP strong converse bound for the classical capacity:

$$C(\mathcal{N}) \le C_{NS\cap PPT}(\mathcal{N}) \le C_\beta(\mathcal{N}) = \log \beta(\mathcal{N}), \tag{5.162}$$

where $\beta(\mathcal{N}) = \min\{\operatorname{Tr} S_B : \ -R \le J_{\mathcal{N}}^{T_B} \le R, -\mathbb{1}_A \otimes S_B \le R^{T_B} \le \mathbb{1}_A \otimes S_B\}$.

(iii) Achieving Matthews-Wehner meta-converse via activated NS codes:

$$C^{(1)}_{NS,a}(\mathcal{N},\varepsilon) = \max_{\rho_{A'}} \min_{\sigma_B} D_H^\varepsilon\left(\mathcal{N}_{A\to B}(\phi_{A'A}) \| \rho_{A'} \otimes \sigma_B\right), \tag{5.163}$$

where $\phi_{A'A}$ is a purification of $\rho_{A'}$.

(iv) For the amplitude damping channel $\mathcal{N}_\gamma^{AD}$, it holds that $\log\left(1+\sqrt{1-\gamma}\right)$ is a strong converse bound for $C\left(\mathcal{N}_\gamma^{AD}\right)$, i.e.,

$$C\left(\mathcal{N}_\gamma^{AD}\right) \le C_{NS\cap PPT}\left(\mathcal{N}_\gamma^{AD}\right) \le C_\beta\left(\mathcal{N}_\gamma^{AD}\right) = \log\left(1+\sqrt{1-\gamma}\right).$$

(v) Meta-converse via constant-bounded subchannels in Section 5.3.1.

(vi) Given quantum erasure channel $\mathcal{E}_p$ with parameter $p$ and input dimension $d$,

$$C^{(1)}\left(\mathcal{E}_p^{\otimes n},\varepsilon\right) = n(1-p)\log d + \sqrt{np(1-p)(\log d)^2}\,\Phi^{-1}(\varepsilon) + O(\log n).$$

### 5.5.2 Outlook

One future direction is to derive better efficiently computable evaluations of classical communication over general quantum channels. Perhaps one could obtain tighter converse bounds via the study of $C_{NS\cap PPT}$. Another direction is to further tighten the one-shot and strong converse bounds by involving the separable constraint [HNW17].

A challenging open problem is the classical capacity of the amplitude damping channel. As we showed in Figure. 5.5, there is still much space between the best known upper and lower bounds. It is of great interest to further improve the bounds from both sides. Or maybe one can try to find an approach to show the additivity of the Holevo capacity in this case.

For the qubit depolarizing channel, the strong converse property of its classical capacity was established in [KW09]. Then one may expect a second-order analysis of its classical capacity. However, this problem remains open and we note that our meta-converse in Theorem. 5.12 cannot lead to a tight second-order analysis.

Moreover, given the fact that the entanglement-assisted capacity allows a single-letter characterization, it is natural to consider a second-order analysis of it. We note that the second-order achievable rate was established in [DTW16], and the remaining direction is to derive a second-order converse bound. Maybe the one-shot $\varepsilon$-error NS-assisted capacity introduced in this chapter may shed some light.

Finally, we close this chapter with a brief overview of the known and open problems in the beyond i.i.d. regime of classical communication over quantum channels.

| | CQ | EB | Erasure | Depolarizing | AD |
|---|---|---|---|---|---|
| $C$ | [Hol98a] | [Sho02a] | [BDS97] | [Kin03] | ? |
| Strong converse | [ON99, Win99] | [WWY14] | [WW14] | [KW09] | ? |
| Second-order | [TT15] | ? | [WFT17] | ? | ? |
| Second-order ($C_E$) | ?[1] | ? | [DTW16] | [DTW16] | ? |

Table 5.1: Table of classical communication capabilities of basic channels (CQ=classical quantum, EB=entanglement breaking, AD=amplitude damping).

[1] One could expect that the second-order for entanglement-assisted capacity will be the same as the un-assisted case in [TT15].

# Chapter 6

# Quantum communication via quantum channels

## 6.1 Introduction

### 6.1.1 Background

Quantum communication refers to the transmission of quantum information via quantum channels: the sender (Alice) has a quantum system whose state she would like to transmit coherently to the receiver (Bob). This requires that an arbitrary quantum state, when encoded and transmitted using a noisy channel, can be recovered by the receiver. The reliable quantum communication via noisy quantum channels is a fundamental problem in quantum information theory as well as a basic technology for quantum internet in the future.

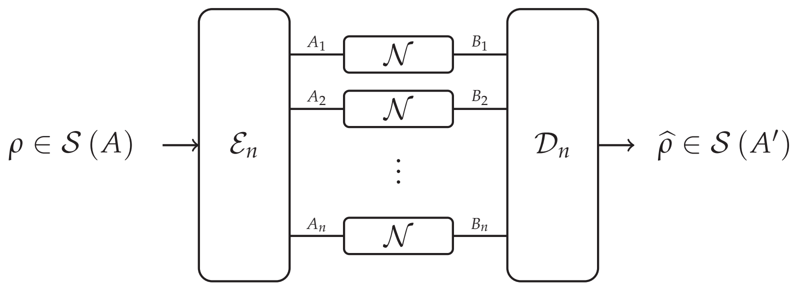


Figure 6.1: The sender (Alice) encodes the states with an encoding operation $\mathcal{E}_n$ and then sends them through the channel $\mathcal{N}^{\otimes n}$ to the receiver (Bob). Bob collects these registers and then applies a decoding operation $\mathcal{D}_n$ which acts collectively on the many outputs of the channels.

**Quantum capacity theorem**

The quantum capacity of a noisy quantum channel is the optimal rate at which it can convey quantum bits (qubits) reliably over asymptotically many uses of the channel. (We refer to Eq. (6.6) for a formal definition.) The work in [Llo97, Sho02b, Dev05] showed that coherent information of $\mathcal{N}$ is an achievable rate for quantum communication while the work in [SN96, BKN00, BNS98] showed the regularized coherent information is also an upper bound on quantum capacity. The above works establish the following quantum capacity theorem, which is one of the most important theorems in quantum Shannon theory.

**Theorem 6.1** (Quantum capacity theorem)**.** *Given a quantum channel $\mathcal{N}$, its quantum capacity is given by the regularized coherent information:*

$$\mathrm{Q}\left(\mathcal{N}\right) = \lim_{n\to\infty} \frac{I_{\mathrm{C}}\left(\mathcal{N}^{\otimes n}\right)}{n} \tag{6.1}$$

*where the coherent information $I_{\mathrm{C}}\left(\mathcal{N}\right)$ is given by*

$$I_{\mathrm{C}}\left(\mathcal{N}\right) = \max_{\rho_A} \; S\left(\mathcal{N}\left(\rho_A\right)\right) - S\left(\mathcal{N}^c\left(\rho_A\right)\right), \tag{6.2}$$

*where $\mathcal{N}^c$ is a complementary channel of $\mathcal{N}$.*

In general, the regularization of coherent information is necessary since the coherent information can be superadditive. The quantum capacity is notoriously difficult to evaluate since it is characterized by a multi-letter, regularized expression and it is not even known to be computable [CEM$^+$15, ES15]. Even for the qubit depolarizing channel, the quantum capacity is still unsolved. (See Section 6.4.2 for discussion.) Our understanding of quantum capacity is quite limited and we even do not know the threshold value of the depolarizing noise for which the quantum capacity vanishes.

**Strong and weak converse bounds**

The converse part of the LSD theorem states that if the rate exceeds the quantum capacity, then the fidelity of any coding scheme cannot approach one in the limit of many channel uses. A strong converse property leaves no room for the trade-off between rate and error, i.e., the error probability vanishes in the limit of many channel uses whenever the rate exceeds the capacity. For quantum communication, the strong converse property was studied in [TWW17] and such property of generalized dephasing channels was established. Given an arbitrary quantum channel, the partial transposition bound was introduced in [HW01] as an efficiently computable upper bound on quantum capacity, and it was proved to be a strong converse bound in [MHRW16]. Recently, the Rains information [TWW17] was established to be a strong converse bound for quantum communication. For the setting of weak converse, there are other known upper bounds for quantum capacity (see e.g., [SSW08, SSWR17, GJL15, BDE+98, WPG07, SS08, LDS17] and most of them require specific settings to be computable and relatively tight.

### 6.1.2 Outline

In this chapter, we investigate the capabilities of quantum channels to convey quantum information and show efficiently computable estimates under both finite blocklength and asymptotic regime. Section 6.2 derives one-shot semidefinite programming (SDP) converse bounds on the amount of quantum information can transmit over a single use of a quantum channel, which improve the previous bound in [TBR16]. Section 6.3 derives an SDP strong converse bound for the quantum capacity of an arbitrary quantum channel, which means the fidelity of any code with a rate exceeding this bound will vanish exponentially fast as the number of channel uses increases. In particular, this SDP strong converse bound is always smaller than or equal to the *partial transposition bound*, and it can be refined as the so-called *max-Rains information*. This SDP strong converse bound is weaker than the Rains information, but it is efficiently computable in general.

## 6.2 One-shot communication capability

### 6.2.1 Task of information processing

In this section, we investigate the finite blocklength regime of quantum communication. Given a noisy channel $\mathcal{N}_{A\to B}$, the aim of quantum communication is to find the optimal encoder and decoder to simulate a noiseless qudit channel. There are different metrics to quantify how well a channel acts as the ideal channel [KW04]. The diamond norm is by no means the only way to evaluate the distance between two channels. But in the case of quantum communication, the channel fidelity [RW05] is a very handy figure of merit since it does not involve an optimization process, and is equivalent to the error criteria based on the diamond norm [KW04].

**Definition 6.2.** For a quantum channel $\mathcal{N}$ from $\mathcal{L}(A')$ to $\mathcal{L}(B)$ with dimension $d_A =$

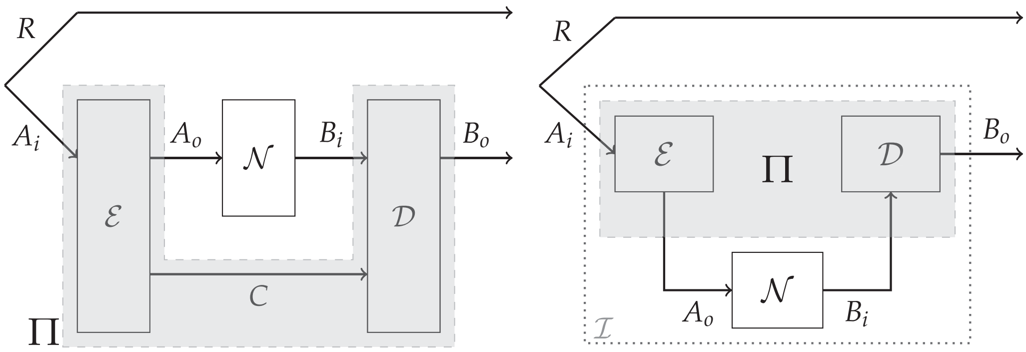


Figure 6.2: General code $\Pi_{A_iB_i\to A_oB_o}$ is equivalently the coding scheme ($\mathcal{E}$,$\mathcal{D}$) with free extra resources $C$, such as entanglement or no-signalling correlations. The goal of the whole operation is to simulate a noiseless quantum channel $\mathcal{I}_{A_i\to B_o}$ using a given noisy quantum channel $\mathcal{N}_{A_o\to B_i}$ and the code Π.

$d_B = m$, the channel fidelity of $\mathcal{N}$ is defined by

$$F_c(\mathcal{N}) := F(\Phi_{BR}, \mathcal{N}_{A'\to B}(\Phi_{A'R})), \tag{6.3}$$

where $|\Phi\rangle = \frac{1}{\sqrt{m}}\sum_{i=0}^{m-1}|ii\rangle$ is the normalized maximally entangled state.

In this following, we use the channel fidelity and focus on optimizing the codes to reliably transmit a state entangled with a reference system from Alice to Bob (also known as entanglement distribution). To be specific, suppose Alice shares a maximally entangled state ($\Phi_{A_iR}$) with a reference system $R$. The goal is to design a quantum coding protocol such that this maximally entangled state can be sent to Bob with high fidelity. To this end, Alice first performs an encoding operation $\mathcal{E}_{A_i\to A_o}$ on system $A_i$ and transmits the prepared state through the channel $\mathcal{N}_{A_o\to B_i}$. The resulting state turns out to be $\mathcal{N}_{A_o\to B_i}\circ\mathcal{E}_{A_i\to A_o}(\Phi_{A_iR})$. Then Bob performs a decoding operation $\mathcal{D}_{B_i\to B_o}$ on system $B_i$, where $B_o$ is some system of the same dimension as $A_i$. The final resulting state will be $\rho_{final} = \mathcal{D}_{B_i\to B_o}\circ\mathcal{N}_{A_o\to B_i}\circ\mathcal{E}_{A_i\to A_o}(\Phi_{A_iR})$. The target of quantum coding is to optimize the fidelity between $\rho_{final}$ and the maximally entangled state $\Phi_{A_iR}$.

One could further imagine the coding protocol as a deterministic super-operator $\Pi_{A_iB_i\to A_oB_o}$, which we refer to as general codes (see Section 2.3 for details). In the following, we will consider quantum communication over quantum channels assisted with Ω codes, where $\Omega \in \{\text{UA}, \text{NS}\cap\text{PPT}, \text{PPT}\}$. We refer to Section 2.3 for more details about the mathematical description of these codes.

**Definition 6.3.** The maximum channel fidelity of $\mathcal{N}$ assisted by the Ω code is defined by

$$F_\Omega(\mathcal{N},k) := \sup_\Pi \operatorname{Tr}\left(\Phi_{B_oR}\cdot \Pi_{A_iB_i\to A_oB_o}\circ \mathcal{N}_{A_o\to B_i}(\Phi_{A_iR})\right), \tag{6.4}$$

where $\Phi_{A_iR}$ and $\Phi_{B_oR}$ are maximally entangled states, $k = \dim|A_i| = \dim|B_o|$ called code size and the supremum is taken over the $\Omega$ codes ($\Omega \in \{\text{UA}, \text{NS}\cap\text{PPT}, \text{PPT}\}$).

**Definition 6.4.** For given quantum channel $\mathcal{N}$ and error tolerance $\varepsilon$, the one-shot $\varepsilon$-error quantum capacity assisted by $\Omega$ codes is defined by

$$Q^{(1)}_\Omega(\mathcal{N},\varepsilon) := \log\sup\{k\in\mathbb{N} : F_\Omega(\mathcal{N},k)\ge 1-\varepsilon\}, \tag{6.5}$$

where $\Omega \in \{\text{UA}, \text{NS}\cap\text{PPT}, \text{PPT}\}$. In the following, we write $Q^{(1)}_{UA}(\mathcal{N},\varepsilon) = Q^{(1)}(\mathcal{N},\varepsilon)$ for simplicity.

The corresponding asymptotic quantum capacity is then defined by

$$Q_\Omega(\mathcal{N}) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}Q^{(1)}_\Omega\left(\mathcal{N}^{\otimes n},\varepsilon\right). \tag{6.6}$$

The authors of [LM15] showed that the maximum channel fidelity assisted with $\text{NS}\cap\text{PPT}$ codes is given by the following SDP:

$$\begin{aligned} F_{NS\cap PPT}(\mathcal{N},k) = \max\ & \operatorname{Tr} J_\mathcal{N} W_{AB} \\ \text{s.t. } & 0\le W_{AB}\le \rho_A\otimes \mathbb{1}_B, \operatorname{Tr}\rho_A = 1, \\ & -k^{-1}\rho_A\otimes\mathbb{1}_B \le W_{AB}^{T_B}\le k^{-1}\rho_A\otimes \mathbb{1}_B, \\ & \operatorname{Tr}_A W_{AB} = k^{-2}\mathbb{1}_B\ \ (\text{NS}). \end{aligned} \tag{6.7}$$

To obtain $F_{PPT}(\mathcal{N},k)$, one only need to remove the NS constraint.

Combining Eqs. (6.5) and (6.7), one can derive the following proposition. It is worth noting that Eq. (6.8) is not an SDP in general, due to the non-linear term $m\rho_A$ and the condition $\operatorname{Tr}_A W_{AB} = m^2\mathbb{1}_B$. But in next subsection, we will derive several semidefinite relaxations of this optimization problem.

**Proposition 6.5.** *For any quantum channel $\mathcal{N}_{A'\to B}$ with Choi-Jamiołkowski matrix $J_\mathcal{N}\in\mathcal{L}(A\otimes B)$ and given error tolerance $\varepsilon$, its one-shot $\varepsilon$-error quantum capacity assisted with PPT codes can be simplified as an optimization problem:*

$$\begin{aligned} Q^{(1)}_{PPT}(\mathcal{N},\varepsilon) = -\log\min\ & m \\ \textit{s.t. } & \operatorname{Tr} J_\mathcal{N} W_{AB}\ge 1-\varepsilon, 0\le W_{AB}\le\rho_A\otimes\mathbb{1}_B, \\ & \operatorname{Tr}\rho_A = 1, -m\rho_A\otimes\mathbb{1}_B\le W_{AB}^{T_B}\le m\rho_A\otimes\mathbb{1}_B. \end{aligned} \tag{6.8}$$

*If the codes are also non-signalling, we can have the same optimization for NS∩PPT codes* *with an additional constraint* $\operatorname{Tr}_A W_{AB} = m^2 \mathbb{1}_B$.

### 6.2.2 SDP converse bounds for quantum communication

To better evaluate the quantum communication rate with finite resources, we introduce several SDP converse bounds for quantum communication with the assistance of PPT (and NS) codes. In Theorem 6.6, we further prove that our SDP bounds are tighter than the one introduced in [TBR16].

Specifically, the authors of [TBR16] established that $-\log f(\mathcal{N},\varepsilon)$ is a converse bound on one-shot $\varepsilon$-error quantum capacity, i.e., $Q^{(1)}(\mathcal{N},\varepsilon) \leq -\log f(\mathcal{N},\varepsilon)$ where

$$\begin{aligned} f(\mathcal{N},\varepsilon) = \min & \operatorname{Tr} S_A \\ \text{s.t.} & \operatorname{Tr} W_{AB} J_{\mathcal{N}} \geq 1-\varepsilon, S_A, \Theta_{AB} \geq 0, \operatorname{Tr}\rho_A = 1, \\ & 0 \leq W_{AB} \leq \rho_A \otimes \mathbb{1}_B, S_A \otimes \mathbb{1}_B \geq W_{AB} + \Theta_{AB}^{T_B}. \end{aligned} \tag{6.9}$$

Here, we introduce a hierarchy of SDP converse bounds on the one-shot $\varepsilon$-error capacity based on the optimization (6.8). If we relax the term $m\rho_A$ to a single variable $S_A$, we obtain $g(\mathcal{N},\varepsilon)$, where

$$\begin{aligned} g(\mathcal{N},\varepsilon) := \min & \operatorname{Tr} S_A \\ \text{s.t.} & \operatorname{Tr} J_{\mathcal{N}} W_{AB} \geq 1-\varepsilon, 0 \leq W_{AB} \leq \rho_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr}\rho_A = 1, -S_A \otimes \mathbb{1}_B \leq W_{AB}^{T_B} \leq S_A \otimes \mathbb{1}_B. \end{aligned} \tag{6.10}$$

Thus, we obtain

$$Q^{(1)}_{\text{PPT}}(\mathcal{N},\varepsilon) \leq -\log g(\mathcal{N},\varepsilon). \tag{6.11}$$

In particular, for the NS condition $\operatorname{Tr}_A W_{AB} = m^2 \mathbb{1}_B$, there are two different ways to get relaxations. The first one is to substitute it with $\operatorname{Tr}_A W_{AB} = t\mathbb{1}_B$ and obtain SDP $\widetilde{g}(\mathcal{N},\varepsilon)$). The second one is to introduce a prior constant $\widehat{m}$ satisfying the inequality

$$Q^{(1)}_{NS\cap PPT}(\mathcal{N},\varepsilon) \leq -\log \widehat{m} \tag{6.12}$$

and then obtain SDP $\widehat{g}(\mathcal{N},\varepsilon)$. Note that the second method can provide a tighter bound, but it requires one more step of calculation since we need to get the prior constant $\widehat{m}$. Successively refining the value of $\widehat{m}$ will result in a tighter bound.

$$\begin{aligned}\widetilde{g}(\mathcal{N},\varepsilon) := \min \operatorname{Tr} S_A & \\ \text{s.t. } & \operatorname{Tr} J_{\mathcal{N}} W_{AB} \geq 1-\varepsilon, 0 \leq W_{AB} \leq \rho_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr} \rho_A = 1, -S_A \otimes \mathbb{1}_B \leq W_{AB}^{T_B} \leq S_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr}_A W_{AB} = t\mathbb{1}_B.\end{aligned} \tag{6.13}$$

$$\begin{aligned}\widehat{g}(\mathcal{N},\varepsilon) := \min \operatorname{Tr} S_A & \\ \text{s.t. } & \operatorname{Tr} J_{\mathcal{N}} W_{AB} \geq 1-\varepsilon, 0 \leq W_{AB} \leq \rho_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr} \rho_A = 1, -S_A \otimes \mathbb{1}_B \leq W_{AB}^{T_B} \leq S_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr}_A W_{AB} = t\mathbb{1}_B, t \geq \widehat{m}^2.\end{aligned} \tag{6.14}$$

**Theorem 6.6.** *For any quantum channel $\mathcal{N}$ and error tolerance $\varepsilon$, the inequality chain holds*

$$\begin{aligned} Q^{(1)}(\mathcal{N},\varepsilon) &\leq Q^{(1)}_{NS\cap PPT}(\mathcal{N},\varepsilon) \\ &\leq -\log \widehat{g}(\mathcal{N},\varepsilon) \leq -\log \widetilde{g}(\mathcal{N},\varepsilon) \leq -\log g(\mathcal{N},\varepsilon) \leq -\log f(\mathcal{N},\varepsilon).\end{aligned}$$

*Proof.* The third and fourth inequalities are easy to obtain since the minimization over a smaller feasible set gives a larger optimal value here.

