2023-03-14

Manipulation of dynamical resources in quantum information theory

Saxena, Gaurav

UNIVERSITY OF CALGARY

Manipulation of dynamical resources in quantum information theory

by

Gaurav Saxena

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHYSICS AND ASTRONOMY

CALGARY, ALBERTA
MARCH, 2023

Abstract

Quantum channels can be regarded as the most fundamental objects in quantum mechanics. With the help of quantum resource theories, it was recently recognized that dynamical quantum systems (described by quantum channels) may exhibit phenomena such as entanglement and coherence, and can be utilized as resources in various operational tasks. In this dissertation, I characterize and quantify the coherence and magic of dynamical quantum systems, formulate interconversion conditions among pairs of channels, and quantify the performance of fixed programmable processors.

Quantum resource theories are governed by constraints arising from physical or practical settings. Considering the absence of coherence and efficient classical simulability (two different notions of classicality) as practical constraints towards achieving quantum advantage, I develop the resource theories of dynamical coherence and dynamical magic, respectively. In developing the resource theory of coherence, the underlying principle I follow is that the free dynamical objects are those that can neither store nor manipulate coherence. This led me to identify classical channels as free elements in this theory. The development of the resource theory of multi-qubit magic channels is motivated by the need to estimate the classical simulation cost of multi-qubit quantum circuits. The set of completely stabilizer preserving operations is the largest known set of operations in the multi-qubit scenario that can be efficiently simulated classically, and as such, they are the perfect candidates for the free channels of this resource theory. In both these resource theories, I quantify the resources using various resource measures, and solve several single-shot resource interconversion problems including different types of resource cost and distillation. I also formulate a classical simulation algorithm to estimate the expectation value of an observable and show that its runtime depends on a dynamical magic monotone.

Besides developing the above resource theories, I generalize Lorenz majorization to the channel domain and use it to find the necessary and sufficient conditions for interconversion among pairs of classical channels. Furthermore, I quantify the performance of a fixed programmable quantum processor and find a trade-off relation between the success probability and the average fidelity error in simulating a target unitary using the processor.

Preface

This thesis is an original work by the author. The results in chapters 3 and 4 are published in peer-reviewed journals. Chapters 5 and 6 present new results that have not yet been published. The following is the list of relevant publications:

1. Gaurav Saxena, Eric Chitambar, and Gilad Gour, 2020, *Dynamical resource theory of quantum coherence*, Phys. Rev. Research 2, 023298

2. Gaurav Saxena and Gilad Gour, 2022, *Quantifying multiqubit magic channels with completely stabilizer-preserving operations*, Phys. Rev. A 106, 042422

Chapters 1, 2, 5, 6, and 7 are original contributions.

Chapter 3 is taken verbatim from Ref.1 above, except sub-section 3.5.6. The result in sub-section 3.5.6 is new and is my contribution in a joint work done with Yunlong Xiao, Xiaoli Hu, Ximing Wang, and Mile Gu, to be published soon.

Chapter 4 is taken verbatim from Ref. 2 above with the following modifications. Minor additions have been made in Sec. 4.1.1 and Sec. 4.6. Sec. 4.1.2 and Sec. 4.2 have been added that provide an overview of previous results.

Acknowledgements

First and foremost, I am indebted to my supervisor, Dr. Gilad Gour, for his able guidance, and for providing me with his time and support throughout my doctoral studies. I thank him for patiently listening to my ideas, for going through my proofs, for improving my scientific writing, for constantly pushing me to be better in my scientific approach, and for giving me the freedom to explore a research direction of my interest.

I would like to thank my supervisory committee members, Dr. Barry Sanders and Dr. David Feder, for their valuable time meeting with me throughout my Ph.D., and for their insightful comments that helped me significantly improve my thesis. Thanks to Dr. Christoph Simon and Dr. Rei Safavi-Naeini for serving on my candidacy-exam committee, and to Dr. Robin Cockett for serving on my thesis-exam committee. I would like to thank Dr. Markus Müller for being the external examiner in my Ph.D. defense.

Next, I would like to thank Carlo Maria Scandolo, Thomas Theurer, and Elia Zanoni in my group who took out time to read and give feedback on my introduction helping me to write it well. I would like to thank them again and other previous group members including Yunlong Xiao, Kuntal Sengupta, Fasiha, Isabelle, and Taylor, for all the discussions we had. Special thanks are due to Yunlong with whom I also shared an office space, who has supported me a lot in multiple ways throughout this academic journey, and with whom collaborations and learning continued even after he left Calgary.

I would like to acknowledge my academic collaborators – Eric Chitambar, Francesco Buscemi, Aditya Nema, Mark Wilde, Yunlong, Sarvagya Upadhyay, Ryuji Takagi, Mile Gu, and Jakob Kottmann – for their amazing ideas and for the privilege of working together, which has helped me expand the horizon of my scientific thinking. Thanks are also due to Bartosz Regula, James Seddon, Ian George, Karol Horodecki, Varun Narasimhachar, and Zi-Wen Liu for various discussions.

I acknowledge the support from the administration staff of the department of physics at the University of Calgary, especially Yanmei Fei, Tracy Korsgaard, Gerri Evans, and Nancy Jing Lu. I would also like to extend my acknowledgement to the University of Calgary's Taylor Family Digital Library for the resources they have provided me in these years. I would like to thank the Government of Alberta for awarding me with the Alberta Graduate Excellence Scholarship (AGES) and the Alberta Innovates award, and the Department

of physics for nominating me for these awards and for all the student excellence awards I have received.

This academic journey would have been super stressful and life in this part of the world would have been utterly boring if it was not for the awesome people I met along the way. Ana, Murali, Sruthi, Saheli, Sameer, Satyam, and Shweta, whenever I'll remember Calgary days, I will always remember you and cherish the time I have spent with you all. You have stayed with me through thick and thin, and I am really grateful to have met you all. Besides, I would like to mention and thank some other friends with whom I had a really fun time in Calgary - Joanna, Kuntal, Janina, Hridya, Abhinaya, Prasoon, Ritul, Akshay, Sarthak, Kunal, Arindam, Anahad, Trishool, Mishty, Pragati, Laura, Sara, and Shawn. Murali, Mishty and Pragati many thanks to you all for giving me comments on my Ph.D. thesis introduction. Laura, Sara, and Shawn, I have known you since my starting days in Calgary, and your company has always been very warm and welcoming. Thank you for introducing me to the Canadian culture and for making me a part of several celebrations. Friendship with you is also special because of the philosophical nature of our conversations and I truly enjoyed each one of those.

This acknowledgement would be incomplete without thanking the person whose love and encouragement has kept me going. Tasha, without you, this journey would have been an emotionally draining one. Thank you for being patient and tolerating me throughout, and for always filling in me a sense of positivity. Also huge thanks for reading my introduction chapter and providing me your critical feedback and comments.

To my parents, thank you for all the unconditional love and support you have provided me, and all the sacrifices you have made for me. Thank you for always showing me the right path in all situations. Thank you for giving me the freedom to pursue my dreams, for believing in me when even I was down. Thank you for motivating me to pursue a Ph.D. in the first place. You have made me whatever I am today. Words fall short to express my gratitude towards you both. This thesis is dedicated to you.

In the end, I thank my (spiritual) guru for initiating me in my spiritual journey and providing me with the energy to face the situations that have come and gone.

To my parents and my teachers

Table of Contents

List of Figures

List of Tables

Notations

In this thesis, we will denote all dynamical systems and their corresponding Hilbert spaces by A, B, C, etc, and all static systems and their corresponding Hilbert spaces by A_1, B_1, C_1, etc. In this setting, the notation for a dynamical system, say A, indicates a pair of systems such that $A = (A_0, A_1) = (A_0 \to A_1)$ where A_0 and A_1 represent the input and output systems, respectively. The choice of notation for the static systems is because all the states can be viewed as channels with trivial input. For a composite system, the notation like A_1B_1 will be used to mean $A_1 \otimes B_1$. To represent the dimension of a system, two vertical lines will be used. For example, the dimension of system A_1 is $|A_1|$. A replica of the same system would be represented by using a tilde symbol. For instance, system $\tilde{A}_1$ is a replica of system A_1, and system $\tilde{A}_1\tilde{B}_1$ is a replica of system A_1B_1 i.e., $|\tilde{A}_1| = |A_1|$ and $|\tilde{A}_1\tilde{B}_1| = |A_1B_1|$.

The set of bounded operators, Hermitian operators, positive operators and density matrices on system A_1 would be denoted by $\mathfrak{B}(A_1)$, $\mathrm{Herm}(A_1)$, $\mathrm{Pos}(\mathrm{A}_1)$, and $\mathfrak{D}(A_1)$, respectively. Note that $\mathfrak{D}(A_1) \subset \mathrm{Pos}(\mathrm{A}_1) \subset \mathrm{Herm}(\mathrm{A}_1) \subset \mathfrak{B}(\mathrm{A}_1)$. Density matrices would be represented by lowercase Greek letters ρ, σ, τ, etc. We will denote the maximally coherent state (or the plus state) for a system B_1 by $\phi^+_{B_1}$, the normalized maximally entangled states by $\phi^+_{A_1B_1}$ (note the subscripts in both) and the unnormalized maximally entangled states by $\Phi^+_{A_1B_1}$ for a bipartite system A_1B_1. The maximally mixed state for a system B_1 will be denoted by u_{B_1}. The set of all stabilizer states in system A_1 will be denoted by $\mathrm{STAB}(A_1)$. For pure stabilizer states in system A_1 we will write $\phi \in \mathrm{STAB}(A_1)$, and notation like $\sigma \in \mathrm{STAB}(A_1)$ will mean a density matrix of a state taken from the stabilizer polytope which is a convex hull of pure stabilizer states.

The set of all linear maps from $\mathfrak{B}(A_0)$ to $\mathfrak{B}(A_1)$ would be denoted by $\mathfrak{L}(A_0 \to A_1)$, the set of all completely positive maps from $\mathfrak{B}(A_0) \to \mathfrak{B}(A_1)$ would be denoted by $\mathrm{CP}(A_0 \to A_1)$ and the set of quantum channels would be denoted by $\mathrm{CPTP}(A_0 \to A_1)$ with $\mathrm{CPTP}(A_0 \to A_1) \subset \mathrm{CP}(A_0 \to A_1) \subset \mathfrak{L}(A_0 \to A_1)$. Throughout this article, we would use calligraphic letters like $\mathcal{E}, \mathcal{F}, \mathcal{M}, \mathcal{N}$, etc, to represent quantum channels. For simplicity, we will denote a quantum channel with a subscript A, like $\mathcal{E}_A$, to denote an element of $\mathrm{CPTP}(A_0 \to A_1)$. The identity channel in $\mathfrak{L}(A_0 \to A_0)$ will be denoted by id_{A_0}.

The notation $\mathfrak{L}(A \to B)$ will be used to denote the set of all maps from $\mathfrak{L}(A_0 \to A_1)$ to $\mathfrak{L}(B_0 \to B_1)$.

Similarly, the set of all maps from $\text{Herm}(A_0 \to A_1)$ to $\text{Herm}(B_0 \to B_1)$ would be denoted by $\text{Herm}(A \to B) \subset \mathfrak{L}(A \to B)$. All linear maps in $\mathfrak{L}(A \to B)$ and $\text{Herm}(A \to B)$ are known as supermaps. We will use capital Greek letters like Θ, Σ, Ω, etc, to denote supermaps. Square brackets will be used to denote the action of supermaps on linear maps. For instance, $\Theta_{A\to B}[\mathcal{E}_A]$ is a linear map in $\mathfrak{L}(B_0 \to B_1)$ obtained by the action of a supermap $\Theta \in \mathfrak{L}(A \to B)$ on a map $\mathcal{E} \in \mathfrak{L}(A_0 \to A_1)$. The set of supermaps that map quantum channels to quantum channels (even when tensored with the identity supermap, i.e., even when acting on part of quantum channels) are called superchannels and would be represented by $\mathfrak{S}(A \to B)$. Identity superchannel in $\mathfrak{S}(A \to A)$ would be denoted by $\mathbb{1}_A$. Lastly, we reserve the symbol Δ to represent a dephasing superchannel. Such a superchannel converts any channel to a classical channel.

The Choi matrix of a channel $\mathcal{N} \in \text{CPTP}(A_0 \to A_1)$ is defined as $J_A^{\mathcal{N}} := \mathcal{N}_A\left(\Phi^+_{A_0\tilde{A}_0}\right)$, the Choi matrix of a superchannel $\Theta \in \mathfrak{S}(A \to B)$ will be denoted in bold as $\mathbf{J}^{\Theta}_{AB}$. To denote normalized Choi matrix of a channel $\mathcal{N}_A$, we will use tilde symbol over J as $\tilde{J}_A^{\mathcal{N}}$.

List of symbols and abbreviations

Symbol	Definition
$\mathfrak{L}(A_0, A_1)$	Set of all linear operators from system A_0 to system A_1
$\mathfrak{L}(A_1)$	Set of all linear operators in system A_1
$\mathfrak{B}(A_1)$	Set of all bounded operators in systems A_1
$\mathrm{Herm}(A_1)$	Set of all Hermitian operators in systems A_1
$\mathrm{Pos(A_1)}$	Set of all positive semidefinite operators in systems A_1
$\mathfrak{D}(A_1)$	Set of all density matrices in systems A_1
$\mathrm{STAB}(A_1)$	Set of all stabilizer states in system A_1
$\mathfrak{L}(A_0 \to A_1)$	Set of all linear maps from linear operators in $\mathfrak{L}(A_0)$ to those in $\mathfrak{L}(A_1)$
$\mathrm{CP}(A_0 \to A_1)$	Set of all completely positive maps from $\mathfrak{L}(A_0) \to \mathfrak{L}(A_1)$
$\mathrm{CPTP}(A_0 \to A_1)$	Set of all quantum channels taking density matrices from $\mathfrak{D}(A_0) \to \mathfrak{D}(A_1)$
$\mathrm{CSPO}(A_0 \to A_1)$	Set of all completely stabilizer preserving operations in dynamical system A
$\Phi^+_{A_1\tilde{A}_1}$	Unnormalized maximally entangled state in system $A_1\tilde{A}_1$
$\phi^+_{A_1\tilde{A}_1}$	Normalized maximally entangled state in system $A_1\tilde{A}_1$
$\phi^+_{A_1}$	Normalized maximally coherent state in system A_1
id_{A_1}	Identity channel on system A_1
$\mathbb{1}_A$	Identity superchannel on dynamical system A
$\mathcal{D}_{A_1}$	Completely dephasing channel on system A_1
Δ_A	Completely dephasing superchannel on dynamical system A

Epigraph

There is really only a single postulate of quantum mechanics, and it is that "everything is a quantum channel."

– Gilad Gour and Mark M. Wilde, *Entropy of a quantum channel*

Chapter 1

Introduction

Since the dawn of quantum information (that is, around the 1980s), quantum computers have been projected as machines capable of outperforming their classical counterparts [1, 2, 3, 4, 5, 6]. Quantum computers are not just next-gen computers with higher processing speeds and bigger memory, they are fundamentally different types of machines that use quantum mechanical principles to operate. Such computers promise to efficiently solve certain problems, such as quantum simulations and quantum random sampling, that are considered hard for traditional computers [7, 8, 9, 10, 11, 12, 13, 14, 15].

On 23 November 2019, Google claimed that they had delivered on one such promise. A team of Google, led by John Martinis, announced that they have demonstrated quantum supremacy, a milestone long awaited in the quantum computing community [13, 16, 17]. They demonstrated quantum supremacy by "comparing their quantum processor against state-of-the-art classical computers in the task of sampling the output of a pseudorandom quantum circuit" [13]. This specific problem has limited practical applicability and the experiment itself was performed on a noisy quantum computer [13]. Nonetheless, this achievement sparked excitement in the community and pushed researchers toward the challenge of engineering a universal fault-tolerant quantum computer (i.e., a perfect quantum computer) [18].

Achieving perfect quantum computation (and communication) is, however, a challenging task [19, 20, 21, 22, 23, 24]. The challenge stems from two main drawbacks. Firstly, quantum systems are ultra-fragile in the sense that they decohere almost instantly when interacting with the environment, thus making it hard to encode information. Secondly, the manipulation of these systems by applying quantum logic gates or performing measurements is subject to random errors like bit-flip and phase-flip errors. It is predicted that it might take a decade or more to overcome these challenges [18, 23, 25]. Given the error-prone nature of quantum systems and quantum devices, the question arises as to whether the quantum mechanical properties

of these current imperfect noisy devices can still be harnessed to our advantage, and what resources can help us in achieving this advantage in a given setting.

To address the aforementioned question, several approaches have been taken in the literature [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. For example, one approach is to consider a noise model and determine how a particular task can be accomplished under such a model. In some of these models, a quantum mechanical phenomenon provides a clear advantage, while in others, some special quantum states might be used to reduce the error. As an example, consider decoherence as a noise model [32, 37, 38, 39, 40, 41]. In this case, the preservation of quantum coherence naturally emerges as a resource that can help us to accomplish tasks that would otherwise be difficult due to decoherence. (In Chapter 3, we will see in detail how the ability to preserve quantum coherence is used as a resource in various operational tasks.) Another approach to gain an advantage using error-prone quantum devices is to focus on maximizing the accuracy to simulate a desired operation, given arbitrary errors [42, 43, 44, 45, 46]. Consider the following example for illustration. The quantum computer that Google used to demonstrate quantum supremacy is not a single-purpose, but a programmable one [13]. Programmable quantum computers can execute a variety of computations by using distinct specially designed states known as program states as one of their inputs [47]. Different operations require different program states whose dimensions vary from one operation to another. As a result, to perform an arbitrary computation, the computer is scaled as per the dimension of the optimal program state [47]. Since in practical applications, scaling makes it hard to prevent noise in computation, fixing the input and output dimensions is a physically motivated constraint. Under this constraint, to maximize the accuracy in simulating the desired quantum operation, it is crucial to find the (sub-)optimal quantum state. Furthermore, if there are two processors, it is important to determine which one can more closely approximate our desired operation. In chapter 6, we will see how to measure the performance of such processors and how to find the optimal resources to approximate a unitary channel. Such strategies (like the two presented above) and others have proven useful for dealing with noise and determining optimal resources in various computational and communication tasks [25, 32, 37, 48, 49, 50].

In this process of finding how to gain an advantage in a given setting, we need to understand what resources are useful and how are they encoded in different quantum objects such as quantum states, unitary operations, and quantum measurements. Interestingly, quantum states, unitary operations, measurements, the process of discarding a system, or any kind of combination of these operations can be described with the formalism of quantum channels. With this view, in which quantum channels describe all quantum objects, I have addressed various resource interconversion problems, characterized resourceful quantum channels in specific physical settings, and quantified the resourcefulness of quantum channels in those settings. A more detailed discussion on the motivation for working with quantum channels as resources is provided in

Section 1.2. Since it was with the emergence of quantum information that several quantum mechanical phenomena were treated as resources, in the next section (Sec. 1.1), I have provided a brief account of the development of quantum information.

1.1 The dawn of quantum information

The advent of quantum mechanics in the first quarter of the twentieth century has profoundly changed our fundamental understanding of the world. Quantum mechanical phenomena such as entanglement and nonlocality have been major subjects of debates on the nature of reality, and a great deal of research went into understanding the fundamental properties of matter using quantum mechanics [51, 52, 53].

While debates on the nature of reality and interpretations of quantum mechanics were being put forth in the 1930s and 40s, a beautiful and profound theory of classical information was developed by Claude E. Shannon in 1948 [54]. Classical information theory involves the study of storage, manipulation, and communication of information when the information is encoded using the laws of classical mechanics. In his work, Shannon used probability theory and statistics to operationally quantify the information content in a message. He called this measure 'entropy'. Furthermore, he also mathematically characterized the capacity of a noisy channel, which is the highest rate at which information can be reliably communicated over the channel. This seminal work forms the basis of today's modern technology. In his honor, the field of classical information is also referred to as classical Shannon theory. With the invention and development of the theory of classical information processing, the field of classical communication and computation progressed rapidly in the second half of the twentieth century.

At around the same time, more quantum mechanical phenomena (such as quantum contextuality and nonlocality) were being discovered [55, 56, 57, 58]. Over time, scientists began to explore the possibilities of using these phenomena as physical resources rather than just as mere features of quantum mechanics to model nature [59, 60, 61, 62, 63]. Around the 1970s and 80s, Alexander Holevo, Roman Stanisław Ingarden, Paul Benioff, and Yuri Mannin, independently wrote papers that are considered the foundations of quantum information science [59, 64, 65, 66]. More interest in this field grew after 1982, when Richard Feynman spoke at a conference on the topic "Simulating physics with computers" [1]. In his talk, he proposed the idea of using quantum-mechanical computers to simulate nature. Following this conference, extensive research was undertaken over the next fifteen years, establishing the advantages of a quantum computer. David Deutsch, in 1985, formulated a description of a quantum Turing machine [67]. In 1992, David Deutsch and Richard Jozsa formulated the Deutsch-Jozsa algorithm and showed that it is exponentially faster than any possible deterministic classical algorithm to find if a function (promised to be constant or balanced, with binary input

of n bits and 0 and 1 as output) is constant or balanced [68]. In 1993, Charles Bennett and others came up with the idea of quantum teleportation [69]. Then in 1995, with Peter Shor's invention of a quantum algorithm to factor numbers exponentially faster than any known classical algorithm, significant interest was ignited in the field of quantum computation and information [70]. Presently, quantum information processing is attracting a lot of interest for commercial applications [71]. Research at both industrial and academic levels is being carried out extensively on the applications of quantum information processing, including quantum algorithms, quantum biology, quantum cryptography, etc. Thus, it is of crucial interest to understand the limitations and boundaries of what can be achieved by harnessing the power of quantum mechanical systems as resources. In a quest to understand these limitations, in this dissertation, I explore the quantification, manipulation, and control of dynamical quantum resources under various practical constraints in quantum information processing.

1.2 What are dynamical quantum resources?

A common theme in physics is the unification of physical phenomena. An example of such a unification is the unification of fundamental forces in which different forces correspond to different aspects of a single unifying force. In a similar manner, quantum channels mathematically unify different quantum objects in quantum mechanics [72]. For example, quantum measurements, quantum states, and quantum instruments, can be viewed as special types of quantum channels, thus making it easier to investigate and analyze the resource character of all quantum objects under a unified framework. However, the primary reason for defining such a mathematical entity (quantum channels) was rather different.

In the early stages of the development of quantum information theory, ideas from classical information theory were being generalized to the quantum case [73, 74]. The natural generalization of a classical communication channel was a quantum communication channel, or a quantum channel in short [59]. In this scenario, a quantum channel mathematically characterizes a physical medium used to transmit quantum information, and as such, models the evolution of the state as a result of any noise applied to the system during communication [48, 59, 75, 76]. Since information is encoded in quantum states, the transmission of quantum information essentially means transmitting a quantum state from point A to point B via some noisy channel. If a quantum channel can model a noisy transmission of a quantum system, it can certainly model the evolution of the state of the system over the period of time when the system is not physically transmitted somewhere. Whether a system is isolated (closed system), interacts with its surroundings (open system), or undergoes a measurement, any kind of change in the state of the system (in space and/or time), can be described by a quantum channel. To put it succinctly, quantum channels describe the most general

evolution of a quantum system. Abstractly, we can think of it as a box with an input and an output system. When viewed in this light, even a quantum state can be regarded as a quantum channel with trivial input (i.e., no input) and the state itself as the output. The reverse of this process, i.e., a quantum input and a trivial output, is a quantum channel that describes the discarding process. (See Fig. 1.1 for a schematic diagram.)

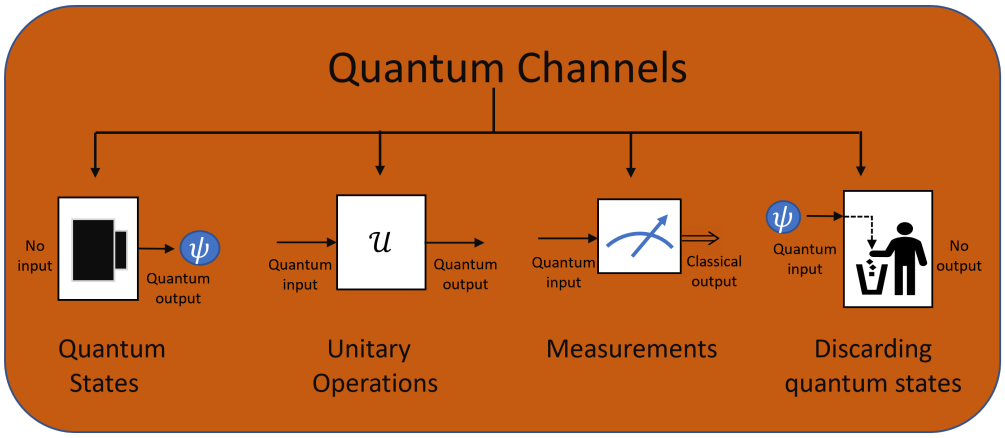

Figure 1.1: Quantum states, unitary operations, measurements, and discarding quantum systems can all be regarded as quantum channels

Studying quantum mechanical phenomena such as entanglement and superposition has helped us to gain a fundamentally better understanding of nature [77, 78]. Besides their interest from a fundamental point of view, these phenomena have also been recognized as resources. Quantum states possessing these phenomena can be used to circumvent certain restrictions, enabling tasks like quantum teleportation and superdense coding, which were otherwise impossible [32, 48, 69, 73, 79, 80, 81, 82]. With time, the broader scientific community, including computer scientists and applied mathematicians became more interested in using these phenomena as resources, resulting in the development of faster algorithms, better cryptography protocols, etc. [3, 63, 70, 73, 83, 84, 85, 86, 87]. The success of this recognition of quantum mechanical properties as resources at the state level brings forward the question of whether we can utilize similar properties of other quantum objects as resources and transcend the practical limitations posed by static quantum systems. These limitations arise from factors such as faulty preparations and difficulty in sustaining the coherence or entanglement of several states together.

Quantum mechanical properties, such as entanglement and coherence, were believed to be contained only in static systems described by quantum states. However, it was only in the last few years and with the help of quantum resource theories that it was recognized that dynamical quantum systems described by quantum channels may possess resources such as entanglement, nonlocality, coherence, athermality, and so on [37, 88, 89, 90, 91, 92, 93, 94, 95, 96]. Studying these dynamical resources and exploring how we can use

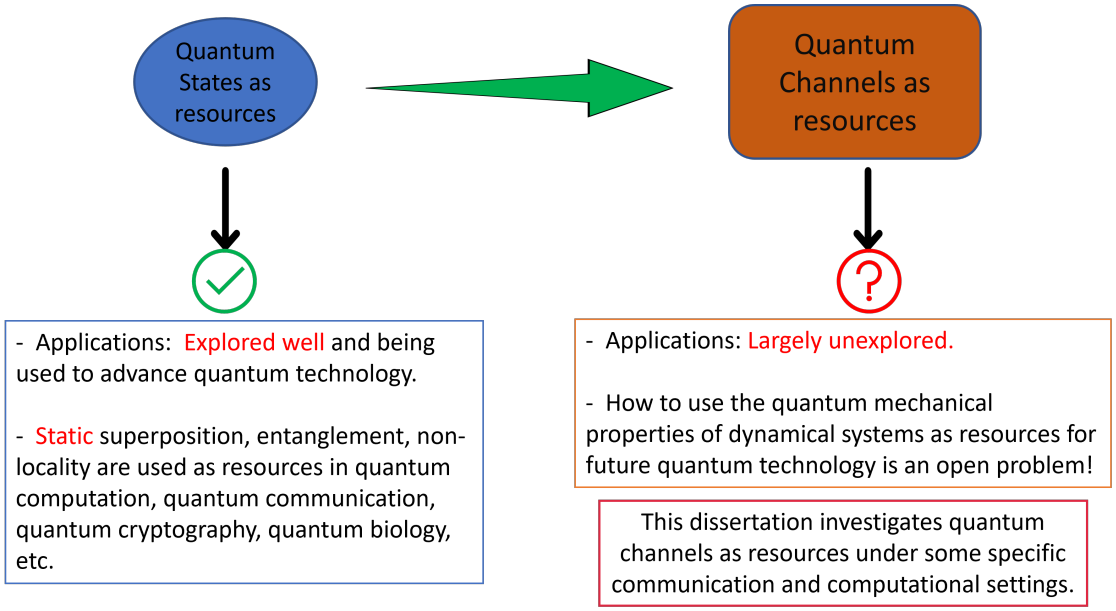

Figure 1.2: Quantum channels as resources

their properties to overcome the limitations of physical systems is largely unexplored, and research in this direction might lead us to find efficient methods to distribute resources, devise new protocols, better control quantum devices, and improved computation. In fact, it is one of the central aims of quantum information to precisely understand the limitations posed by the laws of quantum mechanics and to find the most efficient ways to take advantage of these laws. Therefore, it is important to go beyond static resources in all quantum mechanical phenomena and to understand how to quantify and manipulate dynamical resources, and to find protocols that consume and use such resources in the most efficient way. With this in mind, the primary focus of this dissertation is to examine how quantum channels can be used as resources and to understand when dynamical resources can provide an advantage over existing technologies (see Fig. 1.2).

To harness the resource characteristics of quantum systems, the resource-theoretic framework has proved to be highly versatile and powerful [82]. In a quantum resource theory, quantum states and channels are classified as free or resources based on a given setting. After this partition of states and channels as free and resources, a resource theory then systematically studies what tasks become possible in that setting. Viewed from this lens, the quantum information theory itself can be regarded as a theory of interconversion among various quantum resources. (A detailed description of quantum resource theories is provided in Section 2.5.) In this dissertation, I have used the resource theoretic formalism to study the resource characteristics of quantum channels in various quantum communication and quantum computational settings. The following section provides an overview of these research problems, and the details are discussed in chapters 3-6.

1.3 Research objectives and organization of thesis

In this thesis, I have studied the resource character of quantum channels in various quantum computational and communication scenarios. Chapter 2 provides the mathematical preliminaries that are used throughout this thesis. It also includes a discussion of the mathematics of superchannels – the maps that describe the evolution of quantum channels. Chapter 2 concludes with an overview of the resource theoretic framework and the mathematical structure of quantum resource theory at the level of quantum states. Chapters 3, 4, 5, and 6 cover different problems, and the necessary background for each chapter is provided at the beginning of each chapter.

In chapter 3, I have considered the scenario where quantum computation is error-free but quantum communication is noisy. Any information communicated from a sender to a receiver is completely decohered. That means that if the quantum information encoded by the sender is transferred to the receiver via a communication channel, then the information will be completely lost to the environment. In such a case, the receiver only receives classical information. This error is known as decoherence. Such channels, which completely kill the coherence of the transmitted quantum information and convert it to a classical state, are the classical channels and have the ability to perfectly transmit classical information from a sender to a receiver. These channels do not have the ability to store, create, detect, and transmit quantum information. The resources then are all quantum channels which can generate, manipulate, store, transmit, or detect coherence of an input quantum state. With this partition of channels with classical channels as free and quantum channels as resources, I developed a resource theory of coherence present in quantum channels. I quantified the coherence of quantum channels in terms of various resource measures and showed that some of these measures have operational interpretation in terms of simulating an arbitrary quantum channel from a given coherent resource using a free superchannel. I also show that it is computationally efficient to determine whether a particular quantum channel conversion is possible under the free superchannels or not.

In chapter 4, I considered a scenario in quantum computation. In quantum computation, fault-tolerant quantum computation can be achieved by using a very restricted subset of quantum states and operations known as the stabilizer states and operations. Circuits made from stabilizer operations are called stabilizer circuits. To promote fault-tolerant quantum computation to universal quantum computation, non-stabilizer states or operations (also known as magic states and operations) are injected into stabilizer circuits. Furthermore, the output of stabilizer circuits with stabilizer states as inputs can be efficiently simulated on a classical computer. Thus, stabilizer states and operations can neither perform universal quantum computation nor can they provide any computational advantage. To achieve universal quantum computation using stabilizer circuits and to gain quantum computational advantage, magic states and operations are used

as resources. With this identification of free and resource states and channels to achieve universality, it is crucial to quantify magic resources and understand how a magic resource can be converted to another. To solve these resource conversion problems among magic resources, two branches have emerged – one for the odd-d dimensional case and the other for the multi-qubit case. For the odd d-dimensional case, the discrete Wigner representation is used to classify the useless and the useful elements [90, 97]. For the multi-qubit case however, the classification of useful and useless elements using the Wigner function is hard, and alternative approaches need to be taken. In my work, I use the resource theoretic framework to quantify dynamical magic resources. I solved interconversion problems among qubits and formulated bounds for converting qubit states to channels and vice versa. Lastly, I formulated a classical simulation algorithm to estimate the expectation value from any given circuit and connected it to a magic monotone.

In chapter 5, I considered interconversion among pairs of quantum channels. In 1953, it was shown that a pair of probability distributions can be converted to another pair if and only if a convex region (known as the testing region) defined for the pair of distributions always contains the testing region of the other pair [98]. The definition of the testing region was generalized in 2017 for a pair of quantum states, and interconversion conditions among pairs of quantum states were formulated for certain resource theories. I generalized the definition of the testing region for a pair of quantum channels and mathematically characterized it in terms of the diamond norm and Hilbert α-divergence. Using this characterization, I was able to formulate interconversion conditions among pairs of classical channels.

In chapter 6, I considered the problem of simulating a unitary channel from a fixed quantum processor, i.e., a processor with fixed input and output dimensions. Fixing the dimensions is a physically motivated restriction because scaling makes it hard to prevent noise in computation. We consider the processor to be a programmable bipartite quantum channel shared between Alice and Bob. Bob has limited quantum computational ability and wants to apply some unitary channel on a quantum state using the processor, and conveys this information to Alice. Since the dimensions are fixed, any arbitrary unitary channel cannot be deterministically simulated. So, with the information of the target unitary, Alice prepares a special quantum state, known as the program state, such that using this program state, the processor can approximately simulate the unitary channel on Bob's side. Alice can also perform some post-processing by performing measurements or discarding her system to further improve the probability of approximating the unitary channel. Given this setting, I formulated the trade-off between the maximum success probability of simulating a unitary channel and the approximation error between the resultant channel and the target unitary.

Chapter 7 consists of a summary of the results of each project and lists some open problems related to these projects.

Chapter 2

Background and Preliminaries

In this chapter, I cover the preliminary mathematical tools that have been used throughout this thesis. I start by providing a brief discussion on information and operational tasks in quantum information science in Sec. 2.1. In Sec. 2.2, basic elements of quantum mechanics like quantum states, channels, measurements, etc., are discussed. Sec. 2.3 provides details about the distance measures, and Sec. 2.4 covers the convex analysis tools used in this thesis. I conclude the chapter by giving a non-technical introduction to the resource-theoretic approach followed by the mathematical structure of quantum resource theories for quantum states. Resource theory is a remarkable tool for quantifying resources and solving resource manipulation problems. I have used the resource-theoretic framework to quantify dynamical resources in some specific quantum communication and computational scenarios.

2.1 Overview of classical and quantum information

2.1.1 What is information?

Information, simply put, refers to that which *informs*. Abstractly, information can be thought of as the resolution of uncertainty [54]. In common usage, information is typically any meaningful message. To convey information from one to another, humans have developed the concept of language. A language consists of words used in a systematic, structured, and organized way dictated by the grammar of the language. Using such a system of communication, two people can interact with each other and exchange information. In the case of communication of information over a noisy channel, the abstract notion of *information* can be thought of as a set of possible messages to be encoded and sent over the noisy channel by the sender. The receiver has to decode the received information and reconstruct the message (thereby resolving the uncertainty of

what message was sent) with a low probability of error despite the channel noise. Thus, the transmission of information requires a sender, a receiver, a communication medium, and a storage medium.

The scientific study of the encoding, storing, and transmission of information from a sender to a receiver via some communication medium is called information theory [54]. Such a study heavily relies on concepts and ideas from probability theory, statistics, statistical mechanics, computer science, and electrical engineering.

With the advancements in science and technology, researchers have identified various ways to encode and transmit information over communication channels. When the information is stored using the principles of classical mechanics, it is known as classical information, and the science behind its communication is the subject of classical information processing. Likewise, when the information is stored using quantum mechanical principles, it is known as quantum information, and the science behind communicating quantum information is the subject of quantum information processing.

2.1.2 Classical vs. quantum information

In classical information theory, information is encoded and measured using bits. A bit is represented as a binary number "0" or "1". Physically, a bit can be thought of as something which has two states or possibilities with maximal information, like a coin's heads and tails, high and low voltage, a magnet's spin up and down, etc. This idea of using a two-valued quantity to store information was also introduced by Shannon. But Shannon's notion of a bit was more abstract than just its physical interpretation [54]. To understand Shannon's notion of a bit, let us suppose that we toss a fair coin. Until we look at the coin, we are uncertain of its state, and once we see the coin, we know its state with certainty. Shannon's notion of bit measures the amount of information we gained when some uncertainty is resolved. In Shannon's terms, we can say that we learnt one bit of information when we looked at the coin. However, if the coin was biased and there was a higher chance of getting "heads" than "tails", then there won't be much surprise if we see "heads" after tossing the coin. In this case, we will say we learnt less than one bit of information.

Shannon's bit, as described above, as a way to quantify or measure the surprise upon learning an outcome of any such binary experiment is also known as Shannon binary entropy. Mathematically, it is defined for a given probability distribution $(p, 1-p)$ as [54]

$$h_2(p) \equiv -p\log(p) - (1-p)\log(1-p) \tag{2.1}$$

where the logarithm is taken base two. It is this notion of bit - as a measure of information - which is crucial in information theory.

With this formulation of classical bit and by considering information as a set of possible messages [54],

Shannon then gave the two main theorems which formed the foundations of classical communication. The first theorem, known as the noiseless coding theorem or Shannon's source coding theorem concerns data compression. This theorem established fundamental limits on how much we can compress our data before we start losing information. The second theorem is known as the noisy channel coding theorem or Shannon's channel coding theorem. It is concerned with channel capacity or the transmission of information over noisy classical communication channels. This theorem established that for any noisy (i.e., error-prone) communication channel, there exists an encoding and decoding using which the capacity of the channel can be achieved. The capacity of a channel is defined as the highest rate at which information can be reliably communicated through a channel. From this work, Shannon unified various (classical) communication mediums like radio, telephone, television, etc. in one framework.

With Shannon's notion of bit in place, we now discuss the quantum analogue of the classical bit, which is used as a measure of quantum information. The quantum bit, or qubit, in short, is a physical system which can exist in a superposition of two states of maximal information. (We have discussed quantum state in detail in Sec. 2.2.) Some examples of physical realizations of a qubit include an atom with a ground and an excited state, an electron's spin states, a photon's polarization, etc. This physical notion of a qubit is easy to understand from the concepts of quantum mechanics. What is more pressing is the abstract information-theoretic notion of a qubit as a measure of quantum information. Like classical information, we quantify quantum information by the amount of knowledge we gain after learning the state of a qubit. Depending on the initial information about preparation and the final result after measurement, we learn some "qubits" of information. For example, suppose someone prepares a $|\uparrow_z\rangle$ state (spin-up state in the z-direction) and performs a measurement in the z-direction. We know that the final state would be the same and we will not gain any new information. Thus, we will say we learnt 0 qubits of information. Now, instead of the z-direction, if the measurement is performed in the x-direction, the result would be $|\uparrow_x\rangle$ or $|\downarrow_x\rangle$ with equal probability. In this case, we will say that we learned one "qubit" of quantum information. Therefore, how much we learn about a quantum system after a particular measurement depends on a priori knowledge of the quantum system. Putting it differently, we can say that we learn less about the system if the system is not disturbed after measurement [74].

Now that we understand the physical and information-theoretic meaning of a qubit, let us discuss how to store information in classical and quantum bits. In the classical case, one bit is used to store one bit of information. The bit can either be in the state "0" or in the state "1" and to store any information, strings of bits are required. In the quantum case, however, due to superposition, one might be tempted to think that an exponential amount of information can, in principle, be stored in n qubits as compared to n classical bits, and thus, we can have an exponential advantage over the classical method. Holevo, in

1973, showed that despite storing more information using n quantum bits, we can only retrieve as much information from qubits after measurement as we can get from n classical bits [59]. Despite the fact that we can only retrieve the same amount of information from n qubits and n classical bits, we can still gain an advantage over classical computation by using quantum computers. It is an active area of research in quantum computing to find efficient solutions to problems that are considered difficult by classical computing methods [99, 100, 101, 102, 103]. Similarly, research in quantum communication and information attempts to find limitations of what can be achieved, given a particular quantum resource [32, 82, 104, 105, 106].

2.1.3 Operational tasks in quantum information

With the physical and abstract notions of a quantum bit in place, we now briefly discuss the operational tasks in quantum information theory. Quantum information has several resources that can be exploited by a sender and a receiver [73, 74, 82]. These resources can be categorized as noisy or noiseless, static or dynamic, classical or quantum, etc. An example of a noisy resource is a quantum communication channel that introduces some noise or error in the input quantum system while transferring or manipulating it. A noiseless resource on the other hand is an ideal quantum channel or device. A static resource is a quantum system that maintains the same state over time (i.e., ideally unaffected by the environment). A dynamic resource, described by a quantum channel, is a resource that manipulates static quantum systems, for eg., a unitary operation or a measurement. The classical resources include classical states and classical communication channels, and quantum resources include quantum states and quantum channels. Now, let us look at some operational tasks using these resources. Suppose we have access to a noiseless qubit channel. A key task in quantum information is then to use as few of this resource as possible to communicate quantum information generated from some source from a sender to a receiver [79]. Another resource is a shared entangled state between two parties. With a shared entangled state and classical communication, one can perform quantum teleportation [69]. That is, with the help of the shared entangled state, a sender can transmit quantum information to a receiver without requiring a quantum communication channel (with the assumption that both sender and receiver can perform any operation on their local systems). Similarly, if the two parties have access to a noisy quantum channel, then it can be determined how much classical information can be transmitted from a sender to a receiver by using the noisy quantum channel a large number of times [107, 108]. In this scenario, if the two parties also share an entangled state, then it has been shown that entanglement gives a boost to the amount of noiseless classical communication one can generate using a noisy quantum channel [109, 110, 111]. Using quantum entanglement as a resource, tasks such as superdense coding, teleportation, and secure communication (i.e., quantum cryptography) can also be achieved [48]. Likewise,

in quantum communication settings, where the environment acts on a system and decoheres any quantum information encoded in the system, quantum channels that can preserve and transmit quantum coherence act as resources [32]. In quantum computation, pure magic states are the resources that are useful to promote fault-tolerant quantum computation to universal quantum computation [112]. However, it is experimentally very hard to distill pure magic states from impure ones, and finding magic distillation rates and devising protocols to achieve these rates are important tasks in quantum information. Thus, research on using quantum phenomena as resources has established that quantum mechanics allows for tasks like improved sensing, simulation of complex biomolecules, and speedup over many known classical algorithms [73]. So, quantum phenomena such as entanglement, coherence, and magic, are being investigated in deeper detail to find any technological advantage they may offer. These investigations require a proper classification and quantification of resources in a given particular setting to asses which states and channels are valuable in that setting. Quantification of resources helps in answering various resource interconversion problems which lie at the core of all quantum information processing tasks. To this purpose, the resource-theoretic framework is of great significance as it offers an organized and structured way to quantify resources [82].

2.2 Elements of Quantum Mechanics

Quantum mechanics was developed to understand how nature works fundamentally. To do that, the quantum theory provides a way to describe the state of a physical quantum system like an electron or an atom, how they evolve in closed and open environments, and what happens when a measurement is made on these systems. With the emergence of quantum information, counterintuitive quantum mechanical phenomena like quantum entanglement and coherence were used as resources. By encoding information in static quantum systems and by cleverly manipulating the quantum mechanical properties of these systems, tasks like teleportation, superdense coding, etc. were achieved, which were otherwise impossible. In order to understand how quantum systems, both static and dynamic, can be used as resources, it is important to understand how we can mathematically express the state of a system and its evolution. In this section, we cover the mathematical aspects of the basic elements of quantum mechanics - quantum states, unitary operations, measurements, quantum channels, and superchannels (describe the evolution of quantum channels) - which will be used throughout this thesis. I will start by discussing the Hilbert space followed by a discussion on the operators in this space, which will pave the way to the discussion of quantum states and other quantum objects. As a last remark, throughout this thesis, I have used the Dirac notation to represent vectors, matrices, inner products, etc.

2.2.1 Inner product spaces and Hilbert spaces

An inner product space is a vector space A over the field[1] $\mathbb{F}$ of real or complex numbers and equipped with the following map [113, 114]:

$$\langle\,|\,\rangle : A \times A \to \mathbb{F}$$

that satisfies the following three axioms: for all vectors $\psi, \phi \in A$ and all scalars in $\mathbb{F}$, we have

1. Conjugate symmetry: $\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle}$.
2. Linearity in the second argument:$\langle\phi|c_1\psi_1 + c_2\psi_2\rangle = c_1\langle\phi|\psi_1\rangle + c_2\langle\phi|\psi_2\rangle$.
3. Positive definiteness: $\langle\psi|\psi\rangle \geqslant 0$ with equality if and only if $|\psi\rangle = 0$.

The inner product induces a norm:

$$\|\psi\|_2 = \langle\psi|\psi\rangle^{1/2}\,, \tag{2.2}$$

and a metric:

$$d(\psi, \phi) = \|\psi - \phi\|_2\,. \tag{2.3}$$

A norm (denoted $\|\cdot\|$) is a real-valued function defined on the vector space A with the following properties [113, 114]:

1. For all $\psi \in A$, $\|\psi\| \geqslant 0$ with equality iff $|\psi\rangle = 0$.
2. For all $c \in \mathbb{F}$, it holds that $\|c\psi\| = |c|\|\psi\|$.
3. Triangle inequality: $\|\psi + \phi\| \leqslant \|\psi\| + \|\phi\|$.

A vector space equipped with a norm is called a normed space. An example of a norm is a *p-norm*, $\|\cdot\|_p$, defined on all $\psi = (v_1, v_2, \ldots, v_n) \in A$ as

$$\|\psi\|_p := \left(|v_1|^p + |v_2|^p + \cdots + |v_n|^p\right)^{1/p}\,. \tag{2.4}$$

In this thesis, the following two extreme cases with $p = 1$ and $p = \infty$ will be often used:

$$\|\psi\|_1 := |v_1| + |v_2| + \cdots + |v_n|\,, \quad \text{and} \quad \|\psi\|_\infty := \max_{j\in[n]} |v_j|\,.$$

[1] A field is a set together with two operations: addition and multiplication, defined on that set.

where the notation $[n] := \{1, \ldots, n\}$ for an integer $n \in \mathbb{N}$. It is straightforward to see that the 1-norm is the absolute sum of all entries in the vector ψ. To get the expression for the ∞-norm, we can take the vector element with the maximum absolute value as common in Eq. (2.4) and then take the limit $p \to \infty$.

Hilbert Spaces

A Cauchy sequence in an inner product space A is any sequence of vectors $\{\psi_x\}_{x\in\mathbb{N}}$ such that for every positive real number ϵ, there exists an $N \in \mathbb{N}$ such that for all $m, n > N$ it holds that

$$\|\psi_m - \psi_n\| \leqslant \epsilon\,. \tag{2.5}$$

An inner product space A is called complete when all the Cauchy sequences in A converge in A, with respect to the metric induced by the inner product. Complete inner product spaces are called Hilbert spaces [113, 114]. Common examples of Hilbert spaces are $\mathbb{R}^n$ and $\mathbb{C}^n$ (n-dimensional real and complex vector spaces) equipped with the standard notion of the inner product. Another example of a Hilbert space that is relevant in quantum information is the space of $n \times m$ complex matrices denoted $\mathbb{C}^{m\times n}$. In this Hilbert space, the inner product between two elements $M, N \in \mathbb{C}^{m\times n}$ is given by

$$\langle M|N\rangle = \mathrm{Tr}[M^* N] \tag{2.6}$$

where M^* is the adjoint (i.e., the conjugate transpose) matrix of M and $\mathrm{Tr}[\cdot]$ represents the trace of the matrix which is the sum of diagonal entries of the matrix. This inner product is known as the Hilbert-Schmidt inner product and it is sometimes also expressed as $\langle \cdot\,,\, \cdot \rangle_{HS}$.

Hilbert spaces correspond to physical systems operated by parties that will be referred to as Alice, Bob, Charlie, etc. Thus, to represent physical systems in this thesis, I will use letters of the English alphabet. Static quantum systems and their corresponding Hilbert spaces will be denoted by notations like A_0, A_1, B_0, R_0, etc. A vector in Hilbert space A_0 represents the physical state of a system, for example, an electron with a definite spin orientation. These vectors are called pure states, or state vectors, or ket vectors. A detailed description of a general state of a static physical system is provided in Sec 2.2.3. Dynamical systems (i.e., systems that are responsible for the change in the state of a static physical system) and their corresponding Hilbert spaces are denoted by A, B, C, etc. In this setting, the notation for a dynamical system, say A, indicates a pair of systems such that $A = (A_0, A_1) = (A_0 \to A_1)$ where A_0 and A_1 represent the input and output systems, respectively. Note that in this thesis, I have only considered finite-dimensional systems, so we will only refer to finite-dimensional Hilbert spaces.

Two Hilbert spaces can be composed by means of a tensor product. The physical interpretation of the tensor product between two Hilbert spaces is that the subcomponents of the tensor product correspond to individual systems or particles. For a composite system, notation like A_1B_1 will be used to mean $A_1 \otimes B_1$. To represent the dimension of a system, two vertical lines will be used. For example, the dimension of the system A_1 is $|A_1|$. A replica of the same system would be represented by using a tilde symbol. For example, system $\tilde{A}_1$ is a replica of system A_1, and system $\tilde{A}_1\tilde{B}_1$ is a replica of system A_1B_1, ie, $|\tilde{A}_1| = |A_1|$ and $|\tilde{A}_1\tilde{B}_1| = |A_1B_1|$.

2.2.2 Linear Operators in Hilbert Spaces

An operator $M : A_0 \to A_1$ is said to be linear if and only if for all $|\psi\rangle, |\phi\rangle \in A_0$ and $c, d \in \mathbb{F}$

$$M\left(c|\psi\rangle + d|\phi\rangle\right) = cM\left(|\psi\rangle\right) + dM\left(|\phi\rangle\right) \tag{2.7}$$

We will denote the set of all linear operators from system A_0 to A_1 by $\mathfrak{L}(A_0, A_1)$. If the input and output systems are the same, i.e., if $A_0 = A_1$, then we will use the notation $\mathfrak{L}(A_0)$, for brevity. For any operator $M \in \mathfrak{L}(A_0, A_1)$, its kernel, denoted $\mathrm{Ker}(M)$ is the subspace of A_0 consisting of all vectors $|\psi\rangle \in A_0$ such that $M|\psi\rangle = 0$. The image of M, denoted by $\mathrm{Im}(M)$, is the set of vectors $\{M|\psi\rangle\}$ over all vectors $|\psi\rangle \in A_0$. The support of M, denoted $\mathrm{supp}(M)$, is also a subspace of A_0 consisting of all the vectors that are orthogonal to all the elements in $\mathrm{Ker}(M)$. In particular, for any non-zero vector $|\psi\rangle \in \mathrm{supp}(M)$ we have $M|\psi\rangle \neq 0$.

A linear operator $B \in \mathfrak{L}(A_0)$ is called a bounded linear operator if there exists a real number $c > 0$ such that $\|B(|\psi\rangle)\| \leqslant c\||\psi\rangle\|$ for all $|\psi\rangle \in A_0$. For finite dimensions, all linear operators are bounded. The set of all bounded operators on a Hilbert space A_0 will be denoted by $\mathfrak{B}(A_0)$. A linear operator $H \in \mathfrak{L}(A_0)$ is called Hermitian if $H = H^*$. The set of all Hermitian operators on Hilbert space A_0 will be denoted by $\mathrm{Herm}(A_0)$. A linear operator $\rho \in \mathfrak{L}(A_0)$ is positive semi-definite if and only if $\langle\psi|\rho|\psi\rangle \geqslant 0$ for all $|\psi\rangle \in A_0$. This condition implies that ρ must be Hermitian and its eigenvalues must be non-negative. A positive semidefinite operator ρ will be represented as $\rho \geqslant 0$ and the set of all positive semi-definite operators on Hilbert space A_0 will be denoted by $\mathrm{Pos(A_0)}$. In the thesis, I will write $\rho \geqslant \sigma$ to denote $\rho - \sigma \geqslant 0$. Note that $\mathrm{Pos(A_0)} \subset \mathrm{Herm(A_0)} \subset \mathfrak{B}(\mathrm{A_0}) \subset \mathfrak{L}(\mathrm{A_0})$.

Before concluding this subsection on linear operators, let me briefly discuss the norms on linear operators and the isometry between linear operators and bipartite vectors.

Like the inner product of operators in Eq. (2.6), norms, in particular p-norms (Eq. (2.4)), can also be extended from vectors to linear operators. Let $M \in \mathfrak{L}(A_0, A_1)$ for some Hilbert spaces A_0 and A_1. For

$p \in [1, \infty]$, the Schatten p-norm of M is defined as

$$\|M\|_p := \left(\mathrm{Tr}\left[|M|^p\right]\right)^{1/p} \qquad \text{where} \qquad |M| := \sqrt{M^* M}. \tag{2.8}$$

The case $p = 1$ is known as the trace norm which is discussed in Section 2.3.1 and the case $p = \infty$ (understood in terms of the limit $p \to \infty$) is known as the operator norm and is given by the maximum eigenvalue of $|M|$, i.e., $\|M\|_\infty = \lambda_{\max}(|M|)$.

Another property of linear operators that will be used in multiple proofs in this thesis is the isometry between linear operators and bipartite vectors. Let $|\psi\rangle_{A_0 A_1} \in A_0 A_1$ be a bipartite vector in the Hilbert space $A_0 \otimes A_1$. It can be expressed in terms of the orthogonal basis $\{|i\rangle_{A_0} \otimes |j\rangle_{A_1}\}$ as

$$|\psi\rangle_{A_0 A_1} = \sum_{i,j} c_{ij} |i\rangle \otimes |j\rangle \,. \tag{2.9}$$

Let

$$|\Phi^+\rangle_{\tilde{A}_1 A_1} := \sum_j |j\rangle \otimes |j\rangle_{\tilde{A}_1 A_1} \tag{2.10}$$

be a vector in the Hilbert space $\tilde{A}_1 \otimes A_1$. Then, there exists a linear operator $M_\psi \in \mathfrak{L}(\tilde{A}_1,\, A_0)$ such that

$$|\psi\rangle_{A_0 A_1} = M_\psi \otimes I_{A_1} |\Phi^+\rangle_{\tilde{A}_1 A_1} \,. \tag{2.11}$$

This mapping $|\psi\rangle \mapsto M_\psi$ establishes an isometrical isomorphism between the space $A_0 A_1$ and the space $\mathbb{C}^{|A_0| \times |A_1|}$. Below I provide some important properties of this isomorphism which can also be easily proved. Let $|\Phi^+\rangle_{\tilde{A}_1 A_1}$ be defined as above, then:

1. For any linear operator $M \in \mathfrak{L}(A_1)$, it holds that

$$\langle \Phi^+ | M \otimes I_{A_1} | \Phi^+ \rangle = \mathrm{Tr}[M] \,. \tag{2.12}$$

2. Let $M \in \mathfrak{L}(\tilde{A}_1,\, A_0)$ and let its transpose map be denoted as $M^{\mathrm{T}} \in \mathfrak{L}(\tilde{A}_0,\, A_1)$, then it holds that

$$M \otimes I |\Phi^+\rangle_{\tilde{A}_1 A_1} = I \otimes M^{\mathrm{T}} |\Phi^+\rangle_{A_0 \tilde{A}_0} \,. \tag{2.13}$$

3. Let $M, N \in \mathfrak{L}(A_1,\, A_1)$ be invertible matrices, then

$$M \otimes N|\Phi^+\rangle_{A_1\tilde{A}_1} = |\Phi^+\rangle_{A_1\tilde{A}_1} \iff M = \left(N^{-1}\right)^{\mathrm{T}} \tag{2.14}$$

2.2.3 Quantum States

A quantum state is a mathematical description of the physical state of a static quantum system. This description contains complete information about the probability distribution for the outcomes of each possible measurement made on the system [78, 115]. Thus, we can only deduce the probability of a certain outcome when the system interacts with a measurement apparatus, as opposed to a classical system where we can deterministically predict the outcome. This also implies that, if two quantum systems have the same quantum state, we cannot infer that they will behave in the same way when interacting with a measuring apparatus. We can merely state that both systems have the same probability of evolving to a certain state if both systems are treated in the same way.