For the second inequality, suppose the optimal solution of (6.8) for $Q^{(1)}_{NS\cap PPT}(\mathcal{N},\varepsilon)$, is taken at $\{W_{AB}, \rho_A, m\}$. Let $S_A = m\rho_A$, $t = m^2$. Then we can verify that $\{W_{AB}, \rho_A, S_A, t\}$ is a feasible solution to the SDP (6.14) of $\widehat{g}(\mathcal{N},\varepsilon)$. So $\widehat{g}(\mathcal{N},\varepsilon) \leq \operatorname{Tr} S_A = m$, which implies that $Q^{(1)}_{NS\cap PPT}(\mathcal{N},\varepsilon) = -\log m \leq -\log \widehat{g}(\mathcal{N},\varepsilon)$.

For the last inequality, we only need to show that $f(\mathcal{N},\varepsilon) \leq g(\mathcal{N},\varepsilon)$. Suppose the optimal solution of $g(\mathcal{N},\varepsilon)$ is taken at $\{\rho_A, S_A, W_{AB}\}$. Let us choose $\Theta_{AB} = S_A \otimes \mathbb{1}_B - W_{AB}^{T_B}$. Since $S_A \otimes \mathbb{1}_B \geq W_{AB}^{T_B}$, it is clear that $\Theta_{AB} \geq 0$ and $S_A \otimes \mathbb{1}_B = W_{AB} + \Theta_{AB}^{T_B}$. Thus, $\{S_A, \rho_A, W_{AB}, \Theta_{AB}\}$ is a feasible solution to the SDP (6.9) of $f(\mathcal{N},\varepsilon)$ which implies $f(\mathcal{N},\varepsilon) \leq \operatorname{Tr} S_A = g(\mathcal{N},\varepsilon)$. □

### 6.2.3 Example: amplitude damping channel

For the amplitude damping channel $\mathcal{N}_\gamma = \sum_{i=0}^{1} E_i \cdot E_i^\dagger$ with $E_0 = |0\rangle\langle 0| + \sqrt{1-r}|1\rangle\langle 1|$, $E_1 = \sqrt{r}|0\rangle\langle 1|$ $(0 \leq r \leq 1)$, the differences among $-\log f\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$, $-\log g\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$ and $-\log \widetilde{g}\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$, are presented in Figure 6.3.

When $r \in (0.081, 0.094)$, it holds that

$$-\log \widetilde{g}\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right) \leq -\log g\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right) < 1 < -\log f\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right). \tag{6.15}$$

This shows that we cannot transmit a single qubit within error tolerance $\varepsilon = 0.01$ via two uses of amplitude damping channel where parameter $r \in (0.081, 0.094)$. However, this result cannot be obtained via the converse bound $-\log f\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$.

If we consider three uses of the amplitude damping channel, Figure 6.4 shows that we cannot transmit one qubit over $\mathcal{N}_\gamma$ with infidelity 0.01 when the noise parameter is larger than 0.2625.

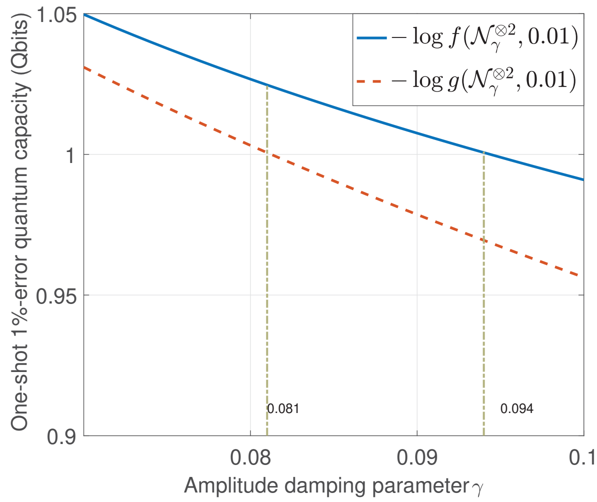


Figure 6.3: This figure demonstrates the differences between the SDP converse bounds (i) $-\log f\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$ (blue solid), (ii) $-\log g\left(\mathcal{N}_\gamma^{\otimes 2}, 0.01\right)$ (red dashed), where $\gamma$ ranges from 0.07 to 0.1.

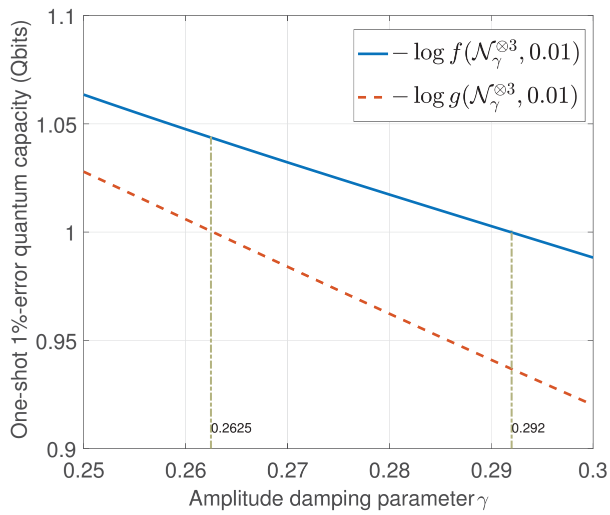


Figure 6.4: This figure demonstrates the differences among the SDP converse bounds (i) $-\log f\left(\mathcal{N}_\gamma^{\otimes 3}, 0.01\right)$ (blue solid), (ii) $-\log g\left(\mathcal{N}_\gamma^{\otimes 3}, 0.01\right)$ (red dashed), where $\gamma$ ranges from 0.25 to 0.3.

## 6.3 Asymptotic communication capability

We now investigate quantum communication under the asymptotic scenario. We first present an SDP strong converse bound, denoted as $Q_\Gamma$, on the quantum capacity for general channels. The proof of this strong converse bound is built on two ingredients: a relationship between the rate and $Q_\Gamma$ as well as the additivity of $Q_\Gamma$. In particular, we also find that $Q_\Gamma$ is a channel analog of SDP entanglement measure $E_W$ in Chapter 3.

### 6.3.1 Quantum capacity

In this section, we introduce an SDP strong converse bound $Q_\Gamma(\mathcal{N}) := \log \Gamma(\mathcal{N})$ to evaluate the quantum capacity for a general quantum channel, where

$$\begin{aligned} \textbf{(Primal)} \quad \Gamma(\mathcal{N}) = \max\ & \operatorname{Tr} J_{\mathcal{N}} R_{AB} \\ \text{s.t. } & R_{AB}, \rho_A \geq 0, \operatorname{Tr} \rho_A = 1, \\ & -\rho_A \otimes \mathbb{1}_B \leq R_{AB}^{T_B} \leq \rho_A \otimes \mathbb{1}_B, \end{aligned} \tag{6.16}$$

$$\begin{aligned} \textbf{(Dual)} \quad \Gamma(\mathcal{N}) = \min\ & \mu \\ \text{s.t. } & Y_{AB}, V_{AB} \geq 0, (V_{AB} - Y_{AB})^{T_B} \geq J_{\mathcal{N}}, \\ & \operatorname{Tr}_B (V_{AB} + Y_{AB}) \leq \mu \mathbb{1}_A. \end{aligned} \tag{6.17}$$

We summarize our strong converse bound with other well-known bounds in Table 6.1. Among those efficiently computable strong converse bound for general channels, we prove that $Q_\Gamma(\mathcal{N})$ is better than the partial transpose bound and remark that it is also strictly tighter than the entanglement-assisted quantum capacity in the case of entanglement-breaking channels with non-zero classical capacity. The relation with Rains information is also obtained.

### 6.3.2 An SDP strong converse bound on quantum capacity

We first establish a relationship between the one-shot PPT-assisted quantum capacity and the bound $Q_\Gamma(\mathcal{N})$ in the following proposition.

**Lemma 6.7.** *For any quantum channel $\mathcal{N}$ and error tolerance $\varepsilon$,*

$$Q_{PPT}^{(1)}(\mathcal{N}, \varepsilon) \leq Q_\Gamma(\mathcal{N}) - \log(1 - \varepsilon). \tag{6.18}$$

*Proof.* Suppose the optimal solution in the optimization (6.8) of $Q_{PPT}^{(1)}(\mathcal{N}, \varepsilon)$ is taken at $\{W_{AB}, \rho_A, m\}$, then $Q_{PPT}^{(1)}(\mathcal{N}, \varepsilon) = -\log m$. Denote $R_{AB} = \frac{1}{m} W_{AB}$ and we can verify that $\{R_{AB}, \rho_A\}$ is a feasible solution to the SDP (6.16). Thus

$$\begin{aligned} Q_\Gamma(\mathcal{N}) &\geq \log \operatorname{Tr} J_\mathcal{N} R_{AB} && (6.19)\\ &= \log \frac{1}{m} \operatorname{Tr} J_\mathcal{N} W_{AB} && (6.20)\\ &\geq \log \frac{1}{m}(1-\varepsilon) && (6.21)\\ &= Q^{(1)}_{PPT}(\mathcal{N}, \varepsilon) + \log(1-\varepsilon). && (6.22)\end{aligned}$$

This concludes the proof. The dual problem can be derived via Lagrange multiplier method. □

Then we prove that the bound $Q_\Gamma$ is additive under tensor products.

**Lemma 6.8.** *For any quantum channel $\mathcal{N}_1$ and $\mathcal{N}_2$, $Q_\Gamma$ is additive, i.e.,*

$$Q_\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2) = Q_\Gamma(\mathcal{N}_1) + Q_\Gamma(\mathcal{N}_2). \quad (6.23)$$

*Proof.* We only need to show that $\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2) = \Gamma(\mathcal{N}_1)\Gamma(\mathcal{N}_2)$. For the primal problem (6.16), suppose the optimal solutions of (6.16) for the channel $\mathcal{N}_1$ and $\mathcal{N}_2$ are taken at $\{R_1, \rho_1\}$ and $\{R_2, \rho_2\}$, respectively. Then we can verify that $\{R_1 \otimes R_2, \rho_1 \otimes \rho_2\}$ is a feasible solution of $\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2)$. Thus,

$$\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2) \geq \operatorname{Tr}(J_{\mathcal{N}_1} \otimes J_{\mathcal{N}_2})(R_1 \otimes R_2) = \Gamma(\mathcal{N}_1)\Gamma(\mathcal{N}_2). \quad (6.24)$$

For the dual problem (6.17), suppose the optimal solutions of (6.17) for the channel $\mathcal{N}_1$ and $\mathcal{N}_2$ are taken at $\{V_1, Y_1, \mu_1\}$ and $\{V_2, Y_2, \mu_2\}$. Let us take

$$\begin{aligned} V &= V_1 \otimes V_2 + Y_1 \otimes Y_2, && (6.25)\\ Y &= V_1 \otimes Y_2 + Y_1 \otimes V_2. && (6.26)\end{aligned}$$

It can be verified that $\{V, Y, \mu_1\mu_2\}$ is a feasible solution of $\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2)$.

Thus,

$$\Gamma(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq \Gamma(\mathcal{N}_1)\Gamma(\mathcal{N}_2). \quad (6.27)$$

□

Finally, utilizing the two lemmas above, we are now able to prove that $Q_\Gamma$ is a strong converse bound for the quantum capacity assisted with PPT codes.

**Theorem 6.9.** *For any quantum channel $\mathcal{N}$,*

$$Q(\mathcal{N}) \leq Q_{PPT}(\mathcal{N}) \leq Q_{\Gamma}(\mathcal{N}). \tag{6.28}$$

*Moreover, $Q_{\Gamma}(\mathcal{N})$ is a strong converse bound. That is, if the rate exceeds $Q_{\Gamma}(\mathcal{N})$, the error probability will approach to one exponentially fast as the number of channel uses increase.*

*Proof.* We first show that $Q_{\Gamma}(\mathcal{N})$ is a converse bound and then prove that it is a strong converse. From Eq. (6.18), take regularization on both sides, we have

$$\begin{aligned} Q_{PPT}(\mathcal{N}) &= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}Q^{(1)}_{PPT}\left(\mathcal{N}^{\otimes n},\varepsilon\right) \\ &\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\left[Q_{\Gamma}\left(\mathcal{N}^{\otimes n}\right)-\log\left(1-\varepsilon\right)\right] \\ &= Q_{\Gamma}(\mathcal{N}). \end{aligned} \tag{6.29}$$

In the last line, we use the additivity of $Q_{\Gamma}$ in Proposition 6.8.

For the $n$-fold quantum channel $\mathcal{N}^{\otimes n}$, suppose its achievable rate is $r$. From Eq. (6.18), we have $nr \leq nQ_{\Gamma}(\mathcal{N}) - \log(1-\varepsilon)$, which implies

$$\varepsilon \geq 1 - 2^{n(Q_{\Gamma}(\mathcal{N})-r)}. \tag{6.30}$$

If $r > Q_{\Gamma}(\mathcal{N})$, the error will exponentially converge to one as $n$ goes to infinity. □

**Remark 6.10.** For $d$-dimensional noiseless quantum channel $\mathcal{I}_d$, we can show

$$Q(\mathcal{I}_d) = Q_{\Gamma}(\mathcal{I}_d) = \log d. \tag{6.31}$$

### 6.3.3 Comparison with other converse bounds

There are several well-known converse bounds on quantum capacity. In this subsection, we compare them with our SDP strong converse bound $Q_{\Gamma}$. Especially, we obtain an inequality chain among the strong converse bound $Q_{\Gamma}$, channel's Rains information $R$ and partial transposition bound $Q_{\Theta}$.

| | Strong converse rate | Efficiently computable |
|---|---|---|
| $Q_\Gamma$ ($R_{\max}$) | ✓ | ✓ |
| $R$ | ✓ | ✗ (max-min) |
| $E_C$ | ✓ | ✗ (regularization) |
| $Q_\Theta$ | ✓ | ✓ |
| $E_{\max}$ | ✓ | ? |
| $Q_E$ | ✓ | ✓ |
| $\varepsilon$-DEG | **?** | ✓ |
| $Q_{ss}$ | **?** | **?** (Unbounded dimension) |

Table 6.1: Comparison of converse bounds on quantum capacity

*Tomamichel et al.* [TWW17] established that the Rains information of any quantum channel is a strong converse rate for quantum communication. To be specific, the Rains information of a quantum channel is defined as [TWW17]:

$$R(\mathcal{N}) := \max_{\rho_A \in \mathcal{S}(A)} \min_{\sigma_{AB} \in \mathrm{PPT}'} D\left(\mathcal{N}_{A' \to B}(\phi_{AA'}) \middle\| \sigma_{AB}\right), \tag{6.32}$$

where $\phi_{AA'}$ is a purification of $\rho_A$ and the set $\mathrm{PPT}' = \{\sigma \in \mathcal{P}(A \otimes B) : \left\|\sigma^{T_B}\right\|_1 \leq 1\}$. We note that our bound $Q_\Gamma$ is weaker than the Rains information (cf. Corollary 6.13.) However, $R(\mathcal{N})$ is not known to be efficiently computable for general quantum channels since it is max-min optimization problem.

An efficiently computable converse bound (abbreviated as $\varepsilon$-DEG) is given by the concept of approximate degradable channel [SSWR17]. This bound usually works very well for approximate degradable quantum channels such as low-noise qubit depolarizing channel. See [LLS18, SWAT18] for some recent works based on this approach. Otherwise, it will degenerate to a trivial upper bound. We can easily show an example that $Q_\Gamma$ can be smaller than $\varepsilon$-DEG bound, e.g., the channel $\mathcal{N}_r$ in Eq. (6.56) with $0 < r < 0.38$. Also, it is unknown whether $\varepsilon$-DEG bound is a strong converse.

Another previously known efficiently computable strong converse bound for general channels is given by the partial transposition bound,

$$Q_\Theta(\mathcal{N}) := \log \|\mathcal{N} \circ T\|_\diamondsuit, \tag{6.33}$$

where $T$ is the transpose map and $\|\cdot\|_\diamondsuit$ is the completely bounded trace norm, which is known to be efficiently computable by SDP in [Wat13].

The entanglement cost of a quantum channel [BBCW13], denoted as $E_C$, is proved to be a strong converse bound. But it is not known to be efficiently computable for general channels, due to its regularization. The entanglement-assisted quantum $Q_E$

is also a strong converse for the quantum capacity [BDH+14, BCR11] and there is a recently developed to efficiently compute it [FF18]. Quantum capacity with symmetric side channels [SSW08], denoted as $Q_{ss}$, is also an important converse bound for general channels. But it is not known to be computable due to the potentially unbounded dimension of the side channel. It is also not known to be a strong converse.

Recently, *Christandl and Müller-Hermes* [CMH17] derived the following strong converse upper bound on the quantum and private communication:

$$E_{\max}(\mathcal{N}) = \max_{\rho_A \in \mathcal{S}(A)} \min_{\sigma_{AB} \in \mathrm{SEP}} D_{\max}\left(\mathcal{N}_{A' \to B}(\phi_{AA'}) \| \sigma_{AB}\right), \tag{6.34}$$

where SEP represents the set of separable states. This bound is known as the max-relative entropy of entanglement of a quantum channel. For quantum communication, $E_{\max}$ improves the partial transposition bound for some channels but is weaker than our bound $Q_\Gamma$ (cf. Proposition 6.12).

**Theorem 6.11.** *For any quantum channel $\mathcal{N}$, it holds that*

$$Q(\mathcal{N}) \leq R(\mathcal{N}) \leq Q_\Gamma(\mathcal{N}) \leq Q_\Theta(\mathcal{N}). \tag{6.35}$$

The first inequality has been proved in [TWW17]. We prove the second inequality in Corollary 6.13 and the third inequality in Proposition 6.14.

In the following proof, we need to introduce an entanglement measure $E_W$ which is defined in Eq. (3.17) in Chapter 3. We will see that the strong converse bound $Q_\Gamma$ is a channel analogue of entanglement measure $E_W$ and can be further reformulated into a similar form as the Rains information.

**Proposition 6.12.** *For any quantum channel $\mathcal{N}$, it holds*

$$Q_\Gamma(\mathcal{N}) = \max_{\rho_A \in \mathcal{S}(A)} E_W\left(\mathcal{N}_{A' \to B}(\phi_{AA'})\right) \tag{6.36}$$

$$= \max_{\rho \in S(A)} \min_{\sigma \in PPT'(A:B)} D_{\max}\left(\mathcal{N}_{A' \to B}(\phi_{AA'}) \| \sigma_{AB}\right), \tag{6.37}$$

*where $\phi_{AA'}$ is a purification of $\rho_A$ and the set $PPT'(A:B) = \left\{\sigma \in \mathcal{P}(A \otimes B) : \left\|\sigma^{T_B}\right\|_1 \leq 1\right\}$.*

*As a consequence, $Q_\Gamma(\mathcal{N}) \leq E_{\max}(\mathcal{N})$.*

*Proof.* Consider purification $\phi_{AA'} = \rho_A^{1/2}\Phi_{AA'}\rho_A^{1/2}\left(= \rho_{A'}^{1/2}\Phi_{AA'}\rho_{A'}^{1/2}\right)$, then

$$\mathcal{N}_{A'\to B}(\phi_{AA'}) = \mathcal{N}_{A'\to B}\left(\rho_A^{1/2}\Phi_{AA'}\rho_A^{1/2}\right) \tag{6.38}$$

$$= \rho_A^{1/2}\mathcal{N}_{A'\to B}(\Phi_{AA'})\rho_A^{1/2} \tag{6.39}$$

$$= \rho_A^{1/2}J_{\mathcal{N}}\rho_A^{1/2}. \tag{6.40}$$

Take $J_{\mathcal{N}} = \rho_A^{-1/2}\mathcal{N}_{A'\to B}(\phi_{AA'})\rho_A^{-1/2}$ into the definition of $Q_\Gamma(\mathcal{N})$ (6.16) and substitute

$$F_{AB} = \rho_A^{-1/2}R_{AB}\rho_A^{-1/2}, \tag{6.41}$$

then we have

$$\begin{aligned} Q_\Gamma(\mathcal{N}) = \log \max \operatorname{Tr} \mathcal{N}_{A'\to B}(\phi_{AA'}) F_{AB} \\ \text{s.t. } F_{AB}, \rho_A \geq 0, \operatorname{Tr}\rho_A = 1, \\ -\mathbb{1}_{AB} \leq F_{AB}^{T_B} \leq \mathbb{1}_{AB} \end{aligned} \tag{6.42}$$

Due to the definition of $E_W$ (3.17), we have

$$Q_\Gamma(\mathcal{N}) = \max_{\rho_A \in \mathcal{S}(A)} E_W(\mathcal{N}_{A'\to B}(\phi_{AA'})). \tag{6.43}$$

On the other hand, by Theorem 3.10, we have that

$$E_W(\rho) = \min_{\sigma \in PPT'(A:B)} D_{\max}(\rho\|\sigma). \tag{6.44}$$

Therefore,

$$Q_\Gamma(\mathcal{N}) = \max_{\rho_A \in \mathcal{S}(A)} E_W(\mathcal{N}_{A'\to B}(\phi_{AA'})) \tag{6.45}$$

$$= \max_{\rho \in S(A)} \min_{\sigma \in PPT'(A:B)} D_{\max}\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\|\sigma_{AB}\right). \tag{6.46}$$

Furthermore, noticing that $\mathrm{SEP}(\mathrm{A}:\mathrm{B}) \subset \mathrm{PPT}'(\mathrm{A}:\mathrm{B})$, we have that

$$Q_\Gamma(\mathcal{N}) \leq \max_{\rho \in S(A)} \min_{\sigma \in SEP(A:B)} D_{\max}\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\|\sigma_{AB}\right) \tag{6.47}$$

$$= E_{\max}(\mathcal{N}). \tag{6.48}$$

□

**Corollary 6.13.** *For any quantum channel $\mathcal{N}$, it holds*

$$R(\mathcal{N}) \leq Q_\Gamma(\mathcal{N}). \tag{6.49}$$

*Proof.* Note that $D(\rho\|\sigma) \leq D_{\max}(\rho\|\sigma)$ [Dat09], we have

$$\begin{aligned} Q_\Gamma(\mathcal{N}) &= \max_{\rho\in S(A)} \min_{\sigma\in \text{PPT}'(\text{A:B})} D_{\max}\left(\mathcal{N}_{A'\to B}(\phi_{A'A})\,\|\sigma_{AB}\right) \\ &\geq \max_{\rho_A\in \mathcal{S}(A)} \min_{\sigma\in \text{PPT}'(A:B)} D\left(\mathcal{N}_{A'\to B}(\phi_{AA'})\,\|\sigma_{AB}\right) \\ &= R(\mathcal{N}). \end{aligned} \tag{6.50}$$

□

**Proposition 6.14.** *For any quantum channel $\mathcal{N}$, it holds*

$$Q_\Gamma(\mathcal{N}) \leq Q_\Theta(\mathcal{N}). \tag{6.51}$$

*Proof.* Assume that the optimal solution of $\Gamma(\mathcal{N})$ is $\{R_{AB}, \rho_A\}$, then

$$\Gamma(\mathcal{N}) = \operatorname{Tr} J_\mathcal{N} R_{AB} = \operatorname{Tr} J_\mathcal{N}^{T_B} R_{AB}^{T_B}. \tag{6.52}$$

Recall the SDP of the diamond norm in Eq. (6.53),

$$\begin{aligned} \|T\circ\mathcal{N}\|_\diamond = \max\ & \frac{1}{2}\operatorname{Tr}\left(J_\mathcal{N}^{T_B} X_{AB}\right) + \frac{1}{2}\operatorname{Tr}\left(J_\mathcal{N}^{T_B} X_{AB}^\dagger\right) \\ \text{s.t.}\ & \begin{pmatrix} \rho_0\otimes\mathbb{1} & X_{AB} \\ X_{AB}^\dagger & \rho_1\otimes\mathbb{1}\end{pmatrix} \geq 0. \end{aligned} \tag{6.53}$$

Let us add two constraints $\rho_0 = \rho_1 = \rho_A$ and $X_{AB} = X_{AB}^\dagger$, then

$$\|J_\mathcal{N}^{T_B}\|_\diamond \geq \max \operatorname{Tr}\left(J_\mathcal{N}^{T_B} X\right) \text{ s.t. } \begin{pmatrix} \rho_A\otimes\mathbb{1}_B & X_{AB} \\ X_{AB} & \rho_A\otimes\mathbb{1}_B\end{pmatrix} \geq 0.$$

Noting that $-\rho_A\otimes\mathbb{1}_B \leq R_{AB}^{T_B} \leq \rho_A\otimes\mathbb{1}_B$, then

$$\begin{aligned} &\begin{pmatrix} \rho_A\otimes\mathbb{1}_B & R_{AB}^{T_B} \\ R_{AB}^{T_B} & \rho_A\otimes\mathbb{1}_B\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} \bigotimes \left(\rho_A\otimes\mathbb{1}_B + R_{AB}^{T_B}\right) + \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} \bigotimes \left(\rho_A\otimes\mathbb{1}_B - R_{AB}^{T_B}\right) \geq 0. \end{aligned} \tag{6.54}$$

Therefore, $R_{AB}^{T_B}$ satisfies the constraint above, which means that

$$\|J_{\mathcal{N}}^{T_B}\|_{\diamond} \geq \operatorname{Tr}\left(J_{\mathcal{N}}^{T_B} R_{AB}^{T_B}\right) = \Gamma\left(\mathcal{N}\right). \tag{6.55}$$

□

In Figure 6.5, we compare the converse bound $Q_\Gamma$ with $Q_\Theta$ in the case of quantum channel

$$\mathcal{M}_r = \sum_{i=0}^{1} E_i \cdot E_i^\dagger,\ (0 \leq r \leq 0.5)\,, \tag{6.56}$$

where

$$E_0 = |0\rangle\langle 0| + \sqrt{r}|1\rangle\langle 1|, \tag{6.57}$$

$$E_1 = \sqrt{1-r}|0\rangle\langle 1| + |1\rangle\langle 2|. \tag{6.58}$$

In the following Figure 6.5, it is clear that $Q_\Gamma$ can be strictly tighter than $Q_\Theta$.

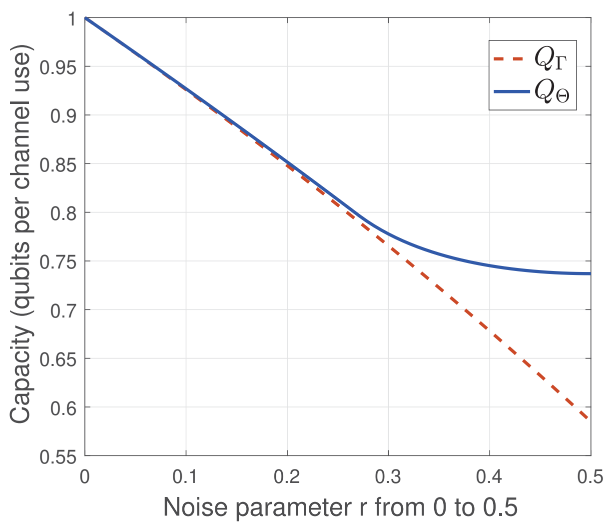


Figure 6.5: This plot demonstrates the difference between converse bounds $Q_\Gamma\left(\mathcal{M}_r\right)$ and $Q_\Theta\left(\mathcal{M}_r\right)$. The dashed line depicts $Q_\Gamma\left(\mathcal{M}_r\right)$ while the solid line depicts $Q_\Theta\left(\mathcal{M}_r\right)$. The noise parameter $r$ ranges from 0 to 0.5.

## 6.4 Discussion

### 6.4.1 Summary

In this chapter, we contributed the semidefinite programs for estimating the quantum communication capability of quantum channels in both the non-asymptotic and asymptotic regimes. We summarize the important results of this chapter in the following box.

**Summary of Chapter 6**

(i) Semidefinite programming converse bounds for quantum communication with finite resources:

$$Q^{(1)}_{\mathrm{PPT}}(\mathcal{N},\varepsilon) \le g(\mathcal{N},\varepsilon) = \min \operatorname{Tr} S_A$$
$$\text{s.t. } \operatorname{Tr} J_{\mathcal{N}} W_{AB} \ge 1-\varepsilon, 0 \le W_{AB} \le \rho_A \otimes \mathbb{1}_B,$$
$$\operatorname{Tr}\rho_A = 1, -S_A \otimes \mathbb{1}_B \le W_{AB}^{T_B} \le S_A \otimes \mathbb{1}_B.$$

There are similar bounds for the NS∩PPT codes in Eq. (6.13) and Eq. (6.14).

(ii) Max-Rains information - an SDP strong converse bound for quantum communication:

$$\begin{aligned} Q_\Gamma(\mathcal{N}) &= \max_{\rho\in S(A)} \min_{\sigma\in \mathrm{PPT}'} D_{\max}\left(\mathcal{N}_{A'\to B}(\phi_{AA'}) \| \sigma_{AB}\right), \\ &= \max \operatorname{Tr} J_{\mathcal{N}} R_{AB} \\ &\quad \text{s.t. } R_{AB}, \rho_A \ge 0, \operatorname{Tr}\rho_A = 1, \\ &\quad -\rho_A \otimes \mathbb{1}_B \le R_{AB}^{T_B} \le \rho_A \otimes \mathbb{1}_B. \end{aligned}$$

Note that $Q_\Gamma$ was recently proved to be a strong converse bound for the LOCC-assisted quantum capacity of an arbitrary quantum channel in [BW18].

(iii) Relationship between several well-known bounds:

$$Q(\mathcal{N}) \le R(\mathcal{N}) \le Q_\Gamma(\mathcal{N}) \le Q_\Theta(\mathcal{N}). \tag{6.59}$$

See Table. 6.1 for a partial overview of the upper bounds on quantum capacity.

### 6.4.2 Outlook

The most fundamental noise is the isotropic noise in a depolarizing channel. But the quantum capacity of this channel is still unsolved despite substantial effort in the past two decades (see e.g., [DSS98, FW08, SS07, SSW08, SSWR17, LDS17, LW17]), and we even do not know at which critical noise the capacity becomes zero. This represents a major gap for us to fully understand the fundamental limits and power of quantum error correction.

Moreover, from the view of strong converse, there is only a pretty strong converse for the degradable channels [MW14b]. The bottleneck is that we do not know whether the strong converse property holds for the 50% quantum erasure channel. Due to a limited understanding on the strong converse property of quantum communication, our understanding of the second-order asymptotics of quantum capacity is also very limited, and the dephasing channel is the only one whose second-order analysis of quantum communication has been fully established [TBR16].

Finally, we end this chapter with a table of the known and open problems in the beyond i.i.d. regime of quantum communication.

| | Degradable | Dephasing | Erasure | Depolarizing |
|---|---|---|---|---|
| $Q(\mathcal{N})$ | [DS05] | [DS05] | [BDS97] | ? |
| Strong converse | ? (Pretty strong [MW14b]) | [TWW17] | ? | ? |
| Second-order | ? | [TBR16] | ? | ? |

Table 6.2: Table of quantum communication capabilities of basic channels

# Part III

# Quantum Zero-error Information

# Chapter 7

# Advancing quantum zero-error information theory

## 7.1 Introduction

### 7.1.1 Background

While the conventional information theory focuses on sending messages with asymptotically vanishing errors [Sha48], Shannon also investigated this problem in the zero-error setting and described the zero-error capacity of a channel as the maximum rate at which it can be used to transmit information with a zero probability of error [Sha56]. The so-called zero-error information theory [Sha56, KO98] concerns the combinatorial problems in the asymptotic regime, most of which are difficult and unsolved.

Recently the zero-error information theory has been studied in the quantum setting and many new phenomena were observed. One remarkable result is the super-activation of the zero-error classical/quantum capacities of quantum channels [DS08, Dua09, CCH11, CS12, SS15]. Another important result is that, for some classical channels, quantum entanglement can be used to improve the zero-error capacity [CLMW10, LMM$^+$12], while there is no such advantage for the normal capacity [Sha48]. Furthermore, there are more kinds of capacities when considering auxiliary resources, such as the shared entanglement [DS08, CLMW10, LMM$^+$12, DSW13, BBL$^+$15], the no-signalling correlations [Mat12, CLMW11, DW16, LM15], and the feedback assistance [Sha56, DSW16]. All of these capacities are only partially understood, and the zero-error information theory of quantum channels seems more complex than that of classical channels.

To study the zero-error communication via quantum channels, the so-called "non-commutative graph theory" was introduced in [DSW13]. The non-commutative graph (an object based on an operator system) associated with a quantum channel fully captures the zero-error communication properties of this channel [DSW13], thus playing a similar role to confusability graph in the classical case. It is well-known that the zero-error capacity is extremely difficult to compute for both classical and quantum channels [BS07]. Nevertheless, the zero-error capacity of a classical channel is upper bounded by the Lovász number [Lov79] while the zero-error capacity of a quantum channel is upper bounded by the quantum Lovász number [DSW13]. Furthermore, the entanglement-assisted zero-error capacity of a classical channel is also upper-bounded by the Lovász number [Bei10, DSW13], and this result can be generalized to quantum channels by introducing the quantum Lovász number [DSW13].

### 7.1.2 Outline

In this chapter, we begin with the basic notations and results of classical and quantum zero-error information theory in Section 7.2. Then we show an approach to separate the entanglement-assisted zero-error capacity and the quantum Lovász number in Section 7.3, which resolves a well-known open problem in the area of zero-error information theory. Furthermore, in Section 7.4, we introduce an activated zero-error communication model and explore its novel properties.

## 7.2 Zero-error capacity of a quantum channel

### 7.2.1 Graphs and their quantum generalizations

#### Confusability graph and bipartite graph

Let us begin with a classical channel $\mathcal{N} = (X, p(y|x), Y)$ with $X$ and $Y$ are the input and output alphabets, respectively. To transmit messages through this channel with no probability of confusion, different messages need to be associated to different inputs $x$ in a way such that the output distributions $p(\cdot|x)$ are disjoint. This motivates the introduction of the confusability graph $G = (X, E)$ of a noisy channel [Sha56], where $X$ is the set of vertices (inputs) and $E$ is the set of edges. An edge $x \sim x'$ exists if $x$ and $x'$ can be confused via the channel, i.e., there exists some $y$ such that $p(y|x)\, p(y|x') > 0$.

The *independence number* $\alpha(G)$ is defined as the maximum size of an independent set in $G$, which is the maximum number of messages that can be transmitted through the channel without any possibility of confusion. For any classical channel with confusability graph $G$, we also denote $\alpha(\mathcal{N}) = \alpha(G)$ as its one-shot zero-error capacity.

### Non-commutative graph

For a quantum channel $\mathcal{N}$ from $\mathcal{L}(A')$ to $\mathcal{L}(B)$, with a Choi-Kraus operator sum representation $\mathcal{N}(\rho) = \sum_k E_k \rho E_k^\dagger$, its non-commutative graph [DSW13] is defined by the operator subspace

$$S := K^\dagger K = \text{span}\{E_j^\dagger E_k : j,k\} < \mathcal{L}(A'),$$

where $S < \mathcal{L}(A')$ means that $S$ is a subspace of $\mathcal{L}(A')$.

Taking the above classical channel $\mathcal{N} = (X, p(y|x), Y)$ as an example, its Choi-Kraus operators may be chosen as $E_{xy} = \sqrt{p(y|x)}|y\rangle\langle x|$. Thus, its non-commutative graph is given by

$$K = \text{span}\{Z : \forall x \not\sim x', \langle x|Z|x'\rangle = 0\}.$$

### Non-commutative bipartite graph

A classical channel $\mathcal{N} = (X, p(y|x), Y)$ also induces a *bipartite graph* $(X, E, Y)$, where $X$ and $Y$ are the input and output alphabets, respectively. And $E \subset X \times Y$ is the set of edges such that $(x, y) \in E$ if and only if the probability $p(y|x)$ is positive. It is worth noting that bipartite graph also plays an important role in the study of zero-error information theory [Sha56] and graph theory.

Given a quantum channel $\mathcal{N}$, its non-commutative bipartite graph (or Choi-Kraus operator space) is denoted by

$$K = K(\mathcal{N}) := \text{span}\{E_k\}. \tag{7.1}$$

Such space can be considered as a quantum analog of the bipartite graph since it determines the zero-error capacity of a quantum channel in the presence of noiseless feedback [DSW16], which plays a similar role to the bipartite graph of a classical channel [Sha56]. We denote $P_{AB}$ as the projection onto the support of the Choi matrix $J_\mathcal{N}$ of the channel, which is just the subspace $(\mathbb{1} \otimes K)|\Phi\rangle$. This indicates that we could use the $P_{AB}$ to characterize the non-commutative graph $K$.

| Classical Channel | Confusability Graph | Bipartite Graph |
|---|---|---|
| Quantum Channel | Non-commutative graph | Non-commutative Bipartite graph |

Table 7.1: Classical graphs and their quantum analogs

Taking the classical channel $\mathcal{N} = (X, p(y|x), Y)$ as an example, its non-commutative bipartite graph is defined by

$$K = \mathrm{span}\{|y\rangle\langle x| : (x,y) \text{ is an edge in } (X,E,Y)\}. \tag{7.2}$$

**Classical-quantum graph**

A classical-quantum (cq) channel $\mathcal{N} : i \to \rho_i$ $(1 \le i \le n)$ is a CPTP map with classical inputs $\{i\}_{i=1}^n$ and quantum outputs $\{\rho_i\}_{i=1}^n$. The non-commutative bipartite graph of a cq channel will be called a cq graph. In this case, the cq graph is given by

$$K = \mathrm{span}\{|\psi\rangle\langle i| : |\psi\rangle \in \mathrm{supp}\,(\rho_i)\}. \tag{7.3}$$

### 7.2.2 Zero-error capacity of a quantum channel

In the quantum world, we do measurements to distinguish the output quantum states of the channel and the *one-shot zero-error capacity of a quantum channel* $\mathcal{N}$ is defined by the maximum number of inputs such that the receiver can perfectly distinguish the corresponding output states. Note that the set of output states can be perfectly distinguished if and only if they are orthogonal.

**Definition 7.1.** This one-shot zero-error capacity can be equivalently defined as the independence number $\alpha\,(S)$ of the non-commutative graph [DSW13] of $\mathcal{N}$, i.e., the maximum size of a set of orthogonal unit vectors $\{|\phi_m\rangle : m = 1, ..., M\}$ such that

$$\forall m \neq m', |\phi_m\rangle\langle\phi'_m| \in S^{\perp}.$$

The zero-error capacity is given by regularization of $\alpha\,(S)$, i.e.,

$$C_0\,(\mathcal{N}) = C_0\,(S) = \sup_{n\to\infty} \frac{1}{n} \log \alpha\left(S^{\otimes n}\right), \tag{7.4}$$

and the sup in Eq. (7.4) can be replaced by lim.

The entanglement-assisted independence number $\widetilde{\alpha}\,(S)$ [DSW13] is motivated by the scenario where sender and receiver share entangled state beforehand and it quantifies the maximum number of distinguishable messages that can be sent via the channel $\mathcal{N}$ with graph $S$ when shared entanglement is free.

**Definition 7.2.** For a quantum channel with non-commutative graph $S$, $\widetilde{\alpha}\,(S)$ is the maximum integer $M$ such that there exist Hilbert spaces $A_0, B_0$ and a state $\sigma \in \mathcal{L}\,(A_0 \otimes B_0)$, and CPTP maps $\mathcal{E}_m : \mathcal{L}\,(A_0) \to \mathcal{L}\,(A)$ $(m = 1, ..., N)$ such that the $N$ output states $\rho_m = (\mathcal{N} \circ \mathcal{E}_m \otimes \mathrm{id}_{B_0})\,\sigma$ are orthogonal. The entanglement-assisted zero-error capacity of $S$ is given by regularization of $\widetilde{\alpha}\,(S)$, i.e.,

$$C_{0E}\,(\mathcal{N}) = C_{0E}\,(S) = \sup_{n\to\infty} \frac{1}{n} \log \widetilde{\alpha}\left(S^{\otimes n}\right). \tag{7.5}$$

**Quantum Lovász number**

The quantum Lovász number introduced by Duan, Severini, and Winter [DSW13] is a quantum analogue of the celebrated Lovász number [Lov79], which upper bounds entanglement-assisted zero-error capacity of the channel. It can be formalized by semidefinite programming (SDP) as follows [DSW13]:

$$\begin{aligned} \widetilde{\vartheta}\,(S) = \max\ & \langle\Phi|\,(\mathbb{1}\otimes\rho + T)\,|\Phi\rangle \\ \text{s.t.}\ & T \in S^{\perp}\otimes\mathcal{L}\left(A'\right), \quad \mathrm{Tr}\,\rho = 1, \\ & \mathbb{1}\otimes\rho + T \geq 0, \quad \rho \geq 0, \end{aligned} \tag{7.6}$$

where $|\Phi\rangle = \sum_i |i\rangle_A |i\rangle_{A'}$.

The dual SDP of $\widetilde{\vartheta}\,(S)$ is given by

$$\begin{aligned} \widetilde{\vartheta}\,(S) = \min\ & \|\,\mathrm{Tr}_A\, Y\|_\infty \\ \text{s.t.}\ & Y \in S\otimes\mathcal{L}\left(A'\right), \quad Y \geq |\Phi\rangle\langle\Phi|. \end{aligned} \tag{7.7}$$

By strong duality, the optimal values of the primal and dual SDPs of $\widetilde{\vartheta}\,(S)$ coincide. In the following, the quantum Lovász number of a channel $\mathcal{N}$ is naturally given by the quantum Lovász number of its non-commutative graph $S$, i.e., $\widetilde{\vartheta}\,(\mathcal{N}) = \widetilde{\vartheta}\,(S)$.