A quantum state is represented using a square matrix, known as the density matrix which is a unit trace, positive semidefinite operator acting on a Hilbert space. That is, a linear operator $\rho \in \mathfrak{L}(A_0)$is a quantum state if

$$\rho \geqslant 0 \qquad \text{and} \qquad \mathrm{Tr}[\rho] = 1\,. \tag{2.15}$$

Any pure, mixed, classical, separable, entangled, coherent, or any other kind of state, can be represented using a density matrix. The set of all density matrices on Hilbert space A_0 will be denoted by $\mathfrak{D}(A_0)$. We will often write a quantum state $\rho \in \mathfrak{D}(A_0)$ as ρ_{A_0} where the subscript denotes the Hilbert space on which it is acting. The density matrix of a pure state $|\psi\rangle \in A_0$ will be denoted by $\psi := |\psi\rangle\langle\psi|_{A_0}$. Also, a given density matrix $\rho \in \mathfrak{D}(A_0)$ is pure if and only if $\rho = \rho^2$. For mixed states, the density matrices will be denoted by Greek lowercase letters like ρ, σ, τ, etc. throughout this thesis.

For bipartite states, i.e., a quantum state of a composite system of two parties, when we discard one part of the system we still have a quantum state. Mathematically, this process is known as a partial trace. When we have a pure bipartite state $|\psi\rangle_{A_0A_1}$, then using the linear map $M_\psi : A_1 \to A_0$ that is isomorphic to it, we can determine the reduced density matrix. Suppose we trace out the system A_1 and denote the reduced

density matrix of $|\psi\rangle\langle\psi|_{A_0A_1}$ as $\rho^{\psi}_{A_0}$, then we get

$$\rho^{\psi}_{A_0} = \mathrm{Tr}_{A_1}[|\psi\rangle\langle\psi|_{A_0A_1}] = \mathrm{Tr}_{A_1}[M_\psi \otimes I_{A_1}(|\Phi^+\rangle\langle\Phi^+|_{\tilde{A}_1A_1})M^*_\psi \otimes I_{A_1}] \tag{2.16}$$

$$= M_\psi M^*_\psi \tag{2.17}$$

Similarly, we can find out the reduced density matrix $\rho^{\psi}_{A_1}$ that we get after tracing out system A_0. Note that both $\rho^{\psi}_{A_0}$ and $\rho^{\psi}_{A_1}$ have the same non-zero eigenvalues. Moreover, if $\rho \in \mathfrak{D}(A_0)$ is the reduced density matrix of a pure bipartite state $|\psi\rangle_{A_0A_1}$, then $|\psi\rangle_{A_0A_1}$ is called the purification of ρ.

2.2.4 Evolution of closed quantum systems

Quantum mechanics postulates that the evolution of a closed or isolated quantum system is described by unitary evolution. Let $|\psi(0)\rangle \in A_0$ be the state of a system at time $t = 0$. If the system is not interacting with the environment, then the state at time t is given by

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \tag{2.18}$$

where $U(t)$ is the unitary matrix that depends on t. From the Schrödinger equation $\frac{d}{dt}|\psi(t)\rangle = -iH|\psi(0)\rangle$, we then get that

$$U(t) = e^{-iHt} \tag{2.19}$$

where H above is assumed to be a time-independent Hamiltonian which is a Hermitian operator corresponding to the energy of the system. (For time-dependent Hamiltonians, a detailed discussion can be found in standard quantum mechanics textbooks.)

For a general mixed state $\rho \in \mathfrak{D}(A_0)$, we can express its evolution under a unitary U as

$$\sigma = \mathcal{U}(\rho) = U\rho U^* \tag{2.20}$$

where $\mathcal{U}$ denotes the unitary channel that acts on a quantum state by conjugation like $U(\cdot)U^*$.

2.2.5 Measurements

Measurements are required to read the information stored in quantum states. A measurement disturbs the state of the quantum system projecting it into one of the eigenstates of the observable being measured. In

P. A. M. Dirac's words [116]: "A measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured. " This can be interpreted as follows: before the measurement of an observable M is made on the system, the state of the system (assumed pure) can be expressed as a linear combination of the eigenkets of M. When the measurement is performed, the system is projected onto one of the eigenstates of M. A simple example is that of the Stern-Gerlach experiment (with an inhomogeneous magnetic field in the z-direction). In this experiment, the incoming electrons (in arbitrary quantum states) hit the screen only at two spots after passing through the Stern-Gerlach apparatus, thus indicating that each electron is in either of the two possible states of the z-component of the spin - the spin up and the spin down states. Such measurements are known as basis measurements [73]. Basis measurements can be extended to projective measurements, where instead of projecting onto a basis or eigenstate, the projection is made to a subspace. However, projective measurement is not the most general measurement that can be performed on a quantum system. To obtain the most general measurement of a quantum system, we need to combine unitary evolution and projective measurement. Such measurements are known as generalized measurements [47, 74]. Explicit mathematical details are provided later in this section.

Sometimes, we may not care about the post-measurement state of a quantum measurement but only care about the probabilities of the outcomes. Measurement of this sort is specified by using a set of operators, and this set of operators is known as a positive operator-valued measure or POVM [47]. Apart from measurements on static systems, we can also perform generalized measurements on dynamical systems. This is done by specifying an input state to the dynamical system and performing a measurement on the output system. The input state and the measurement operators are written together as a tuple to specify a particular dynamical generalized measurement. Similar to the static case, if we are only concerned with the outcome probabilities in the generalized dynamical measurement, then the tuple of the input state and the set of the measurement operators is known as a process-POVM [117]. With this brief introduction, let us now discuss the definition and properties of these measurements in detail.

1. **Projective measurements:** These are the simplest types of measurements that appear in quantum mechanics. The measurement operators are known as projectors, and they project onto a subspace of a Hilbert space. Stern-Gerlach (SG) experiment is an example of projective measurement. Suppose the SG experiment projects quantum states in spin up $|\uparrow_z\rangle$ and spin down $|\downarrow_z\rangle$ states, then the projectors in this case are

$$M_1 = |\uparrow_z\rangle\langle\uparrow_z|, \quad M_2 = |\downarrow_z\rangle\langle\downarrow_z|. \tag{2.21}$$

Formally, projective measurements on a Hilbert space A_0 consist of a collection of mutually orthogonal

projections $\{P_x\}_{x=1}^m$ satisfying $\sum_{x=1}^m P_x = I_{A_0}$ such that for all $x, y \in [m]$ it holds that

$$P_x P_y = \delta_{xy} P_x \,. \tag{2.22}$$

When all the projection operators are basis states, then the projective measurement is also called basis measurement. However, in general P_x need not be a projection onto a state like $|\psi\rangle\langle\psi|$, rather P_x can be a projector onto a subspace. For example, if we have three-level systems with basis states $|0\rangle, |1\rangle$, and $|2\rangle$, then the projectors might be

$$P_1 = |0\rangle\langle 0| + |1\rangle\langle 1| \,, \qquad P_2 = |2\rangle\langle 2| \,. \tag{2.23}$$

Such projectors are used, for instance, in physical systems with degenerate energy levels.

2. **Generalized measurements, POVM and effects:** Projective measurements project a quantum state onto basis states or in some subspace. Since quantum states evolve unitarily (either as a closed system, or together with the environment in open quantum systems - explained in the next section), we can compose the unitary operators with projective measurements to get generalized measurements. When we are only concerned with the probabilities of the outcomes of these generalized measurements, it is called a positive operator-valued measure, or POVM. Thus, a POVM can be thought of as a machine that takes in a quantum system and yields a classical output.

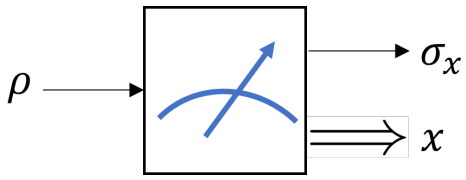

Figure 2.1: Generalized quantum measurement

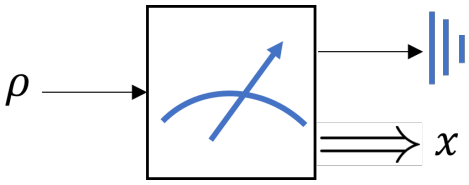

Figure 2.2: POVM

To find the properties of the operators of a generalized measurement, let $|\psi\rangle$ be a quantum state in A_0 that undergoes a unitary evolution by some unitary operator U followed by a projective measurement $\{P_x\}_{x=0}^m$. This process results in a state

$$|\psi_x\rangle := \frac{1}{p_x} P_x U |\psi\rangle \tag{2.24}$$

with probability $p_x = \langle\psi|U^\dagger P_x U|\psi\rangle$. Denoting $P_x U$ as M_x we get

$$|\psi_x\rangle = \frac{1}{p_x} M_x |\psi\rangle \;\text{ with probability }\; p_x = \langle\psi|M_x^\dagger M_x|\psi\rangle \tag{2.25}$$

If ancillary system is also available, then the unitary evolution of the state $|\psi\rangle$ along with the ancilla followed by a projective measurement yields the most general form of measurement known as generalized measurement. Thus, generalized measurements consist of the collection of these complex matrices $\{M_x = P_x U\}$. Formally, a generalized measurement is a collection of m complex matrices $\{M_x\}_{x=1}^m$ such that

$$\sum_{x=1}^{m} M_x^* M_x = I\,. \tag{2.26}$$

To apply the generalized measurements to mixed quantum states, we define $E_x := M_x^* M_x$. Notice that

$$E_x \geqslant 0 \;\text{ and }\; \sum_{x}^{m} E_x = I\,. \tag{2.27}$$

To describe a POVM, we only need to consider the operators $\{E_x\}_x$. These POVM operators E_x, are called effects. Upon applying the generalized measurement $\{M_x\}_{x=1}^m$ to a quantum state ρ, the post-measurement state is

$$\sigma_x = M_x \rho M_x^* \;\text{ with probability }\; p_x = \mathrm{Tr}[M_x^* M_x \rho] = \mathrm{Tr}[E_x \rho]\,. \tag{2.28}$$

Thus, to every generalized measurement, there exists a unique POVM that corresponds to it via the relation $E_x = M_x^* M_x$. However, the converse is not true, that is, for every POVM, there are many quantum measurements associated with it.

3. **Process POVM:** Above, we discussed measurements of quantum states to read quantum information. Since quantum channels are the most general quantum objects, we want to understand and read quantum information from quantum channels[2]. To do this, the idea of POVMs was generalized from states to channels in [117]. Let $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$ be a quantum channel. To perform a process POVM, a state $\rho \in \mathfrak{D}(A_0 R_0)$ is prepared, and the A_0 part of the system is evolved by the action of $\mathcal{N}_A$. Then, a POVM $\{E_x\}$ is applied to the resultant state. Then the probability of some outcome is

[2]This discussion requires the understanding of quantum channels and its representations which is discussed in detail Sec.2.2.6. So, it is suggested to revisit this discussion after going through Sec. 2.2.6.

expressed as

$$p(\mathcal{N}) = \mathrm{Tr}\left[(\mathsf{id} \otimes \mathcal{N})(\rho)\, E_x\right] \tag{2.29}$$

$$= \mathrm{Tr}\left[(\mathcal{M}_\rho \otimes \mathcal{N}_A)\left(\Phi^+\right)(E_x)\right] \tag{2.30}$$

$$= \mathrm{Tr}\left[J_A^{\mathcal{N}}\left(\mathcal{M}_\rho^* \otimes \mathsf{id}(E_x)\right)\right] \tag{2.31}$$

where the first equality follows by construction where the channel $\mathcal{N}$ acts on a part of the input state followed by the action of a POVM $\{E_x\}$. The second equality follows from the fact that for any bipartite vector, say $|\psi_{A_0B_0}\rangle$, there is a corresponding map $M_\psi : \tilde{B}_0 \to A_0$ such that $|\psi_{A_0B_0}\rangle = M_\psi \otimes I_{B_0}|\Phi^+_{\tilde{B}_0B_0}\rangle$. Since we can express any density matrix, say σ, as $\sigma = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ using some pure states $|\phi_i\rangle$ and where $\sum_i p_i = 1$ and $p_i \geqslant 0$, we can write $\sigma = \mathcal{M}_\sigma \otimes \mathsf{id}(|\Phi^+\rangle\langle\Phi^+|)$ where $\mathcal{M}_\sigma(\cdot) = \sum_i p_i M_{\phi_i}(\cdot)M_{\phi_i}$. The third equality follows from the definition of the Choi matrix of a channel. In the last equation, we can define the effects F_x of process POVM as $F_x := (\mathcal{M}_\rho \otimes \mathsf{id}(E_x))$. These effects are positive semidefinite, that is, $F_x \geqslant 0$, and sum to $\rho^T \otimes I$, i.e., $\sum_x F_x = \sum_x \mathcal{M}^* \otimes \mathsf{id}(E_x) = \mathcal{M}^* \otimes \mathsf{id}(\sum_x E_x) = \rho^T \otimes I$.

2.2.6 Quantum channels

Quantum channels describe the most general evolution of a quantum state. Until now (i.e., in Sec. 2.2.4 and 2.2.5), we have seen two kinds of evolution of a static quantum system: the unitary evolution and generalized measurements. Unitary evolution describes the evolution of an isolated quantum system, whereas measurements map a quantum state to a classical state. However, a quantum system can evolve in various other ways. For instance, while transferring quantum information (encoded in a quantum system) from point A to point B, the environment can act on the system, thereby introducing noise and changing the state of the system. This noise can be modeled as a quantum channel. Even noiseless transfer of quantum information is a quantum channel which is the trivial identity channel. Moreover, even a quantum state can be viewed as a quantum channel with a trivial input and a particular quantum state as an output. The reverse of this process, i.e. mapping a quantum system to a trivial system, is also a quantum channel and is known as tracing out a system. So, to account for any kind of change in the state of a quantum system, we need a universal mathematical formulation of quantum channels. Mathematically, quantum channels are linear maps that are completely positive and trace-preserving [47, 74]. Simply put, quantum channels are maps that take density matrices in $\mathfrak{D}(A_0)$ to density matrices in $\mathfrak{D}(A_1)$ in a complete sense, that is, even when acting on part of the input system.

We will denote linear maps that take linear operators in $\mathfrak{L}(A_0)$ to linear operators in $\mathfrak{L}(A_1)$ as $\mathfrak{L}(A_0 \to A_1)$. (Note the difference in the notation with $\mathfrak{L}(A_0, A_1)$ that denotes linear operators taking vectors in A_0 to vectors in A_1). Since the set of density matrices consist of unit trace, positive semi-definite operators, the set of quantum channels that transform density operators in $\mathfrak{D}(A_0)$ to those in $\mathfrak{D}(A_1)$ are a subset of $\mathfrak{L}(A_0 \to A_1)$. In this thesis, I have used calligraphic letters like $\mathcal{E}, \mathcal{F}, \mathcal{M}$, or $\mathcal{N}$ to denote quantum channels. Below, I list the properties that are required for a map to be a quantum channel:

1. **Linearity:** Any quantum channel must be described by a linear map. This is an essential requirement, as we want the action of the quantum channel to be convex[3], and should always give the same result for the evolution of a quantum state no matter what convex combination is used to express the original state. Let $\mathcal{N}$ be a quantum channel and ρ, σ be two density matrices, then the following must hold

$$\mathcal{N}(p\rho + (1-p)\sigma) = p\mathcal{N}(\rho) + (1-p)\mathcal{N}(\sigma) \tag{2.32}$$

   Note that the above holds not just for density operators but for all linear operators.

2. **Complete positivity:** Linear maps that describe a quantum channel must preserve the positivity of any input positive operator even when they act on part of the operator. This property is known as complete positivity. Since quantum states represented by density matrices are positive semidefinite operators, quantum channels preserve positivity. Moreover, given a bipartite state, if a quantum channel acts only on one part of the state, the output is still a quantum state, regardless of the dimension of the other system on which the channel did not act. Let $\mathcal{N} \in \mathcal{L}(A_0 \to A_1)$ and let $\rho \in \mathrm{Pos}(\mathrm{R}_0\mathrm{A}_0)$, then $\mathcal{N}$ is completely positive if

$$\mathsf{id} \otimes \mathcal{N}(\rho) \in \mathrm{Pos}(\mathrm{R}_0\mathrm{A}_1) \tag{2.33}$$

   for any dimension $|R_0|$.

3. **Trace preserving:** Quantum channels map a density matrix to density matrix. Since any Hermitian matrix can be expressed as a combination of density matrices, if we evolve this Hermitian matrix under a quantum channel and take the trace, the trace of the output system is equal to the trace of the input system. This can be extended further for linear operators. Such maps which preserve the trace are

[3] Action of a map being convex means it takes a convex set to a convex set. A convex set is a set where the convex combination of any two points in the set also belongs to the set. A convex combination of elements is a linear combination of elements where the coefficients are non-negative and sum to one. See Sec. 2.4 for details.

called trace preserving. Let $\mathcal{N} \in \mathfrak{L}(A_0 \to A_1)$ and let $\rho \in \mathfrak{L}(A_0)$, then $\mathcal{N}$ is trace preserving if

$$\mathrm{Tr}[\rho] = \mathrm{Tr}[\mathcal{N}(\rho)]\,. \tag{2.34}$$

The set of all such linear maps, which are completely positive (CP) and trace-preserving (TP), constitutes the set of quantum channels. Quantum channels are thus also referred to as linear CPTP maps, or just CPTP maps. In this thesis, I have denoted the set of all quantum channels or CPTP maps taking density matrices in $\mathfrak{D}(A_0)$ to density matrices in $\mathfrak{D}(A_1)$ by $\mathrm{CPTP}(A_0 \to A_1)$ or $\mathrm{CPTP}(A)$, where A denotes the dynamical system mapping one static system to another. The identity channel (which is equivalent to not doing anything) on a quantum state in $\mathfrak{D}(A_0)$ will be denoted by id_{A_0}. Also, we will use subscripts to denote the dynamical physical system of the quantum channel. For instance, $\mathcal{E}_A$ represents a quantum channel $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$. Besides, the set of linear maps that are completely positive (and not trace preserving) will be denoted by $\mathrm{CP}(A_0 \to A_1)$ and it is clear that $\mathrm{CPTP}(A_0 \to A_1) \subset \mathrm{CP}(A_0 \to A_1) \subset \mathfrak{L}(A_0 \to A_1)$. Some examples of quantum channels that are relevant to this thesis include the completely dephasing channel which kills all the coherence of a quantum system with respect to some fixed basis, classical channels which converts one probability distribution to another, completely stabilizer-preserving operations which completely preserve the set of all stabilizer states, etc. All these channels will be discussed in detail in the upcoming chapters. Now, let us briefly discuss the ways to mathematically represent a quantum channel.

Representations of quantum channels

There are three main ways to represent a quantum channel.

1. **The Choi representation:** Also known as the Choi matrix or Choi-Jamiolkowski isomorphism, this representation expresses a quantum channel as a positive semidefinite matrix. The evolution of a quantum state under a quantum channel is represented as a product of matrices followed by a partial trace. For a quantum channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$, its Choi matrix will be represented as $J_A^{\mathcal{N}}$ where the superscript denotes the channel and the subscript denotes the dynamical system. The Choi matrix $J_A^{\mathcal{N}}$ of a channel $\mathcal{N}_A$ is defined as

$$J_A^{\mathcal{N}} := \mathsf{id}_{A_0} \otimes \mathcal{N}_A(\Phi^+_{A_0\tilde{A}_0}) \tag{2.35}$$

where id_{A_0} is the identity channel and $\Phi^+_{A_0\tilde{A}_0} := |\Phi^+\rangle\langle\Phi^+|_{A_0\tilde{A}_0} = \sum_{x,y}^{|A_0|} |x\rangle\langle y|_{A_0} \otimes |x\rangle\langle y|_{\tilde{A}_0}$ is the unnormalized maximally entangled state [75, 118]. For a matrix $J_A^{\mathcal{N}}$ to represent a quantum channel, it must satisfy the following properties:

(a) A linear map is completely positive if and only if its Choi matrix is positive semidefinite. Thus, the Choi matrix representing a channel $\mathcal{N}_A$ must obey

$$J_A^{\mathcal{N}} \geqslant 0\,. \tag{2.36}$$

(b) A linear map $\mathcal{N}_A \in \mathfrak{L}(A_0 \to A_1)$ is trace preserving if and only if the marginal of its Choi matrix is

$$J_{A_0}^{\mathcal{N}} := \mathrm{Tr}_{A_1}[J_A^{\mathcal{N}}] = I_{A_0}\,. \tag{2.37}$$

Due to the nature of the above conditions and the fact that the Choi matrix is a positive semidefinite matrix, the Choi representation is very helpful in characterizing various resources and expressing conditions for optimization problems as a semidefinite program, among other things.

Lastly, the evolution of a quantum state $\rho_{A_0} \mapsto \mathcal{E}_A(\rho)$ under a quantum channel $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ can be expressed as

$$\mathcal{E}_A(\rho_{A_0}) = \mathrm{Tr}_{A_0}\left[J_A^{\mathcal{E}}\left(\rho_{A_0} \otimes I_{A_1}\right)\right]\,. \tag{2.38}$$

2. **The Kraus decomposition:** Also known as the operator-sum representation, it expresses a quantum channel, say $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$, using a set of Kraus operators $\{M_x\}_{x=1}^m$ with $M_x : A_0 \to A_1$ which obey

$$\sum_x M_x^* M_x = I_{A_0}\,. \tag{2.39}$$

The evolution of a quantum state ρ_{A_0} under a channel $\mathcal{N}_A$ using the Kraus representation is given as

$$\mathcal{N}(\rho) = \sum_x M_x \rho M_x^*\,. \tag{2.40}$$

Since the Choi matrix of a quantum channel can be written as

$$J_A^{\mathcal{N}} = \sum_{x=1}^m |\psi_x\rangle\langle\psi_x| \tag{2.41}$$

for some integer m and some (possibly unnormalized) states in $A_0 \otimes A_1$, the Kraus operators can be

found from the isomorphism between bipartite vectors and linear operators, i.e.,

$$|\psi_x\rangle = M_x \otimes I|\Phi^+\rangle\,. \tag{2.42}$$

3. **The unitary representation:** A system together with the environment can be viewed as a closed system and the quantum state of system + environment can then be seen as a pure state [47]. The dynamics of this whole system is now governed by unitary evolution. By choosing an appropriate unitary and tracing out some systems, the action of any quantum channel on the initial system can be simulated. This way of representing a quantum channel is known as the unitary representation. Considering the environment as $|0\rangle\langle 0|_{E_0}$, the action of a quantum channel $\mathcal{N}_A$ on a quantum state ρ_{A_0} can be expressed in the unitary representation as follows

$$\mathcal{N}(\rho) = \mathrm{Tr}_{E_0}\left[U\left(\rho_{A_0} \otimes |0\rangle\langle 0|_{E_0}\right)U^*\right]\,. \tag{2.43}$$

This can be expressed more generally by the Stinespring Dilation theorem. For finite dimensions, this theorem states that a linear map $\mathcal{N} \in \mathrm{CPTP}(A)$ if and only if there exists an environment E_0 with dimension $|E_0| \leqslant |A_0A_1|$ and an isometry $V : A_0 \to A_1 \otimes E_0$ with $V^*V = I_{A_1}$ such that

$$\mathcal{N}(\rho_{A_0}) := \mathrm{Tr}_{E_0}[V\rho V^*]\,. \tag{2.44}$$

To get the Kraus operators from the above equation, we express the above equation in terms of the basis states of system E_0 as

$$\mathcal{N}(\rho_{A_0}) = \sum_{x=1}^{|E_0|}\langle x_{E_0}|V\rho V^*|x_{E_0}\rangle \tag{2.45}$$

Some examples of quantum channels

Let us now look at a few examples of quantum channels which are relevant for this thesis.

1. **Unitary channels:** The evolution of a mixed state $\rho \in \mathfrak{D}(A_0)$ under a unitary U can be expressed as

$$\sigma = \mathcal{U}(\rho) = U\rho U^* \tag{2.46}$$

where $\mathcal{U}$ denotes the unitary channel that acts on a quantum state by conjugation like $U(\cdot)U^*$. (Refer to Sec. 2.2.4 for details of the evolution of a closed or isolated quantum system, and the previous

subsection, Sect. 2.2.6, for the evolution of open quantum systems and how such an evolution can be modeled using a unitary channel.)

2. **Completely dephasing channels:**

   These channels completely remove the coherence of any quantum state with respect to some fixed orthogonal basis. Mathematically, the action of these channels corresponds to the removal of the off-diagonal elements of a density matrix $\rho \in \mathfrak{D}(A_0)$ when the matrix is expressed with respect to some orthogonal basis $\{|x\rangle\}_{x=1}^{|A_0|}$. Throughout this thesis, I will be using $\mathcal{D}$ to denote a completely dephasing channel. Note that $\mathcal{D} \in \mathrm{CPTP}(A_0 \to A_0)$. The action of a completely dephasing channel $\mathcal{D}$ on a state $\rho \in \mathfrak{D}(A_0)$ is given as

$$\mathcal{D}(\rho_{A_0}) = \sum_{x=1}^{|A_0|} |x\rangle\langle x|\rho|x\rangle\langle x| \,. \tag{2.47}$$

   The Choi matrix of a completely dephasing map can be easily found and which is given by

$$J^{\mathcal{D}} = \sum_{x} |x\rangle\langle x| \otimes |x\rangle\langle x| \,. \tag{2.48}$$

   For a more detailed discussion and physical meaning of this channel, refer to Chapter 3.

3. **Classical channels:** Classical channel or a classical communication channel describes a physical transmission medium to convey information signal, for instance, a bit stream, from one or more senders to one or more receivers [74]. (See Fig. 2.3 for a basic illustration portraying multiple senders, a communication channel, and multiple receivers.) A communication channel is usually modeled as a

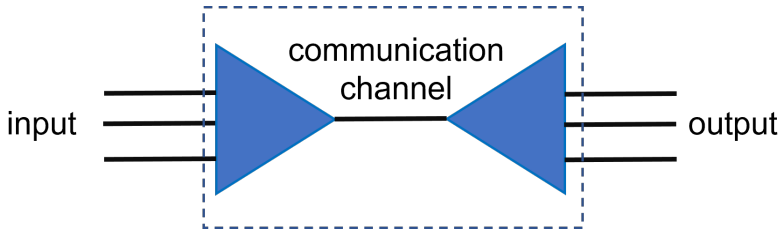

Figure 2.3: A general classical communication scenario

   triple consisting of an input alphabet, an output alphabet, and a transition probability for each pair (i, o) of input-output elements. Using this transition probability, a classical channel $\mathcal{N}$ is represented as a conditional probability distribution involving an input random variable X and an output random

variable Y:

$$\mathcal{N} : p_{Y|X}(y|x)\,, \tag{2.49}$$

which can be expressed as a column stochastic matrix, say N. The dimensions of N depend on the size of the input and output alphabets.

Since quantum channels describe the most general evolution of a quantum system and classical systems are a special type of quantum systems, we can express classical channels in terms of quantum channels. To define classical channels as quantum channels, we need to understand how they act on classical states, i.e., states diagonal in a chosen basis. Since classical channels do not make use of the coherence of any input quantum state (and in fact kill the coherence of the input state) and cannot generate coherence, classical channels remain invariant under conjugation by completely dephasing channels. Thus, we define a classical channel as a quantum channel that remains invariant under conjugation by completely dephasing channels (with appropriate dimensions). Therefore, a channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ is a classical channel if and only if:

$$\mathcal{N}_{A_0\to A_1} = \mathcal{D}_{A_1} \circ \mathcal{N}_{A_0\to A_1} \circ \mathcal{D}_{A_0}\,, \tag{2.50}$$

where the symbol '$\circ$' represents the composition of channels. With this definition, it is easy to see that the Choi matrix $J^{\mathcal{N}}$ of a classical channel $\mathcal{N}$ is diagonal. Furthermore, the correspondence between the column stochastic matrix N of $\mathcal{N}$ and the Choi matrix $J^{\mathcal{N}}$ is that the Choi matrix of $\mathcal{N}$ is just the columns of N stacked one after the other as diagonals of J^N.

4. **POVM channels:** A POVM channel or a quantum to classical channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ obeys the following equation

$$\mathcal{N} = \mathcal{D} \circ \mathcal{N} \tag{2.51}$$

Following some calculations using the above definition, we get the following form for a POVM channel acting on a state ρ

$$\mathcal{N}(\rho) = \sum_{x=1}^{|A_1|} \mathrm{Tr}[E_x\rho]|x\rangle\langle x| \tag{2.52}$$

where $\{E_x\}_x$ is a POVM and thus each $E_x \in \mathrm{Pos(A_0)}$ and $\sum_{x=1}^{|A_1|} E_x = I_{A_1}$

2.2.7 Supermaps and Superchannels

The main focus of this thesis is on the interconversion among dynamical resources or quantum channels. This section establishes the mathematical framework of linear maps, called superchannels, that describe the evolution of a quantum channel. As previously noted that quantum channels can be regarded as the most fundamental objects in quantum mechanics, one can now remark that, since superchannels describe the evolution of quantum channels, why not regard them as more fundamental objects with the same line of reasoning that we adopted for quantum channels. The answer to this lies in the way we realize a superchannel and that is through quantum channels (see Fig. 2.4). Hence, it is nice to have a mathematical model for the most general way to evolve quantum channels, but they cannot be treated as more fundamental, and quantum channels maintain their position of being the most fundamental object in quantum mechanics. Before discussing superchannels, I first discuss the math behind a set of linear maps called supermaps, as superchannels are special types of supermaps (see [119] and references within).

The space $\mathfrak{L}(A_0 \to A_1)$ is equipped with the following inner product

$$\langle \mathcal{N}_A, \mathcal{M}_A \rangle \coloneqq \sum_{i,j} \langle \mathcal{N}_A \left(|i\rangle\langle j|_{A_0} \right), \mathcal{M}_A \left(|i\rangle\langle j|_{A_0} \right) \rangle_{HS} \tag{2.53}$$

where $\langle X, Y \rangle_{HS} \coloneqq \mathrm{Tr}[X^* Y]$ is the Hilbert-Schmidt inner product between the matrices $X, Y \in \mathfrak{B}(A_1)$. The above inner product is independent of the choice of the orthonormal basis $\{|i\rangle\langle j|\} \in \mathfrak{B}(A_0)$, and can be expressed in terms of Choi matrices. The Choi matrix of a channel $\mathcal{N}_A$ is given by

$$J^{\mathcal{N}}_{A_0 A_1} = \mathcal{N}_{\tilde{A}_0 \to A_1} \left(\Phi^+_{A_0 \tilde{A}_0} \right) \tag{2.54}$$

where $\Phi^+_{A_0\tilde{A}_0} \equiv |\Phi^+\rangle\langle\Phi^+|_{A_0\tilde{A}_0}$ is an unnormalized maximally entangled state where $|\Phi^+\rangle_{A_0\tilde{A}_0} \equiv \sum_i^{|A_0|} |i\rangle_{A_0} |i\rangle_{\tilde{A}_0}$. With this notation, the inner product of two channels $\mathcal{N}_A$ and $\mathcal{M}_A$ can be expressed as the inner product of their Choi matrices, i.e.,

$$\langle \mathcal{N}_A, \mathcal{M}_A \rangle = \left\langle J^{\mathcal{N}}_A, J^{\mathcal{M}}_A \right\rangle = \mathrm{Tr}\left[\left(J^{\mathcal{N}}_A \right)^* J^{\mathcal{M}}_A \right]. \tag{2.55}$$

The canonical orthonormal basis of $\mathfrak{L}(A)$ (relative to the above inner product) is given by $\left\{ \mathcal{E}^{ijkl}_A \right\}$ where

$$\mathcal{E}^{ijkl}_A (\rho_{A_0}) = \langle i|\rho_{A_0}|j\rangle \, |k\rangle\langle l|_{A_1} \quad \forall \, \rho_{A_0} \in \mathfrak{B}(A_0). \tag{2.56}$$

The space $\mathfrak{L}(A \to B)$ (where $A = (A_0, A_1)$ and $B = (B_0, B_1)$) is equipped with the following inner

product

$$\langle \Theta_{A\to B}, \Omega_{A\to B}\rangle \coloneqq \sum_{i,j,k,l} \left\langle \Theta_{A\to B}\left[\mathcal{E}_A^{ijkl}\right], \Omega_{A\to B}\left[\mathcal{E}_A^{ijkl}\right]\right\rangle \ , \tag{2.57}$$

where $\Theta_{A\to B},\ \Omega_{A\to B} \in \mathfrak{L}(A\to B)$ are called supermaps, and the inner product on the right-hand side is the inner product between maps as defined in (2.53). The dual of a supermap $\Theta \in \mathfrak{L}(A\to B)$ is a linear map $\Theta^* \in \mathfrak{L}(B\to A)$ with the property

$$\langle \mathcal{N}_B, \Theta\left[\mathcal{M}_A\right]\rangle = \langle \Theta^*\left[\mathcal{N}_B\right], \mathcal{M}_A\rangle \ , \tag{2.58}$$

for all $\mathcal{M}_A \in \mathfrak{L}(A)$ and for all $\mathcal{N}_A \in \mathfrak{L}(B)$.

Similar to how we can express the inner product of two maps by the inner product of their Choi matrices, we can define the inner product of two supermaps as the inner product of their Choi matrices as well. The Choi matrix of a supermap $\Theta_{A\to B}$ is defined as [119]

$$\mathbf{J}_{AB}^{\Theta} = \sum_{i,j,k,l} J_A^{\mathcal{E}^{ijkl}} \otimes J_B^{\Theta[\mathcal{E}^{ijkl}]} \tag{2.59}$$

where $J_A^{\mathcal{E}^{ijkl}}$ and $J_B^{\Theta[\mathcal{E}^{ijkl}]}$ are the Choi matrices of $\mathcal{E}_A^{ijkl}$ and $\Theta_{A\to B}[\mathcal{E}_A^{ijkl}]$, respectively. With this notation, the inner product between two supermaps $\Theta_{A\to B}$ and $\Omega_{A\to B}$ can be expressed as

$$\langle \Theta_{A\to B}, \Omega_{A\to B}\rangle = \left\langle \mathbf{J}_{AB}^{\Theta}, \mathbf{J}_{AB}^{\Omega}\right\rangle_{HS} = \mathrm{Tr}\left[\left(\mathbf{J}_{AB}^{\Theta}\right)^* \mathbf{J}_{AB}^{\Omega}\right] \tag{2.60}$$

We now give three alternative expressions of the Choi matrix of the supermap $\Theta \in \mathfrak{L}(A\to B)$ [119]. First, from its definition, the Choi matrix of a supermap uses the CP map analog of the entangled states which we represent as $\mathcal{P}_{A\tilde{A}}^{+}$ and is given by

$$\mathcal{P}_{A\tilde{A}}^{+} = \sum_{i,j,k,l} \mathcal{E}_{A_0\to A_1}^{ijkl} \otimes \mathcal{E}_{\tilde{A}_0\to \tilde{A}_1}^{ijkl} \ . \tag{2.61}$$

Similar to the properties of the maximally entangled state, the channel $\mathcal{P}_{A\tilde{A}}^{+}$ satisfies the following relation for any $\Theta \in \mathfrak{L}(A\to B)$

$$\Theta_{\tilde{A}\to B}[\mathcal{P}_{A\tilde{A}}^{+}] = \Theta_{\tilde{B}\to A}^{T}[\mathcal{P}_{B\tilde{B}}^{+}] \tag{2.62}$$

where $\Theta^T \in \mathfrak{L}(B\to A)$ is the transpose of the supermap Θ which is defined by its components

$$\left\langle \mathcal{E}_A^{ijkl}, \Theta^T\left[\mathcal{E}_B^{i'j'k'l'}\right]\right\rangle = \left\langle \mathcal{E}_B^{i'j'k'l'}, \Theta\left[\mathcal{E}_A^{ijkl}\right]\right\rangle \quad \forall\ i,j,k,l,i',j',k',l' \tag{2.63}$$

where $\left\{\mathcal{E}_A^{ijkl}\right\}$ and $\left\{\mathcal{E}_B^{i'j'k'l'}\right\}$ are the canonical orthonormal basis of $\mathfrak{L}(A)$ and $\mathfrak{L}(B)$, respectively. Then, the Choi matrix of a supermap $\Theta \in \mathfrak{L}(A \to B)$ can be expressed as

$$\mathbf{J}_{AB}^{\Theta} = \Theta\left[\mathcal{P}_{A\tilde{A}}^{+}\right]\left(\Phi_{A_0\tilde{A}_0}^{+} \otimes \Phi_{B_0\tilde{B}_0}^{+}\right) \tag{2.64}$$

The second way of defining the Choi matrix of a supermap is by its action on the Choi matrices of channels. Let's consider a linear map Θ such that for $\mathcal{M}_A \in \mathfrak{L}(A)$ and $\mathcal{N}_B \in \mathfrak{L}(B)$, $\mathcal{N}_B = \Theta_{A\to B}[\mathcal{M}_A]$. Then the Choi matrices of $\mathcal{M}_A$ and $\mathcal{N}_A$ are related via

$$J_B^{\mathcal{N}} = \mathrm{Tr}_A\left[\mathbf{J}_{AB}^{\Theta}\left(\left(J_A^{\mathcal{M}}\right)^T \otimes I_B\right)\right] \tag{2.65}$$

That is, $\mathbf{J}_{AB}^{\Theta}$ can be interpreted as the Choi matrix of a linear map (say $\mathcal{R}_{A\to B}^{\Theta}$) that converts $J_A^{\mathcal{M}}$ to $J_B^{\mathcal{N}}$ and we can write

$$\mathcal{R}_{A\to B}^{\Theta}(J_A^{\mathcal{M}}) = J_B^{\mathcal{N}}\ . \tag{2.66}$$

For the last representation of the Choi matrix of a supermap, we can view it as a linear map $\mathcal{Q}^{\Theta}$: $\mathfrak{B}(A_1B_0) \to \mathfrak{B}(A_0B_1)$ which is defined by the map satisfying

$$\mathbf{J}_{AB}^{\Theta} \coloneqq \mathcal{Q}_{\tilde{A}_1\tilde{B}_0\to A_0B_1}^{\Theta}(\Phi_{A_1\tilde{A}_1}^{+} \otimes \Phi_{B_0\tilde{B}_0}^{+}). \tag{2.67}$$

We will see that the three representations play a useful role in our study of dynamical resource theory of coherence.

Now, let us define a superchannel. A superchannel is a supermap that takes quantum channels to quantum channels even when tensored with an identity supermap [119, 120, 121, 122, 123, 124, 125, 126]. In other words, a superchannel Θ describes the evolution of a quantum channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ to a target channel $\mathcal{M} \in \mathrm{CPTP}(B_0 \to B_1)$ as

$$\Theta_{A\to B}[\mathcal{N}_A] = \mathcal{M}_B \tag{2.68}$$

and even when acting on part of the channel as

$$\mathbb{1}_R \otimes \Theta_{A\to B}[\mathcal{N}_{AR}] = \mathcal{M}_{BR} \tag{2.69}$$

where $\mathcal{N}_{AR} \in \mathrm{CPTP}(A_0R_0 \to A_1R_1)$, $\mathcal{M}_{BR} \in \mathrm{CPTP}(B_0R_0 \to B_1R_1)$, and $\mathbb{1}_R$ denotes the identity superchannel that takes the dynamical system R to R. For a linear map $\Theta \in \mathfrak{L}(A \to B)$ describing a superchannel, the following are equivalent [119, 121]:

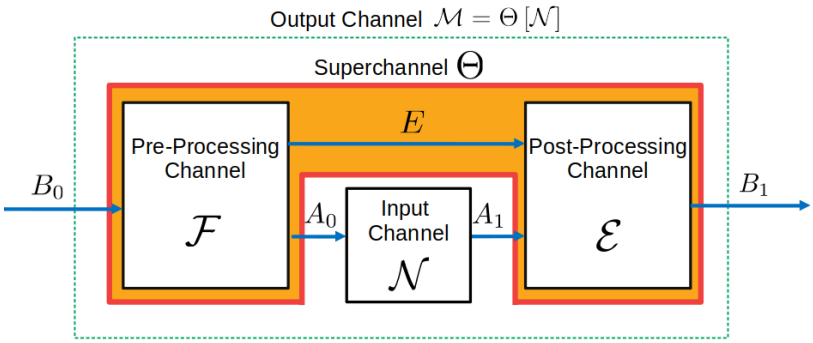

Figure 2.4: Realization of a superchannel in terms of pre- and post-processing channels and its action on an input channel $\mathcal{N}$

1. Θ is a superchannel

2. The Choi matrix $\mathbf{J}^{\Theta}_{AB} \geqslant 0$ with marginals

$$J^{\Theta}_{A_1B_0} = I^{A_1B_0} \quad ; \quad J^{\Theta}_{AB_0} = J^{\Theta}_{A_0B_0} \otimes u_{A_1} \tag{2.70}$$

where $u_{A_1} = \frac{I_{A_1}}{|A_1|}$ is the maximally mixed state for system A_1.

3. The map $\mathcal{R}^{\Theta}_{A\to B}$ in (2.66) is CP, and there exists a unital CP map $\mathcal{R}^{\Theta}_{A_0\to B_0}$ such that the map $\mathcal{R}^{\Theta}_{A\to B_0} \equiv \mathrm{Tr}_{B_1} \circ \mathcal{R}^{\Theta}_{A\to B}$ satisfies

$$\mathcal{R}^{\Theta}_{A\to B_0} = \mathcal{R}^{\Theta}_{A_0\to B_0} \circ \mathrm{Tr}_{A_1} \tag{2.71}$$

Note that a channel is unital iff both marginals of its Choi matrix are equal to the identity matrix.

4. There exists a Hilbert space E, with $|E| \leqslant |A_0B_0|$, and two CPTP maps $\mathcal{F} \in \mathrm{CPTP}(B_0 \to A_0E)$ and $\mathcal{E} \in \mathrm{CPTP}(A_1E \to B_1)$ such that for all $\mathcal{N}_A \in \mathfrak{L}(A_0 \to A_1)$

$$\Theta[\mathcal{N}_A] = \mathcal{E}_{A_1E\to B_1} \circ \mathcal{N}_{A_0\to A_1} \circ \mathcal{F}_{B_0\to A_0E} \tag{2.72}$$

This means that a superchannel can be realized in terms of a pre- and a post-processing channel (see Figure 2.4). Moreover, the transformation of Eq. (2.68) can be expressed using Choi matrices of channels $\mathcal{N}$, $\mathcal{M}$, and the superchannel Θ as

$$J^{\mathcal{M}}_B = \mathrm{Tr}_A \left[\mathbf{J}^{\Theta}_{AB} \left(\left(J^{\mathcal{N}}_A\right)^T \otimes I_B\right)\right] . \tag{2.73}$$

2.3 Distance measures in quantum information

Quantum systems suffer from noise in practice, and the outcome of a protocol might be different from what is desired. Therefore, it is desirable to find out how well a protocol is performing. The easiest way to do that is to compare the output of the ideal protocol with the actual output of the protocol using a *distance measure* of the two states. In this section, I discuss distance measures between quantum states, specifically trace distance, fidelity, and relative entropies. Their generalizations to quantum channels are also mentioned and are discussed in detail in the chapters when I discuss specific resource theories.

2.3.1 Trace Norm, Trace Distance, and Diamond norm

The trace norm or ℓ_1-norm of an element X in the Hilbert space $\mathbb{C}^{m\times n}$ is defined as [47, 74, 127, 128]:

$$\|X\|_1 := \mathrm{Tr}[|X|] = \mathrm{Tr}\left[\sqrt{X^*X}\right] . \tag{2.74}$$

This implies that the trace norm of X is the sum of its singular values. Let U and V be isometries, then from the above definition we get that

$$\|UXV^*\|_1 = \|X\|_1 . \tag{2.75}$$

For any Hermitian matrix $H \in \mathrm{Herm}(A_0)$, its trace norm is the absolute sum of its eigenvalues. Let

$$H = \sum_x \lambda_x |v_x\rangle\langle v_x| \tag{2.76}$$

be the eigenvalue decomposition of H. Then the trace norm of H is

$$\|H\|_1 = \sum_x |\lambda_x| . \tag{2.77}$$

The trace norm of a Hermitian matrix can also be expressed as an optimization problem. To see that, let $H_+ := \sum_{x:\lambda_x \geqslant 0} \lambda_x |v_x\rangle\langle v_x|$ and $H_- := \sum_{x:\lambda_x<0} |\lambda_x||v_x\rangle\langle v_x|$ and the projectors on the positive eigenspace and negative eigenspace of H be defined as $P_+ := \sum_{x:\lambda_x\geqslant 0} |v_x\rangle\langle v_x|$ and $P_- := \sum_{x:\lambda_x<0} |v_x\rangle\langle v_x|$, respectively. Since $H = H_+ - H_-$, the trace norm of H becomes

$$\|H\|_1 = \mathrm{Tr}[\|H\|] = \mathrm{Tr}[H_+] + \mathrm{Tr}[H_-] = \mathrm{Tr}[HP_+] - \mathrm{Tr}[HP_-] = \max_{-I\leqslant P\leqslant I} \mathrm{Tr}[HP] . \tag{2.78}$$

where the maximum is over Hermitian matrices whose eigenvalues are between -1 and 1. This optimization problem belongs to the class of convex optimization problems known as semidefinite programs (discussed in detail in 2.4.2) and can be computed efficiently.

The trace distance between two quantum states $\rho \in \mathfrak{D}(A_0)$ and $\sigma \in \mathfrak{D}(A_0)$ is defined as

$$D(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 \tag{2.79}$$

where the factor $1/2$ is there for normalization so that the trace distance between two quantum states achieves its maximum value when the two states are orthogonal. The following bounds apply to the trace distance between two quantum states, ρ and σ

$$0 \leqslant D(\rho, \sigma) \leqslant 1\,. \tag{2.80}$$

When the trace distance between two states is equal to zero, it implies that the two states are equal. When the trace distance between two states is maximum, i.e., equal to one, it implies that there exists a measurement that can perfectly distinguish ρ from σ. It can be easily shown that for pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$, the trace distance can be expressed as

$$\frac{1}{2}\||\psi\rangle - |\phi\rangle\|_1 = \sqrt{1 - |\langle\psi|\phi\rangle|^2}\,. \tag{2.81}$$

Like Hermitian matrices, the trace distance between two quantum states can also be expressed as an optimization problem. Let $(\rho - \sigma)_+$ be the projection on the positive eigenspace of $\rho - \sigma$ and $(\rho - \sigma)_-$ be the projection on the negative eigenspace, then

$$\mathrm{Tr}[\rho - \sigma] = 0 = \mathrm{Tr}[(\rho - \sigma)_+] - \mathrm{Tr}[(\rho - \sigma)_-]\,. \tag{2.82}$$

Therefore, the trace distance can be written as

$$\frac{1}{2}\|\rho - \sigma\|_1 = \frac{1}{2}\left(\mathrm{Tr}[(\rho - \sigma)_+] + \mathrm{Tr}[(\rho - \sigma)_-]\right) = \mathrm{Tr}[(\rho - \sigma)_+] = \mathrm{Tr}[(\rho - \sigma)P_+]\,. \tag{2.83}$$

Thus, the trace distance between two quantum states can be written as

$$D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1 = \max_{0 \leqslant P \leqslant I} \mathrm{Tr}[(\rho - \sigma)P]\,. \tag{2.84}$$

where the maximum is over all positive semidefinite matrices P whose eigenvalues are less than 1.

The trace distance can be generalized to quantum channels and is known as the diamond norm [128, 129]. The diamond norm measures the distance between two quantum channels. For any two given channels $\mathcal{E}_A \in \mathrm{CPTP}(A_0 \to A_1)$ and $\mathcal{F}_A \in \mathrm{CPTP}(A_0 \to A_1)$, the diamond norm between $\mathcal{E}$ and $\mathcal{F}$ is defined as

$$\|\mathcal{E}_A - \mathcal{F}_A\|_\diamond = \max_\rho \left\| \mathsf{id}_{A_0} \otimes \mathcal{E}_{\tilde{A_0} \to A_1}(\rho_{A_0 \tilde{A_0}}) - \mathsf{id}_{A_0} \otimes \mathcal{F}_{\tilde{A_0} \to A_1}(\rho_{A_0 \tilde{A_0}}) \right\|_1 \tag{2.85}$$

where id_{A_0} is the identity channel on system A_0, $\|\cdot\|_1$ represents the trace norm, and the maximization is over all density matrices $\rho \in \mathfrak{D}(A_0 \tilde{A}_0)$. The diamond distance is then defined as

$$D(\mathcal{E}, \mathcal{F}) := \frac{1}{2} \|\mathcal{E} - \mathcal{F}\|_\diamond \tag{2.86}$$

where the factor $1/2$ is there for normalization.

Lastly, both the trace norm and the trace distance behave monotonically [74, 127, 128]. The trace norm behaves monotonically under positive linear maps. Let $\mathcal{E} \in \mathfrak{L}(A_0 \to A_1)$ be a positive trace non-increasing map and $X \in \mathfrak{L}(A_0)$. Then

$$\|\mathcal{E}(X)\|_1 \leqslant \|X\|_1 \,, \tag{2.87}$$

and the trace distance behaves monotonically under quantum channels. Let $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ and $\rho, \sigma \in \mathfrak{D}(A_0)$, then

$$D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leqslant D(\rho, \sigma) \,. \tag{2.88}$$

2.3.2 Fidelity of quantum states and channels

Fidelity is another distance-like measure between quantum objects [74, 128]. Unlike trace distance, fidelity achieves the maximum value of one when the two states are the same and zero when the states are orthogonal. The fidelity between quantum states $\rho, \sigma \in \mathfrak{D}(A_0)$ is defined as

$$F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1 = \mathrm{Tr}\left[\left|\sqrt{\rho}\sqrt{\sigma}\right|\right] = \mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right] \,. \tag{2.89}$$

It is easy to note that when $\rho = \sigma$, the fidelity $F(\rho, \rho) = \mathrm{Tr}[|\rho|] = 1$. When one of the states is pure, say $\sigma = |\psi\rangle\langle\psi|$, then the fidelity expression can be simplified as

$$F(\rho, |\psi\rangle\langle\psi|) = \mathrm{Tr}\left[\sqrt{|\psi\rangle\langle\psi|\rho|\psi\rangle\langle\psi|}\right] = \sqrt{\langle\psi|\rho|\psi\rangle} \,. \tag{2.90}$$

From the above expression it is easy to see that when both states are pure, then the fidelity becomes

$$F(\psi, \phi) = |\langle\psi|\phi\rangle| = \sqrt{1 - D(\psi, \phi)^2}\,. \tag{2.91}$$

This definition of fidelity can be extended to quantum channels. To understand the fidelity between quantum channels (i.e., to quantify the "closeness" between two quantum channels), we need to see how the fidelity between the output states from the two channels vary when the same input is provided. To get a reasonable measure of the fidelity between channels, we choose it to be the minimum fidelity between the output states. That gives us the worst-case fidelity for any given input. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$, then the fidelity between $\mathcal{N}$ and $\mathcal{M}$ is defined as

$$F(\mathcal{N}, \mathcal{M}) = \min_{\rho_{\tilde{A}_0 A_0}} F(\mathsf{id}_{\tilde{A}_0} \otimes \mathcal{N}_A(\rho), \mathsf{id}_{\tilde{A}_0} \otimes \mathcal{M}_A(\rho))\,. \tag{2.92}$$

By Uhlmann's Theorem, fidelity between two states can be expressed in terms of their purifications. Let $\rho, \sigma \in \mathfrak{D}(A_0)$ and let $|\psi_{A_0 B_0}\rangle$ and $|\phi_{A_0 C_0}\rangle$ be two purifications of ρ and σ, respectively. Then

$$F(\rho_{A_0}, \sigma_{A_0}) = \max_{V: B_0 \to C_0} |\langle\psi_{A_0 B_0}|V^*|\phi_{A_0 C_0}\rangle| \tag{2.93}$$

where the maximum is over all partial isometries $V : B_0 \to C_0$. We say that V is a partial isometry if V is an isometry when restricted to its support.

Like the trace norm, fidelity of states also behaves monotonically under quantum channels. Let $\rho, \sigma \in \mathfrak{D}(A_0)$ and $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$, then

$$F(\rho, \sigma) \leqslant F\left(\mathcal{E}(\rho), \mathcal{E}(\sigma)\right)\,. \tag{2.94}$$

Lastly, fidelity and trace distance between two states ρ, σ satisfy the following inequalities:

$$\sqrt{1 - F(\rho, \sigma)^2} \geqslant D(\rho, \sigma) \geqslant 1 - F(\rho, \sigma)\,. \tag{2.95}$$

2.3.3 Quantum divergences and relative entropies

While distinguishing two quantum states, a key observation is that sending both the states through a quantum channel does not increase the ability to distinguish the states. Let $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ be a quantum channel, and we are trying to distinguish between two states ρ_{A_0} and σ_{A_0}. Then, the states $\mathcal{N}(\rho)$ and $\mathcal{N}(\sigma)$ are never more distinguishable than ρ and σ. So, any measure that quantifies the distinguishability

of quantum states must not increase under any quantum channel. This monotonicity property is known as data processing inequality, and the functions that satisfy this property are called quantum divergences [74, 128, 130]. Formally, a function

$$\mathbb{D} : \bigcup \{\mathfrak{D}(A_0) \times \mathcal{D}(A_0)\} \to \mathbb{R} \cup \{\infty\}$$

that acts on a pair of quantum states is called a quantum divergence if it satisfies the following two conditions:

1. Data Processing Inequality (DPI) [131]:

$$\mathbb{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leqslant \mathbb{D}(\rho\|\sigma)\,.$$

2. Normalization:

$$\mathbb{D}(1\|1) = 0\,.$$

The normalization condition ensures that, for any state ρ, $\mathbb{D}(\rho\|\rho) = 0$.

The above definition of quantum divergence defined for states can be easily generalized to quantum channels if, instead of ρ and σ, we have channels $\mathcal{E}$ and $\mathcal{F}$, and instead of $\mathcal{N}$ in the state case that evolve states, we have a superchannel Θ that evolves channels [130]. The data processing inequality for quantum channels then becomes $\mathbb{D}(\Theta[\mathcal{E}]\|\Theta[\mathcal{F}]) \leqslant \mathbb{D}(\mathcal{E}\|\mathcal{F})$.

Some examples of quantum divergence include trace distance and fidelity, which we discussed in the previous subsections. Other commonly used divergences are the family of functions called Rényi divergences [132]. Rényi divergences are defined for any parameter $\alpha \in [0, \infty]$ and for any probability distributions $\mathbf{p}, \mathbf{q} \in \mathfrak{D}(n)$[4] as

$$D_\alpha(\mathbf{p}\|\mathbf{q}) := \begin{cases} \dfrac{1}{\alpha - 1} \log \sum_{x=1}^{m} p_x^\alpha q_x^{1-\alpha} & \text{if } \operatorname{supp}(\mathbf{p}) \subseteq \operatorname{supp}(\mathbf{q}) \\ \infty & \text{otherwise} \end{cases} \tag{2.96}$$

Quantum divergences that are additive are known as relative entropies. Formally, a function $\mathbb{D} : \bigcup_{A_0} \{\mathfrak{D}(A_0) \times \mathfrak{D}(A_0)\} \to \mathbb{R} \cup \{\infty\}$ acting on a pair of quantum states in $\mathfrak{D}(A_0)$ with $|A_0| < \infty$ is called a relative entropy if it satisfies the following three axioms:

1. Data Processing Inequality [131]:

$$\mathbb{D}(\mathcal{N}(\rho)\|\mathcal{N}(\rho)) \leqslant \mathbb{D}(\rho\|\sigma)\,,$$

where $\rho, \sigma \in \mathfrak{D}(A_0)$ and $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$.