**Theorem 7.3** ([DSW13])**.** *For a quantum channel $\mathcal{N}$ with non-commutative graph $S$, $\widetilde{\vartheta}\,(S)$ is an upper bound of the entanglement-assisted zero-error capacity of the channel, i.e.,*

$$C_0\,(S) \leq C_{0E}\,(S) \leq \log \widetilde{\vartheta}\,(S). \tag{7.8}$$

### 7.2.3 An upper bound on the independence number

In this subsection, we are going to derive an upper bound on the one-shot zero-error capacity of a quantum channel motivated in the same sipirit of Lovász's number [Lov79]. But we do not know whether this bound is efficiently computable or not.

Let us denote

$$\kappa\,(\mathcal{N}) = \min_{\sigma} \max_{\rho} \frac{1}{\mathrm{Tr}\,\sigma P_{\mathcal{N}(\rho)}}, \tag{7.9}$$

where $P_{\mathcal{N}(\rho)}$ denotes the projection onto the support of $\mathcal{N}(\rho)$ and $\sigma$ is a quantum state. As an analog to the geometrical explanation on page 2 of [Lov79], the set $\Pi = \{P_{\mathcal{N}(\rho)} : \forall\rho\}$ can seem as an "umbrella," and we hope to find the "handle" $\sigma$ that minimizes the maximum "angle" between the handle and every rib of the umbrella.

**Proposition 7.4.** *For any quantum channel $\mathcal{N}$, the independence number $\alpha(\mathcal{N})$ is upper bounded by $\kappa(\mathcal{N})$, i.e.,*

$$\alpha(\mathcal{N}) \leq \kappa(\mathcal{N}). \tag{7.10}$$

*Proof.* The idea of this proof follows Lovász's idea in [Lov79]. Suppose that $\alpha(\mathcal{N}) = k$, this means one could find $k$ inputs $\{\rho_i\}_{i=0}^{k-1}$ such that

$$\operatorname{Tr}\left(\mathcal{N}(\rho_i)\,\mathcal{N}(\rho_j)\right) = 0,\ \forall i \neq j, \tag{7.11}$$

and this number $k$ is optimal.

The above Eq. (7.11) just means there are no overlaps between the output states. So let the projection onto $\mathcal{N}(\rho_i)$ be $P_i$ for every $i \in \{0, 1, \ldots, k-1\}$, then Eq. (7.11) is equivalent to

$$\operatorname{Tr}\left(P_i P_j\right) = 0,\ \forall i \neq j, \tag{7.12}$$

which implies that

$$\sum_{i=0}^{k-1} P_i \leq \mathbb{1}. \tag{7.13}$$

Hence, for the optimal $\sigma$ in Eq. (7.9), we have

$$1 \geq \sum_{i=0}^{k-1} \operatorname{Tr} P_i \sigma \geq \frac{\alpha(\mathcal{N})}{\kappa(\mathcal{N})}. \tag{7.14}$$

□

## 7.3 Separating $C_{0E}$ and quantum Lovász number

An intriguing open problem in quantum zero-error information theory is whether the entanglement-assisted zero-error capacity always coincides with the quantum Lovász number for a classical or quantum channel, which is frequently mentioned in [LMM$^+$12, DSW13, Bei10, CLMW11, CMR$^+$14, MSS13]. If they are equal, it will imply that the entanglement-assisted zero-error capacity is additive, while the unassisted case is not [Alo98].

In this section, we resolve the above open problem for quantum channels. To be specific, we construct a class of qutrit-to-qutrit channels for which the quantum Lovász number is strictly larger than the entanglement-assisted zero-error capacity. We utilize the one-shot NS-assisted zero-error capacity and simulation cost to determine the asymptotic NS-assisted zero-error capacity in this case, which is potentially larger than the entanglement-assisted zero-error capacity. An interesting fact is that this class of channels are reversible in a strong sense. To be specific, for this class of channels, the one-shot NS-assisted zero-error capacity and simulation cost are identical. We then give a closed formula for the quantum Lovász number for this class of channels and use it to conclude that there is a strict gap between the quantum Lovász number and the entanglement-assisted zero-error capacity. For this class of channels, we also find that the quantum fractional packing number is strictly larger than the feedback-assisted or NS-assisted zero-error capacity, while these three quantities are equal to each other for any classical channel [CLMW11].

### 7.3.1 Zero-error communication quantities

#### NS-assisted zero-error communication

The no-signalling correlations arises in the research of the relativistic causality of quantum operations [BGNP01, ESW02, PHHH06, OCB12] and Cubitt et al. [CLMW11] first introduced classical no-signalling correlations into the zero-error communication via classical channels and proved that the fractional packing number of the bipartite graph induced by the channel equals to the zero-error capacity of the channel. Recently, quantum no-signalling correlations were introduced into the zero-error communication via quantum channels in [DW16] and the one-shot NS-assisted zero-error classical capability (quantified as the number of messages) was formulated as the following SDP:

$$\begin{aligned} \Upsilon\left(\mathcal{N}\right) = \Upsilon\left(K\right) = \max \operatorname{Tr} R_A & \\ \text{s.t. } 0 \leq U_{AB} \leq R_A \otimes \mathbb{1}_B, & \\ \operatorname{Tr}_A U_{AB} = \mathbb{1}_B, & \\ \operatorname{Tr} P_{AB}\left(R_A \otimes \mathbb{1}_B - U_{AB}\right) = 0, & \end{aligned} \tag{7.15}$$

where $P_{AB}$ denotes the projection onto $(\mathbb{1} \otimes K)\,|\Phi\rangle$. The asymptotic NS-assisted zero-error capacity is given by the regularization:

$$C_{0,\text{NS}}\left(\mathcal{N}\right) = C_{0,\text{NS}}\left(K\right) = \sup_{n\to\infty} \frac{1}{n} \log \Upsilon\left(K^{\otimes n}\right). \tag{7.16}$$

A remarkable feature of NS-assisted zero-error capacity is that one bit noiseless communication can fully activate any classical-quantum channel to achieve its asymptotic capacity [DW15].

### NS-assisted zero-error simulation

A more general problem is the simulation of a channel, which concerns how to use a channel $\mathcal{N}$ from Alice ($A$) to Bob ($B$) to simulate another channel $\mathcal{M}$ also from $A$ to $B$ [KW04]. Shannon's noisy channel coding theorem determines the capability of any noisy channel $\mathcal{N}$ to simulate a noiseless channel [Sha48] and the reverse Shannon theorem was proved in [BSST02]. The quantum reverse Shannon theorem was proved recently [BDH$^+$14, BCR11], which states that any quantum channel can be simulated by an amount of classical communication equal to its entanglement-assisted capacity assisted with free entanglement. In the zero-error setting, there is a kind of reversibility between the zero-error capacity and simulation cost in the presence of no-signalling correlations [CLMW11]. More recently, the no-signalling-assisted (NS-assisted) zero-error simulation cost of a quantum channel was introduced in [DW16].

The zero-error simulation cost of a quantum channel in the presence of quantum no-signalling correlations was introduced in [DW16] and formalized as SDPs. To be specific, for the quantum channel $\mathcal{N}$ with Choi-Jamiołkowski matrix $J_{\mathcal{N}}$, the NS-assisted zero-error simulation cost of $\mathcal{N}$ is given by

$$S_{0,\mathrm{NS}}\left(\mathcal{N}\right) = -H_{\min}\left(A|B\right)_{J_{\mathcal{N}}} := \log \Sigma\left(\mathcal{N}\right), \tag{7.17}$$

where

$$\begin{aligned} \Sigma\left(\mathcal{N}\right) = \min \operatorname{Tr} T_B, \\ \text{s.t. } J_{\mathcal{N}} \leq \mathbb{1}_A \otimes T_B, \end{aligned} \tag{7.18}$$

and $H_{\min}\left(A|B\right)_{J_{\mathcal{N}}}$ is the conditional min-entropy (cf. Eq. (2.61)). By the fact that the conditional min-entropy is additive [KRS09], the asymptotic NS-assisted zero-error simulation cost is given by

$$S_{0,\mathrm{NS}}\left(\mathcal{N}\right) = \log \Sigma\left(\mathcal{N}\right). \tag{7.19}$$

Furthermore, noting that the NS assistance is stronger than the entanglement assistance, the capacities and simulation cost of a quantum channel introduced above obey the following inequality:

$$C_0 \leq C_{0E} \leq C_{0,\mathrm{NS}} \leq C_E \leq S_{0,\mathrm{NS}}, \tag{7.20}$$

where $C_E$ is the entanglement-assisted classical capacity [BSST02].

### 7.3.2 Establishing the gap

In this section, we are going to show the gap between the quantum Lovász number and the entanglement-assisted zero-error capacity. The difficulty in comparing $C_{0E}$ and the quantum Lovász number is that there are few channels whose entanglement-assisted zero-error capacity is known. In fact, $C_{0E}$ is even not known to be computable. The problem whether there exists a gap between them was a prominent open problem in the area of zero-error quantum information theory.

Our approach is to construct a particular class of channels and evaluate its NS-assisted zero-error capacity, which is potentially larger than the entanglement-assisted case.

#### A qutrit-qutrit channel in the spirit of the amplitude damping noise

Let us recall the class of channels $\mathcal{N}_\alpha$ which we established the strong converse property for classical and private communication in Section 5.4.3:

$$\mathcal{N}_\alpha(\rho) = E_\alpha \rho E_\alpha^\dagger + D_\alpha \rho D_\alpha^\dagger \; (0 < \alpha \leq \pi/4), \tag{7.21}$$

where

$$E_\alpha = \sin\alpha |0\rangle\langle 1| + |1\rangle\langle 2|, \tag{7.22}$$
$$D_\alpha = \cos\alpha |2\rangle\langle 1| + |1\rangle\langle 0|. \tag{7.23}$$

This qutrit-qutrit channel $\mathcal{N}_\alpha$ is motivated in the similar spirit of the amplitude damping channel and it exhibits a significant difference from the classical channels.

The Choi-Jamiołkowski matrix of $\mathcal{N}_\alpha$ is given by

$$J_\alpha = \left(1+\sin^2\alpha\right)|u_\alpha\rangle\langle u_\alpha| + \left(1+\cos^2\alpha\right)|v_\alpha\rangle\langle v_\alpha|,$$

where

$$|u_\alpha\rangle = \frac{\sin\alpha}{\sqrt{1+\sin^2\alpha}}|10\rangle + \frac{1}{\sqrt{1+\sin^2\alpha}}|21\rangle, \tag{7.24}$$
$$|v_\alpha\rangle = \frac{\cos\alpha}{\sqrt{1+\cos^2\alpha}}|12\rangle + \frac{1}{\sqrt{1+\cos^2\alpha}}|01\rangle. \tag{7.25}$$

Then, the projection onto the support of $J_\alpha$ is

$$P_\alpha = |u_\alpha\rangle\langle u_\alpha| + |v_\alpha\rangle\langle v_\alpha|. \tag{7.26}$$

**Zero-error capacity and simulation cost of $\mathcal{N}_\alpha$**

We first prove that both NS-assisted zero-error capacity and simulation cost of $\mathcal{N}_\alpha$ are exactly two bits.

**Proposition 7.5.** *For the channel $\mathcal{N}_\alpha$ $(0 < \alpha \leq \pi/4)$,*

$$C_{0,\text{NS}}(\mathcal{N}_\alpha) = C_E(\mathcal{N}_\alpha) = S_{0,\text{NS}}(\mathcal{N}_\alpha) = 2. \tag{7.27}$$

*Proof.* First, we show that Alice can trasmit at least 2 bits prefectly to Bob with a single use of $\mathcal{N}_\alpha$ and the NS-assistance. The approach is to construct a feasible solution of the SDP (7.15) of the one-shot NS-assisted zero-error capacity. To be specific, suppose that $R_A = 2\left(\cos^2\alpha|0\rangle\langle 0| + |1\rangle\langle 1| + \sin^2\alpha|2\rangle\langle 2|\right)$ and

$$\begin{aligned} U_{AB} =& \cos^2\alpha|01\rangle\langle 01| + \sin^2\alpha|21\rangle\langle 21| + |10\rangle\langle 10| + |12\rangle\langle 12| \\ &+ \sin\alpha\left(|10\rangle\langle 21| + |21\rangle\langle 10|\right) + \cos\alpha\left(|01\rangle\langle 12| + |12\rangle\langle 01|\right). \end{aligned}$$

One can simply check that $R_A \otimes \mathbb{1}_B - U_{AB} \geq 0$, $\operatorname{Tr}_A U_{AB} = \mathbb{1}_B$ and $P_\alpha\left(R_A \otimes \mathbb{1}_B - U_{AB}\right) = 0$. Therefore, $\{R_A, U_{AB}\}$ is a feasible solution to SDP (7.15) of $\Upsilon(\mathcal{N}_\alpha)$, which means that

$$C_{0,\text{NS}}(\mathcal{N}_\alpha) \geq \log\Upsilon(\mathcal{N}_\alpha) \geq \log\operatorname{Tr} R_A = 2. \tag{7.28}$$

Second, we prove that the one-shot NS-assisted simulation cost of $\mathcal{N}_\alpha$ is at amost 2 bits. We utilize the SDP (7.18) of one-shot NS-assisted simulation cost and choose

$$T_B = 2\left(\sin^2\alpha|0\rangle\langle 0| + |1\rangle\langle 1| + \cos^2\alpha|2\rangle\langle 2|\right). \tag{7.29}$$

It can be checked that $\mathbb{1} \otimes T_B - J_\alpha \geq 0$. Thus, $T_B$ is a feasible solution to SDP (7.18) of $\Sigma(\mathcal{N}_\alpha)$, which means that

$$S_{0,\text{NS}}(\mathcal{N}_\alpha) \leq \log\Sigma(\mathcal{N}_\alpha) \leq \log\operatorname{Tr} T_B = 2. \tag{7.30}$$

Finally, combining Eq. (7.28), Eq. (7.30) and Eq. (7.20), it is clear that

$$C_{0,\text{NS}}(\mathcal{N}_\alpha) = C_E(\mathcal{N}_\alpha) = S_{0,\text{NS}}(\mathcal{N}_\alpha) = 2. \tag{7.31}$$

□

**Quantum Lovász number of $\mathcal{N}_\alpha$**

We then solve the exact value of the quantum Lovász number of $\mathcal{N}_\alpha$.

**Proposition 7.6.** *For the channel $\mathcal{N}_\alpha$ $(0 < \alpha \le \pi/4)$,*

$$\widetilde{\vartheta}\left(\mathcal{N}_\alpha\right) = 2 + \cos^2\alpha + \cos^{-2}\alpha > 4. \tag{7.32}$$

*Proof.* We first construct a quantum state $\rho$ and an operator $T \in S^\perp \otimes \mathcal{L}\left(A'\right)$ such that $\mathbb{1} \otimes \rho + T$ is positive semidefinite. Then, we use the primal SDP (7.6) of $\widetilde{\vartheta}\left(\mathcal{N}_\alpha\right)$ to obtain the lower bound of $\widetilde{\vartheta}\left(\mathcal{N}_\alpha\right)$.

To be specific, the non-commutative graph of $\mathcal{N}_\alpha$ is $S = \text{span}\{F_1, F_2, F_3, F_4\}$ with

$$F_1 = |0\rangle\langle 0| + \cos^2\alpha |1\rangle\langle 1|, \tag{7.33}$$
$$F_2 = \sin^2\alpha |1\rangle\langle 1| + |2\rangle\langle 2|, \tag{7.34}$$
$$F_3 = |0\rangle\langle 2| \text{ and } F_4 = |2\rangle\langle 0|. \tag{7.35}$$
$$\tag{7.36}$$

Let us choose

$$\rho = \frac{\cos^2\alpha}{1+\cos^2\alpha}|0\rangle\langle 0| + \frac{1}{1+\cos^2\alpha}|1\rangle\langle 1| \tag{7.37}$$

and $T = T_1 \otimes T_2 + R$, where

$$T_1 = \frac{1}{1+\cos^2\alpha}\left(|0\rangle\langle 0| - \frac{1}{\cos^2\alpha}|1\rangle\langle 1| + \frac{\sin^2\alpha}{\cos^2\alpha}|2\rangle\langle 2|\right), \tag{7.38}$$
$$T_2 = \cos^4\alpha|0\rangle\langle 0| - |1\rangle\langle 1|, \tag{7.39}$$
$$R = |00\rangle\langle 11| + |11\rangle\langle 00|. \tag{7.40}$$

It is clear that $\rho \ge 0$ and $\operatorname{Tr}\rho = 1$. Also, it is easy to see that for any matrix $M \in \mathcal{L}\left(A'\right)$ and $j = 1,2,3,4$,

$$\operatorname{Tr} R\left(F_j \otimes M\right) = 0. \tag{7.41}$$

Meanwhile, noticing that $\operatorname{Tr}\left(T_1 F_j\right) = 0$ for $j = 1,2,3,4$, we have

$$T = T_1 \otimes T_2 + R \in S^\perp \otimes \mathcal{L}\left(A'\right). \tag{7.42}$$

Moreover, it is easy to see that

$$\begin{aligned} \mathbb{1} \otimes \rho + T = &\cos^2\alpha|00\rangle\langle 00| + \frac{1}{\cos^2\alpha}|11\rangle\langle 11| + |00\rangle\langle 11| \\ &+ |11\rangle\langle 00| + \frac{\cos^2\alpha - \cos^4\alpha}{1+\cos^2\alpha}|20\rangle\langle 20| \\ &+ \frac{2\cos^2 -1}{(1+\cos^2\alpha)\cos^2\alpha}|21\rangle\langle 21| \geq 0. \end{aligned} \tag{7.43}$$

Then, $\{\rho, T\}$ is a feasible solution to primal SDP (7.6) of $\widetilde{\vartheta}\left(\mathcal{N}_\alpha\right)$. Hence, we have that

$$\begin{aligned} \widetilde{\vartheta}\left(\mathcal{N}_\alpha\right) &\geq \mathrm{Tr}[|\Phi\rangle\langle\Phi| \left(\mathbb{1} \otimes \rho + T\right)] \\ &= \mathrm{Tr}[|\Phi\rangle\langle\Phi| \left(\mathbb{1} \otimes \rho + T_1 \otimes T_2 + R\right)] \\ &= 2 + \cos^2\alpha + \cos^{-2}\alpha. \end{aligned} \tag{7.44}$$

On the other hand, we find a feasible solution to the dual SDP (7.7) of $\widetilde{\vartheta}\left(\mathcal{N}_\alpha\right)$. It is easy to see that

$$S^\perp = \mathrm{span}\{\mathrm{M}_1, \mathrm{M}_2, \mathrm{M}_3, \mathrm{M}_4, \mathrm{M}_5\}, \tag{7.45}$$

where $M_1 = |0\rangle\langle 1|$, $M_2 = |1\rangle\langle 0|$, $M_3 = |1\rangle\langle 2|$, $M_4 = |2\rangle\langle 1|$ and $M_5 = |0\rangle\langle 0| - \cos^{-2}\alpha|1\rangle\langle 1| + \tan^2\alpha|2\rangle\langle 2|$. Let us choose

$$Y = Y_1 \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) + Y_2 \otimes |2\rangle\langle 2| + \frac{1+\cos^2\alpha}{\cos^2\alpha}Y_3 \tag{7.46}$$

with

$$Y_1 = \left(1+\cos^2\alpha\right)\cos^{-2}\alpha|0\rangle\langle 0| + \left(1+\cos^2\alpha\right)|1\rangle\langle 1|, \tag{7.47}$$

$$Y_2 = \left(2-\cos^{-2}\alpha\right)|0\rangle\langle 0| + \left(\cos^{-2}\alpha - \sin^2\alpha\right)|1\rangle\langle 1| \tag{7.48}$$

$$+ \left(1+\cos^2\alpha\right)\cos^{-2}\alpha|2\rangle\langle 2|, \tag{7.49}$$

$$Y_3 = |00\rangle\langle 22| + |22\rangle\langle 00|. \tag{7.50}$$

It is easy to see that for any matrix $V \in \mathcal{L}\left(A'\right)$ and $j = 1,2,3,4,5$, we have that

$$\mathrm{Tr}\, Y_3\left(M_j \otimes V\right) = 0. \tag{7.51}$$

Meanwhile, since $\mathrm{Tr}\left(Y_k M_j\right) = 0$ for $k = 1,2$ and $j = 1,2,3,4,5$, we have that

$$\begin{aligned} Y = &Y_1 \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) + Y_2 \otimes |2\rangle\langle 2| + \frac{1+\cos^2\alpha}{\cos^2\alpha}Y_3 \\ &\in S \otimes \mathcal{L}\left(A'\right). \end{aligned}$$

It is also easy to check that $Y - |\Phi\rangle\langle\Phi| \geq 0$. Thus, $Y$ is a feasible solution to SDP (7.7) of $\widetilde{\vartheta}(\mathcal{N}_\alpha)$. Furthermore, one can simply calculate that

$$\operatorname{Tr}_A Y = \left(2 + \cos^2\alpha + \cos^{-2}\alpha\right) \mathbb{1}_B, \tag{7.52}$$

Therefore,

$$\widetilde{\vartheta}(\mathcal{N}_\alpha) \leq \|\operatorname{Tr}_A Y\|_\infty = 2 + \cos^2\alpha + \cos^{-2}\alpha. \tag{7.53}$$

Finally, combining Eq. (7.44) and Eq. (7.53), we can conclude that

$$\widetilde{\vartheta}(\mathcal{N}_\alpha) = 2 + \cos^2\alpha + \cos^{-2}\alpha.$$

□

### Gap between $\log \widetilde{\vartheta}(\mathcal{N}_\alpha)$ and $C_{0E}(\mathcal{N}_\alpha)$

Now we are able to show a separation between $\log \widetilde{\vartheta}(\mathcal{N}_\alpha)$ and $C_{0E}(\mathcal{N}_\alpha)$.