[4] $\mathfrak{D}(n)$ denotes the set of probability distributions with n components.

2. Additivity: For quantum states $\rho, \sigma \in \mathfrak{D}(A_0)$ and $\rho', \sigma' \in \mathfrak{D}(B_0)$

$$\mathbb{D}(\rho \otimes \rho' \| \sigma \otimes \sigma') = \mathbb{D}(\rho \| \sigma) + \mathbb{D}(\rho' \| \sigma')$$

3. Normalization: For the qubit state $|0\rangle\langle 0|$ and the maximally mixed qubit state $\frac{I_2}{2}$

$$\mathbb{D}\left(|0\rangle\langle 0| \middle\| \frac{I}{2}\right) = 1$$

.

Since relative entropies are quantum divergences, it follows that $\mathbb{D}(\rho\|\rho) = 0$, i.e., they quantify the distance between quantum states and hence their value is zero if the input states are the same. Let me now discuss below some examples which are relevant for this thesis.

Examples of relative entropies

1. Umegaki Relative Entropy [133]: Given two quantum states ρ_{A_0} and σ_{A_0}, the Umegaki relative entropy is defined as

$$D(\rho\|\sigma) := \operatorname{Tr}[\rho(\log \rho - \log \sigma)] \tag{2.97}$$

where the logarithm is taken to base 2.

2. Petz-Rényi relative entropy or quantum Rényi relative entropy [134]: Given two quantum states ρ_{A_0} and σ_{A_0}, the family of quantum Rényi relative entropies is defined as

$$D_\alpha(\rho\|\sigma) := \begin{cases} \dfrac{1}{\alpha - 1} \log \operatorname{Tr}[\rho^\alpha \sigma^{1-\alpha}] & \text{if } \operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma), \text{ or } \alpha < 1 \text{ and } \rho \not\perp \sigma \\ \infty & \text{otherwise} \end{cases} \tag{2.98}$$

where the parameter $\alpha \in (0,1) \cup (1,2)$. The above definition can be extended for $\alpha \to \infty$ but the data-processing inequality only holds for $\alpha \in [0,2]$ and therefore, the Petz-Rényi relative entropy does not have an operational meaning beyond $\alpha = 2$. The cases $\alpha = 0$ and $\alpha = 1$ are defined by taking appropriate limits. For the limit $\alpha \to 1$, this relative entropy converges to Umegaki relative entropy. The Petz-Rényi relative entropy when limit $\alpha \to 0$ is called the min relative entropy and is given by

$$D_{\min}(\rho\|\sigma) := D_0(\rho\|\sigma) = -\log \operatorname{Tr}[\Pi_\rho \sigma] \tag{2.99}$$

where Π_ρ denotes the projection on the support of ρ. The above definition of min relative entropy holds if $\rho\sigma \neq 0$, otherwise $D_{\min}(\rho\|\sigma) = \infty$. Another property of Petz-Rényi relative entropies is that there's an ordering associated with it, i.e., for $\alpha > \beta > 0$, it follows that

$$D_\alpha(\rho\|\sigma) \geqslant D_\beta(\rho\|\sigma)\,.$$

Lastly, the Petz-Rényi relative entropy finds an operational meaning in the context of quantum hypothesis testing.

3. Sandwiched Rényi relative entropy [135, 136]: Given two quantum states ρ_{A_0} and σ_{A_0}, the sandwiched Rényi relative entropy is defined as

$$\tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log \operatorname{Tr}\left[\left(\sigma^{(1-\alpha)/2\alpha}\rho\sigma^{(1-\alpha)/2\alpha}\right)^\alpha\right] \tag{2.100}$$

where the parameter $\alpha \in (0,1)\cup(1,\infty)$ and the cases $\alpha = 0, 1, \infty$ are understood in terms of limits. Like the Petz-Rényi relative entropy, the sandwiched Rényi relative entropy also converges to Umegaki relative entropy when $\alpha \to 1$, and also follows the ordering $\tilde{D}_\alpha(\rho\|\sigma) \geqslant \tilde{D}_\beta(\rho\|\sigma)$ if $\alpha > \beta > 0$.

4. Max quantum relative entropy: Given quantum states ρ_{A_0} and σ_{A_0}, the max quantum relative entropy is defined as

$$D_{\max}(\rho\|\sigma) := \begin{cases} \log\min\{t\in\mathbb{R} : \sigma \geqslant \rho\} & \text{if } \operatorname{supp}(\rho)\subseteq\operatorname{supp}(\sigma) \\ \infty & \text{otherwise} \end{cases} \tag{2.101}$$

For any relative entropy $\mathbb{D}$ and quantum states $\rho_{A_0}, \sigma_{A_0}, \omega_{A_0}$, it holds that

(a) $D_{\min}(\rho\|\sigma) \leqslant \mathbb{D}(\rho\|\sigma) \leqslant D_{\max}(\rho\|\sigma)$.

(b) $\mathbb{D}(\rho\|\sigma) \leqslant \mathbb{D}(\rho\|\omega) + D_{\max}(\omega\|\sigma)$.

More discussion about the two relative entropies (Petz and sandwiched Rényi relative entropies) can be found in [132, 136, 137, 138, 139, 140, 141]. Other generalizations of the Rényi divergence and the quantum Rényi relative entropies are discussed in [142], but their operational meaning is not clear.

For the channel case (i.e., dynamical resources), the relative entropies and divergence have been generalized from the state case (i.e., static resources) and were discussed in [119, 143, 144, 145, 146, 147, 148]. The

channel divergence for two given channels $\mathcal{N}_A, \mathcal{M}_A \in \mathrm{CPTP}(A_0 \to A_1)$ is defined as [119, 146, 147]

$$D(\mathcal{N}_A \| \mathcal{M}_A) = \max_{\phi \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right) \middle\| \mathcal{M}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right)\right) \tag{2.102}$$

where $D(\rho\|\sigma) = \mathrm{Tr}[\rho \log \rho - \rho \log \sigma]$ is the relative entropy.

2.4 Convex analysis tools used in this thesis

Detailed discussions on topics of convex analysis can be found in standard textbooks and notes like [128, 149, 150, 151]. Below, we provide a brief discussion on the topics and terminology that have been used in this thesis.

A set $S \subset \mathbb{R}^n$ is called a convex set if, for any two elements $\mathbf{u}, \mathbf{v} \in S$ and any $p \in [0,1]$, the vector $p\mathbf{u} + (1-p)\mathbf{v} \in S$. As a consequence, if $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in S$ and $p_1, p_2, \ldots, p_n$ are non-negative numbers summing to unity, then

$$\sum_n p_i \mathbf{u}_i \in S\,. \tag{2.103}$$

A function $f : S \to \mathbb{R}$ is called a convex function if

$$f(p\mathbf{u} + (1-p)\mathbf{v}) \leqslant p f(\mathbf{u}) + (1-p) f(\mathbf{v}) \tag{2.104}$$

for all $\mathbf{u}, \mathbf{v} \in S$ and all real numbers $p \in [0,1]$. A function g is concave if $-g$ is convex.

A subset $\mathcal{K}$ of a Hilbert space A is called a cone if for any non-negative number $c \in \mathbb{R}$ and any element $\mathbf{v} \in \mathcal{K}$, $c\mathbf{v} \in \mathcal{K}$. The dual of a cone $\mathcal{K}$ in A is the set $\mathcal{K}^* := \{\mathbf{w} \in A : \mathbf{w} \cdot \mathbf{v} \geqslant 0 \ \forall\ \mathbf{v} \in \mathcal{K}\}$. A convex conical hull of a subset $\mathcal{X} \in \mathbb{R}^n$ is the set $\mathcal{C}(\mathcal{X}) := \{\sum_{i=1}^n t_i \mathbf{u}_i \,:\, n \in \mathbb{N},\, \mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathcal{X}, \text{ and } t_1, \ldots, t_n > 0\}$. One of the most common examples of cones used in quantum information is the cone of positive semidefinite operators on finite-dimensional Hilbert space.

2.4.1 Farkas Lemma and Hyperplane separation theorem

Consider the closed convex cone $\mathcal{C}(A)$ spanned by the columns of the matrix A. Then $\mathcal{C}(A) = \{A\mathbf{x} : x \geqslant 0\}$. So, if a vector $\mathbf{b}$ lies inside the cone $\mathcal{C}(A)$ then there exists an $\mathbf{x} \geqslant 0$ such that $A\mathbf{x} = \mathbf{b}$. On the other hand, if $\mathbf{b}$ lies outside the cone, then there exists a vector $\mathbf{y}$ orthogonal to a hyperplane that separates the cone $\mathcal{C}(A)$ and $\mathbf{b}$, such that $A^{\mathrm{T}}\mathbf{y} \geqslant 0$ and $\mathbf{b}^{\mathrm{T}}\mathbf{y} < 0$. This is known as the Farkas lemma.

Lemma 2.1. *(Farkas Lemma.) Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then exactly one of the following two assertions is true:*

1. *There exists an $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geqslant 0$.*

2. *There exists a $\mathbf{y} \in \mathbb{R}^m$ such that $A^{\mathrm{T}}\mathbf{y} \geqslant 0$ and $\mathbf{b}^{\mathrm{T}}\mathbf{y} < 0$.*

The geometrical interpretation of the Farkas lemma is that if a point lies outside of a convex cone, then there exists a hyperplane separating the point and the convex cone. This idea can be generalized to a pair of convex sets and is known as the hyperplane separation theorem. The hyperplane separation theorem states that two convex sets with empty intersection can always be separated by a hyperplane. The theorem is stated below.

Theorem 2.2. *(Hyperplane separation theorem) Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two disjoint convex subsets of $\mathbb{R}^n$. Then there exist a nonzero vector $\mathbf{u} \in \mathbb{R}^n$ and a real number $c \in \mathbb{R}$ such that*

$$\mathbf{u} \cdot \mathbf{r}_2 \leqslant c \leqslant \mathbf{u} \cdot \mathbf{r}_1 \ \ \forall \mathbf{r}_1 \in \mathcal{C}_1 \text{ and } \forall \mathbf{r}_2 \in \mathcal{C}_2\,. \tag{2.105}$$

That is, $\mathbf{u}$ is the normal vector of the hyperplane $\{\mathbf{v} \in \mathbb{R}^n : \mathbf{u} \cdot \mathbf{v} = c\}$ that separates $\mathcal{C}_1$ and $\mathcal{C}_2$. Moreover, if the sets $\mathcal{C}_1$ and $\mathcal{C}_2$ are also closed and at least one of them is compact, then one can replace the above inequalities with strict inequalities.

Convex hulls and polytopes

Farkas lemma and hyperplane separation theorem are based on convex sets. When forming convex combinations of a set $C \in \mathbb{R}^n$, various types of convex structures can be formed. The smallest convex set in $\mathbb{R}^n$ that contains $\mathcal{C}$ is known as the convex hull of the set $\mathcal{C}$, and is denoted as $\mathrm{Conv}(\mathcal{C})$. If there are finite number of vectors in $\mathcal{C}$, then $\mathrm{Conv}(\mathcal{C})$ is called a convex polytope and contains all the convex combinations of the vectors in $\mathcal{C}$. If $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then

$$\mathrm{Conv}(\mathcal{C}) := \left\{ \sum_{x=1}^{n} p_x \mathbf{v}_x : 0 \leqslant p_x \in \mathbb{R}\,,\ \sum_x p_x = 1 \right\} \tag{2.106}$$

Every convex set has certain extreme points. The extreme points are those points that cannot be expressed as $p\mathbf{v} + (1-p)\mathbf{w}$ for some $0 < p < 1$ and two distinct vectors $\mathbf{v}, \mathbf{w} \in \mathcal{C}$.

2.4.2 Conic linear programming and semidefinite programming

Conic linear programming is a convex optimization problem that can be expressed in terms of two cones. A semidefinite program (SDP) is a type of conic linear programming in which the optimization variable η is positive semidefinite, the objective function is linear in the variable η, and the constraint is an operator inequality featuring a linear function of η [149, 150]. An SDP corresponds to two optimization problems; one is known as the primal problem, and the other is known as the dual problem. Let A_1 and A_2 be two Hilbert spaces, let $\mathcal{C}_1 \subseteq V_1 \subseteq \mathrm{Herm}(A_1)$ and $\mathcal{C}_2 \subseteq V_2 \subseteq \mathrm{Herm}(A_2)$ be two convex cones in two subspaces of Hermitian matrices V_1 and V_2, and let $\mathcal{N} : V_1 \to V_2$ be a linear map between the two vector spaces. Let $H_1 \in V_1$ and $H_2 \in V_2$ be two (fixed) Hermitian matrices. Then the primal problem is defined as

$$\begin{aligned} \alpha := \quad & \inf \ \mathrm{Tr}[\eta H_1] && (2.107)\\ & \text{subject to } \ \mathcal{N}(\eta) - H_2 \in \mathcal{C}_2 \,, \text{ and} && (2.108)\\ & \eta \in \mathcal{C}_1 \,. && (2.109) \end{aligned}$$

The dual of the above primal problem is defined as

$$\begin{aligned} \beta := \quad & \sup \ \mathrm{Tr}[\zeta H_2] && (2.110)\\ & \text{subject to } \ H_1 - \mathcal{N}^*(\zeta) \in \mathcal{C}_1^* \,, \text{ and} && (2.111)\\ & \zeta \in \mathcal{C}_2^* \,. && (2.112) \end{aligned}$$

In the above optimization problems, if there is no optimal solution of the primal problem, then by convention $\alpha = +\infty$ and if there is no optimal solution of the dual problem, then by convention $\beta = -\infty$. Moreover, by weak duality, it holds that $\alpha \geqslant \beta$. When the equality holds, it is known as strong duality.

Many classes of convex optimization problems (including linear programming (LP), second-order cone programming (SOCP), and semidefinite programming (SDP) problems) can be solved efficiently [152]. For SDPs, there exist algorithms to solve them that have an efficient runtime and can efficiently store problem data. By efficient, I mean that the runtime of the algorithm used to solve the SDP and storage required is polynomial in the dimension n of the variable matrix and the number of constraints m [153, 154]. The variable matrix is an $n \times n$ matrix and appears as the optimization variable in the primal or dual problem of the optimization problem [128, 150, 153]. Some of these algorithms are the primal-dual interior point methods, the alternating direction method of multipliers, the HKM method, the bundle method and the augmented Lagrangian method [152, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168].

The primal-dual interior point methods are implemented in most software packages used to numerically solve SDPs. Some such software packages are CSDP, SeDuMi, SDPA, and SDPT3 [169, 170, 171, 172]. (For the details of these algorithms and software packages, an interested reader can refer to the citations provided. The details of these algorithms are out of the scope of this thesis.) In this thesis, I have used the software package CVX to numerically compute the optimal value of an SDP which uses SeDuMi and SDPT3 packages to approximately solve an SDP [173, 174].

The interior point method was first developed by Karmakar in 1984 [155]. Karmakar showed that the interior point methods can solve linear programs in polynomial time. With time, further improvements in the runtime have been made. The most recent results are by Jiang et al. who showed that the runtime of a generic SDP with variable size $n \times n$ and with m constraints is $O(\sqrt{n}(mn^2 + m^\omega + n^\omega)\log(1/\epsilon))$ where ω is the exponent of matrix multiplication and ϵ is the relative accuracy [154]. They have also provided a tabular summary that compares the runtimes of various key SDP solvers (see Table 1.1 and 1.2 in [154]).

Similar to time complexity, SDP solvers are efficient in space usage. In Ref. [163], a detailed analysis of space complexity has been shown for a 64-bit parallel version of CSDP implemented on a shared memory system. The results of this analysis on asymptotic storage requirements are also applicable to SeDuMi and SDPT3 that are used in CVX. It has been shown that the storage required is $O(m^2 + n^2)$ where m is the total number of constraints and the optimization variable in the primal or dual problem shown above is an $n \times n$ matrix. Using this, approximate storage requirements in bytes can also be found (see Eq.6 of [163]). Recently, it was shown that when an SDP is weakly constrained, then the storage required can be further reduced [175]. Thus, with efficient runtime and storage requirements, SDPs become a very useful tool for solving several optimization problems in quantum information, and I have used them to numerically compute the optimal values of several optimization problems in this thesis.

2.5 Resource theories: An Introduction

In any given setting, there are some tasks that can be easily performed. Everything else that cannot be easily accessed is essentially a resource. From economics principles, we can say that something becomes valuable when it cannot be easily obtained. Let us take the simplest example of money. Money is a resource that cannot be freely obtained, and one has to work for it. But once you have money, you can exchange it for any goods and services depending on the amount of money you spend. It is a resource with the maximum value. No other resource is equivalent to it. For instance, for a rice farmer, rice is a free entity with zero value but it is valuable to others. However, the farmer cannot simply exchange rice for any goods or services he desires. In the conventional economic setting, the farmer has to first sell his rice in exchange for some

money that he can then use to buy anything else. Similarly, in our day-to-day settings, we always have some things available to us easily, and for some others, we need to make an effort.

Identifying what is free and what is prohibited in a given scenario is the essence of a resource-theoretic framework. Once this identification of free and prohibited operations is done, it is important to understand the value of the objects that are not free. There might be some resources which do not have any use in what we wish to do, and some resources (like money) are always useful, and thus have the maximum value. After establishing the value of the resources, it is analyzed how one resource can be converted to another desired resource using the allowed operations. Take, for instance, the case of Shannon theory, where two parties (say, Alice and Bob) need to communicate with each other. However, they have at their disposal a noisy channel as a resource through which they communicate. What they want to do is develop a scheme such that they can use as few copies of this resource (i.e., the noisy channel) in order to perfectly communicate with each other. In other words, they want to convert one resource (the noisy channel) into another (a noiseless channel or an identity channel) by using minimum number of copies of the given resource. To do that, Alice and Bob can perform local encoding and decoding freely. Moreover, to quantify different schemes, they can compare the rate at which information is being transmitted using each encoding-decoding scheme. Thus, a resource-theoretic formulation offers a general and structured way to study the interconversion among resources in any given setting. Due to this general approach, resource theories unify different phenomena under the same umbrella [82].

Let us now see how the resource-theoretic framework can be applied to quantum information.

2.5.1 Resources in the quantum world and a sneak-peek on the resources studied in this thesis

The above general resource-theoretic formalism can be applied to various quantum computing and communication scenarios with different constraints. Quantum mechanics offers phenomena like quantum coherence, entanglement, non-locality, magic, etc., which have been used as resources to gain advantage over classical strategies for solving various kinds of problems. For instance, quantum coherence is a key resource for most quantum algorithms, quantum entanglement plays an important role in quantum communication, magic plays a crucial role in achieving universal quantum computation from fault-tolerant quantum computation, etc. Since each of these phenomena have disparate characteristics which can bring advantages under diverse circumstances, it is natural to study these phenomena using a resource theoretic approach. Thus, a quantum resource theory is a framework for studying restricted quantum information processing [82].

Different experiments and operational tasks have different restrictions, and under these constraints, it is

important to identify and analyze resourceful quantum states and channels. Let us understand this through an example. Suppose that there are two spatially separate labs where each experimenter can perform any quantum operation on quantum systems in their labs, and they can both classically communicate with each other. Under such allowed dynamics (of local operations and classical communications or LOCC), the natural resource is a shared entangled state, using which Such tasks can be performed that otherwise cannot be performed using LOCC. For example, to perform quantum teleportation between the two labs, the labs need to have access to a shared entangled state. Thus, in this scenario, the allowed operations are LOCC, the free states are the states that can be generated from LOCC which are the set of separable states, and the resource states are the entangled states. With this partition of states into free and resources, we have the resource theory of entangled states. In the upcoming chapters, we will see other types of restricted settings and the quantum phenomena that act as resources in those settings. For instance, in chapter 3, we have considered decoherence as the noise model in quantum communication. Decoherence is a phenomenon that kills the coherence of the quantum state, leaving behind a classical state. In such a scenario, the preservation of quantum coherence is a resource. Thus, quantum channels that can preserve and transmit coherence are natural resources in this setting. Similarly, in quantum computation, a very restricted set of states and operations known as stabilizer states and operations can be used to perform fault-tolerant quantum computation. Furthermore, any circuit formed from stabilizer operations with stabilizer states as input can be efficiently simulated classically. Thus, to achieve universal quantum computation and to gain quantum computational advantage, non-stabilizer states and operations are used as resources. The free operations are then those operations that preserve the set of stabilizer states. With this bifurcation of states and operations as free and resources, we can define a resource theory, which is the subject of Chapter 4.

With this overview of resource theories, let me now discuss the mathematical structure of quantum resource theory of states.

2.5.2 Basic structure of static quantum resource theories

Given the restrictions in a particular setting, a quantum resource theory models what can be accomplished using the allowed operations in that setting. Let $\mathfrak{F}(A) \subset \mathrm{CPTP}(A)$ where A denotes the dynamical system $(A_0 \to A_1)$ be a mapping that takes any two physical systems A_0 and A_1 to a set of quantum channels. The mapping $\mathfrak{F}(A)$ is called a quantum resource theory if it satisfies the following conditions [82]:

1. Doing nothing is free. In general, doing nothing is represented by the identity channel $\mathsf{id}_{A_0} \in \mathfrak{F}(A_0 \to A_0)$ implying that the state of the system has not changed over time. However, when we consider decoherence, the environment acts on the system and decoheres it. In that case, doing nothing will

mean that no external operation is being acted upon the system, but since environment is decohering the system, we will represent the system as being acted upon by the dephasing channel. This has been made clear in Sec. 3.1.

2. The composition of free operations is free. Let $\mathcal{N} \in \mathfrak{F}(A_0 \to A_1)$ and $\mathcal{M} \in \mathfrak{F}(A_1 \to B_1)$, then $\mathcal{M} \circ \mathcal{N} \in \mathfrak{F}(A_0 \to B_1)$.

3. Discarding a system is free. That is, tracing out a system $\mathrm{Tr}_{A_0} \in \mathfrak{F}(A_0 \to 1)$.

The set of operations $\mathfrak{F}(A_0 \to A_1)$ is called the set of *free operations*, and the states that can be generated from these free operations are called the *free states*.

The first condition above requires that not doing any operation on the quantum state to be free which is equivalent to the identity operation, and is a very natural requirement for most settings. However, in certain special cases, like the setting considered in Chapter 3, we will see that not doing anything is not the same as the identity channel. Since in Chapter 3, we consider the fact that the environment decoheres a system, not doing anything implies that the environment is acting on the system and decohering it. The identity channel (whose output state is the same as the input state) then implies preserving the coherence of a system over a period of time. Since in real-world scenarios it is extremely hard for an experimentalist to maintain the coherence of a quantum system, preservation of coherence is treated as a resource, and thus, the identity channel cannot be considered free in this case, whereas not doing anything (represented by the action of the completely dephasing channel) is still free.

The second condition implies that free operations cannot generate a resource from free inputs. Any kind of composition, serial or parallel, of free operations results in a free operation. Thus, when free states are given as input to free operations, the output is also free. This is the *golden rule of quantum resource theories*.

The third condition implies that discarding a system must be free. This, like the first condition, is a very natural requirement. A consequence of this is that converting any quantum state to a free state is also free.

Apart from the above three basic requirements, there are other conditions that most of the quantum resource theories obey [82]:

4. It is natural to assume that free operations can act on part of composite systems. Since doing nothing is free, we get that a free operation is also completely free, i.e., if $\mathcal{E}_A \in \mathfrak{F}(A)$ then

$$\mathsf{id}_{B_0 \to B_0} \otimes \mathcal{E}_A \in \mathfrak{F}(B_0 A)\,.$$

5. For any composite system, permuting the labels of the systems is free.

6. For a physical system A_0, the set of free states $\mathfrak{F}(A_0)$ is closed.

7. For a physical system A_0, the set of free states $\mathfrak{F}(A_0)$ is convex.

As stated above, the free operations in a quantum resource theory are decided by the physical constraints, and the set of states that can be generated from these operations are the free states. However, in some resource theories, it is difficult to mathematically characterize the set of free operations, which in turn makes it hard to quantify the resources. For example, one such set of free operations is the set of local operations and classical communications in the resource theory of entanglement. To overcome this problem, the set of free operations is enlarged and defined as a set of operations that preserve the set of free states. Since such a set of operations is the largest set of operations that cannot generate a resource from the free states, it is often referred to as the set of *resource non-generating* (RNG) operations. Besides, in some resource theories, for example, in the resource theory of coherence, as we will see in Chapter 3, it is natural to define the set of free states first. Then the *resource non-generating operations* are a natural choice of free operations.

Once the free operations and free states are defined, it is crucial to characterize these elements. Using this characterization, it can be found out whether a given quantum state is free or not. When the characterization of free states is hard and the set of free states is convex, then resource witnesses, a tool from convex analysis, can be used to determine if a state is free or not. Let $\mathcal{F}$ be a quantum resource theory, and A_0 be a physical system. Then the set of operators $W \in \mathrm{Herm}(A_0)$ are called resource witnesses if the following two conditions hold:

1. For any quantum state $\tau \in \mathfrak{F}(A_0)$
$$\mathrm{Tr}[W\tau] \geqslant 0\,.$$

2. There exists a $\rho \in \mathfrak{D}(A_0)$ such that
$$\mathrm{Tr}[W\rho] < 0\,.$$

We will see the use of witnesses in Chapter 4, as the free states in the resource theory of magic are very hard to mathematically characterize. If the set of free states is closed and convex, then it holds that $\sigma \in \mathfrak{F}(A_0)$ if and only if

$$\mathrm{Tr}[W\sigma] \geqslant 0 \tag{2.113}$$

for all resource witnesses $W \in \mathrm{Herm}(A_0)$.

Once we have defined and characterized the free operations, free states, and resources, the next task is to quantify the resources. One way to quantify resources is by using distance measures (some of which

are discussed in the previous section, Sec. 2.3). To use distance measures as resource quantifiers, we find the minimum distance of the given resource from the set of free states. Using this idea, in some cases, the resource value of a state can be cast as an optimization problem where the minimization is over all free states. In general, for any function to be a resource measure, it needs to have the following two properties. First, the function should be a monotone under free operations, i.e., the value of a state cannot increase after it is acted upon by a free operation because free operations cannot generate a resource. Second, it must be faithful, i.e., the resource value of a given state must be zero if and only if it is a free state. After the quantification, we tackle various types of resource interconversion tasks. Different kinds of resource interconversion tasks such as single-shot, probabilistic, asymptotic, or approximate interconversions, become important under various settings.

Chapter 3

Dynamical resource theory of quantum coherence

3.1 Introduction and motivation

All physical systems undergo decoherence. It is an irreversible process, and it can be viewed as the reduction of a general quantum state to an incoherent mixed state due to coupling with the environment [176, 177, 178]. Mathematically, decoherence is represented as the vanishing of the off-diagonal terms of a density matrix. It is because of decoherence that we do not observe quantum mechanical behaviour in everyday macroscopic objects, and in the context of quantum information, it can be viewed as the loss of information from a system into the environment [179].

During the last two decades, interest in quantum information science has shifted towards using quantum mechanical phenomena (like entanglement, nonlocality, coherence, etc.) as resources to achieve something that is otherwise not possible through classical physics (eg., quantum teleportation) [105, 180, 181, 182, 183, 184, 185, 186, 187, 188]. Quantum resource theories (QRTs) use this resource-theoretic approach to exploit the operational advantage of such phenomena and to assess their resource character systematically [82]. The preservation of quantum coherence is crucial for building quantum information devices, since the loss of quantum superposition due to decoherence negates any non-classical effect in a quantum system [176, 189, 190]. Hence from a technological perspective, there is increasing interest in developing a resource theory of coherence [82]. In addition, the resource-theoretic study of quantum coherence can provide new operational and quantitative insights into the differences between classical and quantum physics. Some other examples of quantum resource theories include the QRT of entanglement, thermodynamics, magic states,

Bell non-locality, etc.

As discussed earlier, most quantum resource theories are governed by the constraints arising from physical or practical settings. These constraints then lead to the operations that can be freely performed. For instance, in the static resource theory of quantum entanglement, for any two spatially separated but possibly entangled systems, the spatial separation puts the restriction that only local operations along with classical communications (LOCC) can be performed [105, 180, 191, 192, 193]. Given this restriction, only separable states can be generated using LOCC (when the parties don't already share an entangled state), which makes them the free states of the theory. But unlike entanglement, whose free states (i.e. separable states) are determined from the *a priori* fixed set of operations (i.e. LOCC), coherence theory typically begins by fixing a set of free states. In this case, the free states are the physically-motivated objects, and the free operations are not unique, only being required to satisfy the basic golden rule of a QRT, that the free operations should be completely resource non-generating (CRNG) [82].

In the static (or state-based) resource theory of quantum coherence there is a fixed basis, the so-called classical or incoherent basis, and the set of density matrices that are diagonal in this basis form the free states of the theory. Such states are called incoherent states. The choice of the fixed basis depends on the physics of the system: the basis in which the environment decoheres a quantum system, and this basis usually coincides with the computational basis. The free operations are then some set of quantum channels that map the set of incoherent states to itself. The most well-studied classes of free operations are the maximally incoherent operations (MIO), the incoherent operations (IO), the dephasing-covariant incoherent operations (DIO), and the strictly incoherent operations (SIO) [38, 39, 40, 41, 194, 195]. However, all of these operations cost coherence to be physically implemented in the sense that they do not always admit a free dilation [40, 41, 194, 195]. Therefore, it can be questioned in what sense these operations are truly "free" [40]. On the other hand, as argued in Ref. [196], detecting the presence of resource in a given state should be possible using the free operations, and often this detection requires the consumption of resource. If such a detection is not possible, then both resource and non-resource states are equally useful (or useless) to the experimenter since the two cannot be distinguished, thus begging the question in what sense the former is truly a resource. In general then, having free operations with a nonzero resource cost can still lead to an insightful static resource theory. Indeed, even though MIO/DIO/IO/SIO consume coherence in their implementation, they are still useful for comparing the coherence in two different states based on their convertibility using the given operations. A large amount of fruitful work has been devoted to developing the theory of static coherence under these operations [32].

A consequence of these observations is that the principles for assessing the resourcefulness of quantum states should not necessarily be applied when assessing the resourcefulness of quantum operations. In

particular, the well-known approaches to quantifying the resourcefulness of a quantum channel in terms of its resource cost [197, 198] or its resource-generating power [143, 199, 200, 201, 202] can fall short of fully characterizing its utility in a resource theory. The ability for a channel to generate "resource detectability" is typically something not captured by its resource cost or resource-generation power [196]. Restricting to the resource theory of coherence, a POVM $\{P_m\}_m$ can detect the coherence in a state ρ if $\mathrm{Tr}[\Pi_m \rho] \neq \mathrm{Tr}[\Pi_m \mathcal{D}(\rho)]$ for some outcome m, where $\mathcal{D}(\rho) = \sum_{i=1}^{d} |i\rangle\langle i|\rho|i\rangle\langle i|$ is the completely dephasing map in the incoherent basis $\{|i\rangle\}_{i=1}^{d}$. Since a POVM $\{P_m\}_m$ is unable to detect coherence in some state if (and only if) it is incoherent, i.e. $\mathcal{D}(P_m) = P_m$ for all m, a channel $\mathcal{N}$ fails to generate a detection of coherence if its dual maps any incoherent POVM to an incoherent POVM:

$$\mathcal{N}^\dagger \circ \mathcal{D}(P_m) = \mathcal{D} \circ \mathcal{N}^\dagger \circ \mathcal{D}(P_m) \qquad \forall \{P_m\}_m. \tag{3.1}$$

Maps satisfying Eq. (3.1) have been called detection incoherent in Ref. [196] and nonactivating in Ref. [203], and any map not of this form is a dynamical resource from the coherence-detection perspective. It is not difficult to find channels that satisfy Eq. (3.1) while having a nonzero coherence cost and coherence-generating power (for instance, consider any replacement channel that outputs a coherent state for any input). A full resource theory can then be worked out on the level of channels in which maps having the form of Eq. (3.1) are free and the allowed operations are certain superchannels that act invariantly on the set of detection incoherent channels [196].

In this work, we identify another coherence property of quantum channels that is not captured by coherence cost, coherence-generating power, or coherence detection. We are motivated by the interpretation of a quantum channel as a quantum memory that transmits quantum information from one point in spacetime to another [204, 205, 206]. As a concrete pragmatic scenario, we consider a cloud quantum computer in which

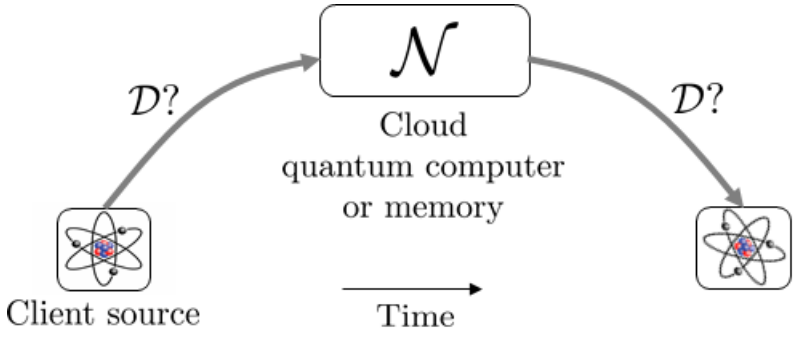

Figure 3.1: A cloud quantum computer offers no computational resource if the upload and download channels are completely dephasing ($\mathcal{D}$). Channels having the form $\mathcal{D} \circ \mathcal{N} \circ \mathcal{D}$ are thus identified as free dynamical objects in the resource theory studied here.

a client uploads and downloads quantum information to a quantum computing processor and memory (see Fig. 3.1). Ideally both the upload and download links are noiseless, and if ρ is the quantum state sent to the cloud computer to perform operation $\mathcal{N}$, the state returned to the client will be $\mathcal{N}(\rho)$. However, in practice the channels connecting client to cloud will be noisy. In the extreme cases, a completely dephasing upload channel has the form $\mathcal{N} = \mathcal{N} \circ \mathcal{D}$ and can be interpreted as a cloud process $\mathcal{N}$ in which the coherences of the input state are not registered and stored, while a completely dephasing download channel has the form $\mathcal{N} = \mathcal{D} \circ \mathcal{N}$ and can be interpreted as a cloud process that fails to output any coherence. Here we consider the worst-case scenario in which both channels are completely dephasing. More precisely, we identify a channel $\mathcal{N}_A$ with input/output space A_0/A_1 to be free if

$$\mathcal{N}_A = \mathcal{D}_{A_1} \circ \mathcal{N}_A \circ \mathcal{D}_{A_0} \ , \tag{3.2}$$

where $\mathcal{D}_{A_0}$ and $\mathcal{D}_{A_1}$ are completely dephasing channels for systems A_0 and A_1 in their respective incoherent bases. Note the similarity between the dynamical free objects defined in Eq. (3.2) and the static free objects in coherence theory. On the level of states, a density operator ρ is incoherent with respect to the fixed basis if

$$\rho = \mathcal{D}_{A_1}(\rho). \tag{3.3}$$

In fact, this can be seen as a special case of Eq. (3.2) when system A_0 is one-dimensional.

Channels satisfying Eq. (3.2) are referred to as classical channels since their action is described entirely by transition probabilities $p(i|j)$ from incoherent states $|j\rangle\langle j|_{A_0}$ to incoherent states $|i\rangle\langle i|_{A_1}$. We will denote the set of classical channels that take system A_0 to A_1 by $\mathfrak{C}(A_0 \to A_1)$,

$$\mathcal{N}_A \in \mathfrak{C}(A_0 \to A_1) \iff \mathcal{N}_A = \mathcal{D}_{A_1} \circ \mathcal{N}_A \circ \mathcal{D}_{A_0} \ . \tag{3.4}$$

In particular, the identity channel $\mathsf{id}_{A_0 \to A_1}$ is not classical as it does not satisfy the above condition (here, A_0 and A_1 correspond to the same system in two different temporal or spatial locations and so, $|A_0| = |A_1|$). Physically, the identity channel corresponds to the preservation of coherence for a certain given amount of time. Even though we refer to the our free channels as being "classical," they are still quantum objects. That is, they represent physical processes acting on quantum systems. A summary of different types of channels relevant to different dynamical QRTs of coherence are given in Table 3.1.

Why study a resource theory of classical channels? Here we describe three different motivations. First, some of the most basic non-classical channels are true resources for quantum information processing. For instance, diagonal unitaries such as the T-gate are essential for universal quantum computing. Even the

Channel type	Definition
Detection Incoherent [196, 203]	$\mathcal{D} \circ \mathcal{N} = \mathcal{D} \circ \mathcal{N} \circ \mathcal{D}$
Creation Incoherent (MIO) [38]	$\mathcal{N} \circ \mathcal{D} = \mathcal{D} \circ \mathcal{N} \circ \mathcal{D}$
Detection and Creation Incoherent (DIO) [40, 41]	$\mathcal{D} \circ \mathcal{N} = \mathcal{N} \circ \mathcal{D}$
Incoherent Storage	$\mathcal{N} = \mathcal{N} \circ \mathcal{D}$
Incoherent Output	$\mathcal{N} = \mathcal{D} \circ \mathcal{N}$
Classical [studied here]	$\mathcal{N} = \mathcal{D} \circ \mathcal{N} \circ \mathcal{D}$

Table 3.1: Different classes of free channels in dynamical resource theories of coherence

identity channel can be considered as a resource since all physical systems undergo decoherence, and the preservation of coherence in a quantum memory (for some given time or some specified distance [1]) should thus be considered a resource. While both the T gate and the identity are detection incoherent, they are identified as dynamical resources when limiting the free channels to be classical. In this regard, all the non-classical channels form the dynamical resources of our theory. The quantum Fourier transform (QFT) channel that can generate maximal coherence from free states (due to its unitarity) and preserve entanglement is the maximally coherent channel. (We have shown this analytically in Appendix A.3 by proving that the QFT channel attains the upper bound of log-robustness of coherence of channels which is a monotone in our resource theory.)

Second, quantum cloud computing scenarios like that depicted in Fig. 3.1 are soon to be physically realized [207]. Having large amounts of noise between the client and cloud computer is to be expected, especially as the spatial separation increases. A highly practical question is then what advantages are possible in the very noisy regime. From this perspective all but the completely dephasing upload/download channels should be deemed as yielding a potential resource for quantum cloud computing. The dynamical QRT we propose here embodies this perspective.

Third, a resource theory in which classical channels constitute the free objects is simple enough that entropic resource-theoretic measures can be analytically derived. Compared to static resource theories, a plethora of new resource measures arise in dynamical theories, and the abstract theory of these measures has been recently developed [72, 89, 119, 143, 145, 148, 196, 197, 198, 204, 206, 208, 209, 210, 211, 212]. Unfortunately, the application of this abstract theory to concrete resource theories is usually quite challenging. Here we provide a rare example of a physically-motivated resource theory in which, for example, channel-divergence resource measures can actually be computed.

It may be challenged that since the set of classical channels is so small, almost all quantum channels are resources and thus the resource theory considered here offers little physical insight into coherence. However,

[1] Note that it is important to consider some specified amount of time or distance, or else one might argue that the preservation of coherence for a longer duration of time (or for a longer distance) is more resourceful than for lesser amount of time (or shorter distance). Since this is not the primary problem addressed in this work, we leave it as an open question to be explored further.

almost all quantum states are not diagonal in the incoherent basis, and so the same argument could be alleged toward the static resource theory of coherence. Arguably some insight into static quantum coherence has been gained by its recent resource-theoretic development, and so we initiate an effort to attain a similar insight into dynamical quantum coherence. With the free dynamical "states" identified in this resource theory, we now turn to the free operations. This will be some collection of superchannels, which are linear maps that map a quantum channel to another quantum channel even if acting on part of the channel. A superchannel can be realized using a pre- and a post-processing channel. The details of supermaps and superchannels have been presented in preliminary section 2.2.7. For the case of dynamical coherence considered here, the set of free superchannels must map the set of classical channels to itself. Since there are many different superchannels having this property, which ones should be identified as being free? In previous works on dynamical coherence [143, 196], the free superchannels were constructed by concatenation of free channels in series or in parallel. In general, this is the most common approach for constructing free superchannels [143, 144, 196, 197, 198, 211, 213, 214]. However, as argued above in the case of static coherence, a free implementation of the allowed operations should not necessarily be required in order to detect or learn about the resource contained in a state. We now apply this principle on the level of superchannels.

For example, like MIO in the QRT of static coherence, we define as one class of free superchannels the set of maximally incoherent superchannels (MISC), which is the entirety of all superchannnels that do not generate non-classical channels from classical ones. Similar to MIO in the static case, MISC cannot be implemented without coherence-generating channels. Indeed, if we take the pre-processing channel to be any detection-incoherent channel (as defined in Eq. (3.1)) and the post-processing channel to be any maximally incoherent channel, then we obtain a superchannel that belongs to MISC but its pre- and post-processing channels are non-classical. Nonetheless, much like the argument in static coherence, since we are interested in quantifying the coherence of a channel (as opposed to the coherence of a superchannel), we can use such superchannels as they cannot generate coherence at the channel level, even if it is tensored with the identity superchannel(i.e. it is CRNG).

The bulk of this chapter is devoted to developing a resource theory of dynamical coherence based on the ideas just described. This requires borrowing a few mathematical tools like the concept of log-robustness, the concept of channel divergence, liberal smoothing, etc. from references like [119, 145] that establish the formal structure of dynamical QRTs. Since our work presented here is more mathematical in nature, we would now like to briefly summarize the broad ideas and problems addressed in various sections of this article. This will also serve to highlight the main results of our work.

In section 3.3, we define four different sets of free superchannels: maximally incoherent superchannels (MISC), dephasing-covariant incoherent superchannels (DISC), incoherent superchannels (ISC), and strictly

incoherent superchannel (SISC), which are the analog of MIO, DIO, IO and SIO, in the static case. We focus specifically on the QRTs of MISC and DISC. Similar to how MIO is defined with respect to the dephasing channel, we define MISC with respect to dephasing superchannel, Δ (whose pre- and post-processing channels are dephasing channels) in the following way

$$\Theta \in \mathrm{MISC}(A \to B) \iff \Delta_B \circ \Theta_{A\to B} \circ \Delta_A = \Theta_{A\to B} \circ \Delta_A \,. \tag{3.5}$$

where $\mathrm{MISC}(A \to B)$ means that the superchannel Θ converts a quantum channel that takes system A_0 to A_1 to another quantum channel that takes system B_0 to B_1. Its illustration is given in figure 3.2.

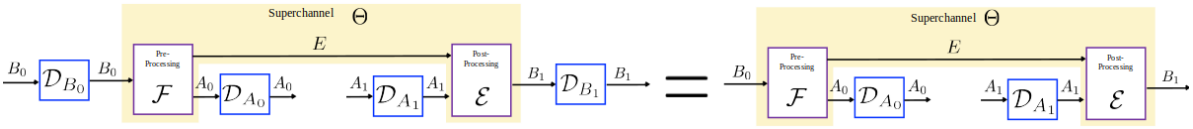

Figure 3.2: MISC

DISC is defined analogously to how DIO is defined in static coherence, i.e.,

$$\Theta \in \mathrm{DISC}(A \to B) \iff \Delta_B \circ \Theta_{A\to B} = \Theta_{A\to B} \circ \Delta_A \tag{3.6}$$

and its illustration is given in figure 3.3. In our work, we provide simple characterization of MISC and DISC in terms of their Choi matrices in (3.57) and (3.65), respectively.

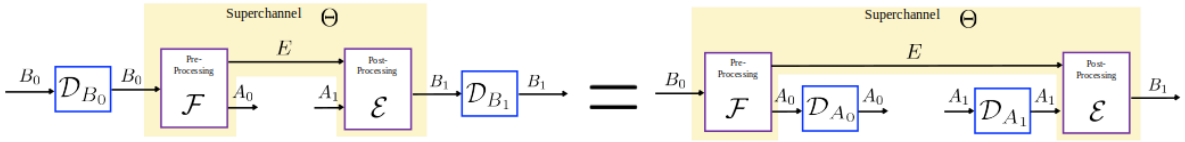

Figure 3.3: DISC

In section 3.4, we study the quantification of dynamical coherence using techniques from QRT of quantum processes [143, 144, 145] and section 3.5 is dedicated to the study of the interconversion of channels (i.e., simulation of one channel with another) under MISC and DISC.

In section 3.4, we first discuss a complete family of monotones under MISC and DISC, and show that these functions can be computed using a semi-definite program. A semi-definite program or an SDP is a subfield of convex optimization. These optimization problems require the variable to be a symmetric matrix which is positive-semidefinite. Section 3.4.2 discusses the relative entropies for the quantification of dynamical coherence, which are relevant information quantities to consider for quantum information tasks. Section 3.4.3 talks about the montones that have an operational interpretation when we discuss the interconversion of

quantum channels. We now list below a few key definitions used in the paper. First, we define the relative entropy of dynamical coherence under MISC to be (for any quantum channel $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$)

$$\begin{aligned} C\left(\mathcal{N}_A\right) &:= \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} D\left(\mathcal{N}_A \middle\| \mathcal{M}_A\right) \\ &:= \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} \max_{\phi \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right) \middle\| \mathcal{M}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right)\right) \end{aligned} \tag{3.7}$$

where $\mathfrak{C}(A_0 \to A_1)$ denotes the set of all classical channels, $\mathfrak{D}(R_0 A_0)$ denotes the set of density matrices on system $R_0 A_0$, and $D(\rho \| \sigma) = \mathrm{Tr}[\rho \log \rho - \rho \log \sigma]$ is the quantum relative entropy. This monotone is faithful, i.e., zero iff $\mathcal{N}_A \in \mathfrak{C}(A_0 \to A_1)$, and does not increase under MISC. For DISC, we define the relative entropy of dynamical coherence to be the function D_Δ, given by

$$D_\Delta(\mathcal{N}_A) := D\left(\mathcal{N}_A \middle\| \Delta_A\left[\mathcal{N}_A\right]\right) \ . \tag{3.8}$$

We show that it is a faithful monotone under DISC.

Similarly, the log-robustness of dynamical coherence is defined as

$$LR_{\mathfrak{C}}(\mathcal{N}_A) := \min_{\mathcal{E} \in \mathfrak{C}(A_0 \to A_1)} D_{\max}\left(\mathcal{N}_A \| \mathcal{E}_A\right) \tag{3.9}$$

and the dephasing log-robustness of dynamical coherence as

$$LR_\Delta(\mathcal{N}_A) := D_{\max}\left(\mathcal{N}_A \middle\| \Delta_A[\mathcal{N}_A]\right) \quad \forall\, \mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1) \ . \tag{3.10}$$

where $D_{\max}(\mathcal{E}_A \| \mathcal{F}_A)$ is the max-relative entropy between two CP maps $\mathcal{E}_A$ and $\mathcal{F}_A$ and is discussed in detail in the preliminary section 3.2.2. We prove that both these quantities are additive under tensor product and have operational interpretations as the exact dynamical coherence costs in the MISC and DISC cases, respectively. We also compute numerically the log-robustness of coherence for qubit channels and show that in the qubit case, the Hadamard channel attains the maximum value.

In section 3.5, the first subsection, section 3.5.1 discusses the general conditions of conversion of one channel using another. This is done by constructing functions called conversion distance such that if the conversion distance from a channel, say $\mathcal{N}_A$ to another channel, say $\mathcal{M}_B$, is zero, then one can perfectly simulate $\mathcal{M}_B$, using $\mathcal{N}_A$ and the free superchannels. Moreover, we use a diamond norm to define the interconversion distance, $d_{\mathfrak{F}}(\mathcal{N}_A \to \mathcal{M}_B) := \min_{\Theta \in \mathfrak{F}(A \to B)} \frac{1}{2} \left\| \Theta_{A \to B}\left[\mathcal{N}_A\right] - \mathcal{M}_B \right\|_\diamond$ between two quantum channels, $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$ and $\mathcal{M}_B \in \mathrm{CPTP}(B_0 \to B_1)$ and show that if the set of free superchannels $\mathfrak{F} =$ MISC or DISC, then $d_{\mathfrak{F}}(\mathcal{N}_A \to \mathcal{M}_B)$ can be computed using a semi-definite program (SDP).

The diamond norm (see (2.85)) used in the equation of the interconversion distance is used to measure the closeness or the distance between two quantum channels. We use this function to find the interconversion distance in the qubit case, numerically. We find that a maximally coherent replacement channel (a channel that outputs the maximally coherent state $\phi^+_{B_1}$) can be simulated by a Hadamard channel using the maximally incoherent superchannels. This was expected because any coherent channel can be simulated using the maximally coherent channel and the free superchannels. More interestingly, we also found that we can simulate a Hadamard channel using two maximal replacement channels. Apart from these numerical calculations, we show in Appendix A.3 that the ratio between the log-robstness of coherence of a maximally coherent channel and the log-robustness of coherence of maximally coherent replacement channel for any dimension is always 2, i.e., using two maximally replacement channels and the free superchannels, we can simulate the maximally coherent channel which is the quantum Fourier transform channel. Using this fact that we just need two copies of the maximal replacement channel (or two maximally coherent states) to simulate the maximally coherent channel, we can find the coherence cost of a channel using the maximally coherent state. Sections 3.5.2, 3.5.3, and 3.5.4, discuss various types of coherence costs of channels. We define the coherence cost of a channel as the minimum amount of the maximally coherent state to be used to simulate the given channel. In section 3.5.2, we calculate the exact coherence cost of a channel under MISC and DISC, i.e., when the free superchannel acts on the maximally coherent state the output is the desired channel. Similarly, in section 3.5.3, we consider the problem of finding the approximate coherence cost of a channel (which we also refer to as the coherence cost of a channel), i.e., the amount of the maximally coherent state used to convert it to a channel that is very close to the desired channel. This is interesting because experimentally it is extremely difficult to convert one channel into another perfectly. We then, in section 3.5.4, compute the liberal asymptotic cost of dynamical coherence (which is the dynamical coherence cost of a channel when the smoothing is "liberal" [145]) under MISC, and show that it is equal to a variant of the regularized relative entropy given by

$$D^{(\infty)}_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \lim_{n\to\infty} \frac{1}{n} \sup_{\varphi\in\mathfrak{D}(RA_0)} \min_{\mathcal{E}\in\mathfrak{C}(A_0^n\to A_1^n)} D\left(\mathcal{N}^{\otimes n}_{A_0\to A_1}\left(\varphi^{\otimes n}_{RA_0}\right) \middle\| \mathcal{E}_{A_0^n\to A_1^n}\left(\varphi^{\otimes n}_{RA_0}\right)\right) \tag{3.11}$$

Lastly, in section 3.5.5, we formulate the one-shot distillable dynamical coherence and compute its value for a few specific channels.

3.2 Preliminaries

3.2.1 Elements of quantum resource theory of static coherence

Coherence of a state is a basis-dependent concept. Hence, a basis is fixed first in the resource theory of static coherence. The density matrices that are diagonal in this basis form the free states of the theory. These states are also called incoherent states. Let us denote this set by $\mathcal{I}_{A_1} \subset \mathfrak{B}(A_1)$ for any system A_1. Hence, all the incoherent density operators $\rho_{A_1} \in \mathcal{I}_{A_1}$ have the following form

$$\rho_{A_1} = \sum_{i=0}^{|A_1|-1} p_i |i\rangle\langle i|_{A_1} \tag{3.12}$$

with probabillities p_i and obey

$$\mathcal{D}_{A_1}(\rho_{A_1}) = \rho_{A_1} \tag{3.13}$$

where $\mathcal{D}_{A_1}$ is the dephasing channel for the system A_1 and is defined as

$$\mathcal{D}_{A_1}(\sigma_{A_1}) = \sum_{i=0}^{|A_1|-1} |i\rangle\langle i|\sigma_{A_1}|i\rangle\langle i| \tag{3.14}$$

for any $\sigma_{A_1} \in \mathfrak{D}(A_1)$. For multi-partite systems, the preferred basis is the tensor product of the preferred basis of each subsystem[215, 216, 217]. Note that it is quite possible that the states are not diagonal in a different basis, but it does not matter. That's because, the dephasing operator is defined using the incoherent basis states. So, incoherent states are just those states which are diagonal in the incoherent basis.

From the golden rule of QRT, the free operations are the set of channels that take the set of incoherent states to itself in the complete sense, i.e., they are completey resource non-generating. Such operations are called incoherent operations. In literature, several types of incoherent operations have been studied. The largest set of incoherent operations is known as the maximally incoherent operations (MIO) [38]. Other incoherent operations include incoherent operations (IO) [39], dephasing-covariant incoherent operations (DIO) [40, 41, 194, 195], strictly incoherent operations (SIO) [217, 218], physically incoherent operations (PIO) [40, 194, 195], translationally-invariant operations (TIO) [219], genuinely incoherent operations (GIO) [220], fully incoherent operations (FIO) [220], etc. In this section, we will briefly discuss about MIO, DIO, IO, and SIO, as we will be defining four sets of free superchannels in the next section taking their analogy.

The maximally incoherent operations (or MIO) [38] are defined as the set of CPTP and non-selective maps $\mathcal{E} \in \mathfrak{L}(A_0 \to A_1)$ such that

$$\mathcal{E}(\rho_{A_0}) \in \mathcal{I}_{A_1} \quad \forall\, \rho_{A_0} \in \mathcal{I}_{A_0}\ . \tag{3.15}$$

Let us denote the set of all channels that follow the above property by $\mathrm{MIO}(\mathrm{A}_0 \to \mathrm{A}_1)$. Any CPTP map $\mathcal{M}_{A_0\to A_1} \in \mathrm{MIO}(\mathrm{A}_0 \to \mathrm{A}_1)$ can be characterized using the dephasing channels in the following way

$$\mathcal{M}_{A_0\to A_1} \in \mathrm{MIO}(\mathrm{A}_0 \to \mathrm{A}_1) \iff \mathcal{D}_{\mathrm{A}_1} \circ \mathcal{M}_{\mathrm{A}_0\to\mathrm{A}_1} \circ \mathcal{D}_{\mathrm{A}_0} = \mathcal{M}_{\mathrm{A}_0\to\mathrm{A}_1} \circ \mathcal{D}_{\mathrm{A}_0} \ . \tag{3.16}$$

Despite the fact that MIO cannot create coherence, these operations do not have a free dilation, i.e., they cost coherence to be implemented [40, 41, 194, 195].

A smaller class of free operations, the incoherent operations (or IO) [39] are defined as the set of CPTP maps $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ having a Kraus operator representation $\{K_n\}$ such that

$$\frac{K_n \rho_{A_0} K_n^\dagger}{\mathrm{Tr}[K_n \rho_{A_0} K_n^\dagger]} \in \mathcal{I}_{A_1} \quad \forall\, n \text{ and } \rho_{A_0} \in \mathcal{I}_{A_0} \ . \tag{3.17}$$

This class of operations also do not have a free dilation [40, 41, 194, 195].

The next class of free operations, the strictly incoherent operations (or SIO) [217, 218] are defined as the set of CPTP maps $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ having a Kraus operator representation $\{K_n\}$ such that

$$K_n \mathcal{D}_{A_0}(\rho_{A_0}) K_n^\dagger = \mathcal{D}_{A_1}\left(K_n \rho_{A_0} K_n^\dagger\right) \quad \forall\, n \ . \tag{3.18}$$

This class of operations also do not have a free dilation [40, 194, 195].

The last class of free operations that is useful to us is the dephasing-covariant incoherent operations (or DIO) [40, 41, 194, 195]. A CPTP map $\mathcal{E}_A$ is said to be DIO if

$$[\mathcal{D}, \mathcal{E}_A] = 0 \ , \tag{3.19}$$

which is equivalent to

$$\mathcal{D}_{A_1}\left(\mathcal{E}_{A_0\to A_1}(\rho_{A_0})\right) = \mathcal{E}_{A_0\to A_1}\left(\mathcal{D}_{A_0}(\rho_{A_0})\right) \quad \forall\, \rho_{A_0} \in \mathfrak{D}(A_0) \ . \tag{3.20}$$

3.2.2 Max-relative entropy for channels

The max-relative entropy is defined on a pair (ρ, σ) with $\rho \in \mathfrak{D}(A_1)$ and $\sigma \in \mathrm{Pos}(\mathrm{A}_1)$ of a state ρ with respect to a positive operator σ is given by

$$D_{\max}(\rho\|\sigma) \coloneqq \log \min\{t : t\sigma \geqslant \rho\} \tag{3.21}$$

where the inequality sign means that the difference between l.h.s. and r.h.s. is a positive operator. Similarly for channels, the maximum relative entropy between two CP maps $\mathcal{N}$ and $\mathcal{E}$ is given by

$$D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) \coloneqq \log\min\{t : t\mathcal{E}_A \geqslant \mathcal{N}_A\} \tag{3.22}$$

where the inequality sign means that the difference between l.h.s. and r.h.s. is a CP map. Denoting the Choi matrix of $\mathcal{E}_A$ by $J_A^{\mathcal{E}}$ and that of $\mathcal{N}_A$ by $J_A^{\mathcal{N}}$, (3.22) can be rewritten as

$$D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) = \log\min\left\{t : tJ_A^{\mathcal{E}} \geqslant J_A^N,\, t \geqslant 0\right\} \tag{3.23}$$

The channel max-relative entropy ($D_{\max}(\mathcal{N}_A\|\mathcal{E}_A)$) can be expressed in a simple closed form as a function of the Choi matrices of the maps $\mathcal{N}_A$ and $\mathcal{E}_A$ [148, 221]. This implies that it is also additive under tensor products. For completeness, we give the following proof.

Lemma 3.1. *The max-relative entropy for channels is additive under tensor product, i.e.,*

$$D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'}) = D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'}) \tag{3.24}$$

Proof. For the proof of the inequality $D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'}) \leqslant D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})$, let

$$D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) = \log\{t_1 : t_1\mathcal{E}_A \geqslant \mathcal{N}_A\}\,, \tag{3.25}$$

$$D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'}) = \log\{t_2 : t_2\mathcal{F}_{A'} \geqslant \mathcal{M}_{A'}\}\,. \tag{3.26}$$

We can rewrite $D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'})$ as

$$\begin{aligned} D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'}) &= \log\min\{t : t\,(\mathcal{E}_A \otimes \mathcal{F}_{A'}) \geqslant \mathcal{N}_A \otimes \mathcal{M}_{A'}\} \\ &= \log\min\{t : \frac{t}{t_1t_2}\,(t_1\mathcal{E}_A \otimes t_2\mathcal{F}_{A'}) \geqslant \mathcal{N}_A \otimes \mathcal{M}_{A'}\} \end{aligned} \tag{3.27}$$

From this, we can clearly see

$$\log\min\{t : t\,(\mathcal{E}_A \otimes \mathcal{F}_{A'}) \geqslant \mathcal{N}_A \otimes \mathcal{M}_{A'}\} \leqslant \log(t_1t_2) \tag{3.28}$$

Hence,

$$D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'}) \leqslant D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'}) \tag{3.29}$$

For the proof of $D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}\|\mathcal{E}_A \otimes \mathcal{F}_{A'}) \geqslant D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})$, note that $D_{\max}$ in

(3.22) and (3.23) can be computed using an SDP and its dual is given by

$$D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) = \log \max\left\{\mathrm{Tr}[\beta_A J_A^{\mathcal{N}}] : \mathrm{Tr}\left[\beta_A J_A^{\mathcal{E}}\right] \leqslant 1\,,\ \beta_A \geqslant 0\right\}\ . \tag{3.30}$$

Now let β_A^1 and $\beta_{A'}^2$ be optimal for $D_{\max}(\mathcal{N}_A\|\mathcal{E}_A)$ and $D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})$, respectively. Therefore,

$$2^{D_{\max}(\mathcal{N}_A\|\mathcal{E}_A)} = \mathrm{Tr}[\beta_A^1 J_A^{\mathcal{N}}]\,, \tag{3.31}$$

$$2^{D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})} = \mathrm{Tr}[\beta_{A'}^2 J_{A'}^{\mathcal{M}}]\,. \tag{3.32}$$

Using (3.30), we can express $2^{D_{\max}(\mathcal{N}_A\otimes\mathcal{M}_{A'}\|\mathcal{E}_A\otimes\mathcal{F}_{A'})}$ as

$$2^{D_{\max}(\mathcal{N}_A\otimes\mathcal{M}_{A'}\|\mathcal{E}_A\otimes\mathcal{F}_{A'})} = \max\left\{\mathrm{Tr}\left[\beta_{AA'}\left(J_A^{\mathcal{N}}\otimes J_{A'}^{\mathcal{M}}\right)\right] : \mathrm{Tr}\left[\beta_{AA'}\left(J_A^{\mathcal{E}}\otimes J_{A'}^{\mathcal{F}}\right)\right] \leqslant 1\,,\ \beta_{AA'} \geqslant 1\right\} \tag{3.33}$$

Since the choice of $\beta_{AA'} = \beta_A^1 \otimes \beta_{A'}^2$ satisfies the above constraint, we can say

$$\begin{aligned} 2^{D_{\max}(\mathcal{N}_A\otimes\mathcal{M}_{A'}\|\mathcal{E}_A\otimes\mathcal{F}_{A'})} &\geqslant \mathrm{Tr}\left[\beta_A^1 J_A^{\mathcal{N}}\right]\ \mathrm{Tr}\left[\beta_{A'}^2 J_{A'}^{\mathcal{M}}\right] \\ &\geqslant 2^{D_{\max}(\mathcal{N}_A\|\mathcal{E}_A)} 2^{D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})} \end{aligned} \tag{3.34}$$

which implies

$$D_{\max}(\mathcal{N}_A\otimes\mathcal{M}_{A'}\|\mathcal{E}_A\otimes\mathcal{F}_{A'}) \geqslant D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'}) \tag{3.35}$$

From (3.29) and (3.35), we can conclude that the max rel-entropy for channels is additive under tensor products, i.e., $D_{\max}(\mathcal{N}_A\otimes\mathcal{M}_{A'}\|\mathcal{E}_A\otimes\mathcal{F}_{A'}) = D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}\|\mathcal{F}_{A'})$. ∎

Lastly, I have also used the ϵ-smooth max-relative entropy in this chapter (which is discussed in detail in [144, 198, 222]) to find approximate coherence cost. It is defined in the following way

$$D_{\max}^{\epsilon}(\mathcal{N}_A\|\mathcal{M}_A) \coloneqq \inf_{\mathcal{N}_A'\in B_\epsilon(\mathcal{N}_A)} D_{\max}(\mathcal{N}_A'\|\mathcal{M}_A) \tag{3.36}$$

where

$$B_\epsilon(\mathcal{N}_A) = \left\{\mathcal{N}_A' \in \mathrm{CPTP}(A_0\to A_1) : \frac{1}{2}\|\mathcal{N}_A' - \mathcal{N}_A\|_\diamond \leqslant \epsilon\right\}\ . \tag{3.37}$$

where $\|\cdot\|_\diamond$ is the diamond norm (as defined in Eq. (2.85)) that measures the distance between two quantum channels.

3.3 The set of free superchannels

As discussed in the introduction Sec. 3.1, the set of free channels in the theory of dynamical coherence are classical channels. Therefore, a free superchannel consists of a pre-processing classical channel and a post-processing classical channel (see Fig. 3.4). However, such a free superchannel always destroys completely any resource; that is, it converts all channels (even coherent ones) into classical channels. This means that the resource theory is in a sense "degenerate" and no interesting consequences can be concluded from such a theory.

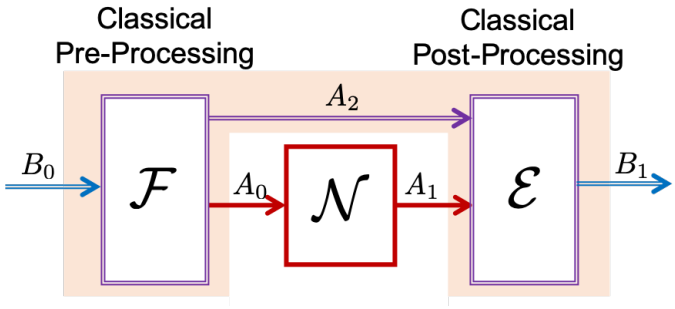

Figure 3.4: The action of a classical superchannel on a quantum channel.

This above type of degeneracy also occurs with the resource theory of coherence in the *state domain*. There, the only free operations that are physically consistent are PIO [40], which are very restricted and cannot provide much insight into the phenomenon of coherence in quantum systems. Therefore, almost all the enormous amount of work in recent years on the QRT of coherence was devoted to the study of coherence under much larger sets of operations, such as MIO, DIO, IO, and SIO. While these larger sets of operations cannot be implemented without a coherence cost, they do not generate coherence, and as such they can be used for the study of coherence of states. However, since MIO, DIO, IO, and SIO all have a coherence cost, they cannot be used as the "free operations" in a resource theory that aims to quantify the coherence of quantum channels.

Instead, for a dynamical QRT of coherence, one can define free superchannels that form a larger set than classical superchannels. Similar to what happens in the state domain, there is a coherent cost to implement such superchannels, however, they do not generate dynamical coherence, and therefore can be used in a dynamical resource theory of coherence. As it happens in the state domain, there are several natural sets of free superchannels that we can define.

3.3.1 Maximally Incoherent Superchannels (MISC)

In any quantum resource theory, free operations cannot generate a resource. Taking this principle to the level of superchannels, we define the maximal incoherent superchannels (MISC) as follows.

Definition 3.2. Given two dynamical systems A and B, a superchannel $\Theta \in \mathfrak{S}(A \to B)$ is said to be MISC if

$$\Theta_{A\to B}[\mathcal{N}_A] \in \mathfrak{C}(B_0 \to B_1) \quad \forall \mathcal{N}_A \in \mathfrak{C}(A_0 \to A_1)\ . \tag{3.38}$$

In words, a superchannel is said to be MISC if for every input classical channel, the output is also a classical channel. We denote the set of all superchannels that have the above property by $\mathrm{MISC}(A \to B)$.

Remark 1. Similar to the characterization of MIO channels with the dephasing channel, the condition that Θ is in $\mathrm{MISC}(A \to B)$ can be characterized with the dephasing superchannels Δ_A and Δ_B. (A dephasing superchannel can be realized using dephasing channels as pre- and post-processing channels.) Specifically, we have that

$$\Theta \in \mathrm{MISC}(A \to B) \iff \Delta_B \circ \Theta_{A\to B} \circ \Delta_A = \Theta_{A\to B} \circ \Delta_A\ . \tag{3.39}$$

That is, for any input quantum channel $\mathcal{E}_A \in \mathrm{CPTP}(A_0 \to A_1)$, if a superchannel $\Theta_{A\to B}$ obeys the equation on rhs ($\Delta_B \circ \Theta_{A\to B} \circ \Delta_A[\mathcal{E}_A] = \Theta_{A\to B} \circ \Delta_A[\mathcal{E}_A]$) then, that superchannel belongs to MISC and vice-versa. Refer figure 3.2 for an illustrative diagram. To explain it further, notice that the dephasing superchannel converts any input to a classical channel. So, for any input quantum channel, say $\mathcal{E}_A$, the dephasing superchannel Δ_A first converts it to a classical channel, $\Delta_A[\mathcal{E}_A] = \mathcal{N}_A \in \mathfrak{C}(A_0 \to A_1)$ which goes as input to the superchannel Θ. So for a classical channel $\mathcal{N}_A$, the rhs of the above condition can be written as $\Delta_B \circ \Theta_{A\to B}[\mathcal{N}_A] = \Theta_{A\to B}[\mathcal{N}_A]$. The lhs of this equation again has a dephasing channel which implies that whatever the output is after the action of the superchannel Θ on the classical channel $\mathcal{N}$, the output would still be classical.

As stated earlier, the maximally incoherent superchannel need not be realized using classical pre- and post-processing channel. For instance, if we use the detection incoherent channels (as defined in [196]) and maximally incoherent operations (MI0) as the pre- and post-processing channels, the resultant superchannel is a maximally incoherent superchannel or MISC. Refer to figure 3.5 below as an illustration.

To show that the above realization is really a MISC, we use the fact that an operation $\mathcal{E}_A \in \mathrm{CPTP}(A_0 \to A_1)$ is called detection incoherent operation iff $\mathcal{D}_{A_1} \circ \mathcal{E}_A = \mathcal{D}_{A_1} \circ \mathcal{E}_A \circ \mathcal{D}_{A_0}$ [196] where $\mathcal{D}$ is the completely dephasing channel for the given system. Recall that an operation $\mathcal{F}_A \in \mathrm{CPTP}(A_0 \to A_1)$ is called maximally incoherent operation if it follows (3.16). Using these definitions, we can see that the above realization

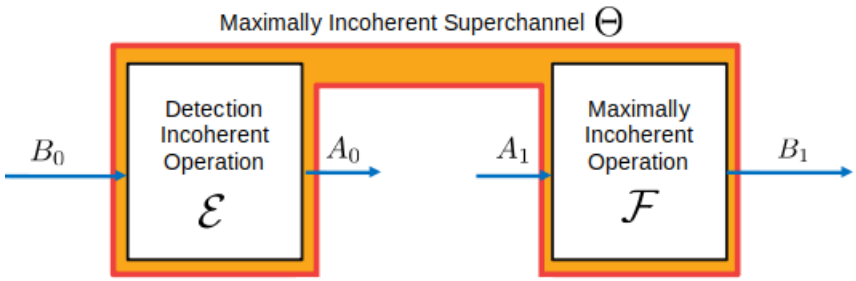

Figure 3.5: An example of realization of maximally incoherent superchannel (MISC)

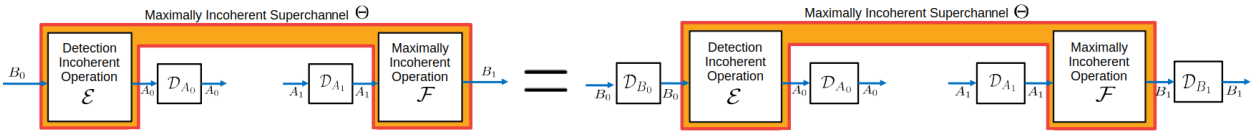

Figure 3.6: Illustration to show that this particular realization is a maximally incoherent superchannel (MISC)

follows (3.39) as illustrated in figure 3.6. One of the key properties of any resource theory is that the free operations are "completely free". This is a physical requirement that a free channel (or superchannel) can act on a subsystem. In the following theorem we show that $\mathrm{MISC}(A \to B)$ is completely free. That is, in the QRT we consider here, there is no difference between RNG and completely RNG.

Theorem 3.3. *Let A and B be two dynamical systems, and let $\Theta \in \mathrm{MISC}(A \to B)$. Then, for any dynamical system R, the superchannel $\mathbb{1}_R \otimes \Theta$ is free; i.e. $\mathbb{1}_R \otimes \Theta \in \mathrm{MISC}(RA \to RB)$.*

Proof. Let $\mathcal{N}_{RA} \in \mathfrak{C}(R_0 A_0 \to R_1 A_1)$ be a classical channel satisfying

$$\Delta_{RA}\left[\mathcal{N}_{RA}\right] = \Delta_R \otimes \Delta_A\left[\mathcal{N}_{RA}\right] = \mathcal{N}_{RA}\ . \tag{3.40}$$

Then,

$$\begin{aligned}
\Delta_{RB} \circ \left(\mathbb{1}_R \otimes \Theta_{A\to B}\right)\left[\mathcal{N}_{RA}\right] &= \Delta_R \otimes \left(\Delta_B \circ \Theta_{A\to B}\right)\left[\mathcal{N}_{RA}\right] && (3.41)\\
&= \mathbb{1}_R \otimes \left(\Delta_B \circ \Theta_{A\to B}\right)\left[\mathcal{N}_{RA}\right] && (3.42)\\
&= \mathbb{1}_R \otimes \left(\Delta_B \circ \Theta_{A\to B} \circ \Delta_A\right)\left[\mathcal{N}_{RA}\right] && (3.43)\\
&= \mathbb{1}_R \otimes \left(\Theta_{A\to B} \circ \Delta_A\right)\left[\mathcal{N}_{RA}\right] && (3.44)\\
&= \mathbb{1}_R \otimes \Theta_{A\to B}\left[\mathcal{N}_{RA}\right] && (3.45)
\end{aligned}$$

where the first equality follows from the equality $\Delta_{RA} = \Delta_R \otimes \Delta_A$, the second equality from the fact that $\mathcal{N}_{RA}$ is classical and in particular $\Delta_R\left[\mathcal{N}_{RA}\right] = \mathcal{N}_{RA}$, the third equality from the similar equality

$\Delta_A[\mathcal{N}_{RA}] = \mathcal{N}_{RA}$, the fourth equality from (3.39), and the last equality follows again from $\Delta_A[\mathcal{N}_{RA}] = \mathcal{N}_{RA}$. Hence, $\mathbb{1}_R \otimes \Theta_{A\to B}[\mathcal{N}_{RA}]$ is classical so that $\mathbb{1}_R \otimes \Theta \in \text{MISC}(RA \to RB)$. This completes the proof. ∎

The theorem above indicates that MISC can be viewed as the set of completely resource non-generating superchannels in the theory of dynamical coherence. We next consider the characterization of the set MISC. Recall that in the state domain, we can determine if a channel $\mathcal{E}_A$ belongs to $\text{MIO}(A_0 \to A_1)$ simply by checking if all the states $\mathcal{E}_A(|x\rangle\langle x|_{A_0})$ are diagonal for all $x = 1, ..., |A_0|$. This simplicity of MIO implies that all state conversions in the single-shot regime can be determined with SDP. In the channel domain, however, the characterization of MISC is slightly more complex.

The Choi matrix of any classical channel $\mathcal{N}_A \in \text{CPTP}(A_0 \to A_1)$ can be expressed as a column stochastic matrix. Recall that the action of this classical channel $\mathcal{N}_A$ on any input quantum state ρ_{A_0} can be viewed as the action of this column stochastic matrix on a vector whose components are the diagonal entries of the input quantum state ρ_{A_0}. The output vector's components then form the diagonal entries of the output state and the off-diagonal entries of this output state are zero. The set of all extreme points of the set of $|A_0| \times |A_1|$ column stochastic matrices consists of matrices that in each column has $|A_0| - 1$ zeros and 1 one. Therefore the number of extreme points is given by $|A_0|^{|A_1|}$. This may give the impression that in order to check if $\Theta \in \text{MISC}(A_0 \to A_1)$ one has to check if the channel $\Theta[\mathcal{E}_A]$ is classical for all the $|A_0|^{|A_1|}$ extreme classical channels. Since the number of conditions is exponential in $|A_1|$ it may give the impression that the problem of deciding if a superchannel belongs to MISC cannot be solved with SDP. However, we show now that this problem can be solved with polynomial (in $|A_0A_1|$) number of constraints. It can be seen from the relationship between the Choi matrix of $\Theta_{A\to B}$ and that of $\Theta_{A\to B} \circ \Delta_A$ and $\Delta_B \circ \Theta_{A\to B}$.

Lemma 3.4. *Let A and B be two dynamical systems, $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel, and $\Delta_A \in \mathfrak{S}(A \to A)$ and $\Delta_B \in \mathfrak{S}(B \to B)$ be the completely dephasing superchannels. Then, the Choi matrices of $\Theta_{A\to B}$, $\Theta_{A\to B} \circ \Delta_A$, and $\Delta_B \circ \Theta_{A\to B}$, satisfy the relations*

$$\mathbf{J}_{AB}^{\Theta\circ\Delta_A} = \mathcal{D}_A\left(\mathbf{J}_{AB}^{\Theta}\right) \quad \text{and} \quad \mathbf{J}_{AB}^{\Delta_B\circ\Theta} = \mathcal{D}_B\left(\mathbf{J}_{AB}^{\Theta}\right) \tag{3.46}$$

Proof. The Choi matrix of a superchannel Θ can be expressed as the Choi matrix of the bipartite channel $\Theta_{\tilde{A}\to B}\left[\mathcal{P}^+_{A\tilde{A}}\right]$ [119]. Similarly, the Choi matrix of the superchannel $\Theta \circ \Delta_A$ can be expressed as the Choi matrix of the bipartite channel $\Theta_{\tilde{A}\to B} \circ \Delta_{\tilde{A}}\left[\mathcal{P}^+_{A\tilde{A}}\right]$ and that of the superchannel $\Delta_B \circ \Theta$ as the Choi matrix of $\Delta_B \circ \Theta_{\tilde{A}\to B}\left[\mathcal{P}^+_{A\tilde{A}}\right]$.

Denoting $\Theta_{\tilde{A}\to B}[\mathcal{P}^+_{A\tilde{A}}]$ as $\mathcal{N}_{AB}$, the Choi matrix of the superchannel $\Delta_B \circ \Theta_{A\to B}$ can be written as

$$\mathbf{J}^{\Delta_B\circ\Theta}_{AB} = J^{\Delta_B[\mathcal{N}_{AB}]}_{AB} \tag{3.47}$$

$$= \mathcal{D}_{B_1} \circ \mathcal{N}_{\tilde{A}_0\tilde{B}_0\to A_1B_1} \circ \mathcal{D}_{\tilde{B}_0}\left(\phi^+_{A_0\tilde{A}_0} \otimes \phi^+_{B_0\tilde{B}_0}\right) \tag{3.48}$$

where the second equality follows from the definition of the Choi matrix of a channel. Now using the fact that for a given channel $\mathcal{M} \in \mathrm{CPTP}(R_0 \to R_1)$, we have $\mathcal{M}_{\tilde{R}_0\to R_1}|\phi^+_{R_0\tilde{R}_0}\rangle = \mathcal{M}^T_{\tilde{R}_1\to R_0}|\phi^+_{\tilde{R}_1R_1}\rangle$, we can rewrite (3.48) as

$$\mathbf{J}^{\Delta_B\circ\Theta}_{AB} = (\mathcal{D}_{B_0} \otimes \mathcal{D}_{B_1}) \circ \mathcal{N}_{\tilde{A}_0\tilde{B}_0\to A_1B_1}\left(\phi^+_{A_0\tilde{A}_0} \otimes \phi^+_{B_0\tilde{B}_0}\right) \tag{3.49}$$

$$= \mathcal{D}_B\left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.50}$$

To find $\mathbf{J}^{\Theta\circ\Delta_A}$, note that for any superchannel $\Omega \in \mathfrak{S}(A \to B)$ we have [119]

$$\mathbb{1}_A \otimes \Omega_{\tilde{A}\to B}[\mathcal{P}^+_{A\tilde{A}}] = \Omega^T_{\tilde{B}\to A} \otimes \mathbb{1}_B[\mathcal{P}^+_{\tilde{B}B}] \ . \tag{3.51}$$

From this, it can be calculated that for the dephasing superchannel, $\Delta^T = \Delta$. Therefore, we have

$$\Theta_{\tilde{A}\to B} \circ \Delta_{\tilde{A}}\left[\mathcal{P}^+_{A\tilde{A}}\right] = \Theta_{\tilde{A}\to B} \circ \Delta^T_A\left[\mathcal{P}^+_{A\tilde{A}}\right] \tag{3.52}$$

$$= \Theta_{\tilde{A}\to B} \circ \Delta_A\left[\mathcal{P}^+_{A\tilde{A}}\right] \tag{3.53}$$

$$= \Delta_A \circ \Theta_{\tilde{A}\to B}\left[\mathcal{P}^+_{A\tilde{A}}\right] \tag{3.54}$$

$$= \Delta_A \circ \mathcal{N}_{AB} \tag{3.55}$$

where the third equality arises because the superchannels Θ and Δ are acting on different subsystems (notice the subscripts with and without tilde). So, the Choi matrix of $\Theta_{\tilde{A}\to B} \circ \Delta_{\tilde{A}}[\mathcal{P}^+_{A\tilde{A}}]$ is equal to finding the Choi matrix of $\Delta_A \circ \mathcal{N}_{AB}$. From the calculation of the Choi matrix of $\Delta_B \circ \mathcal{N}_{AB}$ above, we can conclude that

$$\mathbf{J}^{\Theta\circ\Delta_A}_{AB} = \mathcal{D}_A\left(\mathbf{J}^{\Theta}_{AB}\right) \tag{3.56}$$

∎

With this lemma at hand we get the following characterization for the set $\mathrm{MISC}(A \to B)$.

Theorem 3.5. *Let A and B be two dynamical systems, and $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel. Then,*

$\Theta \in \mathrm{MISC}(A \to B)$ *if and only if*

$$\mathcal{D}_{AB}\left(\mathbf{J}^{\Theta}_{AB}\right) = \mathcal{D}_A \otimes \mathsf{id}_B\left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.57}$$

Proof. The characterization of MISC in (3.39) can be equivalently written in terms of the Choi matrices of the channels in lhs and rhs. Using this property and the lemma above we have that

$$\mathbf{J}^{\Delta_B \circ \Theta \circ \Delta_A}_{AB} = \mathbf{J}^{\Theta \circ \Delta_A}_{AB} = \mathcal{D}_A \otimes \mathsf{id}_B\left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.58}$$

Similarly, using both results from Lemma 3.4, we can write

$$\mathbf{J}^{\Delta_B \circ \Theta \circ \Delta_A}_{AB} = \mathsf{id}_A \otimes \mathcal{D}_B\left(\mathbf{J}^{\Theta \circ \Delta_A}_{AB}\right) = \mathcal{D}_{AB}\left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.59}$$

Equating (3.58) and (3.59), we get

$$\mathcal{D}_{AB}\left(\mathbf{J}^{\Theta}_{AB}\right) = \mathcal{D}_A \otimes \mathsf{id}_B\left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.60}$$

This completes the proof. ∎

Note that for any Hermitian matrix $Z_{AB} \in \mathrm{Herm}(AB)$ we have

$$\mathrm{Tr}\left[\left(\mathcal{D}_{AB}\left(\mathbf{J}^{\Theta}_{AB}\right) - \mathcal{D}_A \otimes \mathsf{id}_B\left(\mathbf{J}^{\Theta}_{AB}\right)\right) Z_{AB}\right] = \mathrm{Tr}\left[\mathbf{J}^{\Theta}_{AB}\left(\mathcal{D}_{AB}\left(Z_{AB}\right) - \mathcal{D}_A \otimes \mathsf{id}_B\left(Z_{AB}\right)\right)\right] \tag{3.61}$$

Therefore, the theorem above implies that $\Theta \in \mathrm{MISC}(A \to B)$ if and only if

$$\mathrm{Tr}\left[\mathbf{J}^{\Theta}_{AB} X_{AB}\right] = 0 \quad \forall X_{AB} \in \mathfrak{K}_{\mathrm{MISC}} \tag{3.62}$$

where $\mathfrak{K}_{\mathrm{MISC}}$ is a subspace of $\mathrm{Herm}(AB)$ defined as

$$\mathfrak{K}_{\mathrm{MISC}} \coloneqq \left\{\mathcal{D}_{AB}\left(Z_{AB}\right) - \mathcal{D}_A \otimes \mathsf{id}_B\left(Z_{AB}\right) \ : \ Z_{AB} \in \mathrm{Herm}(AB)\right\} . \tag{3.63}$$

Since the dimension of the subspace $\mathfrak{K}_{\mathrm{MISC}}$ is $|AB|(|B|-1)$, it is sufficient to restrict X_{AB} in (3.62) to the $|AB|(|B|-1)$ elements of some fixed basis of $\mathfrak{K}_{\mathrm{MISC}}$. Note also that the condition above is equivalent to the inclusion $\mathbf{J}^{\Theta}_{AB} \in \mathfrak{K}^{\perp}_{\mathrm{MISC}}$, where $\mathfrak{K}^{\perp}_{\mathrm{MISC}}$ is the orthogonal complement of $\mathfrak{K}_{\mathrm{MISC}}$ in $\mathrm{Herm}(AB)$.

3.3.2 Dephasing Incoherent Superchannels (DISC)

In the QRT of static coherence, the dephasing channel plays a major role, and in particular, leading to the definition of DIO. Here, the dephasing superchannel defined by $\Delta_A[\mathcal{N}_A] = \mathcal{D}_{A_1} \circ \mathcal{N}_A \circ \mathcal{D}_{A_0}$ plays a similar role, as we have already seen in the definition of MISC. We use here the dephasing superchannel to define the set of dephasing incoherent superchannels.

Definition 3.6. Let A and B be two dynamical systems, and let $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel. Then, Θ is said to be a dephasing incoherent superchannel (DISC) if and only if

$$\Delta_B \circ \Theta_{A\to B} = \Theta_{A\to B} \circ \Delta_A \ . \tag{3.64}$$

Moreover, the set of all such superchannels that satisfy the above relation is denoted by $\mathrm{DISC}(A \to B)$.

Clearly, from its definition $\mathrm{DISC}(A \to B)$ is a subset of $\mathrm{MISC}(A \to B)$, and in particular, it is completely free. Now, from Lemma 3.4 it follows that a superchannel $\Theta \in \mathrm{DISC}(A \to B)$ if and only if

$$\mathcal{D}_A \otimes \mathsf{id}_B \left(\mathbf{J}^{\Theta}_{AB}\right) = \mathsf{id}_A \otimes \mathcal{D}_B \left(\mathbf{J}^{\Theta}_{AB}\right) \ . \tag{3.65}$$

Moreover, similar to the considerations above, since the map $\mathcal{D}_A \otimes \mathsf{id}_B - \mathsf{id}_A \otimes \mathcal{D}_B$ is self adjoint, it follows that $\Theta \in \mathrm{DISC}(A \to B)$ if and only if

$$\mathrm{Tr}\left[\mathbf{J}^{\Theta}_{AB} Y_{AB}\right] = 0 \quad \forall Y_{AB} \in \mathfrak{K}_{\mathrm{DISC}} \tag{3.66}$$

where

$$\mathfrak{K}_{\mathrm{DISC}} \coloneqq \left\{ \mathsf{id}_A \otimes \mathcal{D}_B \left(Z_{AB}\right) - \mathcal{D}_A \otimes \mathsf{id}_B \left(Z_{AB}\right) \ : \ Z_{AB} \in \mathrm{Herm}(AB) \right\} . \tag{3.67}$$

Since the dimension of the subspace $\mathfrak{K}_{\mathrm{DISC}}$ is $|AB|(|A|+|B|-1)$ it is sufficient to restrict Y_{AB} in (3.66) to the $|AB|(|A|+|B|-1)$ elements of some fixed basis of $\mathfrak{K}_{\mathrm{DISC}}$. Note also that the condition above is equivalent to the inclusion $\mathbf{J}^{\Theta}_{AB} \in \mathfrak{K}^{\perp}_{\mathrm{DISC}}$, where $\mathfrak{K}^{\perp}_{\mathrm{DISC}}$ is the orthogonal complement of $\mathfrak{K}_{\mathrm{DISC}}$ in $\mathrm{Herm}(AB)$.

3.3.3 Incoherent superchannels (ISC) and strictly incoherent superchannels (SISC)

Any superchannel $\Theta \in \mathfrak{S}(A \to B)$ has a Kraus decomposition i.e. an operator sum representation

$$\Theta_{A\to B} = \sum_{x=1}^{n} \Theta^{x}_{A\to B} \tag{3.68}$$

where the Choi matrix of each $\Theta^x_{A\to B} \in \mathfrak{L}(A \to B)$ has rank one. We use this property to define two other sets of free operations that we call incoherent superchannels (ISC) and strictly incoherent superchannels (SISC).

Definition 3.7. Let A and B be two dynamical systems, and let $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel. Then, Θ is said to be an incoherent superchannel (ISC) if and only if it has a Kraus decomposition $\{\Theta^x_{A\to B}\}_{x=1}^n$ as in (3.68) that satisfies

$$\Delta_B \circ \Theta^x_{A\to B} \circ \Delta_A = \Theta^x_{A\to B} \circ \Delta_A \quad \forall\, x = 1, \ldots, n. \tag{3.69}$$

Moreover, the set of all such superchannels that satisfy the above relation is denoted by $\mathrm{ISC}(A \to B)$.

Definition 3.8. Let A and B be two dynamical systems, and let $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel. Then, Θ is said to be a strictly incoherent superchannel (SISC) if and only if it has a Kraus decomposition $\{\Theta^x_{A\to B}\}_{x=1}^n$ as in (3.68) that satisfies

$$\Delta_B \circ \Theta^x_{A\to B} = \Theta^x_{A\to B} \circ \Delta_A \quad \forall\, x = 1, \ldots, n. \tag{3.70}$$

Moreover, the set of all such superchannels that satisfy the above relation is denoted by $\mathrm{SISC}(A \to B)$.

3.4 Quantification of dynamical coherence

Before we discuss the conversion of one quantum channel into another using free superchannels, it is important to talk quantitatively about the coherence present in quantum channels. This is done by defining functions on quantum channels that map a channel to a real number. A function is called a monotone if it follows the condition of monotonicity which means that the value of the channel should not increase after the channel is acted upon by a free superchannel. To make this function meaningful one more condition is added such that the free channels have the least value. This condition is faithfulness, i.e., the value of such a function should be zero iff the input is a free channel (in our case, the value should be zero for all classical channels). In the problem of interconversion of two resources (using the free superchannels), a resource monotone is specifically useful to tell if one resource can or cannot be converted to another.

Many types of monotones have been defined in literature [82] including general distance-based monotones, entropic monotones, etc. In this section we find the set of monotones to quantify the coherence present in quantum channels. Our work is confined to the case of MISC and DISC. We have divided this section into three sub-sections. The first sub-section gives a complete family of monotones, i.e., to check for the

convertibility of one channel to another, it is sufficient to check if the value obtained from all the monotones of this set is greater for one channel than other. The second sub-section is based on the relative entropies for channels that form a monotone under the free superchannels of our theory. We show that, of the six relative entropies for channels mentioned in [145], only three relative entropies form a monotone in the case of dynamical coherence. The last subsection is dedicated to the monotones that have an operational interpretation in the resource theory of dynamical coherence, i.e., these monotones become meaningful when we find the coherence cost of a channel. One such monotone is known as the log-robustness of coherence of quantum channels. We show that this function can be formulated as an SDP. Using this SDP formulation, we find the upper bound of the log-robustness of coherence of channels in appendix A.3 and show that it is achieved by the quantum Fourier transform channel.

3.4.1 A complete family of monotones

In recent works [89, 119, 143, 144, 208, 210, 211, 223], various resource measures have been formulated for a general resource theory of channels and for the dynamical resource theory of entanglement. A complete set of monotones for both the general resource theory of channels and the resource theory of entanglement of channels was presented in [89]. For the conversion of one quantum channel to another using the free superchannels of the theory, it is sufficient to check if all the monotones of this (complete) set acting on one channel are greater than the other. It was shown that the complete family of monotones for the dynamical resource theory of NPT entanglement can be computed using an SDP (which otherwise for LOCC-based entanglement is known to be NP-hard [224]). In general, for a given quantum resource theory, it is not obvious if these functions are computable. In our work, we find a complete set of monotones under the free superchannels MISC and DISC, and show that for the dynamical resource theory of coherence, these functions can be computed using an SDP.

For a general convex quantum resource theory, we can define the following complete set of non-negative resource measures using any quantum channel $\mathcal{P}_B \in \mathrm{CPTP}(B_0 \to B_1)$ such that these measures take the value zero on free channels [89]

$$G_{\mathcal{P}}(\mathcal{N}_A) \coloneqq \max_{\Theta \in \mathrm{FREE}(A\to B)} \langle \mathcal{P}_B, \Theta\left[\mathcal{N}_A\right]\rangle \; - \max_{\mathcal{M}_B \in \mathfrak{G}(B_0\to B_1)} \langle \mathcal{P}_B, \mathcal{M}_B\rangle \quad \forall\, \mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)\,. \tag{3.71}$$

where $\mathfrak{G}(B_0 \to B_1)$ denotes the set of free channels for the given resource theory. Recall that using the free superchannels, one can transform a free channel to any other free channel. Therefore using some free superchannel $\Theta \in \mathrm{FREE}(A \to B)$, a given free channel $\mathcal{N}_A$ can be converted to the optimal free channel $\mathcal{M}'_B$ that gives the maximum value for the inner product $\langle \mathcal{P}_B, \mathcal{M}'_B\rangle$ and hence, the value of $G_{\mathcal{P}}(\mathcal{N}_A)$ is zero

for all free channels.

To construct the complete set of resource monotones for the dynamical resource theory of coherence, we first define a function $f_{\mathcal{P}}(\mathcal{M}_A)$ using a quantum channel $\mathcal{P}_B \in \text{CPTP}(B_0 \to B_1)$ and superchannel $\Theta \in \mathfrak{F}(A \to B)$ where $\mathfrak{F} = \text{MISC or DISC}$, as

$$f_{\mathcal{P}}(\mathcal{M}_A) = \max_{\Theta \in \mathfrak{F}(A\to B)} \langle \mathcal{P}_B, \Theta[\mathcal{M}_B] \rangle \quad \forall\, \mathcal{M}_A \in \text{CPTP}(A_0 \to A_1)\ . \tag{3.72}$$

Note that (3.72) can be expressed as an SDP for a given channel $\mathcal{M}_A \in \text{CPTP}(A_0 \to A_1)$ in the following manner

$$f_{\mathcal{P}}(\mathcal{M}_A) = \max \left\{ \text{Tr}\left[\mathbf{J}^{\Theta}_{AB} \left((J^{\mathcal{M}}_A)^T \otimes J^{\mathcal{P}}_B \right) \right] \right\} \tag{3.73}$$

where the maximum is subject to

$$\mathbf{J}^{\Theta}_{AB} \geqslant 0\,,\ \mathbf{J}^{\Theta}_{AB_0} = \mathbf{J}^{\Theta}_{A_0B_0} \otimes u_{A_1}\,,\ \mathbf{J}^{\Theta}_{A_1B_0} = I_{A_1B_0}\ , \tag{3.74}$$

$$\text{Tr}[\mathbf{J}^{\Theta}_{AB} X^i_{AB}] = 0 \ \forall\ i = 1, \ldots, n\ . \tag{3.75}$$

The conditions in (3.74) are the conditions on the Choi matrix of the linear map Θ to be a superchannel. The set of matrices $\{X^i_{AB}\}_{i=1}^n$ in (3.75) denote the basis of the subspace $\mathfrak{K}_{\mathfrak{F}}$ as defined in (3.63) and (3.67) for $\mathfrak{F} = \text{MISC}$ and DISC, respectively. Accordingly, for MISC, $n \equiv |AB|(|B| - 1)$ and for DISC, $n \equiv |AB|(|A| + |B| - 1)$.

Similar to (3.71) and using any channel $\mathcal{P}_B \in \text{CPTP}(B_0 \to B_1)$, we can define the complete family of monotones for dynamical coherence in the following way

$$G_{\mathcal{P}}(\mathcal{N}_A) \coloneqq \max \text{Tr}\left[\mathbf{J}^{\Theta}_{AB} \left(\left(J^{\mathcal{N}}_A\right)^T \otimes J^{\mathcal{P}}_B \right) \right] - \max \text{Tr}[J^{\mathcal{M}}_B J^{\mathcal{P}}_B]\ , \tag{3.76}$$

where we have expressed the inner product between channels as the inner product of the Choi matrices of the respective channels. The first maximum in (3.76) is subject to the constraints given in (3.74) and (3.75) and the second maximum is over all classical channels $\mathcal{M}_B \in \mathfrak{C}(B_0 \to B_1)$. The family $\{G_{\mathcal{P}}\}$ over all $\mathcal{P} \in \text{CPTP}(B_0 \to B_1)$ forms a complete set of monotones, that is, there exists a $\Theta \in \mathfrak{F}(A \to B)$ where $\mathfrak{F} = \text{MISC or DISC}$, that can convert a channel $\mathcal{E}_A \in \text{CPTP}(A_0 \to A_1)$ to another channel $\mathcal{F}_B \in \text{CPTP}(B_0 \to B_1)$ if and only if

$$G_{\mathcal{P}}(\mathcal{E}_A) \geqslant G_{\mathcal{P}}(\mathcal{F}_B) \tag{3.77}$$

Remark 2. For the qubit case we calculated the values of the monotone $G_{\mathcal{P}}(\mathcal{N}_A)$ under MISC for a few

channels(or a class of channels) by plugging into CVX. This required construction of 48 basis elements (Eq. (3.63)). The value of $G_{\mathcal{P}}(\mathcal{N}_A)$ for all classical channels is 0 for all $\mathcal{P}_B$. We found that for a fixed $\mathcal{P}_B$, the value of all unitary channels is the same and they attain the maximum value of 2 when $\mathcal{P}_B$ is the identity channel. If we fix $\mathcal{P}_B$ to be the identity channel, we see that for a replacement channel that outputs a plus state ($|+\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle$), the value of $G_{\mathsf{id}}(\mathcal{N}_A)$ is equal to 2. For any other replacement channel and any depolarizing channel, $G_{\mathsf{id}}(\mathcal{N}_A)$ is less than 2.

Remark 3. Since there are an infinite number of monotones in the above complete set $G_{\mathcal{P}}$, it might give an impression that the conversion of a channel $\mathcal{N}_A \in \text{CPTP}(A_0 \to A_1)$ to another channel $\mathcal{M}_B \in \text{CPTP}(B_0 \to B_1)$ using a superchannel $\Theta \in$ MISC or DISC, is very hard or impractical, but in section 3.5 we show that the problem of interconversion of two quantum channels using a superchannel belonging to MISC or DISC can be computed using an SDP.

3.4.2 Relative entropies of dynamical coherence

A measure of distinguishability or divergence $D(\cdot\|\cdot)$ of two states is a function $D : \mathfrak{D}(A_1) \times \mathfrak{D}(B_1) \to \mathbb{R}$ such that it obeys data-processing inequality and is zero on the set of free states. The data-processing inequality states that the distinguishability between two quantum states cannot be increased when both the states are acted upon by a quantum operation. One example of such a function is Rényi divergence [132]. Its two generalizations which have been given an operational interpretation are "Sandwiched" Rényi relative entropy (also known as quantum Rényi divergence) and Petz-Rényi relative entropy. "Sandwiched" Rényi relative entropy (or quantum Rényi divergence) was introduced and discussed in [135, 136, 137, 138]. It is based on a parameter $\alpha \in [0, \infty]$ and reduces to the relative von Neumann entropy for $\alpha = 1$, to the relative max entropy for $\alpha = \infty$, and closely related to the fidelity $\text{Tr}\left(\sigma^{1/2}\rho\sigma^{1/2}\right)^{1/2}$ for $\alpha = 1/2$, where ρ and σ are two density matrices which are input to the entropy function. Also, the "Sandwiched" Rényi relative entropy is a monotone under quantum operations, i.e., satisfies the data-processing inequality for $\alpha \in [1/2, \infty]$ [138]. The Petz-Rényi relative entropy was introduced and studied in [139, 140, 141] and it finds an operational interpretation in the context of quantum hypothesis testing. Therefore, these two are relevant information quantities to consider for quantum information theory. Other generalizations of the Rényi divergence and the quantum Rényi relative entropies are discussed in [142] but their operational meaning is not clear.

For the channel case (i.e., dynamical resources), the relative entropies and divergence have been generalized from the state case (i.e., static resources) and were discussed in [119, 143, 144, 145, 146, 147, 148]. The

channel divergence for two given channels $\mathcal{N}_A, \mathcal{M}_A \in \text{CPTP}(A_0 \to A_1)$ is defined as [119, 146, 147]

$$D(\mathcal{N}_A \| \mathcal{M}_A) = \max_{\phi \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right) \middle\| \mathcal{M}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right)\right) \tag{3.78}$$

where $D(\rho\|\sigma) = \text{Tr}[\rho \log \rho - \rho \log \sigma]$ is the relative entropy. We take the relative entropies listed in [145] and find the following three relative entropies to be clearly forming a monotone under MISC

$$C_1\left(\mathcal{N}_A\right) = \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} \max_{\phi \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right) \middle\| \mathcal{M}_{A_0 \to A_1}\left(\phi_{R_0 A_0}\right)\right) \tag{3.79}$$

$$C_2\left(\mathcal{N}_A\right) = \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} \sup_{\rho, \sigma \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_A\left(\rho_{R_0 A_0}\right) \middle\| \mathcal{M}_A\left(\sigma_{R_0 A_0}\right)\right) - D\left(\rho_{R_0 A_0} \middle\| \sigma_{R_0 A_0}\right) \tag{3.80}$$

$$C_3\left(\mathcal{N}_A\right) = \max_{\rho \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_A\left(\rho_{R_0 A_0}\right)\right) - D\left(\rho_{R_0 A_0}\right) \tag{3.81}$$

where $D(\rho) = \min_{\mathcal{D}(\sigma)=\sigma} D\left(\rho\|\sigma\right)$. The proof that the above relative entropies form a monotone under MISC is similar to the proof for relative entropies forming a monotone for a general resource theory of quantum processes as given in [145] and are quite straightforward. For completeness, we give the proof of monotonicity for one of the above monotones C_1 and the proof for the other two follows likewise. Let $\Theta_{A \to B} \in \text{MISC}(A \to B)$. Then the relative entropy of the channel $\Theta[\mathcal{N}_A] \in \text{CPTP}(B_0 \to B_1)$ for some channel $\mathcal{N}_A \in \text{CPTP}(A \to B)$ can be written as

$$\begin{aligned}
C_1\left(\Theta\left[\mathcal{N}_A\right]\right) &= \min_{\mathcal{M}' \in \mathfrak{C}(B_0 \to B_1)} \max_{\phi \in \mathfrak{D}(R_0 B_0)} D\left(\Theta_{A \to B}\left[\mathcal{N}_A\right]\left(\phi_{R_0 B_0}\right) \middle\| \mathcal{M}'_{B_0 \to B_1}\left(\phi_{R_0 B_0}\right)\right) \\
&\leqslant \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} \max_{\phi \in \mathfrak{D}(R_0 B_0)} D\left(\Theta_{A \to B}\left[\mathcal{N}_A\right]\left(\phi_{R_0 B_0}\right) \middle\| \Theta_{A \to B}\left[\mathcal{M}_A\right]\left(\phi_{R_0 B_0}\right)\right) \\
&\leqslant \min_{\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)} \max_{\rho \in \mathfrak{D}(R_0 A_0)} D\left(\mathcal{N}_A\left(\rho_{R_0 A_0}\right) \middle\| \mathcal{M}_A\left(\rho_{R_0 A_0}\right)\right) \\
&= C_1\left(\mathcal{N}_A\right)
\end{aligned} \tag{3.82}$$

where the first inequality follows because the relative entropy of the channel $\Theta\left[\mathcal{N}_A\right]$ is the minimum taken over all classical channels in system B and the second inequality follows because of the data-processing inequality. Note that the relative entropies $C_1(\mathcal{N}_A)$ and $C_2(\mathcal{N}_A)$ are faithful, i.e., they take the value zero iff $\mathcal{N}_A \in \mathfrak{C}(A_0 \to A_1)$. The relative entropy $C_3(\mathcal{N}_A)$ is a state-based relative entropy and involves no optimization over the classical channels.

In [145], there are three other relative entropies that are defined similar to C_1, C_2, and C_3 where the optimization is taken over the set of free states instead of all density matrices. There, the proof relies on the pre-processing channel to be completely resource non-generating. Since, we cannot make this assumption (because in the dynamical resource theory of quantum coherence, such a channel would completely destroy

any resource), hence, we cannot say anything about the monotonicity of the relative entropies where the optimization is over the incoherent states.

To define the resource monotones for DISC, we first define the function $D_\Delta : \text{CPTP} \to \mathbb{R}_+$ as follows for any channel divergence D

$$D_\Delta(\mathcal{N}_A) \coloneqq D\left(\mathcal{N}_A \middle\| \Delta_A\left[\mathcal{N}_A\right]\right) \ , \tag{3.83}$$

and for the choice $D = D_{\max}$ we call it the dephasing log-robustness and denote it by $D_\Delta \equiv LR_\Delta$.

Lemma 3.9. *The function D_Δ is a dynamical resource monotones under* DISC.

Proof. Lets $\Theta \in \text{DISC}(A \to B)$ and $\mathcal{N} \in \text{CPTP}(A_0 \to A_1)$. Then,

$$\begin{aligned} D_\Delta(\Theta_{A\to B}[\mathcal{N}_A]) &= D\left(\Theta_{A\to B}[\mathcal{N}_A] \middle\| \Delta_B \circ \Theta_{A\to B}[\mathcal{N}_A]\right) \\ &= D\left(\Theta_{A\to B}[\mathcal{N}_A] \middle\| \Theta_{A\to B} \circ \Delta_A\left[\mathcal{N}_A\right]\right) \\ &\leqslant D\left(\mathcal{N}_A \middle\| \Delta_A\left[\mathcal{N}_A\right]\right) \\ &= D_\Delta(\mathcal{N}_A) \ . \end{aligned} \tag{3.84}$$

This completes the proof. ∎

For the case that $D(\rho\|\sigma) = \text{Tr}[\rho \log \rho] - \text{Tr}[\rho \log \sigma]$ is the relative entropy, we call D_Δ the dephasing relative entropy of coherence.

3.4.3 Operational Monotones

This subsection is dedicated to the discussion of the operationally relevant monotones for the resource theory of dynamical coherence. The monotones discussed here are operationally meaningful in a sense that these monotones are useful for calculating various costs of coherent channels. Hence, this subsection can be treated as a preliminary to the next section where we talk about the interconversion of two coherent channels. We will see that the monotones which are based on $D_{\max}$, like various types of log-robustness, play a major role in the calculation of coherence cost of channels.

The log-robustness of entanglement for states was introduced and investigated in [225, 226, 227, 228]. It was shown that it is an entanglement monotone and its operational significance for the manipulation of entanglement was also discussed. The log-robustness of coherence for states was similarly defined in [229] and it was shown that it is a measure of coherence. For a state $\rho_{A_1} \in \mathfrak{D}(A_1)$, it is defined as follows

$$LR_{\mathcal{I}}(\rho_{A_1}) \coloneqq \min_{\sigma_{A_1} \in \mathcal{I}_{A_1}} D_{\max}\left(\rho_{A_1} \| \sigma_{A_1}\right) \tag{3.85}$$

where $\mathcal{I}_{A_1}$ is the set of incoherent states for system A_1. The log-robustness of channels for a general resource theory was introduced and discussed in [143, 144, 145]. It was shown that the log-robustness of channels satisfy necessary conditions for the resource measure of channels, i.e., it is both faithful (means, it gives the value zero for free channels, and is greater than zero otherwise) and a monotone under tensoring, and left and right compositions with free channels [144]. Likewise, we can define the log-robustness of coherence of channels, which is a monotone under MISC, in the following way

$$LR_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{E}\in\mathfrak{C}(A_0\to A_1)} D_{\max}\big(\mathcal{N}_A\|\mathcal{E}_A\big) \ , \tag{3.86}$$

where the minimum is taken over all the classical channels. The proof that it is a monotone is very similar to the proof presented for the relative entropy C_1 (see (3.82)) and follows easily using the data-processing inequality. Besides, it can be computed with an SDP. To see why, note that

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\min\Big\{t\geqslant 0 \ : \ t\mathcal{E}_A\geqslant\mathcal{N}_A \ , \ \Delta_A[\mathcal{E}_A]=\mathcal{E}_A \ , \ \mathcal{E}\in\mathrm{CPTP}(A_0\to A_1)\Big\} \tag{3.87}$$

where the first condition arises from the definition of $D_{\max}$ for channels, the second and third from the requirement of $\mathcal{E}$ to be a classical channel. Denoting by ω_A the Choi matrix of $t\mathcal{E}_A$ we get that (recall that we are using u to denote the maximally mixed state)

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\min\Big\{\frac{1}{|A_0|}\mathrm{Tr}[\omega_A] \ : \ \omega_A\geqslant J_A^{\mathcal{N}} \ , \ \mathcal{D}_A[\omega_A]=\omega_A \ , \ \omega_{A_0}=\mathrm{Tr}[\omega_A]u_{A_0} \ , \ \omega_A\geqslant 0\Big\} \tag{3.88}$$

which is an SDP optimization problem. As such it has a dual given by (see appendix for details)

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\max\Big\{\mathrm{Tr}[\eta_A J_A^{\mathcal{N}}] \ : \ \mathcal{D}_A(\eta_A)=\mathcal{D}_{A_0}\left(\eta_{A_0}\right)\otimes u_{A_1} \ , \ \mathcal{D}_{A_1}[\eta_{A_1}]=I_{A_1} \ , \ \eta_A\geqslant 0\Big\} \tag{3.89}$$

Remark 4. For the qubit case, we calculated the log-robustness of coherence of few channels. For any classical channel, the log-robustness of coherence is equal to 0. For the identity channel it is equal to 1. For any replacement channel and depolarizing channel, its value is between 0 and 1. If the replacement channel is the one that outputs the plus state ($|+\rangle = \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}|i\rangle$), the log-robustness is equal to 1. For any unitary channel, we found that the value of log-robustness of coherence is between 1 and 2. We found that the value obtained for the Hadamard gate is the maximum and is equal to 2. In these examples, we can see that a quantum channel can have the ability to preserve and/or generate coherence. For instance, replacement channels cannot preserve coherence but can only generate coherence whereas the identity channel can only preserve coherence but not generate coherence. The unitary channels can do both, i.e., they can preserve

coherence as well as they can generate coherence. This fact explains the higher value of the log-robustness of coherence of the unitary channels as compared to the replacement channels. Also, as calculated in appendix A.3, the quantum Fourier transform channel has the maximum value for this monotone. This can be explained from the fact that a QFT channel can generate the maximally coherent state from a free state and has the ability to preserve entanglement. (Note that the replacement channels are entanglement-breaking channels.)

Next, we show the additivity of log-robustness of coherence of channels under tensor products. This result is useful when we go to the asymptotic limit.

Lemma 3.10. *The log-robustness of coherence of a channel is additive under tensor products, i.e.,*

$$LR_{\mathfrak{C}}(\mathcal{N}_A \otimes \mathcal{M}_{A'}) = LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'}) \tag{3.90}$$

Proof. For the proof of the inequality $LR_{\mathfrak{C}}(\mathcal{N}_A \otimes \mathcal{M}_{A'}) \leqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'})$, let $LR_{\mathfrak{C}}(\mathcal{N}_A) = D_{\max}(\mathcal{N}_A||\mathcal{E}_A)$ and $LR_{\mathfrak{C}}(\mathcal{M}_{A'}) = D_{\max}(\mathcal{M}_{A'}||\mathcal{E}_{A'})$ for some optimal $\mathcal{E}_A$ and $\mathcal{E}_{A'}$. Then, we have

$$\begin{aligned} LR_{\mathfrak{C}}(\mathcal{N}_A \otimes \mathcal{M}_{A'}) &\leqslant D_{\max}(\mathcal{N}_A \otimes \mathcal{M}_{A'}||\mathcal{E}_A \otimes \mathcal{E}_{A'}) && (3.91)\\ &= D_{\max}(\mathcal{N}_A||\mathcal{E}_A) + D_{\max}(\mathcal{M}_{A'}||\mathcal{E}_{A'}) && (3.92)\\ &= LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'}) \ . && (3.93) \end{aligned}$$

The first inequality follows trivially from the definition of log-robustness and the second equality follows from the additivity of $D_{\max}$.