**Theorem 7.7.** *For the channel $\mathcal{N}_\alpha$ $(0 < \alpha \leq \pi/4)$, the quantum Lovász number is strictly larger than the entanglement-assisted zero-error capacity (or even with no-signalling assistance), i.e.,*

$$\log \widetilde{\vartheta}(\mathcal{N}_\alpha) > C_{0,\text{NS}}(\mathcal{N}_\alpha) \geq C_{0E}(\mathcal{N}_\alpha). \tag{7.54}$$

*Proof.* It is easy to see this result from Proposition 7.5 and Proposition 7.6. To be specific, we have

$$\begin{aligned}
\log \widetilde{\vartheta}(\mathcal{N}_\alpha) &= \log\left(2 + \cos^2\alpha + \cos^{-2}\alpha\right) && (7.55)\\
&> 2 && (7.56)\\
&= C_{0,\text{NS}}(\mathcal{N}_\alpha) && (7.57)\\
&\geq C_{0E}(\mathcal{N}_\alpha). && (7.58)
\end{aligned}$$

□

### Gap between quantum fractional packing number and feedback-assisted or NS-assisted zero-error capacity

Shannon first introduced the feedback-assisted zero-error capacity [Sha56]. To be precise, his model has noiseless instantaneous feedback of the channel output back to the sender, and it requires some arbitrarily small rate of forward noiseless communication. For any classical channel with a positive zero-error capacity, he showed that the feedback-assisted zero-error capacity $C_{0F}$ of a classical channel $\mathcal{N}$ is given by the fractional packing number of its bipartite graph [Sha56]:

$$\alpha^* (\Gamma) = \max \left\{ \sum_x v_x \ \text{ s.t. } \sum_x v_x p\,(y|x) \le 1 \forall y, 0 \le v_x \le 1\ \forall x \right\}.$$

For any classical bipartite graph, the fractional packing number also gives the NS-assisted zero-error classical capacity and simulation cost [CLMW11], i.e.,

$$C_{0,\text{NS}}\,(K) = S_{0,\text{NS}}\,(K) = \log \alpha^* (\Gamma)\,.$$

The quantum generalization of fractional packing number in [DW16] was suggested by Harrow as

$$\begin{aligned} \mathsf{A}\,(K) &= \max\{\operatorname{Tr} R_A : 0 \le R_A, \operatorname{Tr}_A P_{AB}\,(R_A \otimes \mathbb{1}_B) \le \mathbb{1}_B\}, \\ &= \min\{\operatorname{Tr} T_B \text{ s.t. } 0 \le T_B, \operatorname{Tr}_B P_{AB}\,(\mathbb{1}_A \otimes T_B) \ge \mathbb{1}_A\}. \end{aligned} \tag{7.59}$$

This quantum fractional packing number $\mathsf{A}\,(K)$ has nice mathematical properties such as additivity under tensor product [DW16].

For any bipartite graph $\Gamma$, quantum fractional packing number also reduces to the fractional packing number, i.e.,

$$\mathsf{A}\,(K) = \alpha^* (\Gamma)\,. \tag{7.60}$$

Furthermore, for a classical-quantum channel with non-commutative bipartite graph $K$, it also holds that [DW16]

$$C_{0,\text{NS}}\,(K) = \log \mathsf{A}\,(K)\,. \tag{7.61}$$

However, if we consider general quantum channels, this quantum fractional packing number will exceed the NS-assisted zero-error capacity as well as the feedback-assisted zero-error capacity. An example is the class of channels $\mathcal{N}_\alpha$ and the proof is in the following Proposition 7.9. For $\mathcal{N}_\alpha$, it is easy to see that the set of linear operators $\{E_i^\dagger E_j\}$ is linearly independent, which means that $\mathcal{N}_\alpha$ is an extremal channel [Cho75]. Thus, its non-commutative bipartite graph $K_\alpha$ is an extremal graph [DW16], which means that there can only be a unique channel $\mathcal{N}$ such that $K\,(\mathcal{N}) = K_\alpha$.

For a general quantum channel, its feedback-assisted zero-error capacity depends only on its non-commutative bipartite graph. And the feedback-assisted zero-error capacity is always smaller than or equal to the entanglement-assisted classical capacity [DSW16], i.e.,

$$C_{0\text{F}}\,(K) \le C_{\text{minE}}\,(K)\,, \tag{7.62}$$

where $C_{\text{minE}}(K)$ is defined by

$$C_{\text{minE}}(K) := \min\{C_E(\mathcal{N}) : K(\mathcal{N}) < K\}. \tag{7.63}$$

Considering the fact that $C_{0,\text{NS}}(K) \leq C_{\text{minE}}(K) \leq S_{0,\text{NS}}(K)$ [DSW16], it is easy to see that $C_{\text{minE}}(K_\alpha)$ is exactly two bits from Proposition 7.5.

**Lemma 7.8.** *For non-commutative bipartite graph $K_\alpha$ $(0 < \alpha \leq \pi/4)$, the quantum fractional packing number is given by*

$$\mathsf{A}(K_\alpha) = 2 + \cos^2\alpha + \cos^{-2}\alpha. \tag{7.64}$$

*Proof.* Let us choose $R_A = \left(2 - \sin^2\alpha\right)|0\rangle\langle 0| + x|1\rangle\langle 1|$, then

$$\operatorname{Tr}_A P_\alpha\left(R_A \otimes \mathbb{1}_B\right) = \frac{x\sin^2\alpha}{1+\sin^2\alpha}|0\rangle\langle 0| + |1\rangle\langle 1| + \frac{x\cos^2\alpha}{1+\cos^2\alpha}|2\rangle\langle 2|.$$

When $x = 1 + \cos^{-2}\alpha$, it is clear that $\operatorname{Tr}_A P_\alpha\left(R_A \otimes \mathbb{1}_B\right) \leq \mathbb{1}_B$. Therefore, $R_A$ is a feasible solution to the primal SDP of $\mathsf{A}(\mathcal{N}_\alpha)$, which means that

$$\mathsf{A}(\mathcal{N}_\alpha) \geq \operatorname{Tr} R_A = 2 + \cos^2\alpha + \cos^{-2}\alpha. \tag{7.65}$$

Similarly, it is easy to check that $T_B = \left(2 - \sin^2\alpha\right)|1\rangle\langle 1| + \left(1 + \cos^{-2}\alpha\right)|2\rangle\langle 2|$ is a feasible solution to the dual SDP of $\mathsf{A}(\mathcal{N}_\alpha)$. Therefore,

$$\mathsf{A}(\mathcal{N}_\alpha) \leq \operatorname{Tr} T_B = 2 + \cos^2\alpha + \cos^{-2}\alpha. \tag{7.66}$$

Hence, we have that $\mathsf{A}(\mathcal{N}_\alpha) = 2 + \cos^2\alpha + \cos^{-2}\alpha$. □

Now, we are able to show the separation.

**Proposition 7.9.** *For non-commutative bipartite graph $K_\alpha$ $(0 < \alpha \leq \pi/4)$, we have that*

$$C_{0\text{F}}(K_\alpha) < \log \mathsf{A}(K_\alpha), \tag{7.67}$$

$$C_{0,\text{NS}}(K_\alpha) < \log \mathsf{A}(K_\alpha). \tag{7.68}$$

*Proof.* For general non-commutative bipartite graph $K$, it holds that [DSW16]:

$$C_{0\text{F}}(K) \leq C_{\min \text{E}}(K). \tag{7.69}$$

Then, by Proposition 7.5 and Lemma 7.8, we have

$$C_{0\text{F}}(K_\alpha) \leq C_{\min \text{E}}(K_\alpha) = 2 < \log \mathsf{A}(K_\alpha). \tag{7.70}$$

From Proposition 7.5 and Lemma 7.8, it is also clear that $C_{0,\mathrm{NS}}(K_\alpha) < \log \mathsf{A}(K_\alpha)$. □

## 7.4 Acivated zero-error communication

In this section, we further develop the theory of quantum NS-assisted communication by introducing the activated communication model. The model is introduced in Section 7.4.1 and it considers the additional forward noiseless channel as a catalyst for communication. For a quantum channel $\mathcal{N}$, we can "borrow" a noiseless classical channel $\mathcal{I}$, then we can use $\mathcal{N} \otimes \mathcal{I}$ to transmit information. After the communication finishes we "pay back" the capacity of $\mathcal{I}$. The communication model follows the idea of potential capacities of quantum channels introduced by Winter and Yang [WY15]. In Section 7.4.2, we show a striking result that one bit can even fully activate any cq channel to achieve its asymptotic NS-assisted zero-error capacity (or the fractional packing number). In Section 7.4.3, we further show that there is no activation in the asymptotic regime. We also exhibit a quantum channel to separate the asymptotic NS-assisted zero-error capacity and the semidefinite packing number.

### 7.4.1 Activated one-shot zero-error capacity

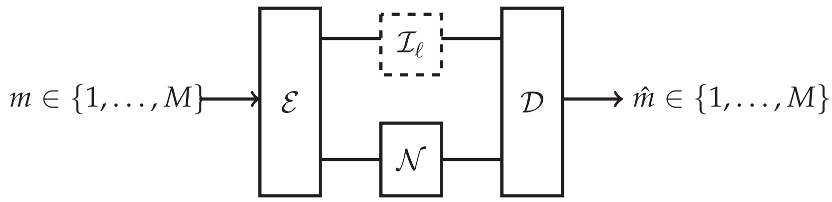


Figure 7.1: Activated classical communication.

**Definition 7.10.** For a quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$, the one-shot activated no-signalling assisted zero-error classical capacity is defined as the following:

$$\mathcal{M}^a_{0,\mathrm{NS}}(\mathcal{N}) = \mathcal{M}^a_{0,\mathrm{NS}}(K) := \sup_{\ell \geq 1}[\mathcal{M}_{0,\mathrm{NS}}(K \otimes \Delta_\ell) - \log \ell], \tag{7.71}$$

where $\Delta_\ell$ is the non-commutative graph of the noiseless channel

$$\mathcal{I}_\ell(\rho) = \sum_{i=0}^{\ell-1} \mathrm{Tr}(\rho|i\rangle\langle i|)\, |i\rangle\langle i|. \tag{7.72}$$

**Definition 7.11.** For a quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$, the asymptotic activated no-signalling zero-error classical capacity is given by the following regularization:

$$C_{0,\mathrm{NS}}^{a}(\mathcal{N}) = C_{0,\mathrm{NS}}^{a}(K) := \sup_{n\geq 1} \frac{1}{n} \mathcal{M}_{0,\mathrm{NS}}^{a}\left(K^{\otimes n}\right). \tag{7.73}$$

To provide a feasible formulation of the activated capacity $\mathcal{M}_{0,\mathrm{NS}}^{a}(\mathcal{N})$, let us first introduce a slightly revised SDP of $\Upsilon(K)$ as follows,

$$\begin{aligned} \widehat{\Upsilon}(K) = \max \operatorname{Tr} S_A & \\ \text{s.t. } & 0 \leq U_{AB} \leq S_A \otimes \mathbb{1}_B, \\ & \operatorname{Tr}_A U_{AB} \leq \mathbb{1}_B, \\ & \operatorname{Tr} P_{AB}\left(S_A \otimes \mathbb{1}_B - U_{AB}\right) = 0. \end{aligned} \tag{7.74}$$

The only difference between $\widehat{\Upsilon}(K)$ and $\Upsilon(K)$ is that now $\operatorname{Tr}_A U_{AB}$ is only required to be less than or equal to $\mathbb{1}_B$, and an equality is not necessary. However, we will see that such a small revision is of crucial importance. The dual SDP of $\widehat{\Upsilon}(K)$ is given by

$$\begin{aligned} \widehat{\Upsilon}(K) = \min \operatorname{Tr} T_B & \\ \text{s.t. } & V_{AB} \leq \mathbb{1}_A \otimes T_B, \\ & \operatorname{Tr}_B V_{AB} \geq \mathbb{1}_A,\ T \geq 0, \\ & (\mathbb{1} - P)_{AB} V_{AB} (\mathbb{1} - P)_{AB} \leq 0. \end{aligned} \tag{7.75}$$

Note that by strong duality, the values of both the primal and the dual SDPs coincide. It is also worth noting that for any given non-commutative bipartite graph $K$, it holds that

$$\widehat{\Upsilon}(K) \geq \Upsilon(K). \tag{7.76}$$

Now we are ready to present the main result.

**Theorem 7.12.** *For any quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$,*

$$\mathcal{M}_{0,\mathrm{NS}}^{a}(\mathcal{N}) = \log \widehat{\Upsilon}(K). \tag{7.77}$$

*Proof.* The intuition of this theorem is that the additional noiseless channel may play the role of a catalyst during the communication task.

To prove the achievable part, it's important to observe that the additional noiseless channel indeed provides a larger solution space of $\Upsilon(K \otimes \Delta_\ell)$. Let us first consider the case $\ell = 2$ and assume that the optimal feasible solution of $\widehat{\Upsilon}(K)$ is $\{S_A, U_{AB}\}$. Let us choose

$$S_{AA'} = S_A \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|)_{A'} \tag{7.78}$$

and

$$U_{AA'BB'} = U_{AB} \otimes (|00\rangle\langle 00| + |11\rangle\langle 11|)_{A'B'} + \bar{U}_{AB} \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A'B'}, \tag{7.79}$$

where $\bar{U}_{AB} = \frac{S_A}{\operatorname{Tr} S_A} \otimes (\mathbb{1}_B - \operatorname{Tr}_A U_{AB})$.

This construction ensures that

$$\operatorname{Tr}_{AA'} U_{AA'BB'} = \operatorname{Tr}_A ((U_{AB} + \bar{U}_{AB}) \otimes \mathbb{1}_{B'}) = \mathbb{1}_{BB'}. \tag{7.80}$$

Moreover, we have

$$S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'} = (S_A \otimes \mathbb{1}_B - U_{AB}) \otimes (|00\rangle\langle 00| + |11\rangle\langle 11|)_{A'B'} \tag{7.81}$$

$$+ (S_A \otimes \mathbb{1}_B - \bar{U}_{AB}) \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A'B'}, \tag{7.82}$$

which directly means that

$$S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'} \geq 0. \tag{7.83}$$

Furthermore, the projection onto the support of the Choi-Jamiołkowski matrix of $\mathcal{N} \otimes \mathcal{I}_2$ is $P_{ABA'B'} = P_{AB} \otimes D_{A'B'}$ with $D_{A'B'} = (|00\rangle\langle 00| + |11\rangle\langle 11|)_{A'B'}$. Therefore, we have that

$$\operatorname{Tr} P_{ABA'B'} (S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'}) \tag{7.84a}$$

$$= \operatorname{Tr} (P_{AB} \otimes D_{A'B'}) [(S_A \otimes \mathbb{1}_B - U_{AB}) \otimes (|00\rangle\langle 00| + |11\rangle\langle 11|)_{A'B'}] \tag{7.84b}$$

$$+ \operatorname{Tr} (P_{AB} \otimes D_{A'B'}) [(S_A \otimes \mathbb{1}_B - \bar{U}_{AB}) \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A'B'}] \tag{7.84c}$$

$$= \operatorname{Tr} P_{AB} (S_A \otimes \mathbb{1}_B - U_{AB}) \times \operatorname{Tr} D_{A'B'} (|00\rangle\langle 00| + |11\rangle\langle 11|)_{A'B'} \tag{7.84d}$$

$$+ \operatorname{Tr} P_{AB} (S_A \otimes \mathbb{1}_B - \bar{U}_{AB}) \times \operatorname{Tr} D_{A'B'} (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A'B'} \tag{7.84e}$$

$$= 0, \tag{7.84f}$$

where the last equality follows from

$$\operatorname{Tr} P_{AB} (S_A \otimes \mathbb{1}_B - U_{AB}) = 0, \tag{7.85}$$

$$\operatorname{Tr} D_{A'B'} (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A'B'} = 0. \tag{7.86}$$

Now we are able to conclude that $\{S_{AA'}, U_{AA'BB'}\}$ is a feasible solution of $\Upsilon(K\otimes\Delta_2)$, which means that

$$\sup_{\ell\geq 2}\frac{\Upsilon(K\otimes\Delta_\ell)}{\ell}\geq\frac{\Upsilon(K\otimes\Delta_2)}{2}\geq\frac{\operatorname{Tr}S_{AA'}}{2}=\widehat{\Upsilon}(K). \tag{7.87}$$

On the other hand, to prove the converse part, we will use the fact that $\widehat{\Upsilon}(K\otimes\Delta_\ell)=\ell\,\widehat{\Upsilon}(K)$, which is provided in the following Lemma 7.13. This fact directly implies that

$$\sup_{\ell\geq 2}\frac{\Upsilon(K\otimes\Delta_\ell)}{\ell}\leq\sup_{\ell\geq 2}\frac{\widehat{\Upsilon}(K\otimes\Delta_\ell)}{\ell}=\widehat{\Upsilon}(K). \tag{7.88}$$

Finally, by Eq. (7.87) and Eq. (7.88), we can conclude that

$$\mathcal{M}^a_{0,\text{NS}}(\mathcal{N})=\mathcal{M}^a_{0,\text{NS}}(K)=\log\widehat{\Upsilon}(K). \tag{7.89}$$

□

A simple but useful property of $\widehat{\Upsilon}$ is shown as follows.

**Lemma 7.13.** *For any non-commutative bipartite graph $K$, we have*

$$\widehat{\Upsilon}(K\otimes\Delta_\ell)=\ell\,\widehat{\Upsilon}(K).$$

*Proof.* On one hand, it is evident from the super-multiplicativity that $\widehat{\Upsilon}(K\otimes\Delta_\ell)\geq\ell\,\widehat{\Upsilon}(K)$. On the other hand, note that an optimal solution for SDP (7.75) for $\Delta_\ell$ is given by $\{\mathbb{1}_{B'},\sum_{i=1}^{\ell}|ii\rangle\langle ii|_{A'B'}\}$, and we assume that the optimal solution of SDP (7.75) for $K$ is $\{T_B, V_{AB}\}$. It is evident that

$$V_{AB}\otimes\sum_{i=1}^{\ell}|ii\rangle\langle ii|_{A'B'}\leq\mathbb{1}_{AA'}\otimes T_B\otimes\mathbb{1}_{B'}. \tag{7.90}$$

Then, it can be checked that $\{V_{AB}\otimes\sum_{i=1}^{\ell}|ii\rangle\langle ii|_{A'B'}, T_B\otimes\mathbb{1}_{B'}\}$ is a feasible solution of SDP(7.75) for $\widehat{\Upsilon}(K\otimes\Delta_\ell)$. Therefore,$\widehat{\Upsilon}(K\otimes\Delta_\ell)\leq\operatorname{Tr}T_B\otimes\mathbb{1}_{B'}=\ell\,\widehat{\Upsilon}(K)$. □

We further discuss the activation via noisy quantum channels.

**Proposition 7.14.** *Let us consider two quantum channels $\mathcal{N}_1$ with non-commutative bipartite graphs $K_1$ and $K_2$, respectively. If $\Upsilon(K_2)-1\geq\frac{1}{\widehat{\Upsilon}(K_1)}$, then*

$$\mathcal{M}_{0,\text{NS}}(K_1\otimes K_2)-\mathcal{M}_{0,\text{NS}}(K_2)\geq\mathcal{M}^a_{0,\text{NS}}(K_1). \tag{7.91}$$

*In other words, $K_2$ can activate $K_1$ if $K_1$ is activatable. In particular, this inequality always holds when $\Upsilon(K_2)\geq 2$.*

*Proof.* Let us assume that the optimal solution to the SDP (7.74) of $\widehat{\Upsilon}(K_1)$ is $\{S_A, U_{AB}\}$ while the optimal solution to the SDP (7.15) of $\Upsilon(K_2)$ is $\{S_{A'}, U_{A'B'}\}$.