To prove the converse, i.e., $LR_{\mathfrak{C}}(\mathcal{N}_A \otimes \mathcal{M}_{A'}) \geqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'})$, we will use the dual of the log-robustness as given in Eq.(3.89). Let η_A and $\eta_{A'}$ be the optimal matrices for the dual of $LR_{\mathfrak{C}}(\mathcal{N}_A)$ and $LR_{\mathfrak{C}}(\mathcal{M}_{A'})$, respectively. Then we have

$$\begin{aligned} 2^{LR_{\mathfrak{C}}(\mathcal{N}_A)} &= \mathrm{Tr}[\eta_A J_A^{\mathcal{N}_A}] \ , \\ 2^{LR_{\mathfrak{C}}(\mathcal{M}_{A'})} &= \mathrm{Tr}[\eta_{A'} J_{A'}^{\mathcal{M}_{A'}}] \ . \end{aligned} \tag{3.94}$$

Since, $LR_{\mathfrak{C}}(\mathcal{N}_A \otimes \mathcal{M}_{A'}) = \log \max \mathrm{Tr}\left[\eta'_{AA'}\left(J_{AA'}^{\mathcal{N}_A \otimes \mathcal{M}_{A'}}\right)\right]$ where the maximum is over all $\eta'_{AA'} \geqslant 0$ satisfying

$$\mathcal{D}_{AA'}(\eta'_{AA'}) = \mathcal{D}_{A_0A'_0}\left(\eta'_{A_0A'_0}\right) \otimes u_{A_1A'_1} \ \ , \ \ \mathcal{D}_{A_1A'_1}[\eta'_{A_1A'_1}] = I_{A_1A'_1} \ . \tag{3.95}$$

For the particular choice of $\eta_{AA'} = \eta_A \otimes \eta_{A'}$, it is easy to verify that it satisfies the above conditions and

therefore for this choice of $\eta_{AA'}$ we have

$$\begin{aligned} 2^{LR_{\mathfrak{C}}(\mathcal{N}_A\otimes\mathcal{M}_{A'})} &\geqslant \mathrm{Tr}\big[\eta_{AA'}\big(J_{AA'}^{\mathcal{N}_A\otimes\mathcal{M}_{A'}}\big)\big] \\ &= 2^{LR_{\mathfrak{C}}(\mathcal{N}_A)}2^{LR_{\mathfrak{C}}(\mathcal{M}_{A'})} \,. \end{aligned} \tag{3.96}$$

In the above equation, the first inequality follows from the dual of the log-robustness of channels and the second equality is because η_A and $\eta_{A'}$ are optimal for $LR_{\mathfrak{C}}(\mathcal{N}_A)$ and $LR_{\mathfrak{C}}(\mathcal{M}_{A'})$. Hence, the above equation implies

$$LR_{\mathfrak{C}}(\mathcal{N}_A\otimes\mathcal{M}_{A'}) \geqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'}) \tag{3.97}$$

This establishes the additivity of the log-robustness of coherence of a quantum channel, i.e., $LR_{\mathfrak{C}}(\mathcal{N}_A\otimes\mathcal{M}_{A'}) = LR_{\mathfrak{C}}(\mathcal{N}_A) + LR_{\mathfrak{C}}(\mathcal{M}_{A'})$ ∎

Another type of log-robustness, the dephasing log-robustness, which will be used to find the exact cost under DISC, is defined by

$$LR_{\Delta}(\mathcal{N}_A) \coloneqq D_{\max}\big(\mathcal{N}_A\big\|\Delta_A[\mathcal{N}_A]\big) \quad \forall\, \mathcal{N}\in \mathrm{CPTP}(A_0\to A_1)\,. \tag{3.98}$$

The dephasing log-robustness is a monotone under DISC which is easy to show using (3.64) and data-processing inequality. While the log-robustness behaves monotonically under any superchannel in MISC, the dephasing log-robustness is in general not monotonic under MISC. Instead, it is monotonic under DISC.

Lemma 3.11. *For any $\mathcal{N}\in\mathrm{CPTP}(A_0\to A_1)$ and $\Theta\in\mathrm{DISC}(A\to B)$ we have*

$$LR_{\Delta}\big(\Theta_{A\to B}[\mathcal{N}_A]\big) \leqslant LR_{\Delta}(\mathcal{N}_A)\,. \tag{3.99}$$

Proof.

$$\begin{aligned} LR_{\Delta}\big(\Theta_{A\to B}[\mathcal{N}_A]\big) &= D_{\max}\big(\Theta_{A\to B}[\mathcal{N}_A]\big\|\Delta_A\circ\Theta_{A\to B}[\mathcal{N}_A]\big) \\ &= D_{\max}\big(\Theta_{A\to B}[\mathcal{N}_A]\big\|\Theta_{A\to B}\circ\Delta_A[\mathcal{N}_A]\big) \\ &\leqslant D_{\max}\big(\mathcal{N}_A\big\|\Delta_A[\mathcal{N}_A]\big) \\ &= LR_{\Delta}(\mathcal{N}_A)\,, \end{aligned} \tag{3.100}$$

where the second equality follows from the commutativity of Θ and Δ, and the inequality follows from the data processing inequality of the channel divergence $D_{\max}$ [119]. ∎

We prove here that the dephasing log-robustness is also additive.

Lemma 3.12. *Let $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ and $\mathcal{M} \in \mathrm{CPTP}(B_0 \to B_1)$ be two channels. Then,*

$$LR_\Delta\big(\mathcal{N}_A \otimes \mathcal{M}_B\big) = LR_\Delta(\mathcal{N}_A) + LR_\Delta(\mathcal{M}_B)\ . \tag{3.101}$$

Proof.

$$\begin{aligned} LR_\Delta\big(\mathcal{N}_A \otimes \mathcal{M}_B\big) &= D_{\max}\big(\mathcal{N}_A \otimes \mathcal{M}_B \big\| \Delta_{AB}\big[\mathcal{N}_A \otimes \mathcal{M}_B\big]\big) \\ &= D_{\max}\big(\mathcal{N}_A \otimes \mathcal{M}_B \big\| \Delta_A\big[\mathcal{N}_A\big] \otimes \Delta_B\big[\mathcal{M}_B\big]\big) \\ &= D_{\max}\big(\mathcal{N}_A \big\| \Delta_A\big[\mathcal{N}_A\big]\big) + D_{\max}\big(\mathcal{M}_B \big\| \Delta_B\big[\mathcal{M}_B\big]\big) \\ &= LR_\Delta(\mathcal{N}_A) + LR_\Delta(\mathcal{M}_B)\ , \end{aligned} \tag{3.102}$$

where the third equality follows from the additivity of $D_{\max}$ for channels. ∎

We also define smoothed log-robustness and asymptotic log-robustness which would be useful in finding the approximate and liberal coherence costs of a channel whose meanings are discussed in detail in the next section. From [145], we know that smoothing maintains monotonicity. Thus, the smoothed log-robustness which is defined below is also a monotone

$$LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{N}' \in B_\epsilon(\mathcal{N}_A)} LR_{\mathfrak{C}}(\mathcal{N}'_A) \tag{3.103}$$

where the minima is taken over the log-robustness of all channels that lie within the ϵ-ball given by the diamond norm. This ϵ-ball, $B_\epsilon(\mathcal{N}_A)$, around the channel $\mathcal{N}_A$ is defined as

$$B_\epsilon(\mathcal{N}_A) = \Big\{\mathcal{N}' \in \mathrm{CPTP}(A_0 \to A_1)\ :\ \frac{1}{2}\|\mathcal{N}'_A - \mathcal{N}_A\|_\diamond \leqslant \epsilon\Big\}\ . \tag{3.104}$$

To obtain the aysmptotic log-robustness, we first regularize the smoothed log-robustness and then take the limit $\epsilon \to 0^+$. Thus, the asymptotic log-robustness is defined as

$$LR^\infty_{\mathfrak{C}}(\mathcal{N}_A) = \lim_{\epsilon \to 0^+} \liminf_{n \to \infty} \frac{1}{n} LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A^{\otimes n}) \tag{3.105}$$

Similarly we define the smoothed dephasing log-robustness and asymptotic dephasing log-robustness, both of which are monotones under DISC. The smoothed dephasing log-robustness is defined by

$$LR^\epsilon_{\Delta}(\mathcal{N}_A) \coloneqq \min_{\mathcal{N}' \in B_\epsilon(\mathcal{N}_A)} LR_{\Delta}(\mathcal{N}'_A) \tag{3.106}$$

and the asymptotic dephasing log-robustness as

$$LR^{\infty}_{\Delta}(\mathcal{N}_A) = \lim_{\epsilon\to 0^+}\lim_{n\to\infty}\frac{1}{n}LR^{\epsilon}_{\Delta}(\mathcal{N}_A^{\otimes n}) \tag{3.107}$$

Now we define the log-robustness with "liberal" smoothing [145] which we find to have an operational meaning. Let

$$LR^{\epsilon,\varphi}_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{N}'\in B^{\varphi}_{\epsilon}(\mathcal{N}_A)} LR_{\mathfrak{C}}\left(\mathcal{N}'_A\right) . \tag{3.108}$$

where

$$B^{\varphi}_{\epsilon}(\mathcal{N}_A) \coloneqq \left\{\mathcal{N}' \in \mathrm{CP}(A_0\to A_1) \ : \ \|\mathcal{N}'_A(\varphi_{RA_0}) - \mathcal{N}_A(\varphi_{RA_0})\|_1 \leqslant \epsilon\right\} , \tag{3.109}$$

and consider its "liberal" smoothing

$$LR^{\epsilon}_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \max_{\varphi\in\mathcal{D}(RA_0)} LR^{\epsilon,\varphi}_{\mathfrak{C}}(\mathcal{N}_A). \tag{3.110}$$

Define also

$$LR^{\epsilon,n}_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \frac{1}{n}\max_{\varphi\in\mathcal{D}(RA_0)} LR^{\epsilon,\varphi^{\otimes n}}_{\mathfrak{C}}(\mathcal{N}_A^{\otimes n}) , \tag{3.111}$$

and

$$LR^{(\infty)}_{\mathfrak{C}}(\mathcal{N}_A) \coloneqq \lim_{\epsilon\to 0^+}\liminf_{n\to\infty} LR^{\epsilon,n}_{\mathfrak{C}}(\mathcal{N}_A) . \tag{3.112}$$

In [145], a new type of regularized relative entropy of a resource given by

$$D^{(\infty)}_{\mathfrak{C}}\left(\mathcal{N}_A\right) \coloneqq \lim_{n\to\infty}\frac{1}{n}\sup_{\varphi\in\mathfrak{D}(RA_0)}\min_{\mathcal{E}\in\mathfrak{C}(A_0^n\to A_1^n)} D\left(\mathcal{N}^{\otimes n}_{A_0\to A_1}\left(\varphi^{\otimes n}_{RA_0}\right)\middle\|\mathcal{E}_{A_0^n\to A_1^n}\left(\varphi^{\otimes n}_{RA_0}\right)\right) \tag{3.113}$$

The quantity $D^{(\infty)}_{\mathfrak{C}}\left(\mathcal{N}_A\right)$ behaves monotonically under completely RNG superchannels and satisfies the following Asymptotic Equipartition Property (AEP)

$$LR^{(\infty)}_{\mathfrak{C}}(\mathcal{N}_A) = D^{(\infty)}_{\mathfrak{C}}\left(\mathcal{N}_A\right) . \tag{3.114}$$

3.5 Interconversions

As the name suggests, in this section we discuss the conversion of one channel into another. Conversion of one resource to another using the set of free operations is one of the major tasks of resource theory [82]. This conversion can be of various types like single shot (exact or approximate) interconversion, asymptotic interconversion, catalytic interconversion, etc. Addressing this leads to two very interesting questions. The

first is the problem of finding the cost, i.e., minimum amount of maximal resource necessary to output a given resource using free operations, and the second is distillation which is the inverse problem of cost, i.e., asking how much maximal resource can be extracted from a given resource using the free operations. We answer these questions in this section for the resource theory of dynamical coherence. Note that since we just require two copies of the maximally coherent replacement channel to simulate the maximally coherent channel (refer to appendix A.3), we use the maximally coherent replacement channel to compute the cost and distillation.

This section is broadly divided into three parts namely, the conversion distance of coherence, cost of a channel, and the problem of distilling an arbitrary channel into pure-state coherence. The first part is discussed in subsection 3.5.1 where we take up the task of conversion of one channel into another using MISC or DISC. For this task, we form a function called the conversion distance, and claim that if the conversion distance is very small, then, we can simulate one channel using the other with the help of free superchannels. We show that for the dynamical resource theory of coherence, the conversion distance $d_{\mathfrak{F}}(\mathcal{N}_A \to \mathcal{M}_B)$ for two given channels $\mathcal{N}_A$ and $\mathcal{M}_B$ can be computed with an SDP. We then take up the problem of finding various costs of coherence which are discussed in subsection 3.5.2, 3.5.3, and 3.5.4. In these three subsections, we calculate the exact, approximate and "liberal" coherence cost of a channel and show that the "liberal" cost of coherence is equal to a variant of the regularized relative entropy. In the last subsection 3.5.5, we calculate one-shot distillable rates and end the subsection by providing examples of the distillable rates for partial depolarizing channel and partial dephasing channel.

3.5.1 The conversion distance of coherence

The conversion distance from a channel $\mathcal{N}_A \in \text{CPTP}(A_0 \to A_1)$ to a channel $\mathcal{M}_B \in \text{CPTP}(B_0 \to B_1)$ is defined as (with $\mathfrak{F}$ standing for either one of the four operations MISC, DISC, ISC, and SISC)

$$d_{\mathfrak{F}}\left(\mathcal{N}_A \to \mathcal{M}_B\right) \coloneqq \min_{\Theta \in \mathfrak{F}(A \to B)} \frac{1}{2} \left\| \Theta_{A \to B}\left[\mathcal{N}_A\right] - \mathcal{M}_B \right\|_{\diamond} . \tag{3.115}$$

Recall that the diamond norm is used to measure the distance of two quantum channels and is defined in (2.85). Therefore, if the conversion distance from a channel $\mathcal{N}_A$ to another channel $\mathcal{M}_B$ is very small, it means that $\Theta_{A \to B}\left[\mathcal{N}_A\right]$ is very close to $\mathcal{M}_B$, which implies that $\mathcal{N}_A$ can be used to simulate $\mathcal{M}_B$ using free superchannels. We now show that for $\mathfrak{F} = \text{MISC}$ or $\mathfrak{F} = \text{DISC}$, this conversion distance can be computed with a semi-definite program (SDP)[2].

[2]The code to calculate the interconversion distance from one channel to another is given in https://github.com/gaurav-iiser/Resource_Theory_of_Dynamical_Coherence.

Theorem 3.13. *Let* $\{X^i_{AB}\}_{i=1}^n$ *be the basis of the subspace* $\mathfrak{K}_{\mathfrak{F}}$ *as defined in* (3.63) *where* $n \equiv |AB|(|B|-1)$ *for the case* $\mathfrak{F} = \text{MISC}$ *and* $n \equiv |AB|(|A|+|B|-1)$ *for the case* $\mathfrak{F} = \text{DISC}$. *Let* α_{AB} *denote the Choi matrix of the superchannel* Θ. *Then,* $d_{\mathfrak{F}}\left(\mathcal{N}_A \to \mathcal{M}_B\right)$, *can be expressed as the following SDP*

$$d_{\mathfrak{F}}\left(\mathcal{N}_A \to \mathcal{M}_B\right) = \min \lambda \tag{3.116}$$

where the minimum is subject to

$$\lambda I_{B_0} \geqslant \omega_{B_0}\ ,\ \omega_B \geqslant 0\ ,\ \alpha_{AB} \geqslant 0\ ,\ \omega_B \geqslant \text{Tr}_A\Big[\alpha_{AB}\left((J^{\mathcal{N}}_A)^T \otimes I_B\right)\Big] - J^{\mathcal{M}}_B\ , \tag{3.117}$$

$$\alpha_{AB_0} = \alpha_{A_0B_0} \otimes u_{A_1}\ ,\ \alpha_{A_1B_0} = I_{A_1B_0}\ , \tag{3.118}$$

$$\text{Tr}[\alpha_{AB}X^i_{AB}] = 0\ \forall\ i = 1\,,\ \ldots\ ,n \tag{3.119}$$

Proof. The proof of the above theorem depends on the fact that the diamond norm can be expressed as an SDP [129] and is presented in detail in appendix section A.2. ∎

Remark 5. For MISC, the interconversion distance from the Hadamard channel to the maximal replacement channel (i.e., the channel that outputs a maximally coherent state $\phi^+_{B_1}$ for any input) is 0, i.e., it is possible to construct a protocol to convert the Hadamard channel to maximal replacement channel using the maximally incoherent superchannel. We also found that the interconversion distance from two maximal replacement channels to a Hadamard channel is also 0 which was expected from the results of appendix A.3.

3.5.2 Exact asymptotic coherence cost

The exact single-shot coherence cost of a channel $\mathcal{N}_A \in \text{CPTP}(A_0 \to A_1)$ is defined as the log of the minimum dimension of the maximally coherent state which can be used to convert to the given channel using a free superchannel and can be expressed as

$$C^0_{\mathfrak{F}}(\mathcal{N}_A) \coloneqq \min\left\{\log|R_1|\ :\ \exists\Theta \in \mathfrak{F}(R_1 \to A)\quad \text{s.t.}\quad \Theta_{R_1\to A}[\phi^+_{R_1}] = \mathcal{N}_A\right\}\ , \tag{3.120}$$

where the zero on the superscript of C on lhs means that the conversion is exact. In our work, we consider the two cases of $\mathfrak{F} = \text{MISC}$ and $\mathfrak{F} = \text{DISC}$. The exact asymptotic coherence cost is defined by regularizing the exact single-shot cost which is given as

$$C^{\text{exact}}_{\mathfrak{F}}(\mathcal{N}_A) = \lim_{n\to\infty}\frac{1}{n}C^0_{\mathfrak{F}}\left(\mathcal{N}^{\otimes n}_A\right) \tag{3.121}$$

We now compute this exact coherence cost for both MISC and DISC.

Exact cost under MISC

Theorem 3.14. *The exact coherence cost of a channel* $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$ *under* $\mathfrak{F} = \mathrm{MISC}$ *is given by*

$$C_{\mathfrak{F}}^{\mathrm{exact}}(\mathcal{N}_A) = LR_{\mathfrak{C}}(\mathcal{N}_A) \tag{3.122}$$

Proof. To prove the above theorem we first prove that

$$LR_{\mathfrak{C}}(\mathcal{N}_A) \leqslant C_{\mathfrak{F}}^{0}(\mathcal{N}_A) \leqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + 1 \tag{3.123}$$

and then regularize the above condition and use the additivity of $LR_{\mathfrak{C}}(\mathcal{N}_A)$ to find the exact asymptotic cost of a channel.

For the proof of $LR_{\mathfrak{C}}(\mathcal{N}) \leqslant C_{\mathfrak{F}}^{0}(\mathcal{N})$, let $\Theta \in \mathrm{MISC}(R_1 \to A)$ be an optimal superchannel satisfying $\Theta_{R_1\to A}[\phi_{R_1}^{+}] = \mathcal{N}_A$ such that $C_{\mathrm{MISC}}^{0}(\mathcal{N}_A) = \log_2 |R_1|$. Therefore,

$$\begin{aligned}
LR_{\mathfrak{C}}(\mathcal{N}_A) &= D_{\max}\left(\mathcal{N}_A \middle\| \mathcal{E}_A\right) && (3.124)\\
&= D_{\max}\left(\Theta_{R_1\to A}\left[\phi_{R_1}^{+}\right] \middle\| \mathcal{E}_A\right) && (3.125)\\
&\leqslant D_{\max}\left(\Theta_{R_1\to A}\left[\phi_{R_1}^{+}\right] \middle\| \Theta_{R_1\to A}\left[\mathcal{D}\left(\phi_{R_1}^{+}\right)\right]\right) && (3.126)\\
&\leqslant D_{\max}\left(\phi_{R_1}^{+} \middle\| \mathcal{D}\left(\phi_{R_1}^{+}\right)\right) && (3.127)\\
&= \log_2 |R_1| && (3.128)\\
&= C_{\mathfrak{F}}^{0}\left(\mathcal{N}_A\right) && (3.129)
\end{aligned}$$

where the first equality follows from the definition of log-robustness of a channel (see (3.86)) for some optimal classical channel $\mathcal{E}_A$. The first inequality above arises because we have chosen the optimal $\mathcal{E}_A$ such that $D_{\max}\left(\mathcal{N}_A \middle\| \mathcal{E}_A\right)$ is minimum. The second inequality, i.e., Eq. (3.127) follows from the data-processing inequality. The equality in Eq. (3.128) can be be easily computed following Eq. (3.21).

To prove $C_{\mathfrak{F}}^{0}(\mathcal{N}_A) \leqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + 1$, first let

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = D_{\max}(\mathcal{N}_A \parallel \mathcal{E}_A) = \log_2 t \tag{3.130}$$

for some optimal t satisfying $t\mathcal{E}_A \geqslant \mathcal{N}_A$. Also, let $m = \lceil t \rceil$, so that $m\mathcal{E}_A \geqslant \mathcal{N}_A$ still holds. Let R_1 be a static

system such that $|R_1| = m$. We now define the following supermap. For any state $\rho_{R_1} \in \mathfrak{D}(R_1)$

$$\Omega_{R_1\to A}\left[\rho_{R_1}\right] \coloneqq \frac{m}{m-1}\Big(\mathrm{Tr}\left[\phi^+_{R_1}\rho_{R_1}\right] - \frac{1}{m}\Big)\mathcal{N}_A \;+\; \frac{m}{m-1}\Big(1 - \mathrm{Tr}\left[\phi^+_{R_1}\rho_{R_1}\right]\Big)\mathcal{E}_A \tag{3.131}$$

Note that the supermap $\Omega_{R_1\to A} \in \mathfrak{F}(R_1 \to A)$ as it can be expressed as

$$\Omega_{R_1\to A}\left[\rho_{R_1}\right] \coloneqq \mathrm{Tr}\left[\phi^+_{R_1}\rho_{R_1}\right]\mathcal{N}_A \;+\; \frac{1}{m-1}\Big(1 - \mathrm{Tr}\left[\phi^+_{R_1}\rho_{R_1}\right]\Big)\left(m\mathcal{E}_A - \mathcal{N}_A\right) \tag{3.132}$$

where $m\mathcal{E}_A - \mathcal{N}_A \geqslant 0$. Also observe that $\Omega_{R_1\to A}(\phi^+_{R_1}) = \mathcal{N}_A$. Hence, such a superchannel implies that

$$C^0_{\mathfrak{F}}(\mathcal{N}_A) \;=\; \log_2 m \;=\; \log_2\lceil t\rceil \;\leqslant\; \log_2 t + 1 \;=\; LR_{\mathfrak{C}}(\mathcal{N}_A) + 1 \tag{3.133}$$

This completes the proof of $LR_{\mathfrak{C}}(\mathcal{N}_A) \leqslant C^0_{\mathfrak{F}}(\mathcal{N}_A) \leqslant LR_{\mathfrak{C}}(\mathcal{N}_A) + 1$.

Therefore, using regularization and the additivity of $LR_{\mathfrak{C}}(\mathcal{N}_A)$, we can conclude

$$C^{\mathrm{exact}}_{\mathfrak{F}}(\mathcal{N}_A) = LR_{\mathfrak{C}}(\mathcal{N}_A) \tag{3.134}$$

∎

Exact cost under DISC

The dephasing log-robustness is given by (3.98)

$$LR_{\Delta}(\mathcal{N}_A) \coloneqq D_{\max}\big(\mathcal{N}_A\big\|\Delta_A[\mathcal{N}_A]\big) \quad \forall\, \mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)\,. \tag{3.135}$$

By definition we have $LR_{\mathfrak{C}}(\mathcal{N}_A) \leqslant LR_{\Delta}(\mathcal{N}_A)$.

Theorem 3.15. *The exact coherence cost of a channel $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$ under $\mathfrak{F} = \mathrm{DISC}$ is given by*

$$C^{\mathrm{exact}}_{\mathfrak{F}}(\mathcal{N}_A) = LR_{\Delta}(\mathcal{N}_A) \tag{3.136}$$

Proof. We first prove that

$$LR_{\Delta}(\mathcal{N}_A) \leqslant C^0_{\mathrm{DISC}}(\mathcal{N}_A) \leqslant LR_{\Delta}(\mathcal{N}_A) + 1 \tag{3.137}$$

and then use the additivity of LR_{Δ}.

For the proof of $LR_{\Delta}(\mathcal{N}_A) \leqslant C^0_{\mathrm{DISC}}(\mathcal{N}_A)$, let $\Theta \in \mathrm{DISC}(R_1 \to A)$ be an optimal superchannel satisfying $\Theta_{R_1\to A}[\phi^+_{R_1}] = \mathcal{N}_A$ such that $C^0_{\mathrm{DISC}}(\mathcal{N}_A) = \log_2|R_1|$. Therefore,

$$LR_{\Delta}(\mathcal{N}_A) = D_{\max}\left(\mathcal{N}_A \middle\| \Delta_A[\mathcal{N}_A]\right) \tag{3.138}$$

$$= D_{\max}\left(\Theta_{R_1\to A}[\phi^+_{R_1}] \middle\| \Delta_A \circ \Theta_{R_1\to A}\left[\phi^+_{R_1}\right]\right) \tag{3.139}$$

$$= D_{\max}\left(\Theta_{R_1\to A}[\phi^+_{R_1}] \middle\| \Theta_{R_1\to A}\left[\mathcal{D}_{R_1}(\phi^+_{R_1})\right]\right) \tag{3.140}$$

$$\leqslant D_{\max}\left(\phi^+_{R_1} \middle\| \mathcal{D}_{R_1}(\phi^+_{R_1})\right) \tag{3.141}$$

$$= \log_2 |R_1| \tag{3.142}$$

$$= C^0_{\mathrm{DISC}}(\mathcal{N}_A)\ . \tag{3.143}$$

For the proof of $C^0_{\mathrm{DISC}}(\mathcal{N}_A) \leqslant LR_{\Delta}(\mathcal{N}_A) + 1$, first let

$$LR_{\Delta}(\mathcal{N}_A) = D_{\max}\left(\mathcal{N}_A \middle\| \Delta_A[\mathcal{N}_A]\right) = \log t \tag{3.144}$$

for some optimal t that satisfies $t\Delta[\mathcal{N}] \geqslant \mathcal{N}$. Also, let $m = \lceil t \rceil$ so that $m\Delta[\mathcal{N}] \geqslant \mathcal{N}$ still holds, and let R_1 be a static system with dimension $|R_1| = m$. We now construct the following supermap. For any state $\rho \in \mathfrak{D}(R_1)$

$$\Omega_{R_1\to A}[\rho_{R_1}] \coloneqq \frac{m}{m-1}\Big(\mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}] - \frac{1}{m}\Big)\mathcal{N}_A \ + \ \frac{m}{m-1}\Big(1 - \mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}]\Big)\Delta_A[\mathcal{N}_A] \tag{3.145}$$

The supermap $\Omega_{R_1\to A}$ has several properties. First, it satisfies $\Delta_A \circ \Omega_{R_1\to A} = \Omega_{R_1\to A} \circ \mathcal{D}_{R_1}$. Indeed, for any density matrix $\rho \in \mathfrak{D}(R_1)$ we have

$$\begin{aligned} \Delta_A \circ \Omega_{R_1\to A}[\rho_{R_1}] &= \frac{m}{m-1}\Big(\mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}] - \frac{1}{m}\Big)\Delta_A[\mathcal{N}_A] \ + \ \frac{m}{m-1}\Big(1 - \mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}]\Big)\Delta_A[\mathcal{N}_A] \\ &= \Delta_A[\mathcal{N}_A]\ , \end{aligned} \tag{3.146}$$

and

$$\begin{aligned} \Omega_{R_1\to A}\big[\mathcal{D}_{R_1}(\rho_{R_1})\big] &= \frac{m}{m-1}\Big(\mathrm{Tr}[\phi^+_{R_1}\mathcal{D}_{R_1}(\rho_{R_1})] - \frac{1}{m}\Big)\mathcal{N}_A \ + \ \frac{m}{m-1}\Big(1 - \mathrm{Tr}[\phi^+_{R_1}\mathcal{D}_{R_1}(\rho_{R_1})]\Big)\Delta_A[\mathcal{N}_A] \\ &= \frac{m}{m-1}\Big(\frac{1}{m} - \frac{1}{m}\Big)\mathcal{N}_A \ + \ \frac{m}{m-1}\Big(1 - \frac{1}{m}\Big)\Delta_A[\mathcal{N}_A] \\ &= \Delta_A[\mathcal{N}_A]\ , \end{aligned} \tag{3.147}$$

so that $\Delta_A \circ \Omega_{R_1\to A} = \Omega_{R_1\to A} \circ \mathcal{D}_{R_1}$. Second, $\Omega_{R_1\to A}$ is a superchannel since the above map can be expressed

as

$$\Omega_{R_1\to A}[\rho_{R_1}] \coloneqq \mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}]\mathcal{N}_A + \frac{1}{m-1}\left(1-\mathrm{Tr}[\phi^+_{R_1}\rho_{R_1}]\right)\left(m\Delta_A[\mathcal{N}_A]-\mathcal{N}_A\right) \tag{3.148}$$

and $m\Delta_A[\mathcal{N}_A]-\mathcal{N}_A \geqslant 0$. Hence, $\Omega \in \mathrm{DISC}(R_1\to A)$. Finally, observe that $\Omega_{R_1\to A}[\phi^+_{R_1}] = \mathcal{N}_A$. Hence, the existence of such Ω implies that

$$C^0_{\mathrm{DISC}}(\mathcal{N}_A) \leqslant \log m = \log\lceil t\rceil \leqslant \log t + 1 = LR_\Delta(\mathcal{N}_A) + 1\ . \tag{3.149}$$

This completes the proof. ∎

3.5.3 Coherence cost of a channel

To find the approximate coherence cost (we will call it coherence cost) of any $\mathcal{N}\in\mathrm{CPTP}(A_0\to A_1)$, we first define the smoothed coherence cost as

$$C^\epsilon_{\mathfrak{F}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{N}'\in B_\epsilon(\mathcal{N})} C^0_{\mathfrak{F}}\left(\mathcal{N}'_A\right) \tag{3.150}$$

where

$$B_\epsilon(\mathcal{N}_A) = \left\{\mathcal{N}'\in\mathrm{CPTP}(A_0\to A_1)\ :\ \frac{1}{2}\|\mathcal{N}'_A-\mathcal{N}_A\|_\diamond \leqslant \epsilon\right\}. \tag{3.151}$$

The coherence cost of the channel $\mathcal{N}_A$ is then given by the regularization of the smoothed coherence cost and taking the limit of $\epsilon\to 0^+$

$$C_{\mathfrak{F}}(\mathcal{N}_A) = \lim_{\epsilon\to 0^+}\lim_{n\to\infty}\frac{1}{n}C^\epsilon_{\mathfrak{F}}(\mathcal{N}_A^{\otimes n}) \tag{3.152}$$

Below we find the coherence cost under MISC and DISC.

The cost under MISC

Theorem 3.16. *For* $\mathfrak{F} = \mathrm{MISC}$

$$C_{\mathfrak{F}}(\mathcal{N}_A) = LR^\infty_{\mathfrak{C}}(\mathcal{N}_A)\ . \tag{3.153}$$

Proof. First, note that from (3.123) it follows that

$$LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A) \leqslant C^\epsilon_{\mathfrak{F}}(\mathcal{N}_A) \leqslant LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A) + 1 \tag{3.154}$$

Hence,

$$\frac{1}{n}LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A^{\otimes n}) \leqslant \frac{1}{n}C^\epsilon_{\mathfrak{F}}(\mathcal{N}_A^{\otimes n}) \leqslant \frac{1}{n}LR^\epsilon_{\mathfrak{C}}(\mathcal{N}_A^{\otimes n}) + \frac{1}{n} \tag{3.155}$$

and the limit $n \to \infty$ concludes the proof. ∎

The cost under DISC

Theorem 3.17. *For* $\mathfrak{F} = \mathrm{DISC}$

$$C_{\mathfrak{F}}(\mathcal{N}_A) = LR^{\infty}_{\Delta}(\mathcal{N}_A) \, . \tag{3.156}$$

Proof. First, note that from (3.137) it follows that

$$LR^{\epsilon}_{\Delta}(\mathcal{N}_A) \leqslant C^{\epsilon}_{\mathfrak{F}}(\mathcal{N}_A) \leqslant LR^{\epsilon}_{\Delta}(\mathcal{N}_A) + 1 \tag{3.157}$$

Hence,

$$\frac{1}{n} LR^{\epsilon}_{\Delta}(\mathcal{N}_A^{\otimes n}) \leqslant \frac{1}{n} C^{\epsilon}_{\mathfrak{F}}(\mathcal{N}_A^{\otimes n}) \leqslant \frac{1}{n} LR^{\epsilon}_{\Delta}(\mathcal{N}_A^{\otimes n}) + \frac{1}{n} \tag{3.158}$$

and the limit $n \to \infty$ concludes the proof. ∎

The lack of AEP for channels motivates us to consider a more liberal method for smoothing.

3.5.4 Liberal coherence cost of a channel

We define the liberal one-shot ϵ-approximate coherence-cost as

$$C^{\epsilon}_{\mathfrak{F}}(\mathcal{N}_A) \coloneqq \max_{\varphi \in \mathfrak{D}(RA_0)} C^{\epsilon,\varphi}_{\mathfrak{F}}(\mathcal{N}_A) \, , \tag{3.159}$$

where

$$C^{\epsilon,\varphi}_{\mathfrak{F}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{N}'_A \in B^{\varphi}_{\epsilon}(\mathcal{N}_A)} C^{0}_{\mathfrak{F}}(\mathcal{N}'_A) \, , \tag{3.160}$$

and

$$B^{\varphi}_{\epsilon}(\mathcal{N}_A) \coloneqq \Big\{ \mathcal{N}' \in \mathrm{CP}(A_0 \to A_1) \ : \ \| \mathcal{N}'_A(\varphi_{RA_0}) - \mathcal{N}_A(\varphi_{RA_0}) \|_1 \leqslant \epsilon \Big\}. \tag{3.161}$$

The liberal coherence cost is defined by regularizing the liberal one-shot ϵ-approximate coherence cost and then taking the limit of $\epsilon \to 0^+$ as follows

$$\begin{aligned} C^{(\infty)}_{\mathfrak{F}}\left(\mathcal{N}_A\right) &\coloneqq \lim_{\epsilon \to 0^+} \lim_{n \to \infty} \max_{\varphi \in \mathfrak{D}(RA)} \frac{1}{n} C^{\epsilon, \varphi^{\otimes n}}_{\mathfrak{F}}\left(\mathcal{N}^{\otimes n}\right) \\ &= \lim_{\epsilon \to 0^+} \lim_{n \to \infty} \max_{\varphi \in \mathfrak{D}(RA)} \min_{\mathcal{N}' \in B^{\varphi^{\otimes n}}_{\epsilon}(\mathcal{N}^{\otimes n})} \frac{1}{n} C^{0}_{\mathfrak{F}}(\mathcal{N}'_{A^n \to B^n}) \end{aligned} \tag{3.162}$$

One can interpret the above cost in the following way. For any pure state $\varphi \in \mathfrak{D}(RA_0)$ (with $|R| = |A_0|$ and

φ is full Schmidt rank) we define a φ-norm

$$\|\mathcal{E}_A\|_\varphi \coloneqq \|\mathcal{E}_A(\varphi_{RA_0})\|_1 \ . \tag{3.163}$$

Then the liberal cost can also be expressed as

$$C_{\mathfrak{F}}^{(\infty)}(\mathcal{N}_A) = \lim_{\epsilon\to 0^+}\lim_{n\to\infty}\max_{\varphi\in\mathfrak{D}(RA_0)}\min_{\|\mathcal{N}'-\mathcal{N}^{\otimes n}\|_{\varphi^{\otimes n}}\leqslant\epsilon}\frac{1}{n}C_{\mathfrak{F}}^{0}(\mathcal{N}'_{A^n\to B^n}) \tag{3.164}$$

That is, we smooth with the $\varphi_{RA_0}^{\otimes n}$-norm and then maximizing over all such norms.

Theorem 3.18. *For* $\mathfrak{F} = \text{MISC}$

$$C_{\mathfrak{F}}^{(\infty)}(\mathcal{N}_A) = D_{\mathfrak{C}}^{(\infty)}(\mathcal{N}_A) \tag{3.165}$$

Proof. From (3.123) it follows that that for any fixed $\varphi \in \mathfrak{D}(RA_0)$ we have

$$LR_{\mathfrak{C}}^{\epsilon,\varphi}(\mathcal{N}_A) \leqslant C_{\mathfrak{F}}^{\epsilon,\varphi}(\mathcal{N}_A) \leqslant LR_{\mathfrak{C}}^{\epsilon,\varphi}(\mathcal{N}_A) + 1 \tag{3.166}$$

From (3.166) it follows that $C_{\mathfrak{F}}^{(\infty)}(\mathcal{N}_A) = LR_{\mathfrak{C}}^{(\infty)}(\mathcal{N}_A)$ so that the theorem follows from the AEP relation (3.114). ∎

3.5.5 One shot deterministic distillation of coherence

We now consider the problem of distilling an arbitrary channel into pure-state coherence using MISC and DISC. Let $\Theta \in \mathfrak{F}(A \to B_1)$ where $\mathfrak{F} = \text{MISC}$ or DISC, such that for any input channel $\mathcal{E}_A$, the output is a state preparation channel $\mathcal{F}_B \in \text{CPTP}(B_0 \to B_1)$ where B_0 is a trivial system. For $\epsilon > 0$ and $n = |B_1|$, define

$$\text{DISTILL}_{\mathfrak{F}}^{\epsilon}(\mathcal{N}_A) = \log\max\{n \ : \ \langle\phi_{B_1}^+|\Theta[\mathcal{N}_A]|\phi_{B_1}^+\rangle > 1-\epsilon, \ \ \Theta \in \mathfrak{F}(A\to B_1)\}, \tag{3.167}$$

which represents the largest coherence attainable by MISC or DISC within ϵ-error. For all $\mathcal{N} \in \text{CPTP}(A_0 \to A_1)$, we can write

$$\begin{aligned}\langle\phi_{B_1}^+|\Theta[\mathcal{N}]|\phi_{B_1}^+\rangle &= \left\langle\phi_{B_1}^+\middle|\left(\text{Tr}_A\left[\mathbf{J}_{AB_1}^{\Theta}\left(\left(J_A^{\mathcal{N}}\right)^T\otimes I_{B_1}\right)\right]\right)\middle|\phi_{B_1}^+\right\rangle \\ &= \text{Tr}\left[\mathbf{J}_{AB_1}^{\Theta}\left(\left(J_A^{\mathcal{N}}\right)^T\otimes\phi_{B_1}^+\right)\right].\end{aligned} \tag{3.168}$$

Note that the space of all operators that are invariant under any permutation in the classical basis, is a linear combination of maximally mixed state, u_{A_1} and maximally coherent state, $\phi_{A_1}^+$. Any operator is permutation

invariant if

$$\Pi_x \, \sigma \, \Pi_x^\dagger = \sigma \;\; \forall \;\; \text{permutation matrices } \Pi_x \tag{3.169}$$

The permutation-twirling operation can be expressed in the following way (see for example [230])

$$\mathcal{T}(\cdot) = \frac{1}{m!} \sum_x \Pi_x(\cdot)\Pi_x^\dagger \;\; \forall \, \Pi_x \tag{3.170}$$

where m is the dimension of the input system. Observe that the output of the above permutation-twirling operation on any state is permutation invariant and so can always be represented as a linear combination of $\phi^+_{A_1}$ and u_{A_1}. Hence, we can express the second equality in (3.168) as

$$\begin{aligned} \text{Tr}\left[\mathbf{J}^\Theta_{AB_1}\left(\left(J^{\mathcal{N}}_A\right)^T \otimes \phi^+_{B_1}\right)\right] &= \text{Tr}\left[\mathbf{J}^\Theta_{AB_1}\left(\left(J^{\mathcal{N}}_A\right)^T \otimes \mathcal{T}\left(\phi^+_{B_1}\right)\right)\right] \\ &= \text{Tr}\left[\left(\mathsf{id}_A \otimes \mathcal{T}\left(\mathbf{J}^\Theta_{AB_1}\right)\right)\left(\left(J^{\mathcal{N}}_A\right)^T \otimes \phi^+_{B_1}\right)\right] \end{aligned} \tag{3.171}$$

where the second equality follows from the fact that $\mathcal{T}$ is self-adjoint in the Hilbert-Schmidt inner product. Hence, without loss of generality we can express the Choi matrix $\mathbf{J}^\Theta_{AB_1}$ in following way

$$\mathbf{J}^\Theta_{AB_1} = \alpha_A \otimes \phi^+_{B_1} + \frac{1}{n-1}\beta_A \otimes (I_{B_1} - \phi^+_{B_1}) \tag{3.172}$$

where $n = |B_1|$ and $\alpha_A, \beta_A \in \text{Herm}(A)$ such that $\mathbf{J}^\Theta_{AB_1} \geqslant 0$, $\mathbf{J}^\Theta_{A_1} = I_{A_1}$, and $\mathbf{J}^\Theta_A = \mathbf{J}^\Theta_{A_0} \otimes u_{A_1}$. In terms of α_A and β_A, we can write these conditions as

$$\alpha_A, \beta_a \geqslant 0, \tag{3.173}$$

$$\text{Tr}(\alpha_A + \beta_A) = |A_1| \tag{3.174}$$

$$\alpha_A + \beta_A = \text{Tr}_{A_1}(\alpha_A + \beta_A) \otimes u_{A_1}. \tag{3.175}$$

From the MISC condition of $\mathcal{D}_{AB}(\mathbf{J}^\Theta_{AB}) = \mathcal{D}_A \otimes \mathsf{id}_B(\mathbf{J}^\Theta_{AB})$, we get

$$\mathcal{D}(\alpha_A)(n-1) = \mathcal{D}(\beta_A). \tag{3.176}$$

Defining $\beta_A = \rho_{A_0} \otimes I_{A_1} - \alpha_A$ where $\rho_{A_0} = \frac{1}{|A_1|}\mathrm{Tr}_{A_1}(\alpha_A + \beta_A)$. Since $\mathrm{Tr}[\rho_{A_0}] = 1$, ρ_{A_0} is a density matrix. So, we can rewrite these constraints as

$$\alpha_A \geqslant 0 \tag{3.177}$$

$$\rho_{A_0} \otimes I_{A_1} \geqslant \alpha_A, \tag{3.178}$$

$$\frac{1}{n}\mathcal{D}(\rho_{A_0}) \otimes I_{A_1} = \mathcal{D}(\alpha_A), \tag{3.179}$$

$$\rho_{A_0} \in \mathfrak{D}(A_0) \tag{3.180}$$

We can also consider imposing the additional DISC constraint of $\mathsf{id}_A \otimes \mathcal{D}_B(\mathbf{J}^{\Theta}_{AB}) = \mathcal{D}_A \otimes \mathsf{id}_B(\mathbf{J}^{\Theta}_{AB})$ which gives

$$\alpha_A + \beta_A = \mathcal{D}(\alpha_A + \beta_A). \tag{3.181}$$

This amounts to replacing Eq. (3.178) with the condition

$$n\mathcal{D}(\alpha_A) \geqslant \alpha_A. \tag{3.182}$$

Next notice that we can always write $\alpha_A = \mathcal{D}_A(\alpha_A) + \gamma_A$ for some γ_A with zeroes on the diagonal. Then, since $\mathrm{Tr}_{A_1}\left[\left(J^{\mathcal{N}}_A\right)^T\right] = I_{A_0}$, we can write

$$\begin{aligned}
\mathrm{Tr}\left[\alpha_A \left(J^{\mathcal{N}}_A\right)^T\right] &= \mathrm{Tr}\left[\left(\mathcal{D}_A\left(\alpha_A\right) + \gamma_A\right)\left(J^{\mathcal{N}}_A\right)^T\right] \\
&= \mathrm{Tr}\left[\mathcal{D}_A\left(\alpha_A\right)\left(J^{\mathcal{N}}_A\right)^T\right] + \mathrm{Tr}\left[\gamma_A \left(J^{\mathcal{N}}_A\right)^T\right] \\
&= \mathrm{Tr}\left[\left(\frac{1}{n}\mathcal{D}(\rho_{A_0}) \otimes I_{A_1}\right)\left(J^{\mathcal{N}}_A\right)^T\right] + \mathrm{Tr}\left[\gamma_A \left(J^{\mathcal{N}}_A\right)^T\right] \\
&= \frac{1}{n} + \mathrm{Tr}\left[\gamma_A \left(J^{\mathcal{N}}_A\right)^T\right]
\end{aligned} \tag{3.183}$$

Hence, we have the following one-shot distillable rates.

Theorem 3.19. *For* $\mathfrak{F} = \mathrm{MISC}$ *or* DISC

$$\mathrm{DISTILL}^{\epsilon}_{\mathfrak{F}}(\mathcal{N}) = \log \max\ n \tag{3.184}$$

such that

$$\mathrm{Tr}\left[\gamma_A \left(J_A^{\mathcal{N}}\right)^T\right] \geqslant 1 - \frac{1}{n} - \epsilon \,, \tag{3.185}$$

$$\mathcal{D}_A(\gamma_A) = 0 \,, \tag{3.186}$$

$$\rho_{A_0} \in \mathfrak{D}(A_0) \,, \tag{3.187}$$

$$\left[\rho_{A_0} - \frac{1}{n}\mathcal{D}_{A_0}\left(\rho_{A_0}\right)\right] \otimes I_{A_1} \geqslant \gamma_A \geqslant -\frac{1}{n}\mathcal{D}_{A_0}\left(\rho_{A_0}\right) \otimes I_{A_1} \quad \textit{(specifically for } \mathfrak{F} = \mathrm{MISC}\textit{)} \,, \tag{3.188}$$

$$\frac{n-1}{n}\mathcal{D}_{A_0}(\rho_{A_0}) \otimes I_{A_1} \geqslant \gamma_A \geqslant -\frac{1}{n}\mathcal{D}_{A_0}\left(\rho_{A_0}\right) \otimes I_{A_1} \quad \textit{(specifically for } \mathfrak{F} = \mathrm{DISC}\textit{)}. \tag{3.189}$$

Remark 6. The first condition in the above theorem, i.e., Eq. (3.185) follows by substituting (3.172) in $\mathrm{Tr}[\mathbf{J}^{\Theta}_{AB_1}((J_A^{\mathcal{N}})^T \otimes \phi^+_{B_1})]$ and using the fact that $\mathrm{Tr}\left[\phi^+_{A_1}\phi^+_{A_1}\right] = \mathrm{Tr}\left[\phi^+_{A_1}\right]$, and then plugging the result in (3.183). The lhs of (3.188) is a consequence of (3.178) and (3.179) whereas the rhs is a consequence of (3.177) and (3.179). Similarly, for DISC, the lhs of (3.189) is a consequence of (3.179) and (3.182) whereas the rhs of (3.189) is a consequence of (3.177) and (3.179).

Remark 7. Note that $D^{\epsilon}_{\mathrm{MISC}}(\mathcal{N}) = D^{\epsilon}_{\mathrm{DISC}}(\mathcal{N})$ when $|A_0| = 1$, and their common rate matches that given in Refs. [231, 232] for distilling coherence from static resources (i.e. states). However for channels, the MISC and DISC distillable coherence can possibly differ. We leave it as an open problem to find channels that have such a property.

Example 1. Let us consider the partially depolarizing channel $\mathcal{N}^{\mathrm{dep}}_{\lambda,d} : \mathcal{B}(A_1) \to \mathcal{B}(A_1)$,

$$\mathcal{N}^{\mathrm{dep}}_{\lambda,d}(\chi) = \lambda\chi + (1-\lambda)\mathrm{Tr}[\chi]u_{A_1}. \tag{3.190}$$

where $d = |A_1|$. The Choi matrix of this channel is given by

$$J^{\mathcal{N}^{\mathrm{dep}}}_{A_1\tilde{A}_1} = \lambda\phi^+_{A_1\tilde{A}_1} + \frac{1-\lambda}{d}I_{A_1\tilde{A}_1}, \tag{3.191}$$

We exploit the symmetry by noting that both $\phi^+_{A_1\tilde{A}_1}$ and $I_{A_1\tilde{A}_1}$ are $U^* \otimes U$ invariant. We restrict our twirling to an average over the group of incoherent unitaries, i.e., each U involves a permutation and/or a change in relative phase. Note that dephasing commutes with this operation so if Eq. (3.179) holds before the twirl, it will also hold after. The action of twirling will convert $\rho_{A_1} \otimes I_{\tilde{A}_1} \to u_{A_1} \otimes I_{\tilde{A}_1}$ while converting α_A into

an operator of the form

$$\begin{aligned}\alpha_{A_1\tilde{A}_1} &= p\sum_{i\neq j}|ij\rangle\langle ij| + q\sum_i |ii\rangle\langle ii| + r\sum_{i\neq j}|ii\rangle\langle jj| \\ &= p\sum_{i\neq j}|ij\rangle\langle ij| + (q-r)\sum_i |ii\rangle\langle ii| + r\phi^+_{A_1\tilde{A}_1}.\end{aligned} \tag{3.192}$$

The eigenvalues of $\alpha_{A_1\tilde{A}_1}$ are easily seen to be $\{p, q-r, q-r+rd\}$, and so equations (3.177) and (3.178) require that $p, q-r \geqslant 0$ and $p, q-r+rd \leqslant \frac{1}{d}$. From equation (3.179), we must also have $p = q = \frac{1}{nd}$. With these constraints in place, our goal is to maximize n such that

$$\mathrm{Tr}\left[\alpha^T_{A_1\tilde{A}_1} J^{\mathcal{N}^{\mathrm{dep}}}_{A_1\tilde{A}_1}\right] = (1-\lambda)\frac{(d-1)}{nd} + \left(\frac{1}{nd} - r\right)(\lambda d + (1-\lambda)) + r(\lambda d^2 + (1-\lambda)). \tag{3.193}$$

This function is strictly increasing w.r.t. r, and the constraints necessitate that $r \leqslant \min\{\frac{n-1}{d-1}\frac{1}{nd}, \frac{1}{nd}\}$. So when $n \leqslant d$, we take $r = \frac{n-1}{d-1}\frac{1}{nd}$ and obtain

$$\begin{aligned}\mathrm{Tr}\left[\alpha^T_{A_1\tilde{A}_1} J^{\mathcal{N}^{\mathrm{dep}}}_{A_1\tilde{A}_1}\right] &= (1-\lambda)\frac{d-1}{nd} + \frac{d-n}{nd(d-1)}(\lambda d + (1-\lambda)) + \frac{n-1}{nd(d-1)}(\lambda d^2 + (1-\lambda)) \\ &= (1-\lambda)\frac{1}{n} + \lambda.\end{aligned} \tag{3.194}$$

Notice that when $\lambda = 1$ we obtain $\mathrm{Tr}\left[\alpha^T_{A_1\tilde{A}_1} J^{\mathcal{N}^{\mathrm{dep}}}_{A_1\tilde{A}_1}\right] = 1$. This says that $\log n$ bits can be perfectly distilled, which is expected: the free superchannel just consists of inputting $\phi^+_{A_1\tilde{A}_1}$ into the given channel and then as post-processing performs a MIO map that converts $\phi^+_{A_1\tilde{A}_1}$ into $\phi^+_{B_1}$. On the other hand, if $n \geqslant d$, we take $r = \frac{1}{nd}$ and Eq. (3.194) becomes

$$\mathrm{Tr}\left[\alpha^T_{A_1\tilde{A}_1} J^{\mathcal{N}^{\mathrm{dep}}}_{A_1\tilde{A}_1}\right] = (1-\lambda)\frac{1}{n} + \frac{d}{n}\lambda. \tag{3.195}$$

Notice also that in this case our optimizer ρ_{A_0} is completely dephased, which means our solution for MISC is also the solution for DISC. We summarize our findings as follows.

Lemma 3.20. *For the partial depolarizing channel $\mathcal{N}^{dep}_{\lambda,d}$ and $0 \leqslant \epsilon < 1$,*

$$\mathrm{DISTILL}^{\epsilon}_{\mathrm{MISC}}\left(\mathcal{N}^{dep}_{\lambda,d}\right) = \mathrm{DISTILL}^{\epsilon}_{\mathrm{DISC}}\left(\mathcal{N}^{dep}_{\lambda,d}\right) = \begin{cases} \log\lfloor\frac{1-\lambda}{1-\lambda-\epsilon}\rfloor & \text{if } \epsilon < \frac{(d-1)(1-\lambda)}{d} \\ \log\lfloor\frac{1-\lambda+\lambda d}{1-\epsilon}\rfloor & \text{if } \epsilon \geqslant \frac{(d-1)(1-\lambda)}{d} \end{cases}. \tag{3.196}$$

Example 2. We next consider the partial dephasing channel $\mathcal{N}^{\Delta}_{\lambda,d} : \mathcal{B}(A_1) \to \mathcal{B}(A_1)$,

$$\mathcal{N}^{\Delta}_{\lambda,d}(\chi) = \lambda\chi + (1-\lambda)\mathcal{D}(\chi). \tag{3.197}$$

The Choi matrix of this channel is given by

$$J^{\mathcal{N}^{\text{dep}}}_{A_1\tilde{A}_1} = \lambda\phi^{+}_{A_1\tilde{A}_1} + (1-\lambda)\sum_{i=1}^{d} |ii\rangle\langle ii|. \tag{3.198}$$

By the same argument as before, we can assume without loss of generality that α_A has the form

$$\alpha_{A_1\tilde{A}_1} = p\sum_{i\neq j} |ij\rangle\langle ij| + (q-r)\sum_{i} |ii\rangle\langle ii| + r\phi^{+}_{A_1\tilde{A}_1}. \tag{3.199}$$

However this time the fidelity with $\phi^{+}_{B_1}$ is given by

$$\text{Tr}\left[\alpha^{T}_{A_1\tilde{A}_1} J^{\mathcal{N}^{\Delta}}_{A_1\tilde{A}_1}\right] = \left(\frac{1}{nd} - r\right) d + r(\lambda d^2 + (1-\lambda)d). \tag{3.200}$$

Again, the constraints of the problem demand $r \leqslant \min\{\frac{n-1}{d-1}\frac{1}{nd}, \frac{1}{nd}\}$. When $n \leqslant d$, it holds that

$$\begin{aligned}\text{Tr}\left[\alpha^{T}_{A_1\tilde{A}_1} J^{\mathcal{N}^{\Delta}}_{A_1\tilde{A}_1}\right] &= \frac{d-n}{n(d-1)} + \frac{n-1}{n(d-1)}(\lambda d + (1-\lambda)) \\ &= \frac{1+(n-1)\lambda}{n}.\end{aligned} \tag{3.201}$$

On the other hand, when $n \geqslant d$, we take $r = \frac{1}{nd}$ to obtain

$$\text{Tr}\left[\alpha^{T}_{A_1\tilde{A}_1} J^{\mathcal{N}^{\Delta}}_{A_1\tilde{A}_1}\right] = \frac{\lambda d + (1-\lambda)}{n}. \tag{3.202}$$

These are the same maximum fidelities as the depolarizing channel, and we therefore have the following conclusion.

Lemma 3.21. *For the partial dephasing channel $\mathcal{N}^{\Delta}_{\lambda,d}$ and $0 \leqslant \epsilon < 1$,*

$$\text{DISTILL}^{\epsilon}_{\text{MISC}}\left(\mathcal{N}^{\Delta}_{\lambda,d}\right) = \text{DISTILL}^{\epsilon}_{\text{DISC}}\left(\mathcal{N}^{\Delta}_{\lambda,d}\right) = \begin{cases} \log\lfloor\frac{1-\lambda}{1-\lambda-\epsilon}\rfloor & \text{if } \epsilon < \frac{(d-1)(1-\lambda)}{d} \\ \log\lfloor\frac{1-\lambda+\lambda d}{1-\epsilon}\rfloor & \text{if } \epsilon \geqslant \frac{(d-1)(1-\lambda)}{d} \end{cases}. \tag{3.203}$$

3.5.6 One shot probabilistic distillation of coherence

In this sub-section, we consider the problem of probabilistic distillation of coherence from quantum channels. For probabilistic conversions, the free superchannels need to be non-deterministic or probabilistic. A completely positive supermap Θ is called a probabilistic superchannel if there exists a superchannel $\Xi \in \mathfrak{S}(AB)$ such that $\Xi \geqslant \Theta$. Here the greater-than-equal notation means that the difference $\Xi - \Theta$ is a completely positive map. Let us denote the set of probabilistic supermaps taking a dynamical system A to a dynamical system B by $\mathfrak{S}_{\text{prob}}(AB)$. It is interesting to remark that, all physical realizable manipulations that send completely positive and trace-preserving map to completely positive and trace non-increasing map are probabilistic superchannels. The concept of probabilistic superchannel is known as measurement on quantum channels in Ref. [125], and probabilistic quantum networks in Ref. [122]. The existence of a superchannel Ξ such that $\Xi \geqslant \Theta$ sets linear constraints for $\Theta \in \mathfrak{S}_{\text{prob}}(AB)$; that are

$$\text{Tr}_{B_1}[J^{\Theta}_{AB}] \leqslant X_{B_0A_0} \otimes I_{A_1}, \tag{3.204}$$

$$\text{Tr}_{A_0}[X_{B_0A_0}] = I_{B_0}, \tag{3.205}$$

where J^{Θ}_{AB} denotes the Choi operator of Θ.