Then we can choose

$$S_{AA'} = S_A \otimes S_{A'}, \tag{7.92}$$
$$U_{AA'BB'} = U_{AB} \otimes U_{A'B'} + \bar{U}_{AB} \otimes V_{A'B'}, \tag{7.93}$$

where $V_{A'B'} = (S_{A'} \otimes \mathbb{1}_{B'} - U_{A'B'}) / (\operatorname{Tr} S_{A'} - 1)$ and $\bar{U}_{AB} = S_A / \operatorname{Tr} S_A \otimes (\mathbb{1}_B - \operatorname{Tr}_A U_{AB})$. This construction ensures that

$$\operatorname{Tr}_{AA'} U_{AA'BB'} = \mathbb{1}_{BB'}. \tag{7.94}$$

With some direct calculation, we have

$$\begin{aligned}
& S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'} && (7.95a)\\
= \; & (S_A \otimes \mathbb{1}_B - U_{AB}) \otimes U_{A'B'} + \left(\operatorname{Tr} S_{A'} - 1 - \frac{1}{\operatorname{Tr} S_A}\right) S_A \otimes \mathbb{1}_B \otimes V_{A'B'} && (7.95b)\\
& + \frac{S_A}{\operatorname{Tr} S_A} \otimes \operatorname{Tr}_A U_{AB} \otimes V_{A'B'}. && (7.95c)
\end{aligned}$$

Then, one can check that the constructed solutions satisfy

$$S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'} \geq 0. \tag{7.96a}$$

Furthermore, we have

$$\begin{aligned}
& \operatorname{Tr}(P_{AB} \otimes P_{A'B'})(S_{AA'} \otimes \mathbb{1}_{BB'} - U_{AA'BB'}) && (7.97a)\\
= \; & \operatorname{Tr} P_{AB}(S_A \otimes \mathbb{1}_B - U_{AB}) \times \operatorname{Tr} P_{A'B'} U_{A'B'} \\
& + \operatorname{Tr} P_{AB}[\left(\operatorname{Tr} S_{A'} - 1 - \frac{1}{\operatorname{Tr} S_A}\right) S_A \otimes \mathbb{1}_B] \times \operatorname{Tr} P_{A'B'} V_{A'B'} \\
& + \operatorname{Tr} P_{AB}\left(\frac{S_A}{\operatorname{Tr} S_A} \otimes \operatorname{Tr}_A U_{AB}\right) \times \operatorname{Tr} P_{A'B'} V_{A'B'} && (7.97b)\\
= \; & 0, && (7.97c)
\end{aligned}$$

where the last equality follows from $\operatorname{Tr} P_{AB}(S_A \otimes \mathbb{1}_B - U_{AB}) = 0$ and $\operatorname{Tr} P_{A'B'} V_{A'B'} = 0$.

Therefore, $\{S_{AA'}, U_{AA'BB'}\}$ is a feasible solution to the SDP (7.15) of $\Upsilon(K_1 \otimes K_2)$, which means that

$$\Upsilon(K_1 \otimes K_2) \geq \widehat{\Upsilon}(K_1)\, \Upsilon(K_2). \tag{7.98}$$

□

If we only consider using the channel $\mathcal{N}$ to activate itself, we have the following result from the above proposition.

For any quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$, if $\Upsilon(K) \geq \frac{1+\sqrt{5}}{2}$, then

$$\frac{\Upsilon(K \otimes K)}{\Upsilon(K)} \geq \widehat{\Upsilon}(K). \tag{7.99}$$

Note that $\Upsilon(K) \geq \frac{1+\sqrt{5}}{2}$ means $\Upsilon(K) - 1 - \frac{1}{\Upsilon(K)} \geq 0$. Thus the result follows directly from Proposition 7.14.

### 7.4.2 Classical-quantum channel

A classical-quantum (cq) channel $\mathcal{N} : i \to \rho_i \ (1 \leq i \leq n)$ is a CPTP map with classical inputs $\{i\}_{i=1}^n$ and quantum outputs $\{\rho_i\}_{i=1}^n$. The non-commutative bipartite graph of a cq channel will be called a cq graph. In this case, the cq graph is given by

$$K = \text{span}\{|\psi\rangle\langle i| : \ |\psi\rangle \in \text{supp}(\rho_i)\}. \tag{7.100}$$

Given a cq channel $\mathcal{N} : i \to \rho_i \ (1 \leq i \leq n)$ with cq graph $K$, its one-shot NS-assisted zero-error capacity (quantified as messages) can be simplified to

$$\begin{aligned} \Upsilon(K) = \max & \sum_i s_i \\ \text{s.t.} \ & 0 \leq s_i, \ 0 \leq R_i \leq s_i(\mathbb{1} - P_i), \\ & \sum_i (s_i P_i + R_i) = \mathbb{1}. \end{aligned} \tag{7.101}$$

where $P_i$ is the projection onto the support of $\rho_i$ for $1 \leq i \leq n$.

Moreover, it was shown in [DW16] that the asymptotic no-signalling assisted zero-error classical capacity of a cq channel is equal to the semidefinite (fractional) packing number.

**Lemma 7.15.** *(Theorem 4 in [DW16]) For any cq channel $\mathcal{N} : i \to \rho_i \ (1 \leq i \leq n)$ with cq graph $K$,*

$$C_{0,\text{NS}}(\mathcal{N}) = \log \mathsf{A}(K), \tag{7.102}$$

*with*

$$\begin{aligned} \mathsf{A}(K) = \max & \sum_i s_i \\ s.t. \ & 0 \leq s_i, \ \sum_i s_i P_i \leq \mathbb{1}. \end{aligned} \tag{7.103}$$

*where $P_i$ is the projection onto the support of $\rho_i$ for $1 \le i \le n$.*

This result is a classical-quantum generalization of the fact that the fractional packing/covering number [Sha56, SU11] of the bipartite graph (induced by the classical channel) is equal to its NS-assisted zero-error capacity [CLMW11]. Moreover, Shannon proved that the feedback-assisted zero-error capacity of a classical channel is also given by the fractional packing number [Sha56].

For any cq channel $\mathcal{N}$ with cq graph $K$, the one-shot activated capacity $\mathcal{M}^a_{0,\mathrm{NS}}(\mathcal{N}) = \log \widehat{\Upsilon}(K)$ can be simplified to

$$\begin{aligned} \widehat{\Upsilon}(K) = \max & \sum_i s_i \\ \text{s.t. } & 0 \le s_i,\ 0 \le R_i \le s_i(\mathbb{1} - P_i), \\ & \sum_i (s_i P_i + R_i) \le \mathbb{1}. \end{aligned} \tag{7.104}$$

**Theorem 7.16.** *For any classical-quantum channel $\mathcal{N}$ with cq graph $K$,*

$$M^a_{0,\mathrm{NS}}(\mathcal{N}) = \log \mathsf{A}(K). \tag{7.105}$$

*In other words, for any cq channel, the asymptotic NS-assisted zero-error capacity (or the semidefinite packing number) can be achieved via activated NS codes in the one-shot regime, i.e.,*

$$C^a_{0,\mathrm{NS}}(\mathcal{N}) = M^a_{0,\mathrm{NS}}(\mathcal{N}) = \log \mathsf{A}(K). \tag{7.106}$$

*Proof.* First, we will show $\mathsf{A}(K) \ge \widehat{\Upsilon}(K)$. Suppose that optimal solution of the SDP (7.104) of $\widehat{\Upsilon}(K)$ is $\{s_i, R_i\}$. Then,

$$\sum_i s_i P_i \le \mathbb{1} - \sum_i R_i \le \mathbb{1}, \tag{7.107}$$

which means that $\{s_i\}$ is a feasible solution for $\mathsf{A}(K)$. So we have $\mathsf{A}(K) \ge \widehat{\Upsilon}(K)$.

Second, let us assume the optimal solution of SDP (7.103) is $\{s_i\}$, let $R_i = 0$ for all $i$. It is easy to check that $\{s_i, R_i\}$ is a feasible solution of SDP (7.104), which means that $\mathsf{A}(K) \le \widehat{\Upsilon}(K)$. Therefore, for any cq graph $K$, it holds that

$$\widehat{\Upsilon}(K) = \mathsf{A}(K). \tag{7.108}$$

□

To see the existence of activation, let us consider an example here.

**Example 7.17.** We begin with the simplest possible cq channel $\mathcal{N}$, which has only two inputs and two pure output states $P_i = |\psi_i\rangle\langle\psi_i|$. Without loss of generality, we assume that $|\psi_0\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|\psi_1\rangle = \alpha|0\rangle - \beta|1\rangle$ with $\alpha \geq \beta = \sqrt{1-\alpha^2}$. In [DW16], it has been shown that $\Upsilon(K) = 1$ and $\mathsf{A}(K) = \frac{1}{\alpha^2}$. Hence, by Theorem 7.16, we know

$$\widehat{\Upsilon}(K) = \frac{\Upsilon(\mathcal{N} \otimes \Delta_2)}{2} = \frac{1}{\alpha^2} > \Upsilon(K) = 1. \tag{7.109}$$

Furthermore, we have

$$C^a_{0,\mathrm{NS}}(\mathcal{N}) = M^a_{0,\mathrm{NS}}(\mathcal{N}) = -2\log\alpha > M_{0,\mathrm{NS}}(\mathcal{N}) = 0. \tag{7.110}$$

### 7.4.3 Asymptotic zero-error capacity

As we find the activation phenomenon of zero-error communication in the one-shot regime, it's natural to wonder whether there exists an activation in the asymptotic regime. In the following theorem, we prove that the answer is negative.

**Theorem 7.18.** *For any quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$ with positive zero-error capacity, let $n_0$ be the smallest integer such that $\Upsilon(K^{\otimes n_0}) \geq 2$. Note that $n_0$ always exists and depends only on $K$. Then for any $n \geq n_0$, we have*

$$2\widehat{\Upsilon}\left(K^{\otimes(n-n_0)}\right) \leq \Upsilon\left(K^{\otimes n}\right) \leq \widehat{\Upsilon}\left(K^{\otimes n}\right). \tag{7.111}$$

*Moreover,*

$$C^a_{0,\mathrm{NS}}(K) = \sup_{n\geq 1}\log\sqrt[n]{\widehat{\Upsilon}(K^{\otimes n})} = \lim_{n\to\infty}\log\sqrt[n]{\widehat{\Upsilon}(K^{\otimes n})} = C_{0,\mathrm{NS}}(K). \tag{7.112}$$

*Proof.* On one hand, from Eq. (7.98) in Proposition 7.14, we have

$$\Upsilon\left(K^{\otimes n}\right) = \Upsilon\left(K^{\otimes(n-n_0)} \otimes K^{n_0}\right) \geq \widehat{\Upsilon}\left(K^{\otimes(n-n_0)}\right)\Upsilon\left(K^{\otimes n_0}\right) \geq 2\widehat{\Upsilon}\left(K^{\otimes(n-n_0)}\right). \tag{7.113}$$

On the other hand, it always holds that $\Upsilon(K^{\otimes n}) \leq \widehat{\Upsilon}(K^{\otimes n})$. Therefore, we obtain Eq. (7.111).

Then,

$$\lim_{n\to\infty}\log\sqrt[n]{\widehat{\Upsilon}(K^{\otimes n})} = \lim_{n\to\infty}\log\sqrt[n]{\Upsilon(K^{\otimes n})}. \tag{7.114}$$

To prove Eq. (7.112), the technique is based on a lemma about the existence of limits in [Fek23]. On one hand, $\log\widehat{\Upsilon}(K^{\otimes n}) \leq 2n\log d$. On the other hand, since $\widehat{\Upsilon}(K)$

is super-multiplicative, then $\log \widehat{\Upsilon}\left(K^{\otimes(mn)}\right) \geq \log \widehat{\Upsilon}\left(K^{\otimes m}\right) + \log \widehat{\Upsilon}\left(K^{\otimes n}\right)$. Therefore,

$$\sup_{n\geq 1} \frac{\log \widehat{\Upsilon}\left(K^{\otimes n}\right)}{n} = \lim_{n\to\infty} \frac{\log \widehat{\Upsilon}\left(K^{\otimes n}\right)}{n} = C_{0,\mathrm{NS}}\left(K\right). \tag{7.115}$$

□

### 7.4.4 Separating $C_{0,NS}$ and semidefinite packing nubmber

As the NS-assisted zero-error capacity of cq channel is given by the semidefinite (or fractional) packing number $\mathsf{A}\left(K\right)$, an interesting question is whether this result also holds for general quantum channels. The semidefinite packing number for a general quantum channel was also introduced in [DW16] as follows:

$$\begin{aligned} \mathsf{A}\left(\mathcal{N}\right) = \mathsf{A}\left(K\right) = \max \operatorname{Tr} S_A& \\ \text{s.t. } 0 \leq S_A, \operatorname{Tr}_A P_{AB}\left(S_A \otimes \mathbb{1}_B\right) \leq \mathbb{1}_B.& \end{aligned} \tag{7.116}$$

To study whether $C_{0,NS}$ equals to $\log \mathsf{A}\left(\mathcal{N}\right)$, the difficulty is that we currently do not know efficient methods to calculate the asymptotic no-signalling zero-error capacity.

In the following, we will exhibit an example to show that $C_{0,NS}$ is not equal to the semidefinite packing number for general quantum channels.

**Proposition 7.19.** *There exists a quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$ such that $\widehat{\Upsilon}\left(K\right) > \mathsf{A}\left(K\right)$. Consequently,*

$$C_{0,\mathrm{NS}}\left(\mathcal{N}\right) \neq \log \mathsf{A}\left(\mathcal{N}\right). \tag{7.117}$$

*Proof.* Let $K$ correspond to the quantum channel $\mathcal{N}\left(\rho\right) = \sum_{i=0}^{2} E_i \rho E_i^\dagger$ with $E_0 = \frac{1}{\sqrt{2}}|0\rangle\langle 0| + \frac{1}{\sqrt{2}}|2\rangle\langle 0|$, $E_1 = \sqrt{\frac{50}{99}}|0\rangle\langle 2| + \sqrt{\frac{1}{99}}|1\rangle\langle 1| + \sqrt{\frac{49}{99}}|2\rangle\langle 2|$ and $E_2 = \sqrt{\frac{98}{99}}|0\rangle\langle 1|$.

By solving SDPs numerically [GB08], we find that

$$\widehat{\Upsilon}\left(\mathcal{N}\right) \approx 1.1767 > 1.1751 > \mathsf{A}\left(\mathcal{N}\right). \tag{7.118}$$

Then, it leads to

$$C_{0,\mathrm{NS}}\left(\mathcal{N}\right) \geq \mathcal{M}_{0,\mathrm{NS}}\left(\mathcal{N}\right) > \log \mathsf{A}\left(\mathcal{N}\right). \tag{7.119}$$

□

## 7.5 Discussion

### 7.5.1 Summary

In this chapter, we investigated the quantum zero-error information theory from several aspects. In particular, we have shown that there is a separation between the quantum Lovász number and the entanglement-assisted zero-error classical capacity.

An overview of the results in this chapter is summarized in the following box.

**Summary of Chapter 7**

(i) An upper bound on independence number:

$$\alpha(\mathcal{N}) \leq \kappa(\mathcal{N}) = \min_{\sigma} \max_{\rho} \frac{1}{\operatorname{Tr} \sigma P_{\mathcal{N}(\rho)}}, \tag{7.120}$$

where $P_{\mathcal{N}(\rho)}$ is the projection onto the support of $\mathcal{N}(\rho)$.

(ii) Separation between quantum Lovász number and $C_{0E}$: $\exists$ non-commutative graph $S$ such that

$$C_{0E}(S) < \log \widetilde{\vartheta}(S). \tag{7.121}$$

(iii) Activated NS-assisted zero-error capacity: for any quantum channel $\mathcal{N}$ with non-commutative bipartite graph $K$,

$$\begin{aligned} \mathcal{M}^a_{0,\mathrm{NS}}(\mathcal{N}) = \log \widehat{\Upsilon}(K) = \max \operatorname{Tr} S_A & \\ \text{s.t. } 0 \leq U_{AB} \leq S_A \otimes \mathbb{1}_B, \operatorname{Tr}_A U_{AB} \leq \mathbb{1}_B, & \\ \operatorname{Tr} P_{AB}(S_A \otimes \mathbb{1}_B - U_{AB}) = 0. & \end{aligned} \tag{7.122}$$

(iv) The one-shot NS-assisted simulation cost of a general non-commutative bipartite graph is not multiplicative.

### 7.5.2 Outlook

Interestingly, for the channel $\mathcal{N}_\alpha$, the quantum fractional packing number is equal to the quantum Lovász number. Let us recall that remarkable fact that the Lovász number of a classical graph $G$ has an operational interpretation [DW16] as

$$\vartheta(G) = \min\{\mathsf{A}(K) : K^\dagger K < S_G\},$$

where the minimization is over classical-quantum graphs $K$ and $S_G$ is non-commutative graph associated with $G$. A natural and interesting question is that for the non-commutative graph $S$, do we have

$$\widetilde{\vartheta}(S) = \min\{\mathsf{A}(K) : K^\dagger K < S\}?$$

The non-commutative bipartite graph of $\mathcal{N}_\alpha$ might be such an interesting example since Proposition 7.6 and Lemma 7.8 imply that $\widetilde{\vartheta}(\mathcal{N}_\alpha) = \mathsf{A}(K_\alpha)$.

The classical zero-error information theory concerns asymptotic combinatorial problems, most of which are difficult and unsolved. It remains unknown whether Lovász number coincides with $C_{0E}$ for every classical channel. For confusability graph $G$, a variant of Lovász number called Schrijver number [Sch79, MRR78] was proved to be a tighter upper bound on the entanglement-assisted independence number than Lovász number [CMR+14] . However, it remains unknown whether Schrijver number will converge to Lovász number in the asymptotic setting, and a gap between the regularized Schrijver number and Lovász number would imply a separation between $C_{0E}(G)$ and $\vartheta(G)$. Moreover, it is also interesting to study how to estimate the regularization of a sequence of semidefinite programs.

Finally, we end this chapter with a table of the known and open problems in quantum zero-error information theory.

| | Classical | Classical-Quantum | Quantum |
|---|---|---|---|
| $C_0$ | ? | ? | ? |
| $C_{0,E}$ | ? | ? | ? |
| $C_{0,NS}$ | [CLMW11] | [DW16] | ? |
| $C_{0,E} < \log \widetilde{\vartheta}$ ? | ? | ? | [WD18] |

Table 7.2: Zero-error capacities of different classes of channels

# Bibliography


[ABDR03] A. Ambainis, H. Buhrman, Y. Dodis, and H. Rohrig. Multiparty quantum coin flipping. In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 250–259. IEEE, (2003).

[ADR$^+$17] A. Acín, R. Duan, D. E. Roberson, A. B. Sainz, and A. Winter. A new property of the Lovász number and duality relations between graph parameters. *Discrete Applied Mathematics* **216**: 489–501 (2017).

[ADVW02] K. Audenaert, B. De Moor, K. G. H. Vollbrecht, and R. F. Werner. Asymptotic relative entropy of entanglement for orthogonally invariant states. *Physical Review A* **66**(3): 032310 (2002).

[AEJ$^+$01] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor. Asymptotic Relative Entropy of Entanglement. *Physical Review Letters* **87**(21): 217902 (2001).

[AFB17] R. Arnon-Friedman and J.-D. Bancal. Device-independent Certification of One-shot Distillable Entanglement. *arXiv preprint arXiv:1712.09369* , (2017).

[AFY17] R. Arnon-Friedman and H. Yuen. Noise-tolerant testing of high entanglement of formation. *arXiv preprint arXiv:1712.09368* , (2017).

[AHW00] G. G. Amosov, A. S. Holevo, and R. F. Werner. On Some Additivity Problems in Quantum Information Theory. *Problemy Peredachi Informatsii* **36**(4): 25–34, (2000).

[Alo98] N. Alon. The Shannon Capacity of a Union. *Combinatorica* **18**(3): 301–310 (1998).

[AMHR13] A. H. Al-Mohy, N. J. Higham, and S. D. Relton. Computing the Fréchet Derivative of the Matrix Logarithm and Estimating the Condition Number. *SIAM Journal on Scientific Computing* **35**(4): C394–C410 (2013).

[APE03] K. Audenaert, M. B. Plenio, and J. Eisert. Entanglement Cost under Positive-Partial-Transpose-Preserving Operations. *Physical Review Letters* **90**(2): 027901 (2003).

[Aud07] K. M. R. Audenaert. A sharp continuity estimate for the von Neumann entropy. *Journal of Physics A: Mathematical and Theoretical* **40**(28): 8127–8136 (2007).

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *International Conference on Computers, Systems & Signal Processing, Bangalore, India, Dec 9-12, 1984*, pages 175–179, (1984).

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters* **70**(13): 1895–1899 (1993).

[BBCW13] M. Berta, F. G. S. L. Brandao, M. Christandl, and S. Wehner. Entanglement Cost of Quantum Channels. *IEEE Transactions on Information Theory* **59**(10): 6779–6795 (2013).

[BBL+15] J. Briet, H. Buhrman, M. Laurent, T. Piovesan, and G. Scarpa. Entanglement-Assisted Zero-Error Source-Channel Coding. *IEEE Transactions on Information Theory* **61**(2): 1124–1138 (2015).

[BBPS96] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Physical Review A* **53**(4): 2046–2052 (1996).

[BCHW15] S. Bäuml, M. Christandl, K. Horodecki, and A. Winter. Limitations on quantum key repeaters. *Nature Communications* **6**: 6908 (2015).

[BCR11] M. Berta, M. Christandl, and R. Renner. The Quantum Reverse Shannon Theorem Based on One-Shot Information Theory. *Communications in Mathematical Physics* **306**(3): 579–615 (2011).

[BD10] F. Buscemi and N. Datta. The Quantum Capacity of Channels With Arbitrarily Correlated Noise. *IEEE Transactions on Information Theory* **56**(3): 1447–1460 (2010).

[BD11a] F. G. S. L. Brandao and N. Datta. One-Shot Rates for Entanglement Manipulation Under Non-entangling Maps. *IEEE Transactions on Information Theory* **57**(3): 1754–1760 (2011).

[BD11b] F. Buscemi and N. Datta. Entanglement Cost in Practical Scenarios. *Physical Review Letters* **106**(13): 130503 (2011).

[BDE+98] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin. Optimal universal and state-dependent quantum cloning. *Physical Review A* **57**(4): 2368–2378 (1998).

[BDF+99] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters. Quantum nonlocality without entanglement. *Physical Review A* **59**(2): 1070–1091 (1999).

[BDH+14] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter. The Quantum Reverse Shannon Theorem and Resource Tradeoffs for Simulating Quantum Channels. *IEEE Transactions on Information Theory* **60**(5): 2926–2959 (2014).

[BDL16] S. Beigi, N. Datta, and F. Leditzky. Decoding quantum information via the Petz recovery map. *Journal of Mathematical Physics* **57**(8): 082203 (2016).

[BDS97] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin. Capacities of Quantum Erasure Channels. *Physical Review Letters* **78**(16): 3217–3220 (1997).

[BDSW96] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A* **54**(5): 3824–3851 (1996).

[BEHY11] F. G. S. L. Brandão, J. Eisert, M. Horodecki, and D. Yang. Entangled Inputs Cannot Make Imperfect Quantum Channels Perfect. *Physical Review Letters* **106**(23): 230502 (2011).

[Bei10] S. Beigi. Entanglement-assisted zero-error capacity is upper-bounded by the Lovász theta function. *Physical Review A* **82**(1): 010303 (2010).

[Bel64] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika* **1**(3): 195–200 (1964).

[BGNP01] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill. Causal and localizable quantum operations. *Physical Review A* **64**(5): 052309 (2001).

[Bha09] R. Bhatia. *Positive definite matrices*. Princeton university press (2009).

[BKN00] H. Barnum, E. Knill, and M. Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory* **46**(4): 1317–1329 (2000).

[BLMW04] B. B. Blinov, D. Leibfried, C. Monroe, and D. J. Wineland. Quantum Computing with Trapped Ion Hyperfine Qubits. *Quantum Information Processing* **3**(1-5): 45–59 (2004).

[BNS98] H. Barnum, M. A. Nielsen, and B. Schumacher. Information transmission through a noisy quantum channel. *Physical Review A* **57**(6): 4153–4175 (1998).

[Boh48] F. Bohnenblust. Joint positiveness of matrices. *Technical report, UCLA* , (1948).

[BP08] F. G. S. L. Brandão and M. B. Plenio. Entanglement theory and the second law of thermodynamics. *Nature Physics* **4**(11): 873–877 (2008).

[BP10] F. G. S. L. Brandão and M. B. Plenio. A Reversible Theory of Entanglement and its Relation to the Second Law. *Communications in Mathematical Physics* **295**(3): 829–851 (2010).

[BS98] C. Bennett and P. Shor. Quantum information theory. *IEEE Transactions on Information Theory* **44**(6): 2724–2742 (1998).

[BS07] S. Beigi and P. W. Shor. On the complexity of computing zero-error and Holevo capacity of quantum channels. *arXiv:0709.2090* , (2007).

[BSST99] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. Entanglement-Assisted Classical Capacity of Noisy Quantum Channels. *Physical Review Letters* **83**(15): 3081–3084 (1999).

[BSST02] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory* **48**(10): 2637–2655 (2002).

[Bur69] D. Bures. An Extension of Kakutani's Theorem on Infinite Product Measures to the Tensor Product of Semifinite w -Algebras. *Transactions of the American Mathematical Society* **135**: 199 (1969).

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press (2004).

[BW92] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters* **69**(20): 2881–2884 (1992).

[BW18] M. Berta and M. M. Wilde. Amortization does not enhance the max-Rains information of a quantum channel. *New Journal of Physics* **20**(5): 053044 (2018).

[BŻ17] I. Bengtsson and K. Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press (2017).

[CA97] N. J. Cerf and C. Adami. Negative Entropy and Information in Quantum Mechanics. *Physical Review Letters* **79**(26): 5194–5197 (1997).

[CCH11] T. S. Cubitt, J. Chen, and A. W. Harrow. Superactivation of the Asymptotic Zero-Error Classical Capacity of a Quantum Channel. *IEEE Transactions on Information Theory* **57**(12): 8114–8126 (2011).

[CD09] E. Chitambar and R. Duan. Nonlocal Entanglement Transformations Achievable by Separable Operations. *Physical Review Letters* **103**(11): 110502 (2009).

[CdOF11] M. F. Cornelio, M. C. de Oliveira, and F. F. Fanchini. Entanglement Irreversibility from Quantum Discord and Quantum Deficit. *Physical Review Letters* **107**(2): 020502 (2011).

[CDP08] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Transforming quantum operations: Quantum supermaps. *EPL (Europhysics Letters)* **83**(3): 30004 (2008).

[CEM$^+$15] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk. Unbounded number of channel uses may be required to detect quantum capacity. *Nature Communications* **6**(1): 6739 (2015).

[CG18] E. Chitambar and G. Gour. Quantum Resource Theories. *arXiv:1806.06107* , (2018).

[CGLM14] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini. Quantum channels and memory effects. *Reviews of Modern Physics* **86**(4): 1203–1259 (2014).

[Cho75] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications* **10**(3): 285–290 (1975).

[Chr06] M. Christandl. The structure of bipartite quantum states-Insights from group theory and cryptography. *arXiv preprint quant-ph/0604183* , (2006).

[CLM$^+$14] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter. Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask). *Communications in Mathematical Physics* **328**(1): 303–326 (2014).

[CLMW10] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Improving Zero-Error Classical Communication with Entanglement. *Physical Review Letters* **104**(23): 230503 (2010).

[CLMW11] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter. Zero-Error Channel Capacity and Simulation Assisted by Non-Local Correlations. *IEEE Transactions on Information Theory* **57**(8): 5509–5523 (2011).

[CMH17] M. Christandl and A. Müller-Hermes. Relative Entropy Bounds on Quantum, Private and Repeater Capacities. *Communications in Mathematical Physics* **353**(2): 821–852 (2017).

[CMR$^+$14] T. Cubitt, L. Mancinska, D. E. Roberson, S. Severini, D. Stahlke, and A. Winter. Bounds on Entanglement-Assisted Source-Channel Coding via the Lovász Number and Its Variants. *IEEE Transactions on Information Theory* **60**(11): 7330–7344 (2014).

[CS12] T. S. Cubitt and G. Smith. An Extreme Form of Superactivation for Quantum Zero-Error Capacities. *IEEE Transactions on Information Theory* **58**(3): 1953–1961 (2012).

[CS17] D. Cavalcanti and P. Skrzypczyk. Quantum steering: a review with focus on semidefinite programming. *Reports on Progress in Physics* **80**(2): 024001 (2017).

[CSUU08] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Perfect Parallel Repetition Theorem for Quantum Xor Proof Systems. *computational complexity* **17**(2): 282–299 (2008).

[CW04] M. Christandl and A. Winter. Squashed entanglement: An additive entanglement measure. *Journal of Mathematical Physics* **45**(3): 829–840 (2004).

[CWY04] N. Cai, A. Winter, and R. W. Yeung. Quantum privacy and quantum wiretap channels. *Problems of Information Transmission* **40**(4): 318–336 (2004).

[Dat09] N. Datta. Min- and Max-Relative Entropies and a New Entanglement Monotone. *IEEE Transactions on Information Theory* **55**(6): 2816–2826 (2009).

[DBW17] S. Das, S. Bäuml, and M. M. Wilde. Entanglement and secret-key-agreement capacities of bipartite quantum interactions and read-only memory devices. *arXiv preprint arXiv:1712.00827* , (2017).

[Dev05] I. Devetak. The Private Classical Capacity and Quantum Capacity of a Quantum Channel. *IEEE Transactions on Information Theory* **51**(1): 44–55 (2005).

[DFJY05] R. Duan, Y. Feng, Z. Ji, and M. Ying. Efficiency of deterministic entanglement transformation. *Physical Review A* **71**(2): 022305 (2005).

[DFW$^+$18] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter. Using and reusing coherence to realize quantum processes. *arXiv:1805.04045* , (2018).

[DH13] N. Datta and M.-H. Hsieh. One-Shot Entanglement-Assisted Quantum and Classical Communication. *IEEE Transactions on Information Theory* **59**(3): 1929–1939 (2013).

[DHS04] N. Datta, A. S. Holevo, and Y. Suhov. Additivity for transpose depolarizing channels. *International Journal of Quantum Information* **04**(01): 85–98 (2004).

[DKF$^+$12] F. Dupuis, L. Kraemer, P. Faist, J. M. Renes, and R. Renner. Generalized Entropies. *XVIIth International Congress on Mathematical Physics* pages 134–153 (2012).

[DS05] I. Devetak and P. W. Shor. The Capacity of a Quantum Channel for Simultaneous Transmission of Classical and Quantum Information. *Communications in Mathematical Physics* **256**(2): 287–303 (2005).

[DS08] R. Duan and Y. Shi. Entanglement between Two Uses of a Noisy Multipartite Quantum Channel Enables Perfect Transmission of Classical Information. *Physical Review Letters* **101**(2): 020501 (2008).

[DSS98] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin. Quantum-channel capacity of very noisy channels. *Physical Review A* **57**(2): 830–839 (1998).

[DSW13] R. Duan, S. Severini, and A. Winter. Zero-Error Communication via Quantum Channels, Noncommutative Graphs, and a Quantum Lovász Number. *IEEE Transactions on Information Theory* **59**(2): 1164–1174 (2013).

[DSW16] R. Duan, S. Severini, and A. Winter. On Zero-Error Communication via Quantum Channels in the Presence of Noiseless Feedback. *IEEE Transactions on Information Theory* **62**(9): 5260–5277 (2016).

[DTW16] N. Datta, M. Tomamichel, and M. M. Wilde. On the second-order asymptotics for entanglement-assisted communication. *Quantum Information Processing* **15**(6): 2569–2591 (2016).

[Dua09] R. Duan. Super-activation of zero-error capacity of noisy quantum channels. *arXiv preprint arXiv:0906.2527* , (2009).

[DW03] I. Devetak and A. Winter. Distillation of secret key and entanglement from quantum states. **461**(2053): 207–235 (2003).

[DW15] R. Duan and X. Wang. Activated zero-error classical communication over quantum channels assisted with quantum no-signalling correlations. *arXiv:1510.05437* , (2015).

[DW16] R. Duan and A. Winter. No-Signalling-Assisted Zero-Error Capacity of Quantum Channels and an Information Theoretic Interpretation of the Lovász Number. *IEEE Transactions on Information Theory* **62**(2): 891–914 (2016).

[Eke91] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters* **67**(6): 661–663 (1991).

[Eld03] Y. Eldar. A semidefinite programming approach to optimal unambiguous discrimination of quantum states. *IEEE Transactions on Information Theory* **49**(2): 446–456 (2003).

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review* **47**(10): 777–780 (1935).

[ES15] D. Elkouss and S. Strelchuk. Superadditivity of Private Information for Any Number of Uses of the Channel. *Physical Review Letters* **115**(4): 040501 (2015).

[ESW02] T. Eggeling, D. Schlingemann, and R. F. Werner. Semicausal operations are semilocalizable. *Europhysics Letters (EPL)* **57**(6): 782–788 (2002).

[EVWW01] T. Eggeling, K. Vollbrecht, R. Werner, and M. Wolf. Distillability via Protocols Respecting the Positivity of Partial Transpose. *Physical Review Letters* **87**(25): 257902 (2001).

[Fan73] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Communications in Mathematical Physics* **31**(4): 291–294 (1973).

[Fek23] M. Fekete. Uber die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Mathematische Zeitschrift* **17**(1): 228–249 (1923).

[FF18] H. Fawzi and O. Fawzi. Efficient optimization of the quantum relative entropy. *Journal of Physics A: Mathematical and Theoretical* **51**(15): 154003 (2018).

[FG11] S. Friedland and G. Gour. An explicit expression for the relative entropy of entanglement in all dimensions. *Journal of Mathematical Physics* **52**(5): 052201 (2011).

[FH04] W. Fulton and J. Harris. *Representation Theory*. Volume 129 of Graduate Texts in Mathematics, Springer New York (2004).

[FSW07] A. S. Fletcher, P. W. Shor, and M. Z. Win. Optimum quantum error recovery using semidefinite programming. *Physical Review A* **75**(1): 012338 (2007).

[Fuk05] M. Fukuda. Extending additivity from symmetric to asymmetric channels. *Journal of Physics A: Mathematical and General* **38**(45): L753–L758 (2005).

[FvdG99] C. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory* **45**(4): 1216–1227 (1999).

[FW08] J. Fern and K. B. Whaley. Lower bounds on the nonzero capacity of Pauli channels. *Physical Review A* **78**(6): 062335 (2008).

[FWL$^+$18] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso. Probabilistic distillation of quantum coherence. *Physical Review Letters* (in press, 2018).

[FWTB18] K. Fang, X. Wang, M. Tomamichel, and M. Berta. Quantum Channel Simulation and the Channel's Max-Information. In *2018 IEEE International Symposium on Information Theory (ISIT)*, (2018).

[FWTD17] K. Fang, X. Wang, M. Tomamichel, and R. Duan. Non-asymptotic entanglement distillation. *arXiv:1706.06221*, (2017).

[GA15] R. Gallego and L. Aolita. Resource Theory of Steering. *Physical Review X* **5**(4): 41008, (2015).

[GB08] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, (2008).
Online: `http://cvxr.com/cvx`.

[GF05] V. Giovannetti and R. Fazio. Information-capacity description of spin-chain correlations. *Physical Review A* **71**(3): 032314 (2005).

[GGF14] M. W. Girard, G. Gour, and S. Friedland. On convex optimization problems in quantum information theory. *Journal of Physics A: Mathematical and Theoretical* **47**(50): 505302 (2014).

[GJL15] L. Gao, M. Junge, and N. LaRacuente. Capacity Bounds via Operator Space Methods. *arXiv:1509.07294* , (2015).

[GKRdW09] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-Error Quantum State Identification and Exponential Separations in Communication Complexity. *SIAM Journal on Computing* **39**(1): 1–24 (2009).

[GLN05] A. Gilchrist, N. K. Langford, and M. A. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A* **71**(6): 062310 (2005).

[GLS93] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Volume 2 of Algorithms and Combinatorics, Springer Berlin Heidelberg (1993).

[GMN$^+$15] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports* **583**: 1–58 (2015).

[GRTZ02] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Reviews of Modern Physics* **74**(1): 145–195 (2002).

[GW95] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM* **42**(6): 1115–1145 (1995).

[GZFG15] M. W. Girard, Y. Zinchenko, S. Friedland, and G. Gour. Erratum: Numerical estimation of the relative entropy of entanglement [Phys. Rev. A 82 , 052336 (2010)]. *Physical Review A* **91**(2): 029901 (2015).

[Har05] A. W. Harrow. *Applications of coherent classical communication and the Schur transform to quantum information theory*. PhD thesis, (2005).

[Has09] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics* **5**(4): 255–257 (2009).

[Hay06] M. Hayashi. *Quantum Information*. Springer (2006).

[Hay09] M. Hayashi. Information Spectrum Approach to Second-Order Coding Rate in Channel Coding. *IEEE Transactions on Information Theory* **55**(11): 4947–4966 (2009).

[Hay17a] M. Hayashi. *A Group Theoretic Approach to Quantum Information*. Springer International Publishing (2017).

[Hay17b] M. Hayashi. *Group Representation for Quantum Theory*. Springer International Publishing (2017).

[Hay17c] M. Hayashi. *Quantum Information Theory*. Springer Berlin Heidelberg (2017).

[Hay17d] M. Hayashi. Role of Hypothesis Testing in Quantum Information. *arXiv preprint arXiv:1709.07701* , (2017).

[Hel76] C. W. Helstrom. *Quantum detection and estimation theory*. Volume 123, Academic press (1976).

[HHH96] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A* **223**(1-2): 1–8 (1996).

[HHH98] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-State Entanglement and Distillation: Is there a Bound Entanglement in Nature? *Physical Review Letters* **80**(24): 5239–5242 (1998).

[HHHH09] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Reviews of Modern Physics* **81**(2): 865–942 (2009).

[HHHO05] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. Secure Key from Bound Entanglement. *Physical Review Letters* **94**(16): 160502 (2005).

[HHT01] P. M. Hayden, M. Horodecki, and B. M. Terhal. The asymptotic entanglement cost of preparing a quantum state. *Journal of Physics A: Mathematical and General* **34**(35): 6891–6898 (2001).

[HJ12] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press (2012).

[HLS07] P. Hoyer, T. Lee, and R. Spalek. Negative weights make adversaries stronger. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing - STOC '07*, number 015848, page 526, New York, New York, USA, (2007). ACM Press.

[HNW17] A. W. Harrow, A. Natarajan, and X. Wu. An Improved Semidefinite Programming Hierarchy for Testing Entanglement. *Communications in Mathematical Physics* **352**(3): 881–904 (2017).

[HOH02] M. Horodecki, J. Oppenheim, and R. Horodecki. Are the Laws of Entanglement Theory Thermodynamical? *Physical Review Letters* **89**(24): 240403 (2002).

[Hol73] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii (Problems of Information Transmission)* **9**(3): 3–11, (1973).

[Hol98a] A. S. Holevo. Quantum coding theorems. *Russian Mathematical Surveys* **53**(6): 1295–1331 (1998).

[Hol98b] A. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory* **44**(1): 269–273 (1998).

[HOW05] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature* **436**(7051): 673–676 (2005).

[HP91] F. Hiai and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics* **143**(1): 99–114 (1991).

[HW01] A. Holevo and R. Werner. Evaluating capacities of bosonic Gaussian channels. *Physical Review A* **63**(3): 032312 (2001).

[Jam72] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics* **3**(4): 275–278 (1972).

[JJUW11] R. Jain, Z. Ji, S. Upadhyay, and J. Watrous. QIP = PSPACE. *Journal of the ACM* **58**(6): 1–27 (2011).

[Kin02] C. King. Additivity for unital qubit channels. *Journal of Mathematical Physics* **43**(10): 4641 (2002).

[Kin03] C. King. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory* **49**(1): 221–229 (2003).

[Kit97] A. Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys* **52**(6): 1191–1249 (1997).

[KO98] J. Korner and A. Orlitsky. Zero-error information theory. *IEEE Transactions on Information Theory* **44**(6): 2207–2229 (1998).

[Kra71] K. Kraus. General state changes in quantum theory. *Annals of Physics* **64**(2): 311–335 (1971).

[KRS09] R. Konig, R. Renner, and C. Schaffner. The Operational Meaning of Min- and Max-Entropy. *IEEE Transactions on Information Theory* **55**(9): 4337–4347 (2009).

[KRT10] J. Kempe, O. Regev, and B. Toner. Unique Games with Entangled Provers Are Easy. *SIAM Journal on Computing* **39**(7): 3207–3229 (2010).

[KSC$^+$15] I. Kogias, P. Skrzypczyk, D. Cavalcanti, A. Acín, and G. Adesso. Hierarchy of Steering Criteria Based on Moments for All Bipartite Quantum Systems. *Physical Review Letters* **115**(21): 210401 (2015).

[KSL08] R. L. Kosut, A. Shabani, and D. A. Lidar. Robust Quantum Error Correction via Convex Optimization. *Physical Review Letters* **100**(2): 020502 (2008).

[KW00] A. Kitaev and J. Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof systems. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing - STOC '00*, pages 608–617, New York, New York, USA, (2000). ACM Press.

[KW04] D. Kretschmann and R. F. Werner. Tema con variazioni: Quantum channel capacity. *New Journal of Physics* **6** (2004).

[KW09] R. König and S. Wehner. A Strong Converse for Classical Channel Coding Using Entangled Inputs. *Physical Review Letters* **103**(7): 070504 (2009).

[KWW12] R. Konig, S. Wehner, and J. Wullschleger. Unconditional Security From Noisy Quantum Storage. *IEEE Transactions on Information Theory* **58**(3): 1962–1984 (2012).

[Lan96] R. Landauer. The physical nature of information. *Physics Letters A* **217**(4-5): 188–193 (1996).

[LDS17] F. Leditzky, N. Datta, and G. Smith. Useful states and entanglement distillation. *IEEE Transactions on Information Theory* pages 1–1 (2017).

[Li14] K. Li. Second-order asymptotics for quantum hypothesis testing. *The Annals of Statistics* **42**(1): 171–189 (2014).

[Lin75] G. Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics* **40**(2): 147–151 (1975).

[LKDW18] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde. Approaches for approximate additivity of the Holevo information of quantum channels. *Physical Review A* **97**(1): 012332 (2018).

[Llo97] S. Lloyd. Capacity of the noisy quantum channel. *Physical Review A* **55**(3): 1613–1622 (1997).

[LLS18] F. Leditzky, D. Leung, and G. Smith. Quantum and Private Capacities of Low-Noise Channels. *Physical Review Letters* **120**(16): 160503 (2018).

[LLSS14] D. Leung, K. Li, G. Smith, and J. A. Smolin. Maximal Privacy without Coherence. *Physical Review Letters* **113**(3): 030502 (2014).