The probabilistic distillation of coherence from channels is the process of transforming a given quantum channel $\mathcal{N}$ to a d-dimensional maximally coherent state $|\phi^+_d\rangle := (1/\sqrt{d}) \sum_{i=0}^{d-1} |i\rangle$. If we also allow the fidelity between the output state and the maximally coherent state to be no less than $1-\epsilon$ under free probabilistic superchannels, then, the maximal probability in distilling ϕ^+_d with error tolerance ϵ is characterized by the following optimization problem

$$\begin{aligned} P^{\epsilon}_{\mathfrak{F}_{\text{prob}}}(\mathcal{N}; d) := \max \quad & p \\ \text{s.t.} \quad & \Theta(\mathcal{N}) = p\sigma, \\ & F(\sigma, \phi^+_d) \geqslant 1-\epsilon,\ \Theta \in \mathfrak{F}_{\text{prob}}, \end{aligned} \tag{3.206}$$

where $\mathfrak{F}_{\text{prob}}$ denotes the set of all free probabilistic superchannels, such as $\text{MISC}_{\text{prob}}$ and $\text{DISC}_{\text{prob}}$. In particular, the elements of $\text{MISC}_{\text{prob}}$ are probabilistic superchannels that are MISC, and the elements of $\text{DISC}_{\text{prob}}$ are probabilistic superchannels that are DISC. Since the target of distillation is a quantum state, the operator $X_{B_0A_0}$ in Eqs. 3.204 and 3.205 can be replaced by some quantum state ρ_{A_0}. By fixing the error tolerance ϵ, the maximal probability of distilling ϕ^+_d under $\text{MISC}_{\text{prob}}$ is characterized as follows.

Proposition 1. Given error tolerance ϵ, the single-shot probabilistic distillation of dynamical coherence for

$\mathcal{N}$ under $\text{MISC}_{\text{prob}}$ is given by the following optimization problem

$$\begin{aligned} P^{\epsilon}_{\text{MISC}_{\text{prob}}}(\mathcal{N};d) = \max \quad & \text{Tr}[Z_A \cdot J^{\mathcal{N}}_A] \\ \text{s.t.} \quad & \text{Tr}[X_A \cdot J^{\mathcal{N}}_A] \geqslant (1-\epsilon)\text{Tr}[Z_A \cdot J^{\mathcal{N}}_A], \\ & 0 \leqslant X_A \leqslant Z_A \leqslant \rho_{A_0} \otimes I_{A_1}, \\ & \mathcal{D}(Z) = d\mathcal{D}(X),\ \text{Tr}[\rho_{A_0}] = 1, \end{aligned} \tag{3.207}$$

where $J^{\mathcal{N}}_A$ represents the Choi matrix of $\mathcal{N}_A$.

Proof. The proofs of the above propositions are provided in Appendix A.4. ∎

By further restricting the capability of manipulating quantum operations, the maximal probability of distilling ϕ^+_d under $\text{DISC}_{\text{prob}}$ is solved by

Proposition 2. Given error tolerance ϵ, the single-shot probabilistic distillation of dynamical coherence for $\mathcal{N}$ under $\text{DISC}_{\text{prob}}$ is given by the following optimization problem

$$\begin{aligned} P^{\epsilon}_{\text{DISC}_{\text{prob}}}(\mathcal{N};d) = \max \quad & d\text{Tr}[\mathcal{D}(X_A) \cdot J^{\mathcal{N}}_A] \\ \text{s.t.} \quad & \text{Tr}[X_A \cdot J^{\mathcal{N}}_A] \geqslant d(1-\epsilon)\text{Tr}[\mathcal{D}(X_A) \cdot J^{\mathcal{N}}_A], \\ & 0 \leqslant X_A \leqslant d\mathcal{D}(X_A) \leqslant \rho_{A_0} \otimes I_{A_1}, \\ & \text{Tr}[\rho_{A_0}] = 1. \end{aligned} \tag{3.208}$$

Proof. The proofs of the above propositions is similar to the one in Appendix A.4 with the difference being in the conditions of DISC instead of MISC. ∎

3.6 Outlook and Conclusions

In this chapter, I have developed the resource theory of dynamical coherence using the classical channels as free channels. In previous works on the quantum resource theory of dynamical coherence [143, 196, 197, 198, 209], the free channels were taken to be the free operations from the QRT of static coherence, like MIO, IO, etc. However, if we consider distributed quantum computing scenarios, one encounters channels with noisy pre- and post-processing links. In this case, a natural candidate for free channels are those with completely dephasing pre- and post- processing. What emerges is a resource theory in which the free objects are the set of classical channels. Note that in such a theory, the T-gate (crucial for quantum computation) is not free

and even the quantum identity channel is not free as the preservation of coherence should be considered a resource.

Similar to the static QRT of coherence where the free operations can have a non-free dilation, in our work on dynamical QRT of coherence, the free superchannels can have a non-free realization. That means, the pre- and post-processing channels need not be classical. The only requirement on the set of free superchannels comes from the golden rule of QRT. This implies that the free superchannels must never generate coherent channels when the input channels are classical, even when tensored with identity, i.e., even when the free superchannel acts on a part of the input classical channel. This enlargement of the set of free superchannels is necessary for a meaningful resource theory of coherence. Take for example the set of free superchannels which can be realized only by classical pre- and post-processing channels. In this case, the output channel is always classical irrespective of the input channel, eliminating all the advantage offered by a quantum channel. Thus, such free superchannels cannot be used to study the resource theory of quantum coherence.

In section 3.3, we start by defining four sets of free superchannels. We name them as maximally incoherent superchannels (MISC), dephasing-covariant incoherent superchannels (DISC), incoherent superchannels (ISC), and strictly incoherent superchannels (SISC). We show that the set of free superchannels in the dynamical resource theory of coherence can be characterized analogous to the free channels in the static resource theory of coherence. We also show that MISC and DISC can be characterized just on the basis of their Choi matrices and dephasing channels which is given in Eq. (3.57) and (3.65) for MISC and DISC, respectively.

Section 3.4 then deals with the quantification of dynamical coherence. In section 3.4.1, we find the complete set of monotones for MISC and DISC. That means, to see if we can convert from one quantum channel to another, it is sufficient to check if all the monotones of this (complete) set acting on one channel are greater than the other. A complete family of monotones for a general resource theory of processes was presented in [89]. It is, in general, a hard problem to compute these functions and in some cases like LOCC-based entanglement, it is even NP-hard. We show that for the resource theory of dynamical coherence, these functions (under MISC and DISC) can be computed using an SDP (Eq. (3.73)). Next, in section 3.4.2, we also find monotones that are based on the relative entropy. In [145], Gour and Winter showed that the generalization of relative entropy from states to channels is not unique. In their work, they listed six relative entropies as measures of dynamical resources. They also introduced a new type of smoothing called "liberal" smoothing. We show in section 3.4.2 that out of these relative entropies defined in [145], three relative entropies clearly form a monotone under MISC. For the case of DISC, we show that the channel divergence for a given channel and the same channel acted on by the dephasing superchannel forms a resource monotone under DISC. We then discuss various types of channel log-robustness of coherence,

which are based on the max-relative entropy of channels $D_{\max}$, and we show that it can be computed with an SDP (Eq. (3.88)). For the qubit case, we calculated the log-robustness of coherence for classical channels, identity channel, replacement channel, depolarizing channels, and unitary channels. We also show that the log-robustness of coherence of channels is additive under tensor product (Lemma 3.10). We then define a "liberally" smoothed log-robustness of coherence which when regularized is equal to a regularized relative entropy introduced in [145] (i.e., it satisfies AEP), and behaves monotonically under completely resource non-generating superchannels. .

The next section is dedicated to the problem of interconversion of one resource into another. In section 3.5.1, we define a conversion distance between two channels (Eq. (3.115)). A given channel can be simulated using another if the interconversion distance is very small. For MISC and DISC, we showed that the interconversion distance can be computed using an SDP (Theorem 3.13). We then calculated the exact, asymptotic, and liberal cost of coherence of a channel and found that the liberal cost of coherence is equal to a variant of regularized relative entropy. Lastly, in this section, we also define the one-shot distillable coherence for MISC and DISC, and calculate it for partial depolarizing and partial dephasing channels.

Due to the realization of a superchannel as a pre- and post-processing channel, there are added complexities in the generalization of a quantum resource theory of states to channels as mentioned in [89]. In our case, we see that the simple generalizations do not work. For example, while calculating coherence costs, we had to introduce the concept of liberal cost (based on liberal smoothing as defined in [145]) to show it to be equal to a relative entropy.

Clearly, this work is just a start of a whole unexplored field of the quantum resource theory of dynamical coherence. For instance, one can solve for interconversion, cost etc. for ISC and SISC. One can define more sets of superchannels analogous to how various free operations are defined in the static case. We also leave as open the problem of finding an example of a channel where the MISC and DISC distillable coherence are different. In section 3.5.5, we worked out the distillable coherence for the partial depolarizing channel and the partial dephasing channel and found no difference for MISC and DISC case.

Chapter 4

Resource theory of multi-qubit magic channels

4.1 Introduction and background

In recent years, several schemes have been developed to achieve fault-tolerant quantum computation, and most of them use the stabilizer formalism [26, 112, 233, 234, 235]. The stabilizer formalism consists of the preparation of stabilizer states, application of Clifford gates, and measurements in the computational basis. Within this formalism, pure non-stabilizer states (popularly known as magic states) are used as a resource to promote fault-tolerant quantum computation to universal quantum computation. This model of quantum computation is known as the magic state model of quantum computation and finding magic distillation rates and estimating classical simulation cost of quantum circuits are active areas of research in this field [97, 112, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261]. While formulating optimal rates promise better distillation protocols, improved classical simulations help benchmark the computational speedups offered by quantum computers [13, 253, 256, 258, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271]. It follows from the Gottesman-Knill theorem that it is possible to efficiently simulate any stabilizer circuit on a classical computer, hence rendering stabilizer states and operations useless for universal quantum computation [272, 273]. For this reason, this model fits the mold of quantum resource theories where all the states and operations that cannot provide any quantum advantage are treated as free [37, 82, 90, 249, 253, 274, 275].

Using the above criterion to define free elements, considerable work has been directed towards developing the resource theory of magic states and channels [90, 97, 247, 251, 253, 258, 259, 260, 270, 271, 276, 277, 278,

279]. In this process, two branches have emerged: one branch deals with odd d-dimensional qudits, and the other branch deals with the practically important case of multi-qubit systems. In the former case, a clear connection between quantum speedup and the negativity of the Wigner representation of the state/channel has been established [90, 97, 245, 246, 247, 280, 281, 282, 283, 284, 285]. However, in the latter case, a discrete phase space approach cannot be cleanly applied without restricting free states to some subset of stabilizer states or excluding some Clifford operations [249, 251, 255, 286, 287, 288, 289, 290]. Thus, to retain all stabilizer states and operations as free elements (in the multiqubit scenario), alternative approaches have been taken [248, 252, 253, 256, 257, 258, 271, 279, 283, 289, 291, 292].

In [253], Howard and Campbell presented a scheme where all density matrices are decomposed as real linear combinations of pure stabilizer states. Borrowing the idea from the resource theory of entanglement [293], they introduced the robustness of magic which is the minimum ℓ_1-norm of all such decompositions. They showed that it is a resource monotone under all stabilizer operations and linked it to the runtime of a classical simulation algorithm, thus giving robustness of magic an operational meaning. Using robustness of magic, they also formulated lower bounds on the cost of synthesizing magic gates. Taking this approach forward, Seddon and Campbell enlarged the set of free operations from stabilizer operations to the set of completely stabilizer preserving operations (CSPOs) and introduced channel robustness of magic for multi-qubit channels [258]. They decomposed a channel as a linear combination of CSPOs and defined channel robustness as the minimum ℓ_1-norm of all such decompositions. They also formulated a classical algorithm and linked its runtime with the channel robustness thus efficiently simulating a circuit consisting of CSPOs.

Since CSPOs cannot provide any quantum advantage, we extend the resource theory of magic states to the channel case by treating CSPOs as free. We introduce two sets of free superchannels, CSPO preserving superchannels and completely CSPO preserving superchannels, to manipulate quantum channels. Since there is no physical restriction over such sets of free superchannels, they are useful in finding fundamental limitations on the ability of a quantum channel to generate magic states. Besides, studying such superchannels gives us no-go results in resource interconversion tasks involving more restricted type of operations such as the set of stabilizer operations.

This work is organized as follows. In section 4.3, we define and characterize the two above-mentioned sets of free superchannels. Then in section 4.4, we generalize the key operational magic monotones defined for states to the channel domain, namely the generalized robustness of magic states and the min relative entropy of magic states. Using these monotones, in section 4.5, we formulate single shot bounds on distilling single qubit magic states from a quantum channel and the magic cost of simulating a channel under the free superchannels. However, due to the complexity in determining whether a state is a stabilizer state or not [253, 280, 294], finding the lower bound on distillation under completely CSPO preserving superchannels

is still an open problem. In section 4.5, we also show that interconversion among single-qubit states under CSPOs is an SDP feasibility problem and hence, can be efficiently solved. As our last result, in section 4.6, we provide an algorithm to classically simulate a general quantum circuit and relate the runtime of this algorithm to the generalized robustness of dynamical magic resources and a desired precision.

4.1.1 Stabilizer formalism

In this subsection, we give a brief overview of the stabilizer formalism as defined for a system of n qubits. For single-qubit systems, the Pauli group consists of Pauli matrices and the identity matrix, together with multiplicative factors $\pm 1, \pm i$. We will denote this group as $\mathcal{P}_1 = (\pm 1, \pm i)\{I, X, Y, Z\}$. For multi-qubit systems, general Pauli group on n-qubits consists of all n-fold tensor products of Pauli matrices (including identity), together with the multiplication factors $\pm 1, \pm i$. We will denote the n-qubit Pauli group as $\mathcal{P}_n$. We say a pure, n-qubit state $|\psi\rangle$ is a stabilizer state if there exists an Abelian subgroup of the Pauli group $\mathcal{S} \subset \mathcal{P}_n$ such that $M|\psi\rangle = |\psi\rangle$ for all $M \in \mathcal{S}$. The elements of the subgroup $\mathcal{S}$ are called stabilizers of $|\psi\rangle$, and the total number of elements in $\mathcal{S}$ is equal to 2^n. One can also define a stabilizer subspace (with more than one stabilizer state) as the span of the elements of the set $V_{\mathcal{S}} = \{|\psi\rangle : M|\psi\rangle = |\psi\rangle \ \forall M \in \mathcal{S}\}$ as the set of simultaneous $+1$-eigenvalues of all the elements of a set $\mathcal{S} \subset \mathcal{P}_n$. From this definition, it is clear that $-I \notin \mathcal{S}$ and which in turn can be used to show that $\mathcal{S}$ will be an Abelian subgroup. Further, from group theory we know that every group has a generator set which is the set of those group elements from which all the elements of the group can be generated. The advantage of using independent generators to describe a group is that they provide a compact means of describing the group. A group G with size $|G|$ has a set of at most $\log(|G|)$ independent generators. In this chapter, when we refer to the set of generators, we will mean the set of independent generators. A stabilizer group $\mathcal{S}$ with 2^l elements can be expressed using l generators. With each generator, the dimension of the stabilizer subspace is halved. For example, the Pauli matrix Z is the stabilizer of state $|0\rangle$ and is the generator of the stabilizer group $\mathcal{S} = \{I, Z\}$. When the stabilizer group is generated by XX and ZZ, the stabilizer subspace is the span of the state $(1/\sqrt{2})(|00\rangle + |11\rangle)$, whereas when the stabilizer group is generated by ZZI and ZIZ, the stabilizer subspace is the span of the states $\{|000\rangle, |111\rangle\}$. Note that in this context, the notation XX or ZZ denote a tensor product and not a matrix multiplication. For a given stabilizer group on n qubits with l generators, the dimension of the stabilizer subspace $V_{\mathcal{S}}$ is equal to 2^{n-l}. For the purposes of this thesis, we are only interested in stabilizer groups $\mathcal{S}$ with $l = n$ independent generators. Then, for different stabilizer groups $\mathcal{S}$, there is a unique stabilizer state $|\phi\rangle$ stabilized by $\mathcal{S}$ and we can specify a stabilizer state by giving an n-element generating set. In this chapter, we will refer to the set of pure n-qubit stabilizer states as the set of all such unique stabilizer

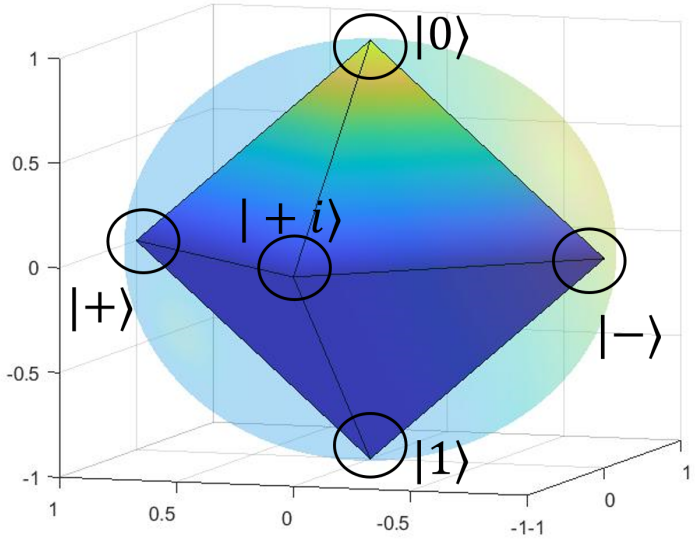

Figure 4.1: The set of qubit stabilizer states. The six pure stabilizer states: $|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle$, form the vertices of the octahedron.

states that can be described using n stabilizers. For single-qubit states, six different stabilizer groups can be formed each corresponding to a pure stabilizer state. Below we list these six pure stabilizer states with their following stabilizers

$$\pm X|\pm\rangle = |\pm\rangle \tag{4.1}$$

$$\pm Y|\pm i\rangle = |\pm i\rangle \tag{4.2}$$

$$Z|0\rangle = |0\rangle \tag{4.3}$$

$$-Z|1\rangle = |1\rangle \ . \tag{4.4}$$

The mixed stabilizer states of a system A_1 are defined as convex combination of pure stabilizer states, i.e., $\text{STAB}(A_1) = \text{conv}\{|\phi\rangle\langle\phi| : |\phi\rangle \in \text{STAB}(A_1)\}$ where $|\phi\rangle \in \text{STAB}(A_1)$ represent the pure stabilizer states. For qubits, the set of stabilizer states form an octahedron inside the Bloch sphere as shown in Fig. 4.1. We can also define the set of stabilizer states using Clifford unitaries which are the unitaries that preserve the Pauli group under conjugation. Let U represent an element of Clifford unitaries such that $UPU^\dagger \in \mathcal{P}_n$ for all $P \in \mathcal{P}_n$. Then the set of stabilizer states can be represented as $\text{conv}\{U|0\rangle\langle 0|U^\dagger : U \in \text{Clifford}\}$. This simply means that all stabilizer states can be generated by applying Clifford unitaries on the $|0\rangle$ state. Moreover, one can show that the Clifford group can be generated by the gate set $\{H, S, CNOT\}$ where H represents the Hadamard gate, S represents the phase gate, and $CNOT$ represents the controlled-NOT gate. This

can be easily verified for the qubit case where all the six pure stabilizer states can be formed by applying different sequences of H and S gates on $|0\rangle$. Clifford unitaries, along with preparation and measurement in computational basis, together constitute the set of stabilizer operations. Evolution of stabilizer states under Clifford unitaries can be efficiently tracked classically; even the measurement of Pauli operators on stabilizer states can be efficiently simulated [272, 273]. This is briefly discussed in Sec.4.1.2. A quantum circuit that comprises stabilizer operations, and classical randomness and conditioning, is known as a stabilizer circuit. The usefulness of the stabilizer formalism comes in quantum error correction and in efficiently simulating stabilizer circuits classically [272]. Below we give a brief sketch of classical simulation of stabilizer circuits as the partition of free and resource states. We will see later in this chapter that the efficient classical simulability of circuits lead to the partition of the free and resources states and channels when we discuss the resource theory of magic.

4.1.2 Classical simulation of stabilizer circuits

Stabilizer states can be easily described using the operators that stabilize them. To check if a state is stabilized by a group $\mathcal{S}$, we only need to check whether the state is stabilized by the generators of $\mathcal{S}$. The key idea behind simulating stabilizer circuits classically lies in the way we track the evolution of the input state under stabilizer operations. So, we need to understand how Clifford unitaries and Pauli measurements transform the stabilizer states [73, 233].

Let $V_{\mathcal{S}}$ denote the stabilizer subspace stabilized by the group $\mathcal{S}$ and let $|\psi\rangle$ be an element of $V_{\mathcal{S}}$. Then, if a unitary operation U is applied to $|\psi\rangle$, for any element $M \in \mathcal{S}$ we can write

$$U|\psi\rangle = UMU^*U|\psi\rangle \tag{4.5}$$

which implies that the state $U|\psi\rangle$ is stabilized by UMU^*. From this we can say that if $M_1, M_2, \ldots, M_l$ generate the group $\mathcal{S}$, then $UM_1U^*, UM_2U^*, \ldots, UM_lU^*$ generate the group $U\mathcal{S}U^* = \{UMU^* : M \in \mathcal{S}\}$. The advantage we get from this is that for certain special unitaries, the transformation of the stabilizer generators take a very simple form. Specifically, we are interested in the special unitaries that are building blocks of Clifford unitaries, i.e., the Hadamard gate, the phase gate, and the controlled-NOT gate. We list in Table 4.1 the transformations of Pauli matrices under these unitaries. Using this table we can find the generators of the output state. For instance, if a state is initially stabilized by X, after being acted upon by H, it will be stabilized by Z.

Next, we need to see how the stabilizers of a stabilizer state transform under Pauli measurement. Suppose we make a measurement of M on the state $|\psi\rangle$. We can assume without loss of generality that M is a product

conjugation by H	conjugation by S	conjugation by $CNOT_{1,2}$
$X \to Z$ $Y \to -Y$ $Z \to X$	$X \to Y$ $Y \to -X$ $Z \to Z$	$X_1 \to X_1X_2$ $X_2 \to X_2$ $Z_1 \to Z_1$ $Z_2 \to Z_1Z_2$

Table 4.1: Pauli evolution under Clifford unitaries

of Pauli matrices with no multiplicative factor. Also let $\{M_1, M_2, \ldots, M_l\}$ be the generators of $|\psi\rangle$. If M commutes with all the generators of $|\psi\rangle$, then, the measurement does not disturb the state and the output is still $|\psi\rangle$. If M anti-commutes with one or more generators, then we can form new generators such that there is only one generator, say M_1, which anti-commutes with M. Then, after measurement, the generators of ψ remain the same except M_1 which is replaced by M or $-M$ with equal probability.

Thus, for a given stabilizer state, it is efficient to track the evolution of the generators of the input state as Clifford operations and Pauli measurements act on it. Hence, any stabilizer circuit with stabilizer input can be efficiently simulated. This result is known as the Gottesman-Knill theorem [272].

4.2 Review of prior art

4.2.1 Resource theory of magic: odd d-dimensional case

For odd d-dimensional systems, it has been shown that quantum states and channels that have a positive Wigner representation cannot provide any quantum advantage. That is, circuits made from operations with positive Wigner representation can be classically simulated when the input states also have a positive Wigner representation. These states and operations cannot also be used to distill such magic states that can promote fault-tolerant quantum computation to universal quantum computation [90, 97]. Moreover, the set of states and operations that have a positive Wigner representation strictly contain the set of stabilizer states and operations. Thus, quantum states and operations that can provide any quantum advantage or from which useful magic states can be distilled become a resource for the purpose of quantum computation, and the states and operations that cannot provide any advantage are identified as free. Further, it has also been shown that odd d-dimensional states with negative Wigner representation are also contextual, hence establishing contextuality as the physical phenomena responsible for quantum advantage for the odd d-dimensional case [249]. With this identification of free and resource states and operations, static and dynamic resource theories of magic have been developed [90, 97, 247, 253, 276]. To quantify magic in both static and dynamic resource theories, the negativity of the Wigner representation has been used. The negativity of Wigner representation has been used to formulate bounds on various resource interconversion tasks including magic

distillation rates. Besides, it has also been connected to the runtime of classical simulation algorithms, thus giving it an operational interpretation [90, 97].

4.2.2 Resource theory of magic: multi-qubit case

Wigner representation is very successful in classifying free and resource elements in the resource theory of magic for the odd case as is clear from the previous subsection. However, for the practically important case of multi-qubit systems, no such function has been defined that clearly classifies the set of states or operations that cannot provide any quantum advantage while also containing the set of stabilizer states and operations. If the approach that is used for odd-dimensional qudits is applied, we get a Wigner function that is positive for some non-stabilizer states and negative for some stabilizer states. The existing approaches for which a Wigner function is well-defined do not cover the full set of stabilizer states and operations [251, 254, 255]. Other approaches using Wigner distributions have the problem of efficiently simulating the evolution under Clifford gates [246, 286, 287]. Recently, another multi-qubit phase-space representation was proposed which covered the full set of stabilizer operations and states [288]. This representation lacks a nice property that the class of positively represented states should be closed under tensor products. So, to date, we don't have a representation that clearly classifies the classically simulable states and operations while containing the set of stabilizer states and operations.

To retain all multi-qubit stabilizer states and operations as free, alternative approaches need to be taken. Howard and Campbell came up with an approach where they decomposed a density matrix as a real linear combination of pure stabilizer state projectors [253]. Such a decomposition is not unique. This decomposition can be viewed as a quasi-probability distribution where the ℓ_1 norm of non-stabilizer states is strictly greater than 1. ℓ_1 norm is the sum of the absolute values of the coefficients used in the decomposition. They introduced a monotone called the robustness of magic which behaves monotonically under stabilizer operations and showed its resource-theoretic applications. The robustness of magic was defined as the minimum over the ℓ_1 norm obtained for each decomposition of a state. Besides, they devised a classical simulation algorithm for stabilizer circuits with arbitrary input, and connected its runtime with the robustness of input magic states, thus giving robustness of magic an operational interpretation. Recently, in 2020, three other monotones using the same idea were proposed in [271] by varying the decomposition of the magic state. Further, Seddon and Campbell in [258] used the same idea and defined a monotone called channel robustness. To do that, they introduced and characterized a new set of operations called completely stabilizer preserving operations (CSPOs). To find the robustness of any arbitrary channel, they decomposed the channel using CSPOs. They proved that the set of CSPOs strictly contains the set of

stabilizer operations. Moreover, they gave channel robustness an operational interpretation by connecting it with the runtime of a classical simulation algorithm and showed that any circuit made from CSPOs with stabilizer states as inputs can also be classically simulated. Using this property, we partition the set of channels as free and resources, and develop a resource theory of multi-qubit magic channels.

4.3 Free elements of resource theory of multi-qubit magic operations

4.3.1 Completely stabilizer preserving operations (CSPO)

The set of completely stabilizer preserving operations or CSPOs was introduced in [258] and comprises of all the quantum operations that preserve stabilizer states in a complete sense. The set of completely stabilizer preserving operations taking system A_0 to system A_1 will be denoted by $\mathrm{CSPO}(A_0 \to A_1)$ or $\mathrm{CSPO}(A)$. Let $\mathcal{E}_A \in \mathrm{CPTP}(A)$. Then $\mathcal{E}_A$ is a completely stabilizer preserving operation if for any system R_0 it holds that

$$\mathcal{E}_A(\rho_{R_0A_0}) \in \mathrm{STAB}(R_0A_1) \;\; \forall\, \rho_{R_0A_0} \in \mathrm{STAB}(R_0A_0)\,. \tag{4.6}$$

These operations can alternatively be defined using their Choi matrices as follows

$$\mathcal{E}_A \in \mathrm{CSPO}(A) \iff \frac{J_A^{\mathcal{E}}}{|A_0|} \in \mathrm{STAB}(A)\,. \tag{4.7}$$

In [258], it was also shown that the action of CSPOs on a stabilizer state can be efficiently simulated classically. This set is the largest known set of operations in the multi-qubit scenario that do not provide any quantum advantage and as such they are perfect candidates for the free channels of a dynamical resource theory of magic. To manipulate quantum channels, we choose the two natural sets of superchannels – namely, the set of CSPO preserving superchannels and the set of completely CSPO preserving superchannels – as the set of free superchannels in our work. We will denote the set of CSPO preserving superchannels taking dynamical system A to dynamical system B by $\mathfrak{F}_1(A \to B)$ and the set of completely CSPO preserving superchannels taking dynamical system A to dynamical system B by $\mathfrak{F}_2(A \to B)$. In the following two subsections we define and characterize the two sets of free superchannels.

4.3.2 CSPO preserving superchannels

Definition 4.1. Given two dynamical systems A and B, a superchannel $\Theta \in \mathfrak{S}(A \to B)$ is said to be CSPO preserving superchannel if

$$\Theta_{A\to B}[\mathcal{N}_A] \in \text{CSPO}(B) \;\; \forall\, \mathcal{N}_A \in \text{CSPO}(A)\,. \tag{4.8}$$

Let $\{W_j\}$ be the set of stabilizer witnesses for system B_0B_1. Then, using the above definition and the set of stabilizer witnesses, we can characterize the set of CSPO preserving superchannels using their Choi matrices as follows. The Choi matrix of a superchannel $\Theta \in \mathfrak{F}_1(A \to B)$ must satisfy the following conditions

$$\mathbf{J}^{\Theta}_{AB} \geqslant 0\,, \tag{4.9}$$

$$\mathbf{J}^{\Theta}_{AB_0} = \mathbf{J}^{\Theta}_{A_0B_0} \otimes \frac{I_{A_1}}{|A_1|}\,, \tag{4.10}$$

$$\mathbf{J}^{\Theta}_{A_1B_0} = I_{A_1B_0}\,, \tag{4.11}$$

$$\text{Tr}\left[\mathbf{J}^{\Theta}_{AB}(\phi_i \otimes W_j)\right] \geqslant 0 \quad \forall\, \phi_i \in \text{STAB}(A_0A_1), W_j\,. \tag{4.12}$$

In the above, the first three conditions follow from the requirement of Θ to be a superchannel [119]. The condition in equation (4.12) simply uses the fact that if a CSPO preserving superchannel takes the extreme points of the stabilizer polytope to a stabilizer state, then it will also take any convex combination of them to a stabilizer state. However, finding all stabilizer witnesses is a hard problem, but for small dimensions, they can be found and the above characterization can be used as a set of conditions in resource interconversion tasks formulated as conic optimization problems.

4.3.3 Completely CSPO preserving superchannels

Definition 4.2. Given two dynamical systems A and B, a superchannel $\Theta \in \mathfrak{S}(A \to B)$ is said to be completely CSPO preserving if

$$\Theta_{A\to B}[\mathcal{N}_{AR}] \in \text{CSPO}(BR) \quad \forall\, \mathcal{N} \in \text{CSPO}(AR) \tag{4.13}$$

In other words, a superchannel is completely CSPO preserving if, for every input CSPO, the output is also CSPO, even if the superchannel acts only on a subsystem of the input channel.

Theorem 4.3. *Let $\Theta \in \mathfrak{S}(A \to B)$. Then $\Theta \in \mathfrak{F}_2(A \to B)$ if and only if*

$$\frac{1}{|A_1B_0|}\mathbf{J}^{\Theta}_{AB} \in \text{STAB}(AB) \tag{4.14}$$

Proof. We first prove that if Θ is a completely CSPO preserving superchannel (i.e., belongs to $\mathfrak{F}_2(A \to B)$), then its normalized Choi matrix is a stabilizer state. For the other side, we show that if a superchannel Θ is not a completely CSPO preserving superchannel, then its normalized Choi matrix is not a stabilizer state.

Let $\Theta \in \mathfrak{S}(A \to B)$ be a completely CSPO preserving superchannel. By definition, a superchannel can be realized using a pre-processing channel $\mathcal{E} \in \text{CPTP}(B_0 \to E_1 A_0)$ and a post-processing channel $\mathcal{F} \in \text{CPTP}(E_1 A_1 \to B_1)$ [119]. The normalized Choi matrix of the superchannel can be expressed in terms of these pre- and post-processing channels in the following way:

$$\frac{1}{|A_1 B_0|}\mathbf{J}^{\Theta}_{AB} = \mathsf{id}_{A_1 B_0} \otimes (\mathsf{id}_{A_0} \otimes \mathcal{F}_{E_1 A_1 \to B_1}) \circ (\mathsf{id}_{A_1} \otimes \mathcal{E}_{B_0 \to A_0 E_1}) \left(\phi^{+}_{A_1 \tilde{A}_1} \otimes \phi^{+}_{B_0 \tilde{B}_0}\right) , \tag{4.15}$$

where $\phi^{+}_{A_1 \tilde{A}_1}$ ($\phi^{+}_{B_0 \tilde{B}_0}$) represents the maximally entangled state in the system $A_1 \tilde{A}_1$ ($B_0 \tilde{B}_0$). Eq.(4.15) can be diagrammatically illustrated using Fig.4.2.

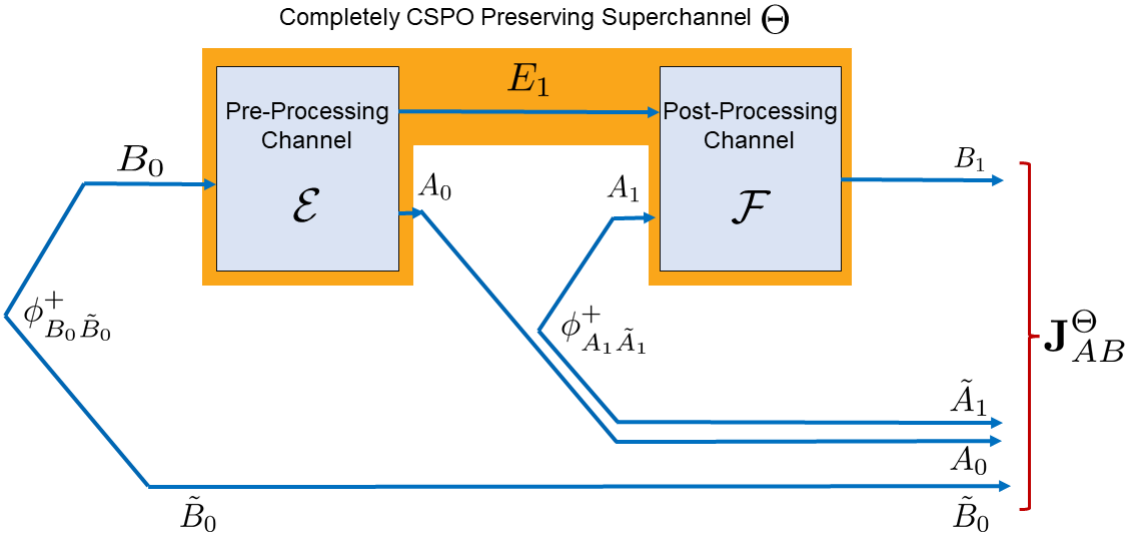

Figure 4.2: Normalized Choi matrix of a superchannel

Define $\mathcal{N} \in \text{CPTP}(A_0 \to \tilde{A}_0 A_1 \tilde{A}_1)$ such that

$$\mathcal{N}(\rho_{A_0}) := \rho_{A_0} \otimes \phi^{+}_{A_1 \tilde{A}_1} , \tag{4.16}$$

for any input density matrix in A_0. Note that the normalized Choi matrix of $\mathcal{N}$ is a stabilizer state. Therefore, $\mathcal{N}$ is a completely stabilizer preserving operation [258]. Using such a channel we can view the Choi matrix of a superchannel as shown in Fig. 4.3.

Since Θ is a completely CSPO preserving superchannel, and $\mathcal{N}$ is a CSPO as defined in Eq.(4.16), the output channel $\Theta[\mathcal{N}]$ is a CSPO and so, $\Theta[\mathcal{N}](\phi^{+}_{B_0 \tilde{B}_0})$ is a stabilizer state.

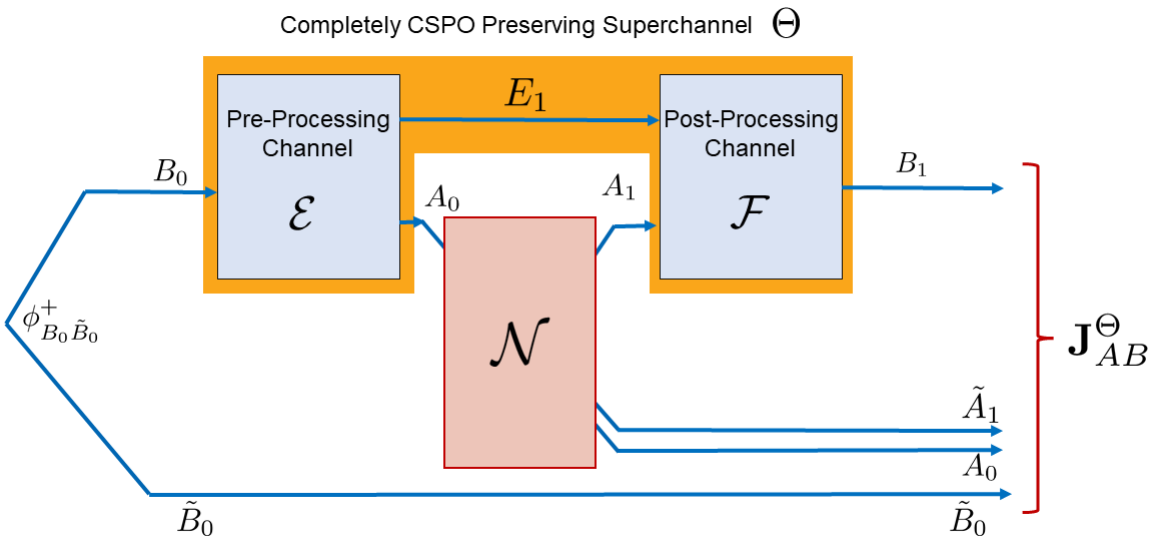

Figure 4.3: Choi matrix of a completely CSPO preserving superchannel viewed as a CSPO

Hence, the normalized Choi matrix of a completely CSPO preserving superchannel is a stabilizer state.

For the other side of the proof, let $\Theta \in \mathfrak{S}(A \to B)$ be a superchannel that is not completely CSPO preserving. Then there exists a CPTP map $\mathcal{E} \in \text{CSPO}(A_0R_0 \to A_1R_1)$ such that $\Theta_{A\to B}[\mathcal{E}_{AR}] \notin \text{CSPO}(B_0R_0 \to B_1R_1)$. Therefore, for some stabilizer witness $W_{B\tilde{R}}$, it holds that

$$\text{Tr}\left[W_{B\tilde{R}}\text{Tr}_{AR}\left[\frac{|A_0|}{|B_0|}\mathbf{J}^{\mathbb{1}\otimes\Theta}_{ARB\tilde{R}}\left(\frac{J^{\mathcal{E}}_{AR}}{|A_0R_0|}\otimes I_{B\tilde{R}}\right)\right]\right] < 0\,. \tag{4.17}$$

After some algebraic manipulations, the above inequality reduces to

$$\text{Tr}\left[\left(\frac{J^{\mathcal{E}}_{AR}}{|A_0R_0|}\otimes W_{B\tilde{R}}\right)\frac{|A_0|}{|B_0|}\mathbf{J}^{\mathbb{1}\otimes\Theta}_{ARB\tilde{R}}\right] < 0\,. \tag{4.18}$$

Since the normalized Choi matrix of $\mathcal{E}$ is a stabilizer state, the following inequality

$$\text{Tr}\left[(|\phi\rangle_{AR}\langle\phi|\otimes W_{B\tilde{R}})\frac{1}{|A_1B_0R_0R_1|}\mathbf{J}^{\mathbb{1}\otimes\Theta}_{ARB\tilde{R}}\right] < 0 \tag{4.19}$$

must hold for some pure stabilizer state $|\phi\rangle_{AR}$. From [258] we know that $(|\phi\rangle_{AR}\langle\phi|\otimes W_{B\tilde{R}})$ is a valid stabilizer witness. Hence,

$$\frac{1}{|A_1B_0R_0R_1|}\mathbf{J}^{\mathbb{1}\otimes\Theta}_{ARB\tilde{R}} \notin \text{STAB}(ARB\tilde{R}) \tag{4.20}$$

which is equivalent to

$$\frac{1}{|A_1B_0R_0R_1|}\mathbf{J}^{\mathbb{1}}_{R\tilde{R}} \otimes \mathbf{J}^{\Theta}_{AB} \notin \mathrm{STAB}(ARB\tilde{R}) \tag{4.21}$$

and that implies

$$\frac{1}{|A_1B_0|}\mathbf{J}^{\Theta}_{AB} \notin \mathrm{STAB}(AB)\ . \tag{4.22}$$

Therefore, we can conclude that the normalized Choi matrix of a superchannel is a stabilizer state if and only if the superchannel is completely CSPO preserving. ∎

4.4 Magic measures

In this section, we quantify magic states and channels. We extend the generalized robustness and the min-relative entropy magic measures from the state to the channel domain [257, 271]. These quantifiers arise from the standard resource theoretic techniques and are related to the channel divergences which have been studied recently in detail in [119, 145, 146, 147, 295, 296, 297, 298]. Next, we formally define the geometric measure for magic states which to the best of our knowledge has not been defined earlier. We couldn't find any operational use of this monotone and leave it as an open problem. Note that, we will denote the (free) robustness of magic as R, the generalized robustness of magic as R_g, the min relative entropy of magic states as $D^{\mathrm{STAB}}_{\min}$, the hypothesis testing relative entropy of magic states as $D^{\epsilon,\,\mathrm{STAB}}_{\min}$, and the min relative entropy of magic channels as $D^{\mathrm{CSPO}}_{\min}$. For completeness, we have briefly discussed robustness of magic and hypothesis testing relative entropy of magic states in Appendix B.3 and B.4, respectively.

4.4.1 Generalized robustness of dynamical magic resources

The generalized robustness for magic states was defined in [271]. Below we generalize it for the channel case and define the log of generalized robustness for a magic channel $\mathcal{N}_A \in \mathrm{CPTP}(A_0 \to A_1)$ as

$$\mathrm{LR}_g(\mathcal{N}_A) = \min_{\mathcal{E}\in\mathrm{CSPO}(A_0\to A_1)} D_{\max}(\mathcal{N}_A\|\mathcal{E}_A) \tag{4.23}$$

$$= \log\min\{\lambda : \lambda\,\mathcal{E} \geqslant \mathcal{N};\ \mathcal{E} \in \mathrm{CSPO}(A_0 \to A_1)\} \tag{4.24}$$

This optimization problem can be expressed in terms of Choi matrices as

$$\begin{aligned} \mathrm{LR}_g(\mathcal{N}_A) = \log \min \lambda \\ \text{s.t.} : \lambda\, J_A^{\mathcal{E}} \geqslant J_A^{\mathcal{N}}\ , \\ J_{A_0}^{\mathcal{E}} = I_{A_0}\ , \\ \frac{J_A^{\mathcal{E}}}{|A_0|} \in \mathrm{STAB}(A_0 A_1) \end{aligned} \tag{4.25}$$

which can be simplified as

$$\begin{aligned} \mathrm{LR}_g(\mathcal{N}_A) = \log \min \frac{\mathrm{Tr}[\omega_A]}{|A_0|} \\ \text{s.t.} : \omega_A \geqslant J_A^{\mathcal{N}}\ , \\ \omega_{A_0} = \mathrm{Tr}[\omega_A]\frac{I_{A_0}}{|A_0|}\ , \\ \frac{\omega_A}{\mathrm{Tr}[\omega]} \in \mathrm{STAB}(A_0 A_1)\,. \end{aligned} \tag{4.26}$$

The dual of the above primal problem can be written as

$$\begin{aligned} \mathrm{LR}_g(\mathcal{N}_A) = \log \sup \mathrm{Tr}[\alpha_A J_A^{\mathcal{N}}] \\ \text{s.t.} : \mathrm{Tr}\left[\phi_i \left(\alpha_A + \beta_{A_0} \otimes I_{A_1} - \mathrm{Tr}[\beta_{A_0}]\frac{I_A}{|A_0|}\right)\right] \leqslant \frac{1}{|A_0|}\ \ \forall\ \phi_i \in \mathrm{STAB}(A_0 A_1), \\ \alpha_A \geqslant 0,\ \beta_{A_0} \in \mathrm{Herm}(A_0)\,. \end{aligned} \tag{4.27}$$

Some properties of the generalized robustness of magic channels are listed below.

1. *Faithfulness.* $\mathrm{LR}_g(\mathcal{N}_A) = 0 \iff \mathcal{N} \in \mathrm{CSPO}(A_0 \to A_1)$. The proof is similar to the state case.

2. *Monotonicity.* $\mathrm{LR}_g(\Theta[\mathcal{N}]) \leqslant \mathrm{LR}_g(\mathcal{N})$ for any free superchannel $\Theta \in \mathfrak{F}_1(A \to B)$ or $\Theta \in \mathfrak{F}_2(A \to B)$. The proof follows from the data-processing inequality as

$$\begin{aligned} \mathrm{LR}_g(\Theta[\mathcal{N}]) &= \min_{\mathcal{F} \in \mathrm{CSPO}(B)} D_{\max}(\Theta[\mathcal{N}] \| \mathcal{F}) \\ &\leqslant \min_{\mathcal{E} \in \mathrm{CSPO}(A)} D_{\max}(\Theta[\mathcal{N}] \| \Theta[\mathcal{E}]) \\ &\leqslant \min_{\mathcal{E} \in \mathrm{CSPO}(A)} D_{\max}(\mathcal{N} \| \mathcal{E})\,. \end{aligned} \tag{4.28}$$

3. *Sub-additivity.* $\mathrm{LR}_g(\mathcal{N} \otimes \mathcal{M}) \leqslant \mathrm{LR}_g(\mathcal{N}) + \mathrm{LR}_g(\mathcal{M})$. The proof easily follows from equation (4.23).

Remark 8. Eq.(4.24) can be rewritten (without the log) as

$$\mathrm{R}_g(\mathcal{N}_A) = \min\Big\{\lambda \geqslant 1 : \frac{\mathcal{N} + (\lambda-1)\mathcal{M}}{\lambda} \in \mathrm{CSPO}(A_0 \to A_1)\,, \mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)\Big\}\,. \tag{4.29}$$

Hence, for any $\lambda \geqslant \mathrm{R}_g(\mathcal{N}_A)$, a channel $\mathcal{N}_A$ can then be expressed as

$$\mathcal{N}_A = \lambda\mathcal{E} - (\lambda-1)\mathcal{M} \tag{4.30}$$

for some $\mathcal{E} \in \mathrm{CSPO}(A_0 \to A_1)$ and some $\mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$.

4.4.2 Min-relative entropy of magic resources

Below, we present another monotone, the min relative entropy of magic states and channels. The min-relative entropy of a magic state ρ is defined as

$$D_{\min}^{\mathrm{STAB}}(\rho) \coloneqq \min_{\sigma\in\mathrm{STAB}} D_{\min}(\rho\|\sigma) \tag{4.31}$$

$$= \min_{\sigma\in\mathrm{STAB}} \left(-\log_2 \mathrm{Tr}[P_\rho\sigma]\right) \tag{4.32}$$

$$= -\log_2 \max\ \mathrm{Tr}[P_\rho\sigma] \tag{4.33}$$

$$\text{s.t.}: \sigma \in \mathrm{STAB}\,,$$

$$= -\log_2 \max\ \mathrm{Tr}[P_\rho\phi] \tag{4.34}$$

$$\text{s.t.}: \phi \in \mathrm{STAB}\,.$$

where P_ρ denotes the projection onto the support of ρ. Similarly, the min-relative entropy of a magic channel $\mathcal{N}$ can be defined as

$$D_{\min}^{\mathrm{CSPO}}(\mathcal{N}_A) \coloneqq \min_{\mathcal{E}\in\mathrm{CSPO}(A)} D_{\min}\left(\mathcal{N}\|\mathcal{E}\right) \tag{4.35}$$

$$= \min_{\mathcal{E}\in\mathrm{CSPO}(A)} \sup_{\psi\in\mathfrak{D}(R_0A_0)} D_{\min}\left(\mathcal{N}(\psi)\|\mathcal{E}\left(\psi\right)\right)\,. \tag{4.36}$$

Below we list some of the properties of the min-relative entropy of magic states and channels.

1. *Faithfulness.* The min-relative entropy of magic states and channels is faithful, i.e.,

$$D_{\min}^{\mathrm{CSPO}}(\mathcal{N}_A) = 0 \iff \mathcal{N} \in \mathrm{CSPO}(A_0 \to A_1) \tag{4.37}$$

$$D_{\min}^{\mathrm{STAB}}(\rho_{A_0}) = 0 \iff \rho \in \mathrm{STAB}(A_0)\,. \tag{4.38}$$

2. *Monotonicity.* The min-relative entropy is a magic monotone under CSPOs for the state case, and under CSPO preserving and completely CSPO preserving superchannels for the channel case. Thus, for any state $\rho \in \mathfrak{D}(A_0)$ it follows that $D_{\min}^{\mathrm{STAB}}(\mathcal{E}(\rho)) \leqslant D_{\min}^{\mathrm{STAB}}(\rho)$ for any $\mathcal{E} \in \mathrm{CSPO}$, and for any channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$, it follows that $D_{\min}^{\mathrm{CSPO}}(\Theta[\mathcal{N}]) \leqslant D_{\min}^{\mathrm{CSPO}}(\mathcal{N})$ for any $\Theta \in \mathfrak{F}_1(A \to B)$ or $\Theta \in \mathfrak{F}_2(A \to B)$. The proof for the state case is given below which follows from the data-processing inequality as

$$\begin{aligned} D_{\min}^{\mathrm{STAB}}(\mathcal{E}(\rho)) &= \min_{\sigma \in \mathrm{STAB}} D_{\min}(\mathcal{E}(\rho)\|\sigma) \\ &\leqslant \min_{\sigma \in \mathrm{STAB}} D_{\min}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) \\ &\leqslant \min_{\sigma \in \mathrm{STAB}} D_{\min}(\rho\|\sigma)\,. \end{aligned} \tag{4.39}$$

   Proof for the channel case follows similarly.

3. *Sub-additivity.* Sub-additivity holds for min-relative entropies of both static and dynamic magic resources, i.e., $D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) \leqslant D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2)$ for any two density matrices ρ_1 and ρ_2, and $D_{\min}^{\mathrm{CSPO}}(\mathcal{N} \otimes \mathcal{M}) \leqslant D_{\min}^{\mathrm{CSPO}}(\mathcal{N}) + D_{\min}^{\mathrm{CSPO}}(\mathcal{M})$ for any two quantum channels $\mathcal{N}$ and $\mathcal{M}$. Moreover, for single qubit states, the min-relative entropy of magic states is additive, i.e., $D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) = D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2)$ [257]. The proof of this is provided in Appendix B.2.

4.4.3 Geometric magic measure for static resources

In this subsection, we formally define the geometric magic measure for states which to the best of our knowledge has not been defined before. Inspired from the geometric measure of entanglement [299], we define the geometric magic measure for pure states as

$$g(\psi) = 1 - \max_{\phi \in \mathrm{STAB}} \mathrm{Tr}[\psi\phi] \tag{4.40}$$

For general mixed states, we can extend the above measure using fidelity as

$$g(\rho) = 1 - \max_{\sigma \in \mathrm{STAB}} F^2(\rho, \sigma) \tag{4.41}$$

where $F(\rho,\sigma) := \mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]$ is the fidelity between two states ρ and σ. Below we list the properties of this measure:

1. *Faithfulness:* $g(\rho) = 0$ if and only if $\rho \in \mathrm{STAB}$.

2. *Monotonicity:* $g(\mathcal{E}(\rho)) \leqslant g(\rho)\ \forall\ \mathcal{E} \in \mathrm{CSPO}$. The proof is similar to the proof of monotonicity of geometric measures in [82].

3. *Subadditivity:* $g(\rho_1 \otimes \rho_2) \leqslant g(\rho_1) + g(\rho_2)$. This follows easily if we let σ_1 and σ_2 be the respective optimal stabilizer states such that $g(\rho_1) = 1 - F^2(\rho_1,\sigma_1)$ and $g(\rho_2) = 1 - F^2(\rho_2,\sigma_2)$. Then

$$\max_{\sigma\in\mathrm{STAB}} F(\rho_1\otimes\rho_2,\sigma) = \max_{\sigma} \mathrm{Tr}\left[\sqrt{\sqrt{\sigma}(\rho_1\otimes\rho_2)\sqrt{\sigma}}\right] \tag{4.42}$$

$$\geqslant \mathrm{Tr}\left[\sqrt{(\sqrt{\sigma_1}\rho_1\sqrt{\sigma_1})\otimes(\sqrt{\sigma_2}\rho_2\sqrt{\sigma_2})}\right] \tag{4.43}$$

$$= F(\rho_1,\sigma_1)F(\rho_2,\sigma_2) \tag{4.44}$$

   where the inequality follows by choosing $\sigma = \sigma_1 \otimes \sigma_2$.

4.5 Interconversions

Resource interconversion is one of the central themes of resource theory. In this section, we discuss the conditions for qubit interconversions under CSPOs in 4.5.1, and the conversion of magic states to channels and vice-versa under CSPO preserving and completely CSPO preserving superchannels in 4.5.2. We also formulated the interconversion distance which is given in Appendix B.1.

4.5.1 Qubit interconversion under CSPOs

For the resource theory of magic, any pure magic state can be used as a resource to achieve universal quantum computation [236]. The procedure involves distilling a pure magic state from a given magic state and then using a few copies of this pure magic state to perform any quantum computation. Experimentally, it of interest to distill single qubit magic states, and the common choices are that of the $|T\rangle$ state or the $|H\rangle$ state where:

$$|T\rangle\langle T| = \frac{1}{2}\left(I + (X+Y)/\sqrt{2}\right)\ , \tag{4.45}$$

$$|H\rangle\langle H| = \frac{1}{2}\left(I + (X+Y+Z)/\sqrt{3}\right)\ . \tag{4.46}$$

Here, we are interested in a more general problem of finding whether a given single qubit magic state can be converted to another by repeated application of CSPOs. Equivalently, we want to find out which set of states on the Bloch sphere can be reached by restricting ourselves to the application of CSPOs on a single qubit magic state. For multiqubit systems, this problem is an NP-hard problem because the number of stabilizer states increases super-exponentially as we increase the dimension. For the qubit case, we use geometry to our advantage and provide the following theorem for the conversion of a state ρ into a state σ. We show that this interconversion problem can be cast as a linear programming feasibility problem. For the purpose of this theorem, let us define $C(\rho) := \{U\rho U^\dagger : U \in \text{Clifford}\}$ as the set of Clifford equivalent states of ρ. We show in the proof of the theorem below that for a single qubit state ρ, the set $C(\rho)$ contains 24 elements unless the state has additional symmetry, in which case the number of elements are less than 24. For instance, $C(|0\rangle\langle 0|)$ contains only 6 elements which are all the pure single-qubit stabilizer states.