[LM15] D. Leung and W. Matthews. On the Power of PPT-Preserving and Non-Signalling Codes. *IEEE Transactions on Information Theory* **61**(8): 4486–4499 (2015).

[LMM$^+$12] D. Leung, L. Mancinska, W. Matthews, M. Ozols, and A. Roy. Entanglement can increase asymptotic rates of zero-error classical communication over classical channels. *Communications in Mathematical Physics* **311**(1): 97–111 (2012).

[LMR$^+$11] T. Lee, R. Mittal, B. W. Reichardt, R. Palek, and M. Szegedy. Quantum Query Complexity of State Conversion. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 344–353. IEEE, (2011).

[Lov79] L. Lovasz. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory* **25**(1): 1–7 (1979).

[LQWD18] Y. Li, Y. Qiao, X. Wang, and R. Duan. Tripartite-to-Bipartite Entanglement Transformation by Stochastic Local Operations and Classical Communication and the Structure of Matrix Spaces. *Communications in Mathematical Physics* **358**(2): 791–814 (2018).

[LRW$^+$18] L. Lami, B. Regula, X. Wang, R. Nichols, A. Winter, and G. Adesso. Gaussian quantum resource theories. *arXiv preprint arXiv:1801.05450* , (2018).

[LV16] M. Laurent and F. Vallentin. Semidefinite optimization. *Lecture notes* , (2016).

[LW17] D. Leung and J. Watrous. On the complementary quantum capacity of the depolarizing channel. *Quantum* **1**: 28 (2017).

[LWD17] Y. Li, X. Wang, and R. Duan. Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario. *Physical Review A* **95**(5): 052346 (2017).

[LWZ$^+$17] S. Liu, X. Wang, L. Zhou, J. Guan, Y. Li, Y. He, R. Duan, and M. Ying. QSI: A Quantum Programming Environment. *arXiv:1710.09500* , (2017).

[LY16] D. Leung and N. Yu. Maximum privacy without coherence, zero-error. *Journal of Mathematical Physics* **57**(9): 092202 (2016).

[LY17] Y. Li and M. Ying. Algorithmic analysis of termination problems for quantum programs. *Proceedings of the ACM on Programming Languages* **2**(POPL): 1–29 (2017).

[Mat12] W. Matthews. A Linear Program for the Finite Block Length Converse of Polyanskiy-Poor-Verdú Via Nonsignaling Codes. *IEEE Transactions on Information Theory* **58**(12): 7036–7044 (2012).

[Mat13] MathWorks. Matrix logarithm function in Matlab, (2013).
Online: `http://au.mathworks.com/help/matlab/ref/logm.html`.

[MD09] M. Mosonyi and N. Datta. Generalized relative entropies and the capacity of classical-quantum channels. *Journal of Mathematical Physics* **50**(7): 072104 (2009).

[MHRW16] A. Müller-Hermes, D. Reeb, and M. M. Wolf. Positivity of linear maps under tensor powers. *Journal of Mathematical Physics* **57**(1): 015202 (2016).

[MI08] A. Miranowicz and S. Ishizaka. Closed formula for the relative entropy of entanglement. *Physical Review A* **78**(3): 032310 (2008).

[MLDS+13] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. On quantum Rényi entropies: a new generalization and some properties. *Journal of Mathematical Physics* **54**(12): 122203 (2013).

[MRR78] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey Jr. The Lovász bound and some generalizations. *J. Combin. Inform. System Sci* **3**(3): 134–152, (1978).

[MSS13] L. Mancinska, G. Scarpa, and S. Severini. New Separations in Zero-Error Channel Capacity Through Projective KochenSpecker Sets and Quantum Coloring. *IEEE Transactions on Information Theory* **59**(6): 4025–4032 (2013).

[MW08] W. Matthews and A. Winter. Pure-state transformations and catalysis under operations that completely preserve positivity of partial transpose. *Physical Review A* **78**(1): 012317 (2008).

[MW14a] W. Matthews and S. Wehner. Finite Blocklength Converse Bounds for Quantum Channels. *IEEE Transactions on Information Theory* **60**(11): 7317–7329 (2014).

[MW14b] C. Morgan and A. Winter. Pretty Strong Converse for the Quantum Capacity of Degradable Channels. *IEEE Transactions on Information Theory* **60**(1): 317–333 (2014).

[Nag01] H. Nagaoka. Strong converse theorems in quantum information theory. In *ERATO Workshop on Quantum Information Science*, page 33, (2001).

[Nat16] Nathaniel Johnston. QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9, (2016). `DOI: 10.5281/zenodo.14186`.

[NC10] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge university press (2010).

[NO00] H. Nagaoka and T. Ogawa. Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Transactions on Information Theory* **46**(7): 2428–2433 (2000).

[NPA07] M. Navascués, S. Pironio, and A. Acín. Bounding the Set of Quantum Correlations. *Physical Review Letters* **98**(1): 010401 (2007).

[NPA08] M. Navascués, S. Pironio, and A. Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics* **10**(7): 073013 (2008).

[NST15] A. Nayak, J. Sikora, and L. Tuncel. Quantum and classical coin-flipping protocols based on bit-commitment and their point games. *arXiv:1504.04217* , (2015).

[OCB12] O. Oreshkov, F. Costa, and Č. Brukner. Quantum correlations with no causal order. *Nature Communications* **3**: 1092 (2012).

[ON99] T. Ogawa and H. Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Transactions on Information Theory* **45**(7): 2486–2489 (1999).

[ON07] T. Ogawa and H. Nagaoka. Making Good Codes for Classical-Quantum Channel Coding via Quantum Hypothesis Testing. *IEEE Transactions on Information Theory* **53**(6): 2261–2266 (2007).

[OP04] M. Ohya and D. Petz. *Quantum entropy and its use*. Springer Science & Business Media (2004).

[Pau02] V. Paulsen. *Completely bounded maps and operator algebras*. Volume 78, Cambridge University Press (2002).

[Per96] A. Peres. Separability Criterion for Density Matrices. *Physical Review Letters* **77**(8): 1413–1415 (1996).

[PHHH06] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki. Properties of quantum nonsignaling boxes. *Physical Review A* **74**(1): 012305 (2006).

[Ple05a] M. B. Plenio. Logarithmic Negativity: A Full Entanglement Monotone That is not Convex. *Physical Review Letters* **95**(9): 090503 (2005).

[Ple05b] M. B. Plenio. Problem 20. *Problem 20 of the list https://qig.itp.uni-hannover.de/qiproblems/Open_Problems* , (2005).

[PLOB17] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi. Fundamental limits of repeaterless quantum communications. *Nature Communications* **8**: 15043 (2017).

[PPV10] Y. Polyanskiy, H. V. Poor, and S. Verdu. Channel Coding Rate in the Finite Blocklength Regime. *IEEE Transactions on Information Theory* **56**(5): 2307–2359 (2010).

[PV07] M. B. Plenio and S. S. Virmani. An introduction to entanglement measures. *Quant. Inf. Comp.* **7**(1): 1–51, (2007).

[PV16] C. Palazuelos and T. Vidick. Survey on nonlocal games and operator space theory. *Journal of Mathematical Physics* **57**(1): 015220 (2016).

[PW15] M. Piani and J. Watrous. Necessary and Sufficient Quantum Information Characterization of Einstein-Podolsky-Rosen Steering. *Physical Review Letters* **114**(6): 060404 (2015).

[Rai97] E. M. Rains. Entanglement purification via separable superoperators. *arXiv:quant-ph/9707002* , (1997).

[Rai99] E. M. Rains. Bound on distillable entanglement. *Physical Review A* **60**(1): 179–184 (1999).

[Rai01] E. Rains. A semidefinite program for distillable entanglement. *IEEE Transactions on Information Theory* **47**(7): 2921–2933 (2001).

[Ras02] A. E. Rastegin. Relative error of state-dependent cloning. *Physical Review A* **66**(4): 042304 (2002).

[Ras03] A. E. Rastegin. A lower bound on the relative error of mixed-state cloning and related operations. *Journal of Optics B: Quantum and Semiclassical Optics* **5**(6): S647–S650 (2003).

[Ras06] A. E. Rastegin. Sine distance for quantum states. *arXiv:quant-ph/0602112* , (2006).

[RBL18] D. Rosset, F. Buscemi, and Y.-C. Liang. Resource Theory of Quantum Memories and Their Faithful Verification with Minimal Assumptions. *Physical Review X* **8**(2): 021033 (2018).

[Reg18] B. Regula. Convex geometry of quantum resource quantification. *Journal of Physics A: Mathematical and Theoretical* **51**(4): 045303 (2018).

[Rei11] B. W. Reichardt. Reflections for quantum query algorithms. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, number 1, pages 560–569. Society for Industrial and Applied Mathematics (2011).

[Rén60] A. Rényi. On measures of entropy and information. *Fourth Berkeley Symposium on Mathematical Statistics and Probability* **1**(c): 547–561, (1960).

[Ren05] R. Renner. Security of quantum key distribution. *PhD thesis, ETH Zurich* **6**(01): 1–127 (2005).

[RFWA18] B. Regula, K. Fang, X. Wang, and G. Adesso. One-Shot Coherence Distillation. *Physical Review Letters* **121**(1): 010401 (2018).

[RR11] J. M. Renes and R. Renner. Noisy Channel Coding via Privacy Amplification and Information Reconciliation. *IEEE Transactions on Information Theory* **57**(11): 7377–7385 (2011).

[RV13] O. Regev and T. Vidick. Quantum XOR Games. In *2013 IEEE Conference on Computational Complexity*, number 0844626, pages 144–155. IEEE, (2013).

[RW05] M. Reimpell and R. F. Werner. Iterative Optimization of Quantum Error Correcting Codes. *Physical Review Letters* **94**(8): 080501 (2005).

[SAP17] A. Streltsov, G. Adesso, and M. B. Plenio. Colloquium : Quantum coherence as a resource. *Reviews of Modern Physics* **89**(4): 041003 (2017).

[SBPC$^+$09] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev. The security of practical quantum key distribution. *Reviews of Modern Physics* **81**(3): 1301–1350 (2009).

[Sch05] I. Schur. *Neue begründung der theorie der gruppencharaktere*. (1905).

[Sch35] E. Schrodinger. Die gegenwartige Situation in der Quantenmechanik. *Die Naturwissenschaften* **23**(48): 807–812 (1935).

[Sch79] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory* **25**(4): 425–429 (1979).

[Sha48] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal* **27**(3): 379–423 (1948).

[Sha56] C. E. Shannon. The zero error capacity of a noisy channel. *IRE Trans. Inf. Theory* **2**(3): 8–19 (1956).

[Sho02a] P. W. Shor. Additivity of the classical capacity of entanglement-breaking quantum channels. *Journal of Mathematical Physics* **43**(9): 4334–4340 (2002).

[Sho02b] P. W. Shor. The quantum channel capacity and coherent information. In *lecture notes, MSRI Workshop on Quantum Computation*, (2002).

[Sio58] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics* **8**(1): 171–176 (1958).

[SN96] B. Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Physical Review A* **54**(4): 2629–2635 (1996).

[SS07] G. Smith and J. A. Smolin. Degenerate Quantum Codes for Pauli Channels. *Physical Review Letters* **98**(3): 030501 (2007).

[SS08] G. Smith and J. A. Smolin. Additive extensions of a quantum channel. In *Proceedings of IEEE Information Theory Workshop (ITW)*, pages 368–372. IEEE, (2008).

[SS15] M. E. Shirokov and T. Shulman. On Superactivation of Zero-Error Capacities and Reversibility of a Quantum Channel. *Communications in Mathematical Physics* **335**(3): 1159–1179 (2015).

[SSMR16] D. Sutter, T. Sutter, P. Mohajerin Esfahani, and R. Renner. Efficient Approximation of Quantum Channel Capacities. *IEEE Transactions on Information Theory* **62**(1): 578–598 (2016).

[SSW08] G. Smith, J. A. Smolin, and A. Winter. The Quantum Capacity With Symmetric Side Channels. *IEEE Transactions on Information Theory* **54**(9): 4208–4217 (2008).

[SSWR17] D. Sutter, V. B. Scholz, A. Winter, and R. Renner. Approximate Degradable Quantum Channels. *IEEE Transactions on Information Theory* **63**(12): 7832–7844 (2017).

[Ste12] B. Steinberg. *Representation Theory of Finite Groups*. Springer New York (2012).

[Sti55] W. F. Stinespring. Positive Functions on C -Algebras. *Proceedings of the American Mathematical Society* **6**(2): 211 (1955).

[SU11] E. R. Scheinerman and D. H. Ullman. *Fractional graph theory: a rational approach to the theory of graphs*. Courier Corporation (2011).

[SW97] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A* **56**(1): 131–138 (1997).

[SW01] B. Schumacher and M. D. Westmoreland. Optimal signal ensembles. *Physical Review A* **63**(2): 022308 (2001).

[SWAT18] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka. Bounding the energy-constrained quantum and private capacities of phase-insensitive bosonic Gaussian channels. *New Journal of Physics* **20**(6): 063025 (2018).

[TBR16] M. Tomamichel, M. Berta, and J. M. Renes. Quantum coding with finite resources. *Nature Communications* **7**: 11419 (2016).

[TCR10] M. Tomamichel, R. Colbeck, and R. Renner. Duality Between Smooth Min- and Max-Entropies. *IEEE Transactions on Information Theory* **56**(9): 4674–4681 (2010).

[TGW14] M. Takeoka, S. Guha, and M. M. Wilde. The squashed entanglement of a quantum channel. *IEEE Transactions on Information Theory* **60**(8): 4987–4998, (2014).

[TH13] M. Tomamichel and M. Hayashi. A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks. *IEEE Transactions on Information Theory* **59**(11): 7693–7710 (2013).

[Tod01] M. J. Todd. Semidefinite optimization. *Acta Numerica* **10** (2001).

[Tom12] M. Tomamichel. A framework for non-asymptotic quantum information theory. *arXiv preprint arXiv:1203.2142* (2012).

[Tom16] M. Tomamichel. *Quantum Information Processing with Finite Resources*. Volume 5 of SpringerBriefs in Mathematical Physics, Springer International Publishing (2016).

[TT15] M. Tomamichel and V. Y. F. Tan. Second-Order Asymptotics for the Classical Capacity of Image-Additive Quantum Channels. *Communications in Mathematical Physics* **338**(1): 103–137 (2015).

[TWW17] M. Tomamichel, M. M. Wilde, and A. Winter. Strong Converse Rates for Quantum Communication. *IEEE Transactions on Information Theory* **63**(1): 715–727 (2017).

[Uhl76] A. Uhlmann. The transition probability in the state space of a -algebra. *Reports on Mathematical Physics* **9**(2): 273–279 (1976).

[Uhl77] A. Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics* **54**(1): 21–32 (1977).

[VB96] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review* **38**(1): 49–95 (1996).

[VC01a] G. Vidal and J. I. Cirac. Irreversibility in asymptotic manipulations of a distillable entangled state. *Physical Review A* **65**(1): 012323 (2001).

[VC01b] G. Vidal and J. I. Cirac. Irreversibility in Asymptotic Manipulations of Entanglement. *Physical Review Letters* **86**(25): 5803–5806 (2001).

[VDC02] G. Vidal, W. Dür, and J. I. Cirac. Entanglement Cost of Bipartite Mixed States. *Physical Review Letters* **89**(2): 027901 (2002).

[VHGE14] V. Veitch, S. A. Hamed Mousavian, D. Gottesman, and J. Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics* **16**(1): 013009 (2014).

[vN28] J. v. Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen* **100**(1): 295–320 (1928).

[VP98] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Physical Review A* **57**(3): 1619–1633 (1998).

[VPJK97] V. Vedral, M. B. Plenio, K. Jacobs, and P. L. Knight. Statistical inference, distinguishability of quantum states, and quantum entanglement. *Physical Review A* **56**(6): 4452–4455 (1997).

[VPRK97] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying Entanglement. *Physical Review Letters* **78**(12): 2275–2279 (1997).

[VW02] G. Vidal and R. F. Werner. Computable measure of entanglement. *Physical Review A* **65**(3): 032314 (2002).

[VWW04] K. G. H. Vollbrecht, R. F. Werner, and M. M. Wolf. Irreversibility of entanglement distillation for a class of symmetric states. *Physical Review A* **69**(6): 062304 (2004).

[Wat13] J. Watrous. Simpler semidefinite programs for completely bounded norms. *Chicago Journal of Theoretical Computer Science* **19**(1): 1–19 (2013).

[Wat18] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press (2018).

[WD16a] X. Wang and R. Duan. A semidefinite programming upper bound of quantum capacity. In *2016 IEEE International Symposium on Information Theory (ISIT)*, Volume 2016-Augus, pages 1690–1694. IEEE, (2016).

[WD16b] X. Wang and R. Duan. Improved semidefinite programming upper bound on distillable entanglement. *Physical Review A* **94**(5): 050301 (2016).

[WD16c] X. Wang and R. Duan. On the quantum no-signalling assisted zero-error classical simulation cost of non-commutative bipartite graphs. In *2016 IEEE International Symposium on Information Theory (ISIT)*, Volume 2016-Augus, pages 2244–2248. IEEE, (2016).

[WD17a] X. Wang and R. Duan. Irreversibility of Asymptotic Entanglement Manipulation Under Quantum Operations Completely Preserving Positivity of Partial Transpose. *Physical Review Letters* **119**(18): 180506 (2017).

[WD17b] X. Wang and R. Duan. Nonadditivity of Rains' bound for distillable entanglement. *Physical Review A* **95**(6): 062322 (2017).

[WD18] X. Wang and R. Duan. Separation Between Quantum Lovász Number and Entanglement-Assisted Zero-Error Classical Capacity. *IEEE Transactions on Information Theory* **64**(3): 1454–1460 (2018).

[Weh06] S. Wehner. Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities. *Physical Review A* **73**(2): 022110 (2006).

[WFD17] X. Wang, K. Fang, and R. Duan. Semidefinite programming converse bounds for quantum communication. *arXiv:1709.00200* , (2017).

[WFT17] X. Wang, K. Fang, and M. Tomamichel. On converse bounds for classical communication over quantum channels. *arXiv:1709.05258* , (2017).

[Wil16] M. M. Wilde. Squashed entanglement and approximate private states. *Quantum Information Processing* **15**(11): 4563–4580 (2016).

[Wil17] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press (2017).

[Win99] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory* **45**(7): 2481–2485 (1999).

[Wol78] J. Wolfowitz. Coding theorems of information theory. *Mathematics of Computation* (1978).

[WPG07] M. M. Wolf and D. Pérez-García. Quantum capacities of channels with small environment. *Physical Review A* **75**(1): 012303 (2007).

[WR12] L. Wang and R. Renner. One-Shot Classical-Quantum Capacity and Hypothesis Testing. *Physical Review Letters* **108**(20): 200501 (2012).

[WSV00] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming*. Volume 27 of International Series in Operations Research & Management Science, Springer US (2000).

[WTB17] M. M. Wilde, M. Tomamichel, and M. Berta. Converse Bounds for Private Communication Over Quantum Channels. *IEEE Transactions on Information Theory* **63**(3): 1792–1817 (2017).

[WW13] M. M. Wilde and A. Winter. Strong converse for the classical capacity of the pure-loss bosonic channel. *Problems of Information Transmission* **50**(2): 117–132 (2013).

[WW14] M. M. Wilde and A. Winter. Strong Converse for the Quantum Capacity of the Erasure Channel for Almost All Codes. *9th Conference on the Theory of Quantum Computation, Communication and Cryptography* page 52 (2014).

[WWY14] M. M. Wilde, A. Winter, and D. Yang. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Communications in Mathematical Physics* **331**(2): 593–622 (2014).

[WXD18] X. Wang, W. Xie, and R. Duan. Semidefinite programming strong converse bounds for classical capacity. *IEEE Transactions on Information Theory* **64**(1): 640–653 (2018).

[WY15] A. Winter and D. Yang. Potential capacities of quantum channels. *IEEE Transactions on Information Theory* **62**(3): 1415–1424 (2015).

[XFWD17] W. Xie, K. Fang, X. Wang, and R. Duan. Approximate broadcasting of quantum correlations. *Physical Review A* **96**(2): 022302 (2017).

[XWD17] W. Xie, X. Wang, and R. Duan. Converse bounds for classical communication over quantum networks. *arXiv:1712.05637* , (2017).

[YDY14] N. Yu, R. Duan, and M. Ying. Distinguishability of Quantum States by Positive Operator-Valued Measures With Positive Partial Transpose. *IEEE Transactions on Information Theory* **60**(4): 2069–2079 (2014).

[YHHSR05] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke. Irreversibility for All Bound Entangled States. *Physical Review Letters* **95**(19): 190501 (2005).

[Yur03] F. Yura. Entanglement cost of three-level antisymmetric states. *Journal of Physics A: Mathematical and General* **36**(15): L237–L242 (2003).

[YYW17] M. Ying, S. Ying, and X. Wu. Invariants of quantum programs: characterisations and generation. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages - POPL 2017*, pages 818–832, New York, New York, USA, (2017). ACM Press.

[ZFG10] Y. Zinchenko, S. Friedland, and G. Gour. Numerical estimation of the relative entropy of entanglement. *Physical Review A* **82**(5): 052336 (2010).

[ZLY+18] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and X. Ma. One-Shot Coherence Dilution. *Physical Review Letters* **120**(7): 070403 (2018).