Theorem 4.4. *Let A be a (4×30) matrix with first 24 columns being the Bloch vectors of the elements of $C(\rho)$, the next 6 columns being the Bloch vectors of the pure qubit stabilizer states, and the last row being all ones. Let $\mathbf{b}$ be the (4×1) vector with the first three entries being the Bloch vector corresponding to the state σ and the fourth entry being equal to 1. Then, the state ρ can be converted to the state σ using CSPOs if there exists an $\mathbf{x} \in \mathbb{R}_+^{30}$ such that $A\mathbf{x} = \mathbf{b}$.*

Remark 9. The problem of finding $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geqslant 0$ is known as an SDP feasibility problem and can be solved using standard techniques in convex analysis [173, 174]. It also has a dual given by the Farkas lemma. Using the dual of the above feasibility problem, we can say that the state ρ cannot be converted to σ if there exists a $\mathbf{y} \in \mathbb{R}^3$ such that $A^{\mathrm{T}}\mathbf{y} \geqslant 0$ and $\mathbf{b} \cdot \mathbf{y} < 0$.

Proof. From [258] and Eq. (4.7), we know that the normalized Choi matrix of any CSPO is a stabilizer state. Let $\mathcal{E}_{A_0 \to A_1} \in \text{CSPO}(A)$ such that both A_0 and A_1 are single qubit systems. If we denote a pure two qubit maximally entangled stabilizer state as ψ^{ent} and a single qubit stabilizer state as ϕ, we can write the action of $\mathcal{E}_A$ on any input $\rho \in \mathfrak{D}(A_0)$ as

$$\mathcal{E}(\rho_{A_0}) = \mathrm{Tr}_{A_0}\left[J_A^{\mathcal{E}}\left(\rho_{A_0} \otimes I_{A_1}\right)\right] \tag{4.47}$$

$$= |A_0|\left(\sum_i p_i \mathrm{Tr}_{A_0}\left[\psi_i^{ent}\left(\rho_{A_0} \otimes I_{A_1}\right)\right] + \sum_{j,k} p_{j,k} \mathrm{Tr}_{A_0}\left[\left(\left(\phi_j\right)_{A_0} \otimes \left(\phi_k\right)_{A_1}\right)\left(\rho_{A_0} \otimes I_{A_1}\right)\right]\right) \tag{4.48}$$

$$= \sum_i p_i \mathcal{U}_i(\rho_{A_0}) + |A_0| \sum_{j,k} p_{j,k} \mathrm{Tr}\left[\left(\phi_j\right)_{A_0} \rho_{A_0}\right]\left(\phi_k\right)_{A_1} \tag{4.49}$$

$$= \sum_i p_i \mathcal{U}_i(\rho_{A_0}) + \sum_k q_k \phi_k\,, \tag{4.50}$$

where $q_k = |A_0| \sum_j p_{j,k} \mathrm{Tr}\left[\phi_j \rho\right]$. In the above, the second equality follows because any two-qubit stabilizer state can be expressed as a convex combination of pure two-qubit entangled and pure two-qubit separable stabilizer states. From the above equations, we see that the action of a (qubit input and output) CSPO on a qubit can be represented as a convex combination of the action of completely stabilizer preserving unitary operations and stabilizer replacement channels. (An alternative proof can also be found in [300]). Note that for two-qubit states, there are a total of 60 pure stabilizer states of which only 24 are entangled [280]. Hence there are only 24 single-qubit unitary gates that are completely stabilizer preserving. These unitary gates are listed in Appendix B.6 and are Clifford unitaries. Therefore, any state can be transformed to at the most 24 states (including itself) on the Bloch sphere using these unitary gates. For a single qubit state, which can be expressed as a vector $(r_1, r_2, r_3)^T$ in the Bloch sphere, its transformations using these unitary gates are given in Appendix B.6. Furthermore, if we view the Bloch sphere as been divided into 8 octants according to $(\pm X, \pm Y, \pm Z)$ and each octant to be further subdivided into three subsets such that for one subset it holds that $|\langle X \rangle| \leqslant |\langle Y \rangle|, |\langle Z \rangle|$, for second subset it holds that $|\langle Y \rangle| \leqslant |\langle X \rangle|, |\langle Z \rangle|$, and for the third subset we have $|\langle Z \rangle| \leqslant |\langle X \rangle|, |\langle Y \rangle|$, then using table B.2 in Appendix B.6, we can say that any arbitrary state in some subset (of an octant) is Clifford equivalent to a state in any other subset. Therefore, we can conclude from the equations and the arguments above that the set of states that can be generated from a given state under the action of CSPOs must belong to a convex polytope in the Bloch sphere, the extreme points of which are the Clifford-equivalent states of the given state and the stabilizer states. Further, if we let $\{\mathbf{r}_i\}$ denote the set of Bloch vectors of the Clifford equivalent state of ρ, $\{\mathbf{s}_k\}$ denote the Bloch vectors of the pure single qubit stabilizer states, and $\mathbf{b}$ as the Bloch vector of $\mathcal{E}(\rho)$, then from Eq (4.50), we can write the Bloch vector $\mathbf{b}$ as

$$\mathbf{b} = \sum_i p_i \mathbf{r}_i + \sum_k q_k \mathbf{s}_k \,. \tag{4.51}$$

We can now express the above in the form of the equation $A\mathbf{x} = \mathbf{b}$, where the matrix A is a (3×30) matrix consisting of $\mathbf{r}_i$'s and $\mathbf{s}_k$'s as column vectors, and $\mathbf{x}$ is the (30×1) vector consisting of non-negative numbers summing to one. We can include this last condition on $\mathbf{x}$ by inserting a $(1, 1, \ldots, 1)$ row in A thus making A, a (4×30) matrix. Therefore, we can now say that a state ρ can be converted to a state σ with Bloch vector $\mathbf{b}$ if there exists a vector $\mathbf{x} \in \mathbb{R}^{30}$ such that

$$A\mathbf{x} = \mathbf{b}, \text{ and} \tag{4.52}$$

$$\mathbf{x} \geqslant 0 \,. \tag{4.53}$$

Geometrical interpretation of Theorem 4.4

The above interconversion conditions can be also be expressed and visualized on a Bloch sphere. To find whether a qubit can be converted to another using CSPOs, from Eq. (4.50) we get that it is enough to check whether the target state (or any of its Clifford equivalent state) lies outside the facets of the convex polytope (generated by the original state) that together cover any subset of any octant. For convenience, let us choose this subset to be the positive octant $(+X, +Y, +Z)$ for which $|\langle X\rangle| \leqslant |\langle Y\rangle|, |\langle Z\rangle|$ and denote it by P_X. Hence, it is enough to find only those extreme points of the convex polytope which are used to form the facets that together cover P_X. Using the hyperplane separation theorem, we can then find whether the target state lies inside this convex polytope. Now, let the Bloch vector corresponding to ρ (or its Clifford equivalent state) belonging to P_X be denoted by $\mathbf{r_1} = (r_x, r_y, r_z)$. We denote the neighbouring Clifford equivalent states which are used to form the facets of the convex polytopes as

$$\begin{aligned}
\mathbf{r_2} &= (r_z, r_x, r_y)\,, \\
\mathbf{r_3} &= (r_y, r_z, r_x)\,, \\
\mathbf{r_4} &= (-r_x, r_z, r_y)\,, \\
\mathbf{r_5} &= (-r_y, r_x, r_z)\,, \\
\mathbf{r_6} &= (r_y, -r_x, r_z)\,, \\
\mathbf{r_7} &= (0, 0, 1)\,, \\
\mathbf{r_8} &= (0, 1, 0)\,, \\
\mathbf{r_9} &= (-r_z, r_y, r_x)\,, \\
\mathbf{r_{10}} &= (r_z, r_y, -r_x)\,.
\end{aligned} \tag{4.54}$$

Now depending on the location of $\mathbf{r_1}$ in P_X, there are three possible ways to form a convex polytope. Since we are only interested in the facets of these polytopes that cover P_X, we list below the set of vectors which, for each possible polytope, form a facet partially covering P_X :

Possibility 1: $(\mathbf{r_1}, \mathbf{r_6}, \mathbf{r_7}), (\mathbf{r_1}, \mathbf{r_7}, \mathbf{r_5}), (\mathbf{r_1}, \mathbf{r_5}, \mathbf{r_4}), (\mathbf{r_1}, \mathbf{r_4}, \mathbf{r_3}), (\mathbf{r_1}, \mathbf{r_3}, \mathbf{r_2}), (\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_6}), (\mathbf{r_3}, \mathbf{r_4}, \mathbf{r_8})$

Possibility 2: $(\mathbf{r_1}, \mathbf{r_3}, \mathbf{r_2}), (\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_7}), (\mathbf{r_1}, \mathbf{r_7}, \mathbf{r_4}), (\mathbf{r_1}, \mathbf{r_4}, \mathbf{r_8}), (\mathbf{r_1}, \mathbf{r_8}, \mathbf{r_3})$

Possibility 3: $(\mathbf{r_1}, \mathbf{r_{10}}, \mathbf{r_3}), (\mathbf{r_1}, \mathbf{r_3}, \mathbf{r_2}), (\mathbf{r_1}, \mathbf{r_2}, \mathbf{r_4}), (\mathbf{r_1}, \mathbf{r_4}, \mathbf{r_9}), (\mathbf{r_1}, \mathbf{r_9}, \mathbf{r_8}), (\mathbf{r_1}, \mathbf{r_8}, \mathbf{r_{10}}), (\mathbf{r_4}, \mathbf{r_2}, \mathbf{r_7})$

In figures 4.4 and 4.5, we have marked the location of the points in possibility 1 and possibility 2, respectively, highlighted (with red arcs) the subset they belong to, and connected the points in the way they are connected

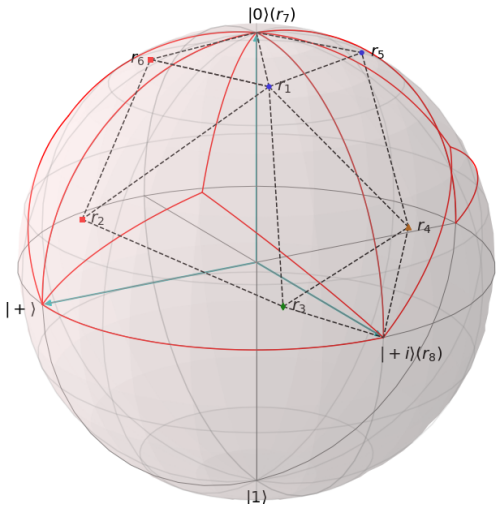

Figure 4.4: Points corresponding to possibility 1

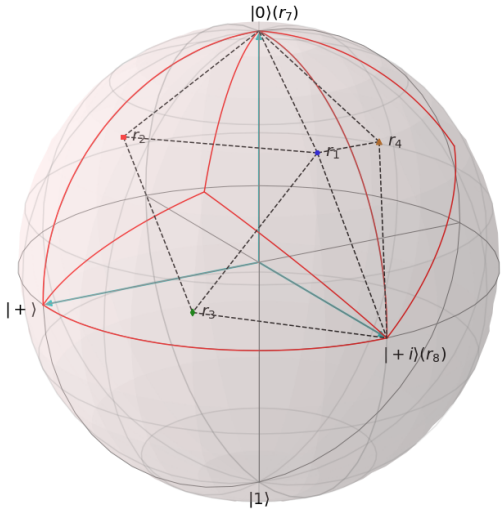

Figure 4.5: Points corresponding to possibility 2

in the convex polytope for a particular possibility.

Using these set of vectors for each possible convex polytope, it is straightforward to find the vector (say $\mathbf{v}$) perpendicular to each facet such that the inner product of $\mathbf{v}$ with all vectors lying inside that facet is less than or equal to the inner product of $\mathbf{v}$ with one of the vectors on the surface of the facet. Let's call this inner product as v. All the vectors on the other side of this facet will then give a value more than v when their inner product is calculated with $\mathbf{v}$. Therefore, by finding all such vectors perpendicular to each facet, we find the conditions to verify whether a vector lies inside or outside the facets. Hence, a state ρ can be converted to a state σ using completely stabilizer preserving operations if and only if

$$\begin{aligned} \mathbf{s}\cdot\mathbf{u_i} &\leqslant u_i,\ \forall\, i=1,\ldots,7 \quad \text{or} \\ \mathbf{s}\cdot\mathbf{v_j} &\leqslant v_j,\ \forall\, j=1,\ldots,5 \quad \text{or} \\ \mathbf{s}\cdot\mathbf{w_k} &\leqslant w_k,\ \forall\, k=1,\ldots,7 \end{aligned} \tag{4.55}$$

where $\mathbf{s}$ corresponds to the Bloch vector of the Clifford equivalent state of σ in P_X. The vectors $\mathbf{u_i}$'s, $\mathbf{v_j}$'s, and $\mathbf{w_k}$'s are the vectors perpendicular to the facets of the respective possible polytopes, and u_i's, v_j's, and w_k's are the constants which can be calculated from the inner product of $\mathbf{u}_i$, $\mathbf{v}_j$, and $\mathbf{w}_k$ with any vector lying on the surfaces of the respective facets of the possible polytopes.

Remark 10. The code for the above interconversion has been uploaded in a public git repository and can be freely accessed using the link in [1]. In the same link, we have also provided a code to construct a convex polytope from a given state. The code can also be used to construct convex polytopes for various states at

[1] https://github.com/gaurav-iiser/Resource-Theory-of-multiqubit-magic-channels

the same time, and hence can be used to check whether a convex polytope corresponding to some state lies inside another convex polytope or not.

4.5.2 Cost and Distillation bounds under CSPO preserving and completely CSPO preserving superchannels

In this subsection, we find bounds on the cost of converting a magic state to a multi-qubit magic channel and the bounds on distilling magic states from a quantum channel using both CSPO preserving and completely CSPO preserving superchannels. For the case of distillation, we focus on distilling pure single qubit magic states because a pure magic state is enough for achieving universality in the magic state model of quantum computation. Besides, due to the complexity involved in verifying whether a state is a stabilizer state, we leave the problem of finding the upper bound of cost and lower bound of distillation using completely CSPO preserving superchannels as open.

Since any pure magic state can be used as a resource to perform universal quantum computation, we define the dynamical magic cost of converting a pure magic state to a channel $\mathcal{N} \in \mathrm{CPTP}(B_0 \to B_1)$ under CSPO preserving superchannels or completely CSPO preserving superchannels as

$$\mathrm{COST}_{\mathfrak{F}_{1(2)}}(\mathcal{N}_B) = \min \log\{|A_1| : \Theta[\psi_{A_1}] = \mathcal{N}_B,\, \psi \in \mathfrak{D}(A_1),\, \Theta \in \mathfrak{F}_{1(2)}(A_1 \to B)\}\,. \tag{4.56}$$

If we want the cost of simulating a channel in terms of a particular magic state $\psi \in \mathfrak{D}(A_1)$, we define cost as

$$\mathrm{COST}_{\psi,\,\mathfrak{F}_{1(2)}}(\mathcal{N}_B) = \min\big\{n : \Theta[\psi^n] = \mathcal{N}_B, \\ \Theta \in \mathfrak{F}_{1(2)}(A_1 \to B)\big\}\,. \tag{4.57}$$

Distillation of a pure single qubit magic state ψ from a channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ using CSPO preserving or completely CSPO preserving superchannels is defined as

$$\mathrm{DISTILL}^{\epsilon}_{\psi,\mathfrak{F}_{1(2)}}(\mathcal{N}_A) = \max\{n : F\left(\Theta[\mathcal{N}], \psi^n\right) \geqslant 1-\epsilon,\, \Theta \in \mathfrak{F}_{1(2)}(A \to B_1)\}\,. \tag{4.58}$$

Proposition 3. $\mathrm{COST}_{\mathfrak{F}_1}(\mathcal{N}) \leqslant \log(|A_1|)$ if for some system A_1, we have

$$\max_{\psi \in \mathfrak{D}(A_1)} D^{\mathrm{STAB}}_{\min}(\psi_{A_1}) \geqslant \mathrm{LR}(\mathcal{N}_B) \tag{4.59}$$

where $\mathrm{LR}(\mathcal{N}_B)$ is the log of the robustness of $\mathcal{N}_B$. If ψ is a given single qubit magic state, then it follows

that

$$\mathrm{COST}_{\psi,\mathfrak{F}_1}(\mathcal{N}) \leqslant \left\lceil \frac{\mathrm{LR}(\mathcal{N})}{D_{\min}^{\mathrm{STAB}}(\psi)} \right\rceil . \tag{4.60}$$

Proof. Let for some $\psi \in \mathfrak{D}(A_1)$, the following is satisfied

$$D_{\min}^{\mathrm{STAB}}(\psi_{A_1}) \geqslant \mathrm{LR}(\mathcal{N}_B) . \tag{4.61}$$

Now consider the following superchannel $\Theta \in \mathfrak{S}(A_1 \to B)$ whose action on any input state $\rho \in \mathfrak{D}(A_1)$ is given as

$$\Theta[\rho] \coloneqq \mathrm{Tr}[\psi\rho]\mathcal{N} + (1 - \mathrm{Tr}[\psi\rho])\mathcal{M} , \tag{4.62}$$

where $\mathcal{M}$ is the optimal CSPO chosen from the definition of the channel robustness, $\mathrm{R}(\mathcal{N})$. It is easy to verify that $\Theta[\psi] = \mathcal{N}$. From Eq. (4.61), we also get that

$$-\log \mathrm{Tr}[\psi\sigma] \geqslant \log(1 + \mathrm{R}(\mathcal{N})) \quad \forall \sigma \in \mathrm{STAB}(A_1) . \tag{4.63}$$

Hence, for any $\sigma \in \mathrm{STAB}(A_1)$, it holds that $\mathrm{Tr}[\psi\sigma] \leqslant \frac{1}{1+\mathrm{R}(\mathcal{N})}$, implying that $\Theta \in \mathfrak{F}_1(A_1 \to B)$. Thus, the cost of converting a pure magic state to a magic channel $\mathcal{N}_B$ using CSPO preserving superchannels is no greater than $\log(|A_1|)$ if $\max_{\psi\in\mathfrak{D}(A_1)} D_{\min}^{\mathrm{STAB}}(\psi_{A_1}) \geqslant \mathrm{LR}(\mathcal{N})$.

Further, if ψ is a given single qubit pure magic state, then using the additivity of min-relative entropy of single qubit magic states, we get

$$\mathrm{COST}_{\psi,\mathfrak{F}_1}(\mathcal{N}_B) \leqslant \left\lceil \frac{\mathrm{LR}(\mathcal{N})}{D_{\min}^{\mathrm{STAB}}(\psi)} \right\rceil . \tag{4.64}$$

∎

Remark 11. We numerically verify that the bound in Eq. (4.60) is not trivial. As an example, we use the $|T\rangle$ state ($D_{\min}^{\mathrm{STAB}}(|T\rangle\langle T|) = 0.2284$) to calculate the upper bound of cost of creating some magic states. We present the comparison of the upper bound of our results of cost with the lower bound obtained in [253] as a table (see table 4.2). Note that in [253] the free operations were stabilizer operations. In the table, a general resource state $|U\rangle = U|+\rangle$ where $|+\rangle$ is the maximally coherent state and U is some unitary gate. Also, some special states include the $|H\rangle$ state which is the single-qubit state with Bloch vector $(1,1,1)/\sqrt{3}$ and has robustness $\sqrt{3}$, $|\chi\rangle$ state is the two-qubit state with maximum robustness of $\sqrt{5}$ for two-qubit states, and $|\mathrm{Hoggar}\rangle$ state is the three-qubit state which maximizes robustness for three-qubit states and has robustness 3.8.

State	upper bound from our work	lower bound from [253]
$\|H\rangle$	2	2
$\|CS_{1,2}\rangle$	3	3
$\|T_{1,2,3}\rangle$	4	3
$\|\chi\rangle$	4	4
$\|CCZ\rangle$	4	4
$\|CS_{12,13}\rangle$	4	4
$\|T_1CS_{2,3}\rangle$	5	4
$\|T_1CS_{12,13}\rangle$	5	5
$\|\text{Hoggar}\rangle$	6	6

Table 4.2: Comparison of magic costs

Remark 12. We would like to emphasize here that we provide a general result for the case of channels by giving a precise formula to find the upper bound on the cost that depends on the log-robustness of the magic channel and the min-relative entropy of the single-qubit magic state.

Proposition 4. The cost of converting a pure magic state $\psi_{A_1} \in \mathfrak{D}(A_1)$ to a target channel $\mathcal{N}_B \in \text{CPTP}(B_0 \to B_1)$ using CSPO preserving or completely CSPO preserving superchannels is lower bounded by

$$\frac{\text{LR}_g(\mathcal{N}_B)}{\text{LR}_g(\psi_{A_1})} \leqslant \text{COST}_{\psi,\mathfrak{F}_{1(2)}}(\mathcal{N}_B)\,. \tag{4.65}$$

Proof. The proof follows from the standard resource theoretic methods and can be seen as a special case of theorem 1 of [301] together with the sub-additivity of generalized robustness of magic resources. ∎

Proposition 5. Given a channel $\mathcal{N} \in \text{CPTP}(A_0 \to A_1)$ and a single qubit state ψ, the following holds

$$\text{DISTILL}_{\psi,\mathfrak{F}_{1(2)}}(\mathcal{N}_A) \leqslant \frac{D_{\min}^{\text{CSPO}}(\mathcal{N}_A)}{D_{\min}^{\text{STAB}}(\psi)}\,. \tag{4.66}$$

Proof. The proof of the above proposition also follows from standard resource theoretic methods [301, 302] and the additivity of min-relative entropy of single-qubit magic states. For completeness, we provide the proof in Appendix B.5. ∎

Proposition 6. The lower bound on distilling a single qubit pure magic state ψ from a channel $\mathcal{N} \in \text{CPTP}(A_0 \to A_1)$ using a CSPO preserving superchannel is given by

$$\text{DISTILL}^{\epsilon}_{\psi,\mathfrak{F}_1}(\mathcal{N}_A) \geqslant \left\lfloor \frac{D_{\min}^{\epsilon,\,\text{STAB}}(\tilde{J}_A^{\mathcal{N}})}{\text{LR}(\psi)} \right\rfloor\,, \tag{4.67}$$

where $\tilde{J}_A^{\mathcal{N}}$ is the normalized Choi matrix of the channel $\mathcal{N}$, and $D_{\min}^{\epsilon,\,\text{STAB}}(\cdot)$ represents the hypothesis testing relative entropy of magic states which we have defined in Appendix B.4.

Proof. Let n be the largest non-negative integer such that $D_{\min}^{\epsilon,\mathrm{STAB}}(\tilde{J}_A^{\mathcal{N}}) \geqslant n\mathrm{LR}(\psi)$. Then, we can construct the following superchannel $\Theta \in \mathfrak{S}(A \to B_1)$ such that for any input channel $\mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$

$$\Theta[\mathcal{M}] \coloneqq \mathrm{Tr}[\tilde{J}_A^{\mathcal{M}} E]\psi^n + (1 - \mathrm{Tr}[\tilde{J}_A^{\mathcal{M}} E])\sigma\,, \tag{4.68}$$

where $\sigma \in \mathrm{STAB}(B_1)$ is chosen from the definition of $R(\psi^n)$, and E is the optimal POVM element chosen in the definition of hypothesis testing relative entropy of magic states, $D_{\min}^{\epsilon,\mathrm{STAB}}(\tilde{J}^{\mathcal{N}})$. We first notice that for such a superchannel

$$F(\Theta[\mathcal{N}], \psi^n) \geqslant \mathrm{Tr}[\Theta[\mathcal{N}]\psi^n] \tag{4.69}$$

$$\geqslant \mathrm{Tr}[\tilde{J}^{\mathcal{N}} E] \tag{4.70}$$

$$\geqslant 1 - \epsilon \tag{4.71}$$

where the last inequality comes from the fact that E is optimal in $D_{\min}^{\epsilon,\mathrm{STAB}}(\tilde{J}^{\mathcal{N}})$.

Since $D_{\min}^{\epsilon,\mathrm{STAB}}(\tilde{J}_A^{\mathcal{N}}) \geqslant n\mathrm{LR}(\psi)$, we get

$$-\log \mathrm{Tr}[E\sigma] \geqslant \log(1 + \mathrm{R}(\psi))^n \geqslant \log(1 + \mathrm{R}(\psi^n)) \tag{4.72}$$

for all $\sigma \in \mathrm{STAB}(A_0A_1)$. Therefore, if the input $\mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$ is a CSPO, then $-\log \mathrm{Tr}[E\tilde{J}_A^{\mathcal{M}}] \geqslant \log(1 + \mathrm{R}(\psi^n))$ which implies that

$$\mathrm{Tr}[E\tilde{J}_A^{\mathcal{M}}] \leqslant \frac{1}{1 + \mathrm{R}(\psi^n)}\,.$$

Hence, Θ is a CSPO preserving superchannel. Thus, we can distill atleast n copies of the single qubit state ψ from the channel $\mathcal{N}$ where n satisfies $D_{\min}^{\epsilon,\mathrm{STAB}}(\tilde{J}_A^{\mathcal{N}}) \geqslant n\mathrm{LR}(\psi)$. ∎

4.6 Classical simulation algorithm for circuits

The goal of a classical simulation algorithm is to estimate Born rule probabilities or to find the expectation value of an observable. To this purpose, a class of algorithms, known as the quasiprobability simulation techniques, have recently been developed that make use of the quasiprobability decomposition of magic states or channels [245, 250, 253, 258, 271]. The runtime of these algorithms has been shown to be of the order of the square of the robustness [253, 258], or the square of another similar monotone, the dyadic negativity [271]. In [271], another simulation technique, the constrained path simulator for states was introduced with the idea to reduce the runtime of the simulation. This simulation technique offers constant

runtime by compromising with the precision in estimating the expected value. Before discussing the dynamic constrained path simulator, we give a brief sketch of the quasiprobability simulator for completeness as part of our algorithm also depends on it.

The quasiprobability simulator for multi-qubit systems was first presented in [253] and two other variations of this quasiprobability simulators were presented in [271]. The runtimes of all these quasiprobability simulators depend on the product of the squares of the robustness of input magic states. The idea of the quasiprobability simulator for estimating the expectation value, P_ρ, of an observable P given a stabilizer operation $\mathcal{E}$ and an arbitrary input state ρ, as given in [253], is as follows. The input magic state is first decomposed as a linear combination of pure stabilizer states. Let ρ be the input state decomposed as $\rho = \sum_i x_i \phi_i$, where ϕ_i's are pure stabilizer states and x_i's are quasiprobabilities. These quasiprobabilities can be used to form a probability distribution $p_i = |x_i|/\sum_i |x_i|$. Now using this probability distribution, an i value is sampled and Gottesman-Knill theorem is used to obtain an eigenvalue $m = \pm 1$ for the measurement of P on $\mathcal{E}(\phi_i)$. With this eigenvalue, we obtain another value $M = \text{sign}(x_i) m \sum_i |x_i|$ where $\text{sign}(x_i)$ is equal to 1 for $x_i > 0$ and -1, otherwise. By repeating this sampling process many times, the mean value of M can be found which is an "unbiased estimator" of P_ρ. Then, for random variables bounded in the interval $[-\sum_i |x_i|, +\sum_i |x_i|]$, Hoeffding inequalities are used to show that N samples will estimate the mean to within δ of the actual mean with probability exceeding $1-\epsilon$ where $\epsilon = 2\exp[-N\delta^2/2(\sum_i |x_i|)^2]$. Put simply, the number of samples required are of the order of $(\sum_i |x_i|)^2$. This quantity is the least when we used the most optimal decomposition and is then known as the robustness of magic. Thus, the number of samples and the runtime scales quadratically in the robustness. Now, using the same idea the expectation value of an observable given an arbitrary circuit can also be estimated. In this case, the runtime similarly depends on the channel robustness of each channel used in the circuit [258]. The question then is whether the runtime of the algorithm can be reduced any further. To do that, we allowed for some error in the expectation value, and modify the static constrained path algorithm presented in [271] for the general case of a circuit composed of a sequence of channels acting on an initial stabilizer state and ending with a measurement of some Pauli observable. We modify the static constrained path simulator algorithm such that we achieve the estimate with a precision more than or equal to a desired precision. With this modification, the runtime of the algorithm is not a constant but depends on the desired precision (or the desired error). For any non-zero error, the runtime never exceeds that of a quasiprobability simulator for channels. Moreover, if there is no bound on the error/precision, the algorithm achieves a constant runtime.

The overall idea of the constrained path simulator for states is as follows. A magic state $\rho \in \mathfrak{D}(A_1)$ can be decomposed as $\rho = t\sigma_+ - (t-1)\rho_-$ for some $t \geqslant 1$, $\sigma_+ \in \text{STAB}(A_1)$, and $\rho_- \in \mathfrak{D}(A_1)$. The constrained path simulator for states works by constraining the quasiprobability decomposition of a state to the positive

part, i.e., by making the approximation $\rho \approx t\sigma_+$. Then, the algorithm estimates $t\text{Tr}[E\mathcal{O}(\sigma_+)]$ upto ϵ error using a Clifford simulator such as the quasiprobability simulator. Here, E is some Pauli observable, and $\mathcal{O}$ is a CSPO. This estimate is then used to obtain the expectation value $\text{Tr}[E\mathcal{O}(\rho)]$ and the estimation error. The runtime of the algorithm is decided by the Clifford simulator used. By defining ϵ as the product of a constant c and t, the algorithm was shown to have a constant runtime.

Constrained path simulator for channels. Let $\mathcal{N}$ be a circuit composed of a sequence of n channels and let the i^{th} circuit element be denoted by $\mathcal{N}_i$. As mentioned previously in remark 8, the circuit element $\mathcal{N}_i$ can be decomposed using some CSPO $\mathcal{E}_i$ and some other channel $\mathcal{M}_i$ such that $\mathcal{N}_i = \lambda_i\mathcal{E}_i - (\lambda_i - 1)\mathcal{M}_i$ where λ_i is the generalized robustness of $\mathcal{N}_i$. Then, for the whole circuit we can write

$$\begin{aligned}\mathcal{N} = \mathcal{N}_n \circ \cdots \circ \mathcal{N}_1 &= (\lambda_n \cdots \lambda_1)(\mathcal{E}_n \circ \cdots \circ \mathcal{E}_1) + \ldots + ((\lambda_n - 1)\cdots(\lambda_1 - 1))\mathcal{M}_n \circ \cdots \circ \mathcal{M}_1 \\ &= (\lambda_n \cdots \lambda_1)(\mathcal{E}_n \circ \cdots \circ \mathcal{E}_1) + ((\lambda_n \cdots \lambda_1) - 1)\mathcal{M} \\ &= \lambda\mathcal{E} + (\lambda - 1)\mathcal{M}\end{aligned} \tag{4.73}$$

where $\lambda = \lambda_n \cdots \lambda_1$, $\mathcal{E} = \mathcal{E}_n \circ \cdots \circ \mathcal{E}_1$ and $\mathcal{M}$ follows from simple arithmetic manipulation of the first equation and is the probabilistic combination of the sequence of channels where each sequence contains atleast one $\mathcal{M}_i$. The aim of the algorithm is to estimate $\text{Tr}[E\mathcal{N}(|0\rangle\langle 0|)]$ with a precision more than or equal to some target precision and a runtime less than what can be achieved by a quasiprobability simulator.

The algorithm starts by replacing the original circuit $\mathcal{N}$ with another circuit $\mathcal{N}'$ to achieve the mean estimate up to some target error Δ^*. The algorithm first replaces the channel $\mathcal{N}_j$ with $\lambda_j\mathcal{E}_j$ if λ_j, the generalized robustness of $\mathcal{N}_j$, is less than some fixed real number λ^*. Here, $\mathcal{E}_j$ is the optimal CSPO such that $\lambda_j\mathcal{E}_j \geqslant \mathcal{N}_j$. The choice of λ^* ensures that the estimation error never exceeds the target allowed error. Then, using the static Monte Carlo routine introduced in [258] for circuits, the algorithm estimates $\lambda'\text{Tr}[E\mathcal{N}'(|0\rangle\langle 0|)]$ up to ϵ error where λ' is the product of the generalized robustnesses of the replaced channels and the error ϵ equals a constant c multiplied with λ'. Next, using ϵ, λ', and the estimate we obtained above, the algorithm outputs the estimate of the expectation value $\text{Tr}[E\mathcal{N}(|0\rangle\langle 0|)]$ up to error $\Delta \leqslant \Delta^*$ following some trivial steps.

In the static Monte Carlo routine, the runtime of the algorithm is decided by finding the total number N of steps required to achieve the mean estimate up to an additive error ϵ with success probability $1 - p_{\text{fail}}$. The number of steps N that the static Monte Carlo takes is given by

$$N = \left\lceil 2\epsilon^{-2}\|q\|_1^2 \log(2p_{\text{fail}}^{-1})\right\rceil \tag{4.74}$$

where $\|q\|_1 = \prod_j \mathrm{R}(\mathcal{N}_j)$ and $\mathrm{R}(\mathcal{N}_j)$ is the robustness of the circuit element $\mathcal{N}_j$ as defined in [258]. In our hybrid algorithm, since we choose to keep some channels and replace some with CSPOs, the number of steps to estimate $\lambda \mathrm{Tr}[E\mathcal{N}'(|0\rangle\langle 0|)]$ upto ϵ error with success probability $1 - p_{\text{fail}}$ is given by

$$N = \left\lceil 2\epsilon^{-2}\lambda^2 \prod_{j:\lambda_j > \lambda^*} \mathrm{R}(\mathcal{N}_j)^2 \log(2p_{\text{fail}}^{-1}) \right\rceil \tag{4.75}$$

$$= \left\lceil 2c^{-2} \prod_{j:\lambda_j > \lambda^*} \mathrm{R}(\mathcal{N}_j)^2 \log(2p_{\text{fail}}^{-1}) \right\rceil \tag{4.76}$$

where c is a pre-defined small constant. In this sense, the number of steps only depend on the robustness of the channels whose $\lambda_i > \lambda^*$. Note that if all the channels are selected by the algorithm, we essentially have the runtime as that of static Monte Carlo routine. If all the channels are replaced in the initial steps then we get a constant runtime.

Algorithm Dynamic constrained path simulator

Input: (i) Sequence of channels $\mathcal{N}_1, \ldots, \mathcal{N}_n$ such that the target channel $\mathcal{N} = \mathcal{N}_n \circ \cdots \circ \mathcal{N}_1$. (ii) Real numbers $0 < c, p_{\text{fail}} << 1$ and Pauli observable E. (iii) Desired error Δ^*.
Pre-Computation: (i) $\lambda^* = (\Delta^* + 1)^{1/n}$. (ii) For each circuit element, an optimal decomposition in terms of CSPOs is determined.
Output: (i) Born rule probability estimate $\hat{E}$. (ii) Error Δ such that, $|\hat{E} - \mathrm{Tr}\left[E\,\mathcal{N}(|0\rangle\langle 0|)\right]| \leqslant \Delta$, and $\Delta \leqslant \Delta^*$.

for $i \leftarrow 1$ to n **do**
  $\lambda_i \leftarrow \Lambda^+(\mathcal{N}_i)$, and denote the optimal free channel by $\mathcal{E}_i$.
  if $\lambda_i \leqslant \lambda^*$: **then**
    $\mathcal{N}_i \leftarrow \lambda_i \mathcal{E}_i$
  end if
end for
$\mathcal{N}' \leftarrow \left(\prod_{j:\lambda_j \leqslant \lambda^*} \lambda_j\right)(\mathcal{F}_n \circ \cdots \circ \mathcal{F}_1)$, where $\mathcal{N}'$ denotes the new circuit that will be used to find the estimate and $\mathcal{F}_k$'s denote the circuit elements given be

$$\mathcal{F}_k = \begin{cases} \mathcal{E}_k & \text{if } \lambda_k \leqslant \lambda^* \\ \mathcal{N}_k & \text{otherwise} \end{cases} \tag{4.77}$$

$\epsilon \leftarrow c\lambda$ where $\lambda = \prod_{j:\lambda_j \leqslant \lambda^*} \lambda_j$
Let $E_{\mathcal{N}'}$ be an estimate of $\lambda \mathrm{Tr}[E\mathcal{N}'(|0\rangle\langle 0|)]$ upto ϵ error and success probability $1 - p_{\text{fail}}$.
$E_{\max} \leftarrow \min\{1, E_{\mathcal{N}'} + \epsilon + \lambda - 1\}$
$E_{\min} \leftarrow \max\{-1, E_{\mathcal{N}'} - \epsilon - \lambda + 1\}$
$\hat{E} \leftarrow (E_{\max} + E_{\min})/2$
$\Delta \leftarrow (E_{\max} - E_{\min})/2$

Analysis

As with the constrained path simulator for states, the choice of $E_{\max}$ and $E_{\min}$ ensure that for all λ and

$E_{\mathcal{N}'}$, the following inequality holds with probability $1 - p_{\text{fail}}$

$$|\hat{E} - \text{Tr}[E\mathcal{N}(|0\rangle\langle 0|)]| \leqslant \Delta \tag{4.78}$$

To justify the choice of λ^*, let λ^* be the generalized robustness of each channel used in the circuit and λ^* times the optimal CSPO for each channel is considered in the above routine. Then, for any Δ we have

$$\Delta \leqslant \lambda(1 + c) - 1 \tag{4.79}$$

and hence we require $\lambda(1 + c) - 1 \leqslant \Delta^*$. Assuming there are n channels in the circuit, we get

$$\lambda^* \leqslant \left(\frac{\Delta^* + 1}{1 + c}\right)^{1/n} \tag{4.80}$$

$$\approx (\Delta^* + 1)^{1/n} \tag{4.81}$$

Since this is the worst-case analysis, in practical scenarios we will have $\lambda \leqslant \lambda^*$(equality only arising when the circuit consists of just one channel applied n times), and therefore $\Delta \leqslant \Delta^*$.

4.7 Conclusion

In this work, we developed the dynamical multi-qubit resource theory of magic channels by identifying the completely stabilizer preserving operations (CSPOs) as the set of free operations. CSPOs are a perfect candidate for the free channels of a resource theory of magic channels because they form the largest known set of operations that cannot provide any quantum advantage. In previous resource theoretic studies of magic channels, the superchannel approach was only taken in [90] where the authors considered the odd-dimensional qudit case and the free channels were the completely positive Wigner preserving operations (CPWPO). There, the free superchannels were chosen to be the ones that completely preserve the set of CPWPO. In this paper, we defined and characterized two sets of free superchannels - namely, the CSPO preserving superchannels and the completely CSPO preserving superchannels. We characterized completely CSPO preserving superchannels in terms of their Choi matrices, and in particular, we showed that a superchannel is completely CSPO preserving if and only if its normalized Choi matrix is a stabilizer state. We then defined monotones for states and channels which include the generalized robustness of magic channels, the min-relative entropy of magic channels, and the geometric magic measure for states. We also addressed some resource interconversion problems, specifically proving that the qubit interconversion under CSPOs

can be solved with simple linear programming. We then determined a closed formula for the upper and lower bound on both the cost of simulating a channel from a qubit and distilling a qubit magic state from a channel, under CSPO preserving superchannels. We also formulated the lower bound on the qubit cost of simulating a magic channel, and the upper bound on distilling a pure qubit magic state from a magic channel under completely CSPO preserving superchannels using the standard resource theoretic techniques. Finally, we gave a classical simulation algorithm to find expectation values given a general quantum circuit. The algorithm works by selecting and replacing some circuit elements with some CSPO, based on a parameter that depends on the minimum target precision required. Hence, due to this selective replacement algorithm, the runtime of our algorithm also depends on the precision required. If the precision required is too tight, then the runtime reaches that of the static Monte Carlo simulation algorithm given in [258], whereas, if there is no bound on the precision, the algorithm has a constant runtime and can be seen as a generalization of the constrained path simulator introduced in [271] for states. These classical simulation algorithms help benchmark the quantum computational speedup and there is a lot left to explore in the general circuit case. Apart from that, it would be interesting to explore non-deterministic transformations and catalytic transformations under CSPO preserving and completely CSPO preserving superchannels. Lastly, because of the difficulty in verifying whether a state is a stabilizer or not, we were unable to find lower bounds on distilling magic states using completely CSPO preserving operations and leave it as an open problem.

Chapter 5

Lorenz majorization among quantum channels

5.1 Introduction

One of the primary objectives of both classical and quantum information theory is to find conditions under which interconversion among resources is possible. For instance, interconversion among probability distributions using doubly stochastic matrices is characterized by majorization. Majorization is a pre-order on vectors which finds its applications in variety of subjects including quantum information (see [113, 114, 303, 304, 305, 306] and references within). Another pre-order known as relative majorization generalizes the concept of majorization between vectors to a pair of vectors. While majorization concerns interconvertibility among vectors, relative majorization concerns interconvertibility among pairs of vectors [307, 308]. A pair of vectors is said to relatively majorize another pair of vectors if there exists a column stochastic matrix that transforms the former pair to the latter. Relative majorization generalizes the concept of majorization between vectors when either the first or the second vector of both the pairs is the uniform distribution.

Interestingly, relative majorization between pairs of probability distributions has an elegant geometrical characterization in terms of symmetric convex regions called testing regions [98]. By symmetric, we mean that if (x, y) belongs to the testing region, then so does $(1-x, 1-y)$. In 1953, Blackwell established that a pair of probability distributions can be transformed to another pair of distributions if and only if the testing region of the former pair contains that of the latter [98]. These conditions can also be expressed using the upper/lower boundary of the testing region known as the upper/lower Lorenz curves. Then, if the testing

region of one pair contains that of the other, we say that the first pair Lorenz majorizes the other.

Lorenz curves were first introduced in 1905 by Max Lorenz to compare the distribution of income and wealth within an economy [309]. Since then, these curves have been used for comparing distributions in various fields because of the simple geometrical characterization that they offer [82, 98, 310, 311, 312]. After Blackwell established the interconversion conditions for pairs of probability distributions using Lorenz curves in 1953, Alberti and Uhlmann in 1980, extended these results from pairs of probability distributions to pairs of qubits [313]. To gain more insights in quantum states' interconversion and to draw analogies between classical probability theory and quantum theory, the idea of Lorenz curves and testing region defined for pairs of probability distributions was generalized to quantum states recently in 2017 [308, 314, 315]. These generalizations were used as a tool to show the existence of a transformation from a pair of states to another in various resource theories like athermality, thermodynamics, and magic [306, 307, 308, 315, 316, 317, 318].

In this work, we tackle interconversion problems among pairs of channels and find conditions using Lorenz curves. To this purpose, we generalize the concept of Lorenz majorization, quantum testing region, and quantum relative majorization to the channel domain in Sec. 5.3. We also generalize the definition of Hilbert α divergences to quantum channels and find equivalent conditions of Lorenz majorization in its terms. Using this generalization, we find conditions for converting a pair of probability distributions to a pair of classical channels in terms of Lorenz curves. We show equivalence in Lorenz and relative majorization for interconversion among pairs of classical channels when the resultant pair of channels has a two-dimensional output. These interconversion conditions are provided in Sec. 5.4. In the end, we conclude by presenting some open problems.

5.2 Background

5.2.1 Testing region and Lorenz majorization for pairs of states

The testing region $\mathcal{T}(\rho_{A_0}, \sigma_{A_0})$ of a pair of states ρ, $\sigma \in \mathfrak{D}(A_0)$ is defined as the set of achievable points

$$(x, y) = \left(\mathrm{Tr}[\Lambda\sigma], \mathrm{Tr}[\Lambda\rho]\right) \tag{5.1}$$

where $0 \leqslant \Lambda \leqslant I_{A_0}$ is a POVM element. The quantum Lorenz curve is defined to be the upper boundary of $\mathcal{T}(\rho, \sigma)$. A pair of states is said to Lorenz majorize another pair if the testing region of the former contains that of the latter.

The lower Lorenz curve can be expressed as an optimization problem which is closely related to the hypothesis testing relative entropy. The hypothesis testing relative entropy (see [319, 320, 321] and references

within) for a pair of states $\rho, \sigma \in \mathcal{D}(A_0)$, is defined for $0 \leqslant \epsilon \leqslant 1$ as follows:

$$D_H^\epsilon(\rho\|\sigma) := -\log Q^\epsilon(\rho\|\sigma) \tag{5.2}$$

where

$$Q^\epsilon(\rho\|\sigma) := \min \ \mathrm{Tr}[\sigma\Lambda] \tag{5.3}$$

$$\text{s.t. } 0 \leqslant \Lambda \leqslant I_{A_0} \tag{5.4}$$

$$\mathrm{Tr}[\rho\Lambda] \geqslant 1-\epsilon\,. \tag{5.5}$$

The function $Q^\epsilon(\rho\|\sigma)$ for all $0 \leqslant \epsilon \leqslant 1$ defines the lower Lorenz curve of the pair (ρ, σ). Since $Q^\epsilon(\rho\|\sigma)$ is a semi-definite program, it is efficiently computed.

To see how the testing region and the upper and lower Lorenz curve looks like, let us take a simple example of diagonal states. Let ρ and σ be diagonal states with entries $(1/3, 1/4, 1/4, 1/6)$ and $(1/12, 1/6, 1/3, 5/12)$ on their diagonals, respectively. Then the lower Lorenz curve will have following extreme points:

$$(0,0), (1/3, 1/12), (7/12, 3/12), (5/6, 7/12), (1,1)$$

which can be easily computed. Using the symmetry of the testing region we can easily find the extreme points of the upper Lorenz curve, and construct the full testing region corresponding to (ρ, σ).

5.2.2 Hilbert α divergence for states

Hilbert α divergence, a family of divergences, was introduced in [308] as a tool to characterization quantum relative majorization for quantum states. Given a pair of quantum states ρ and σ, for all $\alpha \geqslant 1$, the following is defined:

$$\sup_\alpha(\rho/\sigma) := \sup_{\alpha^{-1}I \leqslant \Lambda \leqslant I} \frac{\mathrm{Tr}[\Lambda\rho]}{\mathrm{Tr}[\Lambda\sigma]} \tag{5.6}$$

and the corresponding divergence is defined as:

$$H_\alpha(\rho\|\sigma) := \frac{\alpha}{\alpha-1}\log_2 \sup_\alpha(\rho/\sigma) \tag{5.7}$$

The properties of the Hilbert α divergence of states can be found in [308].

5.3 Quantum relative Lorenz curves and Hilbert α divergences for channels

We can extend the definition of relative majorization from the state to the channel domain as follows

Definition 5.1. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A)$ and $\mathcal{E}, \mathcal{F} \in \mathrm{CPTP}(B)$ be two pairs of quantum channels. We say that $(\mathcal{N}, \mathcal{M})$ relatively majorize $(\mathcal{E}, \mathcal{F})$ and write

$$(\mathcal{N}, \mathcal{M}) \succ (\mathcal{E}, \mathcal{F}) \iff \mathcal{E} = \Theta[\mathcal{N}] \text{ and } \mathcal{F} = \Theta[\mathcal{M}] \tag{5.8}$$

where $\Theta \in \mathfrak{S}(A \to B)$.

Next, we generalize the notion of a testing region for a pair of channels. Using this, we extend the concept of Lorenz majorization to the channel domain.

Definition 5.2. Given two quantum channels $\mathcal{N}$, $\mathcal{M} \in \mathrm{CPTP}(A)$, the associated testing region, denoted $\mathcal{T}(\mathcal{N}, \mathcal{M})$, is defined as the set of achievable points

$$(x, y) = \left(\mathrm{Tr}\left[\Lambda_{RA_1} \mathcal{M}_{A_0 \to A_1}(\psi_{RA_0})\right], \mathrm{Tr}\left[\Lambda_{RA_1} \mathcal{N}_{A_0 \to A_1}(\psi_{RA_0})\right]\right) \tag{5.9}$$

where $0 \leqslant \Lambda_{RA_1} \leqslant I_{RA_1}$, and $\psi \in \mathfrak{D}(RA_0)$. Equivalently, using the Choi matrices $J_A^{\mathcal{N}}$ and $J_A^{\mathcal{M}}$ of the channels $\mathcal{N}$ and $\mathcal{M}$, and process positive operator values measures (PPOVM) we can write the set of achievable points as

$$(x, y) = \left(\mathrm{Tr}\left[\Lambda'_{A_0 A_1} J_A^{\mathcal{M}}\right], \mathrm{Tr}\left[\Lambda'_{A_0 A_1} J_A^{\mathcal{N}}\right]\right) \tag{5.10}$$

where $0 \leqslant \Lambda'_{A_0 A_1} \leqslant \rho_{A_0} \otimes I_{A_1}$ is a process POVM, and $\rho \in \mathfrak{D}(A_0)$. The quantum Lorenz curve is defined to be the upper boundary of $\mathcal{T}(\mathcal{N}, \mathcal{M})$.

Definition 5.3. Let $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A)$ and $\mathcal{E}, \mathcal{F} \in \mathrm{CPTP}(B)$ be two pairs of quantum channels. We say that $(\mathcal{N}, \mathcal{M})$ Lorenz majorizes $(\mathcal{E}, \mathcal{F})$ and write

$$(\mathcal{N}, \mathcal{M}) \succ_L (\mathcal{E}, \mathcal{F}) \iff \mathcal{T}(\mathcal{N}, \mathcal{M}) \supseteq \mathcal{T}(\mathcal{E}, \mathcal{F}) \tag{5.11}$$

The hypothesis testing relative entropy (see [319, 320, 321] and references within) for a pair of states

$\rho, \sigma \in \mathcal{D}(A_0)$, is defined for $0 \leqslant \epsilon \leqslant 1$ as follows:

$$D_H^\epsilon(\rho\|\sigma) := -\log Q^\epsilon(\rho\|\sigma), \tag{5.12}$$

$$Q^\epsilon(\rho\|\sigma) := \min_{\substack{0\leqslant\Lambda\leqslant I_{A_0} \\ \mathrm{Tr}[\rho\Lambda]\geqslant 1-\epsilon}} \mathrm{Tr}[\sigma\Lambda] \tag{5.13}$$

The channel hypothesis testing relative entropy can similarly be generalized using process POVM as

Definition 5.4. The channel hypothesis testing relative entropy can be defined as

$$\begin{aligned} D_H^\epsilon(\mathcal{N}\|\mathcal{M}) &= \sup_\psi D_H^\epsilon(\mathcal{N}(\psi)\|\mathcal{M}(\psi)) && (5.14)\\ &= \sup_\psi (-\log Q^\epsilon(\mathcal{N}(\psi)\|\mathcal{M}(\psi))) && (5.15)\\ &= -\log \inf_\psi Q^\epsilon(\mathcal{N}(\psi)\|\mathcal{M}(\psi)) && (5.16)\\ &= -\log Q^\epsilon(\mathcal{N}\|\mathcal{M}) && (5.17) \end{aligned}$$

and hence we can define $Q^\epsilon(\mathcal{N}\|\mathcal{M})$ as

$$\begin{aligned} Q^\epsilon(\mathcal{N}\|\mathcal{M}) &:= \inf_\psi Q^\epsilon(\mathcal{N}(\psi)\|\mathcal{M}(\psi)) && (5.18)\\ &= \inf \mathrm{Tr}[\mathcal{M}(\psi)\Lambda] && (5.19)\\ &\quad \text{s.t. } 0 \leqslant \Lambda \leqslant I, && (5.20)\\ &\qquad \mathrm{Tr}[\mathcal{N}(\psi)\Lambda] \geqslant 1-\epsilon, && (5.21)\\ &\qquad \psi \in \mathfrak{D}(RA_0) && (5.22)\\ &= \inf\ \mathrm{Tr}[J^{\mathcal{M}}\Lambda'] && (5.23)\\ &\quad \text{s.t. } : 0 \leqslant \Lambda'_{A_0A_1} \leqslant \rho_{A_0} \otimes I_{A_1}, && (5.24)\\ &\qquad \mathrm{Tr}[\rho] = 1, && (5.25)\\ &\qquad \mathrm{Tr}[J^{\mathcal{N}}\Lambda'] \geqslant 1-\epsilon. && (5.26) \end{aligned}$$

where we have used the fact that any bipartite pure state $|\psi\rangle_{RA_0}$ can be written as $|\psi\rangle_{RA_0} = M_\psi \otimes I_{A_0}|\Phi^+\rangle_{\tilde{A}_0A_0}$ using a linear map $M_\psi : A_0 \to R$ and the unnormalized maximally entangled state $|\Phi^+\rangle_{\tilde{A}_0A_0}$, and $M_\psi^\dagger M_\psi = \rho \in \mathfrak{D}(A_0)$.

We can extend the definition of Hilbert α divergences to quantum channels as follows: Given two quantum

channels $\mathcal{N}$, $\mathcal{M} \in \mathrm{CPTP}(A)$, we defined for all $\alpha \geqslant 1$, the following quantity:

$$\begin{aligned}\sup_\alpha(\mathcal{N}/\mathcal{M}) &:= \sup_{\rho \in \mathfrak{D}(R_0A_0)} \sup_\alpha\left(\mathcal{N}_A(\rho_{R_0A_0})/\mathcal{M}_A(\rho_{R_0A_0})\right) \\ &= \sup_{\psi \in \mathfrak{D}(R_0A_0)} \sup_\alpha(\mathcal{N}_A(\psi_{R_0A_0})/\mathcal{M}_A(\psi_{R_0A_0})) \\ &= \sup_{\substack{\psi_{R_0A_0}, \\ \alpha^{-1}I \leqslant \Lambda_{R_0A_1} \leqslant I}} \frac{\mathrm{Tr}[\Lambda_{R_0A_1}\mathcal{N}_A(\psi_{R_0A_0})]}{\mathrm{Tr}[\Lambda_{R_0A_1}\mathcal{M}_A(\psi_{R_0A_0})]}\,.\end{aligned}$$

Using the above quantity, the Hilbert α divergence can be defined as

$$H_\alpha(\mathcal{N}/\mathcal{M}) := \frac{\alpha}{\alpha-1}\log_2 \sup_\alpha(\mathcal{N}/\mathcal{M})\,. \tag{5.27}$$

Theorem 5.5. *Let $\mathcal{N}$, $\mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$. Then*

1. *for all $\alpha \geqslant 1$, $H_\alpha(\mathcal{N}\|\mathcal{M}) \geqslant 0$ with equality if and only if $\mathcal{N} = \mathcal{M}$;*
2. *for all $\alpha \geqslant 1$, the data-processing inequality holds for any superchannel $\Theta \in \mathfrak{S}(A \to B)$, $H_\alpha[\Theta(\mathcal{N})\|\Theta(\mathcal{M})] \leqslant H_\alpha(\mathcal{N}\|\mathcal{M})$;*
3. *$H_\infty(\mathcal{N}\|\mathcal{M}) = \lim_{\alpha\to\infty} H_\alpha(\mathcal{N}\|\mathcal{M}) = D_{\max}(\mathcal{N}\|\mathcal{M})$*
4. *$H_1(\mathcal{N}\|\mathcal{M}) = \lim_{\alpha\to 1} H_\alpha(\mathcal{N}\|\mathcal{M}) = \frac{1}{2\ln(2)}\|\mathcal{N}-\mathcal{M}\|_\diamond$*

Proof of property 1.

$$\begin{aligned}H_\alpha(\mathcal{N}\|\mathcal{M}) &= \frac{\alpha}{\alpha-1}\log_2\sup_\alpha(\mathcal{N}/\mathcal{M}) \\ &= \frac{\alpha}{\alpha-1}\log_2 \sup_{\psi\in\mathfrak{D}(R_0A_0)} \sup_\alpha(\mathcal{N}(\psi)/\mathcal{M}(\psi)) \\ &= \sup_{\psi_{R_0A_0}} H_\alpha(\mathcal{N}(\psi)\|\mathcal{M}(\psi)) \\ &\geqslant 0\end{aligned}$$

where the second equality follows from the definition of $\sup_\alpha(\mathcal{N}/\mathcal{M})$ and the inequality follows from the fact that for states $\rho, \sigma \in \mathfrak{D}(R_0A_1)$, $H_\alpha(\rho\|\sigma) \geqslant 0$ [308]. It was shown in [308] for the state case that $H_\alpha(\rho\|\sigma) = 0$ if and only if $\rho = \sigma$. Therefore, in the channel case, we can conclude that $H_\alpha(\rho\|\sigma) \geqslant 0$ holds if and only if $\mathcal{N} = \mathcal{M}$. ∎

Proof of property 2. Let $\Theta \in \mathfrak{S}(A \to B)$ be realized using some pre-processing channel $\mathcal{E} \in \mathrm{CPTP}(B_0 \to$

E_0A_0) and post-processing channel $\mathcal{F} \in \mathrm{CPTP}(E_0A_1 \to B_1)$. Then for all $\alpha \geqslant 1$ it holds that

$$\begin{aligned}
\sup_\alpha\left(\Theta[\mathcal{N}]\|\Theta[\mathcal{M}]\right) &= \sup_{\psi_{R_0B_0}} \sup_\alpha\left(\Theta[\mathcal{N}](\psi)/\Theta[\mathcal{N}](\psi)\right) \\
&= \sup_{\substack{\psi_{R_0B_0} \\ \alpha^{-1}I \leqslant \Lambda_{R_0B_1} \leqslant I}} \frac{\mathrm{Tr}\left[\Lambda(\Theta[\mathcal{N}](\psi))\right]}{\mathrm{Tr}\left[\Lambda(\Theta[\mathcal{M}](\psi))\right]} \\
&= \sup_{\substack{\psi_{R_0B_0} \\ \alpha^{-1}I \leqslant \Lambda_{R_0B_1} \leqslant I}} \frac{\mathrm{Tr}\left[\Lambda(\mathcal{F} \circ (I \otimes \mathcal{N}) \circ \mathcal{E}(\psi))\right]}{\mathrm{Tr}\left[\Lambda(\mathcal{F} \circ (I \otimes \mathcal{M}) \circ \mathcal{E}(\psi))\right]} \\
&= \sup_{\substack{\psi_{R_0B_0} \\ \alpha^{-1}I \leqslant \Lambda \leqslant I}} \frac{\mathrm{Tr}\left[\mathcal{F}^*(\Lambda)((I \otimes \mathcal{N}) \circ \mathcal{E}(\psi))\right]}{\mathrm{Tr}\left[\mathcal{F}^*(\Lambda)((I \otimes \mathcal{M}) \circ \mathcal{E}(\psi))\right]} \\
&\leqslant \sup_{\substack{\rho \in \mathfrak{D}(R_0E_0A_0) \\ \alpha^{-1}I \leqslant \Lambda'_{R_0E_0A_1} \leqslant I}} \frac{\mathrm{Tr}\left[\Lambda'(\mathcal{N}(\rho))\right]}{\mathrm{Tr}\left[\Lambda'(\mathcal{M}(\rho))\right]} \\
&= \sup_{\substack{\phi \in \mathfrak{D}(R_0E_0A_0) \\ \alpha^{-1}I \leqslant \Lambda'_{R_0E_0A_1} \leqslant I}} \frac{\mathrm{Tr}\left[\Lambda'(\mathcal{N}(\phi))\right]}{\mathrm{Tr}\left[\Lambda'(\mathcal{M}(\phi))\right]} \\
&= \sup_\alpha(\mathcal{N}/\mathcal{M})\,.
\end{aligned}$$

Therefore, $H_\alpha[\Theta(\mathcal{N})\|\Theta(\mathcal{M})] \leqslant H_\alpha(\mathcal{N}\|\mathcal{M})$ for all $\alpha \geqslant 1$. ∎

Proof of property 3 and 4. By taking the limit $\alpha \to \infty$ we get

$$\begin{aligned}
\lim_{\alpha\to\infty} H_\alpha(\mathcal{N}\|\mathcal{M}) &= \lim_{\alpha\to\infty} \frac{\alpha}{\alpha-1} \log_2 \sup_{\psi \in \mathfrak{D}(R_0A_0)} \sup_\alpha(\mathcal{N}(\psi)/\mathcal{M}(\psi)) && (5.28)\\
&= \sup_{\psi \in \mathfrak{D}(R_0A_0)} \log_2 \left(\sup_{0 \leqslant \Lambda \leqslant I} \frac{\mathrm{Tr}[\Lambda\mathcal{N}(\psi)]}{\mathrm{Tr}[\Lambda\mathcal{M}(\psi)]} \right) && (5.29)\\
&= \sup_{\psi \in \mathfrak{D}(R_0A_0)} \log_2 \inf\{\lambda : \lambda\mathcal{M}(\psi) \geqslant \mathcal{N}(\psi)\} && (5.30)\\
&= \log_2 \inf\{\lambda : \lambda J_A^{\mathcal{M}} \geqslant J_A^{\mathcal{N}}\} && (5.31)\\
&= D_{\max}(\mathcal{N}\|\mathcal{M}) && (5.32)
\end{aligned}$$

Now by taking the limit $\alpha \to 1$ we get

$$H_1(\mathcal{N}\|\mathcal{M}) = \lim_{\alpha\to 1} H_\alpha(\mathcal{N}\|\mathcal{M}) = \lim_{\alpha\to 1}\left(\sup_{\psi\in\mathfrak{D}(R_0A_0)} \frac{\alpha}{\alpha-1}\log_2 \sup_{\alpha^{-1}I\leqslant\Lambda\leqslant I} \frac{\mathrm{Tr}[\Lambda\mathcal{N}(\psi)]}{\mathrm{Tr}[\Lambda\mathcal{M}(\psi)]}\right) \tag{5.33}$$

$$= \sup_{\psi\in\mathfrak{D}(R_0A_0)} \frac{1}{2\ln(2)}\|\mathcal{N}(\psi)-\mathcal{M}(\psi)\|_1 \tag{5.34}$$

$$= \frac{1}{2\ln(2)}\|\mathcal{N}-\mathcal{M}\|_\diamond \tag{5.35}$$

where the second equality follows from Theorem 1 property 4 of [308]. ∎

Theorem 5.6. *Consider two pairs of quantum channels $(\mathcal{N},\mathcal{M})\in \mathrm{CPTP}(A_0\to A_1)$ and $(\mathcal{E},\mathcal{F})\in\mathrm{CPTP}(B_0\to B_1)$ such that $(\mathcal{N},\mathcal{M})\succ_L(\mathcal{E},\mathcal{F})$. The following are equivalent:*

1. *for all $t\geqslant 0$, $\|\mathcal{N}-t\mathcal{M}\|_\diamond \geqslant \|\mathcal{E}-t\mathcal{F}\|_\diamond$;*
2. *for all $\alpha\geqslant 1$,*

$$H_\alpha(\mathcal{N}\|\mathcal{M}) \geqslant H_\alpha(\mathcal{E}\|\mathcal{F})$$
$$H_\alpha(\mathcal{M}\|\mathcal{N}) \geqslant H_\alpha(\mathcal{F}\|\mathcal{E})$$

3. *for all $0\leqslant\epsilon\leqslant 1$, $D_H^\epsilon(\mathcal{N}\|\mathcal{M})\geqslant D_H^\epsilon(\mathcal{E}\|\mathcal{F})$*

The proof of this theorem is broken down into the following lemmas.

Lemma 5.7. *Given two pairs of quantum channels $(\mathcal{N},\mathcal{M})\in\mathrm{CPTP}(A)$ and $(\mathcal{E},\mathcal{F})\in\mathrm{CPTP}(B)$ such that $(\mathcal{N},\mathcal{M})\succ_L(\mathcal{E},\mathcal{F})$, the following are equivalent:*

(i) $\mathfrak{T}(\mathcal{N},\mathcal{M})\supseteq\mathfrak{T}(\mathcal{E},\mathcal{F})$;

(ii) $\|t_1\mathcal{N}+t_2\mathcal{M}\|_\diamond\geqslant\|t_1\mathcal{E}+t_2\mathcal{F}\|_\diamond$ *for all* $t_1,t_2\in\mathbb{R}$;

(iii) $\|\mathcal{N}-t\mathcal{M}\|_\diamond\geqslant\|\mathcal{E}-t\mathcal{F}\|_\diamond$ *for all* $t\geqslant 0$.

Proof. To show the equivalence of property (ii) with (i), let (n,m) and (e,f) be the generic elements of $\mathfrak{T}(\mathcal{N},\mathcal{M})$ and $\mathfrak{T}(\mathcal{E},\mathcal{F})$, respectively. Then, using the separation theorem for convex sets, we get that $\mathfrak{T}(\mathcal{N},\mathcal{M})\supseteq\mathfrak{T}(\mathcal{E},\mathcal{F})$ if and only if, for any $v=(a,b)\in\mathbb{R}^2$,

$$\max_{(n,m)\in\mathfrak{T}(\mathcal{N},\mathcal{M})}[an+bm] \geqslant \max_{(e,f)\in\mathfrak{T}(\mathcal{E},\mathcal{F})}[ae+bf] \tag{5.36}$$

Now,

$$\max_{(n,m)\in\mathfrak{T}(\mathcal{N},\mathcal{M})}[an+bm] = \max_{0\leqslant\Lambda\leqslant I}\max_{\psi\in\mathfrak{D}(R_0A_0)}\{a\mathrm{Tr}[\Lambda\mathcal{N}(\psi)]+b\mathrm{Tr}[\Lambda\mathcal{M}(\psi)]\} \tag{5.37}$$

$$= \max_{0\leqslant\Lambda\leqslant I}\max_{\psi\in\mathfrak{D}(R_0A_0)}\{\mathrm{Tr}[(a\mathcal{N}(\psi)+b\mathcal{M}(\psi))\Lambda]\} \tag{5.38}$$

$$= \max_{\psi\in\mathfrak{D}(R_0A_0)}\mathrm{Tr}[(a\mathcal{N}(\psi)+b\mathcal{M}(\psi))_+] \tag{5.39}$$

$$= \frac{1}{2}\left(\max_{\psi}\left(\|a\mathcal{N}(\psi)+b\mathcal{M}(\psi)\|_1+\mathrm{Tr}[a\mathcal{N}(\psi)+b\mathcal{M}(\psi)]\right)\right) \tag{5.40}$$

$$= \frac{1}{2}\left(\|a\mathcal{N}+b\mathcal{M}\|_\diamond+a+b\right) \tag{5.41}$$

This shows that Eq. (5.36) is satisfied if and only if $\|a\mathcal{N}+b\mathcal{M}\|_\diamond \geqslant \|a\mathcal{E}+b\mathcal{F}\|_\diamond$. To show the equivalence of ii and iii, first note that iii is a special case of ii. To prove that iii implies ii, let us first take the case when $t_1, t_2 \geqslant 0$ or $t_1, t_2 \leqslant 0$, then $\|t_1\mathcal{E}+t_2\mathcal{F}\|_\diamond = \|t_1\mathcal{N}+t_2\mathcal{M}\|_\diamond$ because $\mathcal{E}$, $\mathcal{F}$, $\mathcal{N}$, $\mathcal{M}$ are completely positive maps. Hence, we can consider the case $t_1 > 0 > t_2$ or $t_2 > 0 > t_1$. Either of the choices will give the same result, because for any matrix A it holds that $\|A\| = \|-A\|$ for any norm. Now, by restricting to the former choice, we can rescale both t_1 and t_2 by the positive factor $1/t_1$ to obtain the desired statement. ∎

Lemma 5.8. *For any choice of quantum channels $\mathcal{N}$ and $\mathcal{M}$,*

$$\sup_\alpha(\mathcal{N}/\mathcal{M}) = \inf\{\lambda\geqslant 1 : \frac{\|\lambda\mathcal{M}-\mathcal{N}\|_\diamond}{\lambda-1}\leqslant\frac{\alpha+1}{\alpha-1}\}$$

Proof.

$$\sup_\alpha(\mathcal{N}/\mathcal{M}) = \sup_{\psi\in\mathfrak{D}(R_0A_0)}\sup_{\alpha^{-1}I\leqslant\Lambda\leqslant I}\frac{\mathrm{Tr}[\Lambda\mathcal{N}(\psi)]}{\mathrm{Tr}[\Lambda\mathcal{M}(\psi)]} \tag{5.42}$$

$$= \inf\left\{\lambda : \lambda\geqslant\frac{\mathrm{Tr}[\Lambda\mathcal{N}(\psi)]}{\mathrm{Tr}[\Lambda\mathcal{M}(\psi)]},\ \ \forall\,\psi\in\mathfrak{D}(R_0A_0),\ \alpha^{-1}I\leqslant\Lambda\leqslant I\right\} \tag{5.43}$$

$$= \inf\left\{\lambda : \mathrm{Tr}[\Lambda(\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi))]\geqslant 0,\ \ \forall\,\psi\in\mathfrak{D}(R_0A_0),\ \alpha^{-1}I\leqslant\Lambda\leqslant I\right\} \tag{5.44}$$

$$= \inf\{\lambda\in\mathbb{R} : \alpha^{-1}\mathrm{Tr}[(\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi))_+]\geqslant\mathrm{Tr}[(\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi))_-],\ \psi\in\mathfrak{D}(R_0A_0)\} \tag{5.45}$$

where, in the last equality, we used the decomposition $A = A_+ - A_-$ for Hermitian operators and the choice $\Lambda = \alpha^{-1}P_+ + P_-$ where $P_\pm$ are the projectors onto the positive and negative part of $\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi)$. Then, using the relations

$$\lambda - 1 = \mathrm{Tr}[(\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi))_+] - \mathrm{Tr}[(\lambda\mathcal{M}(\psi)-\mathcal{N}(\psi))_-]$$

for all $\psi \in \mathfrak{D}(R_0 A_0)$, and

$$\begin{aligned}\|\lambda\mathcal{M} - \mathcal{N}\|_\diamond &= \sup_{\psi \in \mathfrak{D}(R_0 A_0)} \|\lambda\mathcal{M}(\psi) - \mathcal{N}(\psi)\|_1 \\ &= \sup_{\psi} \big(\operatorname{Tr}[(\lambda\mathcal{M}(\psi) - \mathcal{N}(\psi))_+] \\ &\qquad + \operatorname{Tr}[(\lambda\mathcal{M}(\psi) - \mathcal{N}(\psi))_-] \big)\end{aligned}$$

we find that

$$\operatorname{Tr}[(\lambda\mathcal{M}(\psi) - \mathcal{N}(\psi))_-] \leqslant \frac{\lambda - 1}{\alpha - 1} \tag{5.46}$$

holds if and only if $\|\lambda\mathcal{M} - \mathcal{N}\|_\diamond \leqslant \frac{\lambda-1}{\alpha-1} + \operatorname{Tr}[(\lambda\mathcal{M}(\psi) - \mathcal{N}(\psi))_+]$. From this, we easily get that

$$\sup\nolimits_\alpha(\mathcal{N}/\mathcal{M}) = \inf\left\{ \lambda \geqslant 1 : \frac{\|\lambda\mathcal{M} - \mathcal{N}\|_\diamond}{\lambda - 1} \leqslant \frac{\alpha + 1}{\alpha - 1} \right\}$$

∎

Lemma 5.9. *For any choice of quantum channels $\mathcal{N}$ and $\mathcal{M}$, the function*

$$f(\lambda) = \frac{\|\lambda\mathcal{M} - \mathcal{N}\|_\diamond}{\lambda - 1} \tag{5.47}$$

is monotonically nonincreasing in the domain $\lambda \geqslant 1$ with $f(1) = \infty$ and $f(\infty) = 1$.

Proof. Since the function $f(\lambda) = \frac{\|\lambda\sigma - \rho\|_1}{\lambda - 1}$ is monotonically nonincreasing for any density matrices ρ, σ, the equation (5.47) will hold for the channel case and $f(1) = \infty$ and $f(\infty) = 1$. ∎

Lemma 5.10. *Consider two pairs of quantum channels $(\mathcal{N}, \mathcal{M})$ and $(\mathcal{E}, \mathcal{F})$. Then, the following are equivalent:*

(i) for all $\alpha \geqslant 1$, $\sup_\alpha(\mathcal{N}/\mathcal{M}) \geqslant \sup_\alpha(\mathcal{E}/\mathcal{F})$ and $\inf_\alpha(\mathcal{N}/\mathcal{M}) \leqslant \inf_\alpha(\mathcal{E}/\mathcal{F})$;

(ii) $\|t\mathcal{M} - \mathcal{N}\|_\diamond \geqslant \|t\mathcal{F} - \mathcal{E}\|_\diamond$ for all $t \geqslant 0$.

Proof. The proof for channels is similar to the state case in [308]. ∎

Lemma 5.11. *Given two pairs of quantum channels $(\mathcal{N}, \mathcal{M})$ and $(\mathcal{E}, \mathcal{F})$ on dynamical system A and B respectively, the following are equivalent:*

(i) $Q^\epsilon(\mathcal{N}\|\mathcal{M}) \leqslant Q^\epsilon(\mathcal{E}\|\mathcal{F})$, for all $\epsilon \in [0, 1]$;

(ii) $\|p\mathcal{N} - \mathcal{M}\|_\diamond \geqslant \|p\mathcal{E} - \mathcal{F}\|_\diamond$ for all $p \geqslant 0$.

Proof. From Equations (5.23) to (5.26), we know that $Q^\epsilon(\mathcal{N}\|\mathcal{M})$ can be expressed as an SDP. It's dual can be expressed as

$$Q^\epsilon(\mathcal{N}\|\mathcal{M}) = \sup\ (1-\epsilon)p - t \tag{5.48}$$

$$\text{s.t. } tI_{A_0} \geqslant \alpha_{A_0} \geqslant 0\,, \tag{5.49}$$

$$\alpha_A \geqslant pJ_A^{\mathcal{N}} - J_A^{\mathcal{M}}\,, \tag{5.50}$$

$$t \in \mathbb{R},\ p \geqslant 0\,. \tag{5.51}$$

To get the supremum in Eq. (5.48), we can first optimize over t by keeping p fixed. Notice that minimizing t subject to Eqs. (5.49) and (5.50) is equal to the half of the diamond norm between $p\mathcal{N}$ and $\mathcal{M}$ [129]. Thus, $Q^\epsilon(\mathcal{N}\|\mathcal{M})$ can be expressed as

$$Q^\epsilon(\mathcal{N}\|\mathcal{M}) = \sup_{p\geqslant 0}\ f_\epsilon(p) \tag{5.52}$$

where

$$f_\epsilon(p) := (1-\epsilon)p - \frac{1}{2}\left\|p\mathcal{N}_A - \mathcal{M}_A\right\|_\diamond \tag{5.53}$$

Using the above function, the equivalence in (i) and (ii) can be shown by following similar steps as given in Lemma 5 of [308]. ∎

5.4 Interconversions

Theorem 5.12. *Let $\mathcal{E},\mathcal{F} \in \mathrm{CPTP}(A_0 \to A_1)$ and $\mathcal{N},\mathcal{M} \in \mathrm{CPTP}(B_0 \to B_1)$. Then $(\mathcal{E},\mathcal{F}) \succ_L (\mathcal{N},\mathcal{M}) \iff (\mathcal{E},\mathcal{F}) \succ (\mathcal{N},\mathcal{M})$ in the following cases:*

I. When $|A_0| = 1$ and A_1, B_0, B_1 are classical systems. That is, a pair of classical states can be converted to a pair of classical channels if and only if the pair of classical states Lorenz majorizes the pair of classical channels.

II. When A_0, A_1, B_0, B_1 are classical systems and $|B_1| = 2$. That is, a pair of classical channels can be converted to a pair of classical channels with two-dimensional output if and only if the former pair Lorenz majorizes the latter.

We break the proof of the above theorem (and its cases) into several lemmas.

Lemma 5.13 (Theorem 5.12 Case I)**.** *Let $\rho, \sigma \in \mathfrak{C}(A_1)$ be classical states in systems A_1 and $\mathcal{N}, \mathcal{M} \in \mathfrak{C}(B)$. Then,*

$$(\rho, \sigma) \succ (\mathcal{N}, \mathcal{M}) \iff (\rho, \sigma) \succ_L (\mathcal{N}, \mathcal{M}) \tag{5.54}$$

Proof. It follows directly from the definitions that $(\rho_{A_1}, \sigma_{A_1}) \succ (\mathcal{N}_B, \mathcal{M}_B) \implies (\rho_{A_1}, \sigma_{A_1}) \succ_L (\mathcal{N}_B, \mathcal{M}_B)$. For the proof of other direction, we just need to show the existence of a superchannel $\Theta \in \mathfrak{S}(A_1 \to B)$ such that $\Theta[\rho] = \mathcal{N}$ and $\Theta[\sigma] = \mathcal{M}$. Since $\mathcal{N}$ and $\mathcal{M}$ are classical channels, their Choi matrices, $J_A^{\mathcal{N}}$ and $J_A^{\mathcal{M}}$, are diagonal and can be expressed as $|A_1| \times |A_0|$ column stochastic matrices. Let us denote the columns of these matrices by $\mathbf{n}_i$ and $\mathbf{m}_i$, respectively, where $i = \{1, \ldots, |A_0|\}$. The testing region associated with $(\mathcal{N}, \mathcal{M})$ is then the convex combination of the testing regions $\mathfrak{T}(\mathbf{n}_i, \mathbf{m}_i)$ and the Lorenz curve associated with $(\mathcal{N}, \mathcal{M})$ is the upper boundary of the convex region $\mathcal{T}(\mathcal{N}, \mathcal{M})$. Then, the condition $(\rho, \sigma) \succ_L (\mathcal{N}, \mathcal{M})$ implies that the Lorenz curve of (ρ, σ) is never below the Lorenz curve of $(\mathbf{n}_i, \mathbf{m}_i)$ for all $i = \{1, \ldots, |A_0|\}$. From [308] we know that if a pair of classical states Lorenz majorizes another pair of classical states then there exists a channel that transforms the former into the latter. Then, for every $i \in \{1, \ldots, |A_0|\}$ there exists a channel $\mathcal{E}_i \in \text{CPTP}(A_1 \to B_1)$ that can transform the pair of states $(\rho_{A_1}, \sigma_{A_1})$ to the pair of classical states in system B_1 with diagonals $(\mathbf{n}_i, \mathbf{m}_i)$. Using these $\mathcal{E}_i$'s, we can now define a superchannel $\Theta \in \mathfrak{S}(A_1 \to B)$ as

$$\mathbf{J}^{\Theta}_{A_1 B} := \sum_i |i\rangle\langle i|_{B_0} \otimes J^{\mathcal{E}_i}_{A_1 B_1} . \tag{5.55}$$

It is straightforward to verify that the above superchannel transforms the pair of state (ρ, σ) to the pair of channels $(\mathcal{N}, \mathcal{M})$.

∎

Lemma 5.14. *Let $\mathcal{N}, \mathcal{M} \in \mathfrak{C}(A_0 \to A_1)$ and $\mathcal{E}, \mathcal{F} \in \mathfrak{C}(B_0 \to B_1)$ where $|A_0| = |A_1| = 2 = |B_0| = |B_1|$. Then,*

$$(\mathcal{N}, \mathcal{M}) \succ (\mathcal{E}, \mathcal{F}) \iff (\mathcal{N}, \mathcal{M}) \succ_L (\mathcal{E}, \mathcal{F}) \tag{5.56}$$

Proof. The proof of $(\mathcal{N}, \mathcal{M}) \succ_R (\mathcal{E}, \mathcal{F}) \implies (\mathcal{N}, \mathcal{M}) \succ_L (\mathcal{E}, \mathcal{F})$ follows from the definition. For the other side, let us look at the Lorenz curves of both pairs of qubit classical channels. Since $(\mathcal{N}, \mathcal{M}) \succ_L (\mathcal{E}, \mathcal{F})$, either the individual Lorenz curves $L(\mathbf{e}_i, \mathbf{f}_i)$ for $i = \{1, 2\}$ lie below either of $L(\mathbf{n}_i, \mathbf{m}_i)$ for $i = \{1, 2\}$ or they do not. In the former case, we know from the results of the state case that there exists a classical channel transforming one pair of incoherent states to another when one Lorenz majorizes the other. Then wlog,

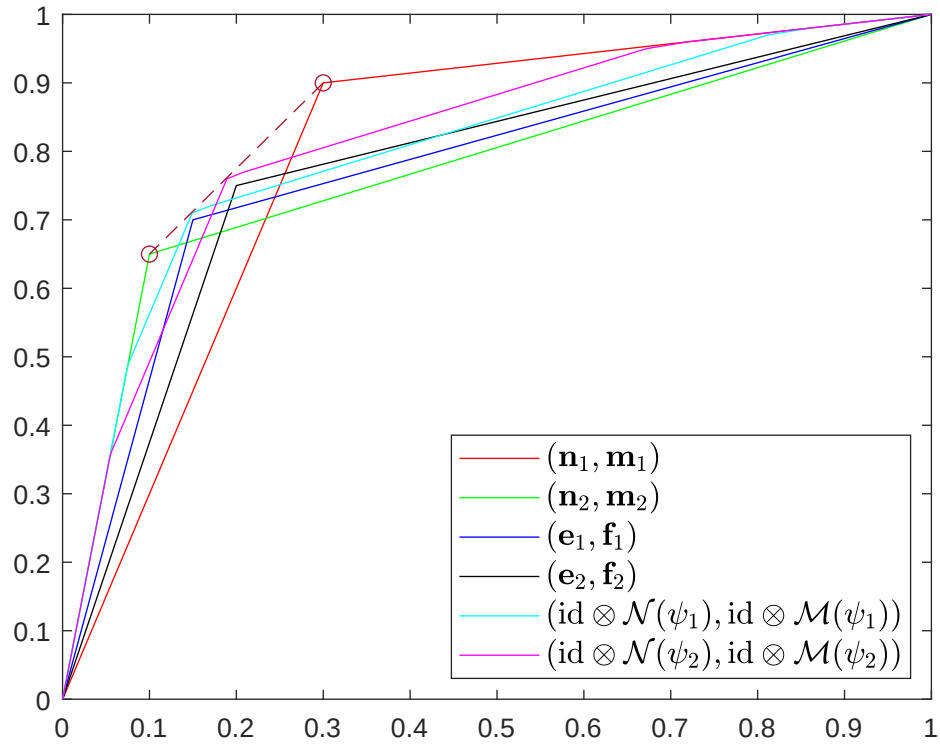

Figure 5.1: Classical Lorenz curve of two pairs of classical channels

suppose $(\mathbf{n}_i, \mathbf{m}_i) \succ_L (\mathbf{e}_i, \mathbf{f}_i)$ for $i = (1, 2)$. Then, there exist classical channels $\mathcal{R}_1$ and $\mathcal{R}_2$ transforming the pairs $(\mathbf{n}_1, \mathbf{m}_1)$ and $(\mathbf{n}_2, \mathbf{m}_2)$ to $(\mathbf{e}_1, \mathbf{f}_1)$ and $(\mathbf{e}_2, \mathbf{f}_2)$, respectively. Then the Choi matrix of the superchannel $\Theta \in \mathfrak{S}(A \to B)$ that transforms the pair of channels $(\mathcal{N}, \mathcal{M})$ to $(\mathcal{E}, \mathcal{F})$ is given by

$$\begin{aligned} J^{\Theta} = &|0\rangle\langle 0|_{A_0} \otimes J^{\mathcal{R}_1}_{A_1 B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &+ |1\rangle\langle 1|_{A_0} \otimes J^{\mathcal{R}_2}_{A_1 B_1} \otimes |1\rangle\langle 1|_{B_0} \end{aligned} \tag{5.57}$$

For the latter part, when the Lorenz curve of $(\mathcal{N}, \mathcal{M})$ lies above that of $\{\mathbf{e}_i, \mathbf{f}_i\}$ for both $i = \{1, 2\}$, but either one or both of these are not contained in $L(\mathbf{n}_i, \mathbf{m}_i)$ for $i = \{1, 2\}$ (refer Fig.5.1), we can always find two pairs of pure bipartite states of the form $\sqrt{p}|00\rangle + \sqrt{1-p}|11\rangle$ whose Lorenz curves have a extreme point on the line connecting the extreme points of $L(\mathbf{n}_1, \mathbf{m}_1)$ and $L(\mathbf{n}_2, \mathbf{m}_2)$ and which always lie above the Lorenz curves $L(\mathbf{e}_i, \mathbf{f}_i)$ for $i = 1, 2$, respectively. This is true because the Lorenz curve is convex. Let $|\psi_1\rangle = \sqrt{p}|00\rangle + \sqrt{1-p}|11\rangle$ be such that

$$L\left(\mathcal{N}_{A_0 \to A_1}\left(\psi_{1_{\tilde{A}_0 A_0}}\right), \mathcal{M}_{A_0 \to A_1}\left(\psi_{1_{\tilde{A}_0 A_0}}\right)\right) \succ_L L(e_1, f_1) \tag{5.58}$$

and let $|\psi_2\rangle = \sqrt{q}|00\rangle + \sqrt{1-q}|11\rangle$ be such that

$$L\left(\mathcal{N}_{A_0\to A_1}\left(\psi_{2\,\tilde{A}_0A_0}\right), \mathcal{M}_{A_0\to A_1}\left(\psi_{2\,\tilde{A}_0A_0}\right)\right) \succ_L L(e_2, f_2)\,. \tag{5.59}$$

Notice that the states $\mathcal{M}(\psi_1)$, $\mathcal{M}(\psi_2)$, $\mathcal{N}(\psi_1)$, and $\mathcal{N}(\psi_2)$ are diagonal. Using these, we can define two classical channels $\mathcal{P}, \mathcal{Q} \in \mathfrak{C}(C_1 \to A_0A_1)/\mathfrak{C}(B_0 \to B_1)$ where $|C_1| = 2$, and with Choi matrices $J^{\mathcal{P}}$ and $J^{\mathcal{Q}}$ as given by

$$\begin{aligned} J^{\mathcal{P}} &= |0\rangle\langle 0| \otimes \mathcal{N}_{A_0\to A_1}\left(\psi_{1\,\tilde{A}_0A_0}\right) + \\ &\qquad |1\rangle\langle 1| \otimes \mathcal{N}_{A_0\to A_1}\left(\psi_{2\,\tilde{A}_0A_0}\right)\,, \\ J^{\mathcal{Q}} &= |0\rangle\langle 0| \otimes \mathcal{M}_{A_0\to A_1}\left(\psi_{1\,\tilde{A}_0A_0}\right) + \\ &\qquad |1\rangle\langle 1| \otimes \mathcal{M}_{A_0\to A_1}\left(\psi_{2\,\tilde{A}_0A_0}\right) \end{aligned} \tag{5.60}$$

Then, it can be easily verified that the following Choi matrix of a superchannel Θ converts the pair of classical channels $(\mathcal{N}, \mathcal{M})$ to another pair of classical channels $(\mathcal{P}, \mathcal{Q})$

$$\begin{aligned} \mathbf{J}^{\Theta}_{AB} &= |0\rangle\langle 0|_{A_0} \otimes J^{\Theta_1}_{A_1B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &\quad + |1\rangle\langle 1|_{A_0} \otimes J^{\Theta_2}_{A_1B_1} \otimes |1\rangle\langle 1|_{B_0} \\ &= |0\rangle\langle 0|_{A_0} \otimes \Big[|0\rangle\langle 0|_{A_1} \otimes \big(p\left(|0\rangle\langle 0| \otimes |0\rangle\langle 0|\right)_{B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &\qquad\qquad + q\left(|0\rangle\langle 0| \otimes |0\rangle\langle 0|\right)_{B_1} \otimes |1\rangle\langle 1|_{B_0}\big) \\ &\qquad + |1\rangle\langle 1|_{A_1} \otimes \big(p\left(|0\rangle\langle 0| \otimes |1\rangle\langle 1|\right)_{B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &\qquad\qquad + q\left(|0\rangle\langle 0| \otimes |1\rangle\langle 1|\right)_{B_1} \otimes |1\rangle\langle 1|_{B_0}\big)\Big] \\ &\quad + |1\rangle\langle 1|_{A_0} \otimes \Big[|0\rangle\langle 0|_{A_1} \otimes \big((1-p)\left(|1\rangle\langle 1| \otimes |0\rangle\langle 0|\right)_{B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &\qquad\qquad + (1-q)\left(|1\rangle\langle 1| \otimes |0\rangle\langle 0|\right)_{B_1} \otimes |1\rangle\langle 1|_{B_0}\big) \\ &\qquad + |1\rangle\langle 1|_{A_1} \otimes \big((1-p)\left(|1\rangle\langle 1| \otimes |1\rangle\langle 1|\right)_{B_1} \otimes |0\rangle\langle 0|_{B_0} \\ &\qquad\qquad + (1-q)\left(|1\rangle\langle 1| \otimes |1\rangle\langle 1|\right)_{B_1} \otimes |1\rangle\langle 1|_{B_0}\big)\Big]\,. \end{aligned} \tag{5.61}$$

Finally, the superchannel to convert the pair of channels $(\mathcal{P}, \mathcal{Q})$ to $(\mathcal{E}, \mathcal{F})$ can be found in a similar way we formulated Eq. (5.57). Thus, we can find superchannels that convert $(\mathcal{N}, \mathcal{M})$ to $(\mathcal{P}, \mathcal{Q})$ and $(\mathcal{P}, \mathcal{Q})$ to $(\mathcal{E}, \mathcal{F})$, making the transformation of $(\mathcal{N}, \mathcal{M})$ to $(\mathcal{E}, \mathcal{F})$ possible.

∎

Lemma 5.15. *Let $\mathcal{N}, \mathcal{M} \in \mathfrak{C}(A_0 \to A_1)$ and $\mathcal{E}, \mathcal{F} \in \mathfrak{C}(B_0 \to B_1)$ be such that $|A_1| = |B_1| = 2$. Then*

$(\mathcal{N},\mathcal{M}) \succ_L (\mathcal{E},\mathcal{F}) \iff (\mathcal{N},\mathcal{M}) \succ (\mathcal{E},\mathcal{F})$.

Proof. When $|A_1| = |B_1| = 2$, the stochastic matrices corresponding to $\mathcal{N}$ and $\mathcal{M}$ are $2 \times |A_0|$ matrices, and the ones corresponding to $\mathcal{E}$ and $\mathcal{F}$ are $2 \times |B_0|$ matrices. The Lorenz curve of $(\mathcal{N},\mathcal{M})$ is the convex curve formed from the extreme points of the Lorenz curves of $(\mathbf{n}_i, \mathbf{m}_i)$ for all $i = \{1,\ldots,|A_0|\}$. Likewise, the Lorenz curve of $(\mathcal{E},\mathcal{F})$ is the convex curve formed from the extreme points of the Lorenz curves of $(\mathbf{e}_j, \mathbf{f}_j)$ for all $j = \{1,\ldots,|B_0|\}$. Since $\mathbf{n}_i, \mathbf{m}_i, \mathbf{e}_j$ and $\mathbf{f}_j$ are two-dimensional probability vectors, we can use Lemma 5.14 to find a state ψ_j such that $(I \otimes \mathcal{N}(\psi_j), I \otimes \mathcal{M}(\psi_j))$ Lorenz majorizes the pair $(\mathbf{e}_j, \mathbf{f}_j)$. Thus, we can find a channel that transforms $(\mathcal{N},\mathcal{M})$ to $(\mathbf{e}_j, \mathbf{f}_j)$ for all j. Using this, we can construct a superchannel that converts $(\mathcal{N},\mathcal{M})$ to $(\mathcal{E},\mathcal{F})$.

∎

Lemma 5.16 (Theorem 5.12 Case III)**.** *Let* $\mathcal{N},\mathcal{M} \in \mathfrak{C}(A_0 \to A_1)$ *and* $\mathcal{E},\mathcal{F} \in \mathfrak{C}(B_0 \to B_1)$*. Then* $(\mathcal{N},\mathcal{M}) \succ_L (\mathcal{E},\mathcal{F}) \iff (\mathcal{N},\mathcal{M}) \succ (\mathcal{E},\mathcal{F})$ *holds if* $|B_1| = 2$.

Proof. With each vertex of the Lorenz curve of $(\mathcal{N},\mathcal{M})$, we can construct a Lorenz curve corresponding to a pair of two-dimensional classical states or probability vectors (p_i, q_i) which can be achieved by evolving some state through $\mathcal{N}$ and $\mathcal{M}$. Thus, using these p_i and q_i, we can form stochastic matrices corresponding to two channels $\mathcal{P},\mathcal{Q} \in \mathfrak{C}(R_1 \to R_2)$ where $|R_1|$ is equal to the number of extreme points of the Lorenz curve of $(\mathcal{N},\mathcal{M})$ and $|R_2| = 2$. Thus, a superchannel can be constructed that converts the pair of channels $(\mathcal{N},\mathcal{M})$ to $(\mathcal{P},\mathcal{Q})$. Since the Lorenz curve of $(\mathcal{E},\mathcal{F})$ lies below that of $(\mathcal{N},\mathcal{M})$ (and hence is below that of $(\mathcal{P},\mathcal{Q})$), from Lemma 5.15 we can find a superchannel that transforms the pair of channels $(\mathcal{P},\mathcal{Q})$ to $(\mathcal{E},\mathcal{F})$. Therefore, $(\mathcal{N},\mathcal{M}) \succ_L (\mathcal{E},\mathcal{F}) \iff (\mathcal{N},\mathcal{M}) \succ (\mathcal{E},\mathcal{F})$ when $\mathcal{E},\mathcal{F} \in \mathfrak{C}(B_0 \to B_1)$ where $|B_1| = 2$. ∎

5.5 Summary

In this chapter, I generalized Lorenz majorization to the channel domain. In Section 5.3, we defined the testing region for a pair of channels by using a process POVM and extended the definition of Hilbert α divergence for a pair of channels. We also formulated equivalent conditions for Lorenz majorization between a pair of channels in terms of Hilbert α divergence that we defined for channels. In section 4.5, we solve some interconversion problems among states and channels and found the conditions for equivalence between relative and Lorenz majorization. In Lemma 5.13, we show that there exists a superchannel to convert a pair of classical states to a pair of classical channels if the pair of classical states Lorenz majorizes the pair of classical channels. We also made progress towards completing the picture of Lorenz and relative majorization equivalence for the classical case. Until now, it was known that Lorenz and relative majorization

are equivalent for pairs of probability distributions and classical states. In Lemma 5.16, we show equivalence in Lorenz and relative majorization for the case when a pair of channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(A_0 \to A_1)$ Lorenz majorizes a pair of channels $\mathcal{E}, \mathcal{F} \in \mathrm{CPTP}(B_0 \to B_1)$ where $|B_1| = 2$. We leave as open the last piece of puzzle in the classical case i.e., the equivalence of Lorenz and relative majorization for classical channels when $|B_1| > 2$. We would like to conclude by giving the following conjecture. Let $(\rho_1, \rho_2) \in \mathfrak{D}(A_0)$ and $(\sigma_1, \sigma_2) \in \mathfrak{D}(A_1)$ be such that $(\rho_1, \rho_2) \succ_L (\sigma_1, \sigma_2)$, and $(\rho_1', \rho_2') \in \mathfrak{D}(B_0)$ and $(\sigma_1', \sigma_2') \in \mathfrak{D}(B_1)$ be such that $(\rho_1', \rho_2') \succ_L (\sigma_1', \sigma_2')$. Then, we conjecture that

$$(\rho_1 \otimes \rho_1', \rho_2 \otimes \rho_2') \succ_L (\sigma_1 \otimes \sigma_1', \sigma_2 \otimes \sigma_2')\,. \tag{5.62}$$

Chapter 6

Limitations on simulating a unitary channel using fixed processor

6.1 Introduction

In quest for automation, programming devices or machines is a vital step. In the language of classical computing, programs (software) allow us to execute any sequence of logical operations on a user-defined data using a fixed processor (hardware). Similarly, one hopes to realize automation in the quantum era.

Quantum computers, powered by the laws of quantum mechanics, promise to process information and perform computation in a manner far more superior than classical computers [25]. A fundamental model to realize universal quantum computation is that of a fixed programmable quantum processor – a concept similar to its classical counterpart. A mathematical abstraction of such a model is a bipartite processor with one of the inputs being a specifically programmed quantum state that helps the processor implement a desired unitary gate on the other input quantum state.

This idea of (universal) quantum programmability was first proposed by Nielsen and Chuang [47]. They showed that in order to simulate distinct unitaries, the program states must be orthogonal. This result implies that the dimension of the programming system must scale up to infinity, thus disproving the existence of a universal quantum processor. Luckily, in an imprecise manner, universal programmability of quantum processors can still be reached, and bounds on the dimension of the program state have been calculated for the following cases:

(i) *Approximate Quantum Programming:* The desired gate is implemented by allowing some errors quantified by distance functions of quantum channels, such as channel fidelity or diamond norm (see

Refs. [322, 323, 324] and references therein).

(ii) *Probabilistic Quantum Programming:* The target gate is simulated probabilistically. Specific protocols include the gate-teleportation presented in Ref. [325] (see Refs. [42, 44, 45, 326, 327] and references therein).

In the above approaches to approximately or probabilistically simulate a unitary gate, the processor needs to be scaled. Considering the fact that scaling makes it hard to prevent errors in computation, a physically motivated constraint is to fix the dimensions of the input and output systems of the processor. With this restriction, several questions emerge, such as how to quantify a programmable processor, how to compare the efficiency of different processors for simulating a given unitary, and what are optimal ways to achieve the peak performance using a particular processor.

To address these questions, we adopt the most general approach of manipulating a quantum processor, i.e., by encoding program states into entangled states and decoding information through joint measurements. We show that these operations (the input program state and the joint measurements) together are completely characterized by the process positive operator valued measure (PPOVM) introduced in the theory of quantum channel measurement [117]. Using this framework, we quantify the performance of a processor with the help of average fidelity between quantum channels. This measure can be used to compare the average performance of different processors. We also derive a trade-off relation between the average fidelity error and the maximum success probability of simulating a target unitary channel. We show that this trade-off is a semidefinite program and, hence, is efficiently computable. We also apply our framework to a practical scenario by constraining the quantum computational ability of implementing only certain types of PPOVMs. In the end, we numerically demonstrate the trade-off between success probability and average error by considering various processors.

6.2 Unitary channel simulation using a fixed quantum processor

6.2.1 Preliminaries

There are two different ways of composition of quantum channels: parallel and sequential. For quantum channels $\mathcal{E} : A_0 \to A_1$ and $\mathcal{F} : B_0 \to B_1$, the Choi operators of their parallel composition $\mathcal{E} \otimes \mathcal{F}$ is given by $J_A^{\mathcal{E}} \otimes J_B^{\mathcal{F}}$, with $J_A^{\mathcal{E}}$ and $J_B^{\mathcal{F}}$ stand for the Choi operator of $\mathcal{E}$ and $\mathcal{F}$, respectively. To investigate the sequential composition of quantum channels, the concept of *link product* is used [120, 122].

Definition 6.1 (Link Product)**.** For two matrices M and N, which have system X in common, their link

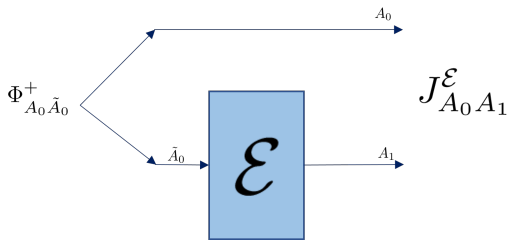

Figure 6.1: Choi operator of point-to-point channel

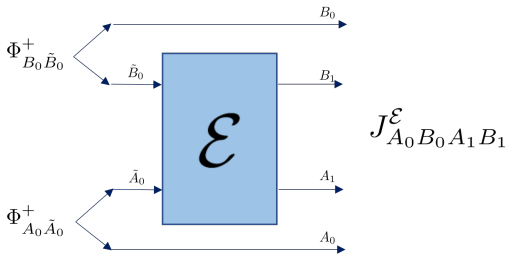

Figure 6.2: Choi operator of bipartite channel

product $M \star N$ is defined as

$$\mathrm{Tr}_X[M^{\mathbf{T}_X} \cdot N], \tag{6.1}$$

where $\mathbf{T}_X$ represents the partial transpose over system X. The system connecting two matrices together, i.e. X, has been "swallowed" by the product $\star$.

Using the link product, the Choi operator of $\mathcal{F} \circ \mathcal{E}$ is given by $J^{\mathcal{F}} \star J^{\mathcal{E}}$, where $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ and $\mathcal{F} \in \mathrm{CPTP}(A_1 \to B_1)$. Note that the link product is commutative up to the SWAP over systems [120, 122]. An illustration of Choi operators of a single-partite and a bipartite channel is given in Fig. 6.1 and 6.2.

6.2.2 Limitations on deterministic protocols

A fixed quantum processor that can be used for simulating any arbitrary channel is known as a programmable processor. A programmable quantum processor, also known as a programmable quantum gate array [47], consists of two parts: a fixed quantum processor $\mathcal{E}$, which is a bipartite quantum channel from A_0B_0 to A_1B_1 shared between a user Alice and an agent Bob, and a family of programs $\mathcal{P}_{\mathcal{U}}$, encoded into quantum states $\{\rho_{\mathcal{U}}\}$. To implement a target gate $\mathcal{U}$ on the user's side, Bob selects the program state $\rho_{\mathcal{U}}$ and inserts it into the quantum processor $\mathcal{E}$, as modeled in Fig. 6.3. Then the action of the induced quantum channel $\mathcal{F}$ (on Alice's side) on a state σ_{A_0} is given by

$$\mathcal{F}(\sigma_{A_0}) = \mathrm{Tr}_{B_1}[\mathcal{E}(\sigma_{A_0} \otimes \rho_{\mathcal{U}})]\,. \tag{6.2}$$

Using the Choi operator, it is straightforward to check that

$$J^{\mathcal{F}} = J^{\mathcal{E}} \star (\rho_{\mathcal{U}} \otimes I_{B_1}). \tag{6.3}$$

To benchmark the performance of simulating quantum gate on Alice's side remotely, we consider the

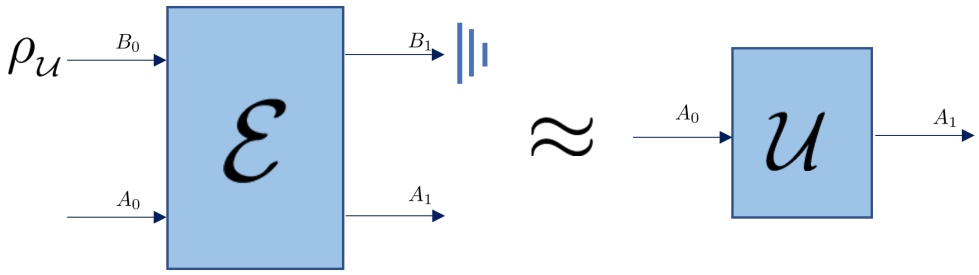

Figure 6.3: Deterministic Universal Quantum Processor.

fidelity between the resultant channel $\mathcal{F}$ and the target gate $\mathcal{U}$. In particular, for distinct quantum channels, we have

Definition 6.2 (Channel Fidelity)**.** Given two channels $\mathcal{E}_1$ and $\mathcal{E}_2$, from A_0 to A_1, the channel fidelity F_{C} is defined as

$$F_{\mathrm{C}}(\mathcal{E}_1, \mathcal{E}_2) := F\left(\frac{1}{|A_0|} J^{\mathcal{E}_1}, \frac{1}{|A_0|} J^{\mathcal{E}_2}\right), \tag{6.4}$$

where F stands for the Uhlmann's fidelity for quantum states, $J^{\mathcal{E}_1}$ and $J^{\mathcal{E}_2}$ are Choi operators of channels $\mathcal{E}_1$ and $\mathcal{E}_2$ respectively.

According to the definition of Uhlmann fidelity, Eq. (6.4) can be rewritten as

$$F_{\mathrm{C}}(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{|A_0|^2} \left(\mathrm{Tr}[\sqrt{\sqrt{J^{\mathcal{E}_1}} J^{\mathcal{E}_2} \sqrt{J^{\mathcal{E}_1}}}]\right)^2. \tag{6.5}$$

If one of the channels considered in $F_{\mathrm{C}}(\mathcal{E}_1, \mathcal{E}_2)$ is a unitary channel, say $\mathcal{E}_2 = \mathcal{U}$, then the above equation can be further simplified to

$$F_{\mathrm{C}}(\mathcal{E}_1, \mathcal{U}) = \frac{1}{|A_0|^2} \mathrm{Tr}[J^{\mathcal{E}_1} \cdot J^{\mathcal{U}}]. \tag{6.6}$$

Using the channel fidelity, we can quantify the distance between the resultant channel $\mathcal{F}$ as given in Eq. (6.3) and the target unitary $\mathcal{U}$ as

$$F_{\mathrm{C}}(\mathcal{F}, \mathcal{U}) = \frac{1}{|A_0|^2} \mathrm{Tr}[J^{\mathcal{E}}_{AB} \star (\rho_u \otimes I_{B_1}) \cdot J^{\mathcal{U}}_{A}], \tag{6.7}$$

with $A := A_0 A_1$ and $B := B_0 B_1$. Our goal here is to find the highest fidelity, i.e. optimal performance, with

respect to all program states $\rho_{\mathcal{U}}$. To do that, let us define a quantity M as

$$M(\mathcal{E},\mathcal{U}) := \frac{1}{|A_0|^2}\mathrm{Tr}_{AB_1}[J^{\mathcal{E}}_{AB}\cdot(J^{\mathcal{U}}_{A}\otimes I_{B_1})], \tag{6.8}$$

which depends only on the fixed quantum processor $\mathcal{E}$ and the target gate $\mathcal{U}$. Then, the channel fidelity $F_{\mathrm{C}}(\mathcal{F},\mathcal{U})$ follows the following condition

$$\begin{aligned} F_{\mathrm{C}}(\mathcal{F},\mathcal{U}) \leqslant \max \quad & \mathrm{Tr}[M(\mathcal{E},\mathcal{U})\cdot\rho] \\ \text{s.t.} \quad & \rho \geqslant 0,\ \mathrm{Tr}[\rho]=1. \end{aligned} \tag{6.9}$$

Hence, in this deterministic case, the performance of universal quantum programming is characterized by the following theorem

Theorem 6.3 (Optimal Performance)**.** *Given a bipartite quantum processor $\mathcal{E}\in\mathrm{CPTP}(A_0B_0\to A_1B_1)$ and a target quantum channel $\mathcal{U}$, the optimal performance of the fixed processor with respect to the target unitary channel is upper-bounded by*

$$F_C(\mathcal{F},\mathcal{U}) \leqslant \lambda_1(M(\mathcal{E},\mathcal{U})), \tag{6.10}$$

where $\lambda_1(\cdot)$ stands for the largest eigenvalue of $(\cdot)$ with $M(\mathcal{E},\mathcal{U})$ defined in Eq. (6.8).

This immediately implies the following limitation on the error in simulating a unitary channel from the fixed processor.

Corollary 6.4. *If there exists a program state $\rho_{\mathcal{U}}$ such that $F_C(\mathrm{Tr}_{B_1}\circ\mathcal{E}(\rho_{\mathcal{U}}),\mathcal{U})\geqslant 1-\epsilon$ for target gate $\mathcal{U}$, then*

$$\epsilon \geqslant 1-\lambda_1(M(\mathcal{E},\mathcal{U})). \tag{6.11}$$

From an experimental point of view, the optimal bound of Eq. (6.9) also indicates that to approximate a desired unitary channel $\mathcal{U}$, Bob should prepare the eigenstate $|\psi\rangle$ of the operator $M(\mathcal{E},\mathcal{U})$ with respect to the eigenvalue of $\lambda_1(M(\mathcal{E},\mathcal{U}))$. Written explicitly, that is,

$$M(\mathcal{E},\mathcal{U})|\psi\rangle = \lambda_1(M(\mathcal{E},\mathcal{U}))|\psi\rangle. \tag{6.12}$$

Thus, given the quantum processor $\mathcal{E}$ and the target gate $\mathcal{U}$, taking the state as $|\psi\rangle$, obeying Eq. (6.12), as

the program state will help Bob achieve the optimal performance.

The question now is can we improve the average performance of the processor for simulating a desired unitary. In other words, can we reduce the error in simulation? To tackle this problem, we need to use probabilistic protocols.

6.2.3 Limitations on probabilistic protocols

From Nielsen and Chuang's no-programming theorem [47], we know that quantum theory prevents us from building universal programmable quantum gate arrays without scaling the dimensions to infinity. Thus, establishing a machine that can execute universal quantum computation is impossible. However, there are two different ways to bypass this no-go theorem:

- Approximate programmable devices: simulating a channel on Alice's side which is very close to the desired unitary $\mathcal{U}$. For more details, see Refs. [322, 323, 324] and references therein.
- Probabilistic programmable devices: probabilistically simulating target unitary $\mathcal{U}$ on Alice's side. For more details, see Refs. [42, 44, 45, 326, 327] and references therein.

In practical settings, however, there is no reason why we should restrict ourselves to approximate or probabilistic protocols alone. Instead, we should consider probabilistic programmable devices with some error tolerance.

To that end, we employ the concept of process positive operator-valued measure (PPOVM) for quantum channels [117] which has been discussed in Sec. 2.2.5. Denote the Choi operator of PPOVM as J^{Θ_x} (where $\Theta_x := M_x \circ \rho_{B_0R}$, and ρ_{B_0R} is the input state to the channel and M_x is the measurement done on the output), it is straightforward to check that it satisfies the following conditions:

$$J^{\Theta_x} \geqslant 0, \tag{6.13}$$

$$\sum_x J^{\Theta_x} = \rho_{B_0} \otimes I_{B_1}. \tag{6.14}$$

Here, the state ρ_{B_0} is the reduced state of ρ_{B_0R}, i.e. $\rho_{B_0} := \mathrm{Tr}_R[\rho_{B_0R}]$.

Given a quantum processor $\mathcal{E} \in \mathrm{CPTP}(A_0B_0 \to A_1B_1)$, and a PPOVM $\{\Theta_x\}_x$ with each $\Theta_x : (1 \to B_0R)\,\mathrm{to}\,(RB_1 \to \mathbb{C})$, as illustrated in Fig. 6.4, the resultant quantum operation $\Theta_x[\mathcal{E}] \in \mathrm{CP}(A_0 \to A_1)$ is completely positive (CP) and trace-non-increasing (TNI). Its action on state σ_{A_0} is characterized by

$$J_B^{\Theta_x} \star J_{AB}^{\mathcal{E}} \star \sigma_{A_0} = p_x J_A^{\mathcal{F}_x} \star \sigma_{A_0}, \tag{6.15}$$

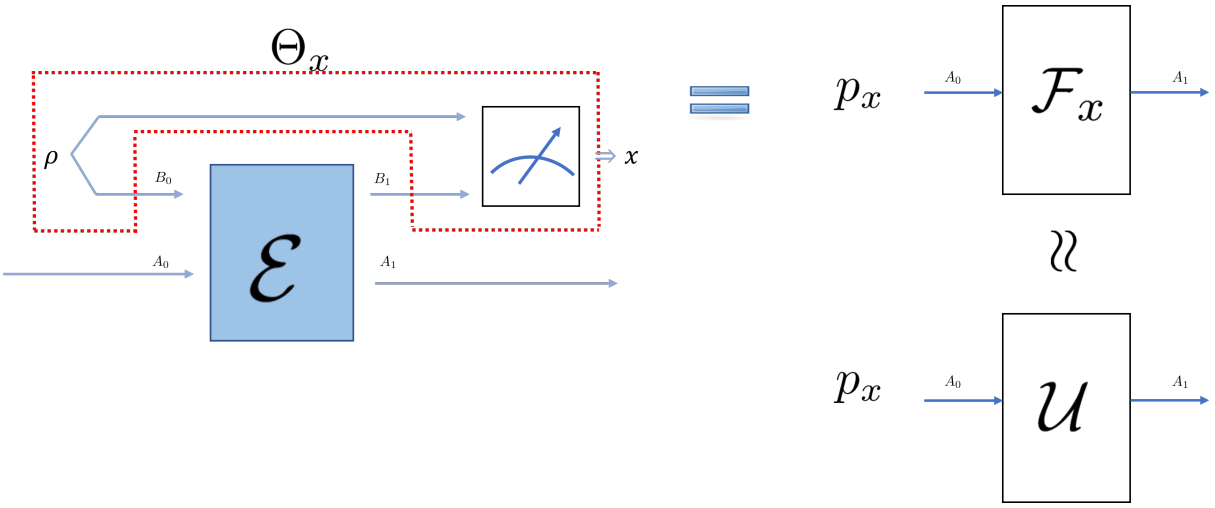

Figure 6.4: Probabilistic Universal Quantum Processor.

where p_x is the probability of obtaining the measurement outcome x by implementing Θ_x, and $\mathcal{F}_x$ represents the quantum channel conditioned on receiving x.

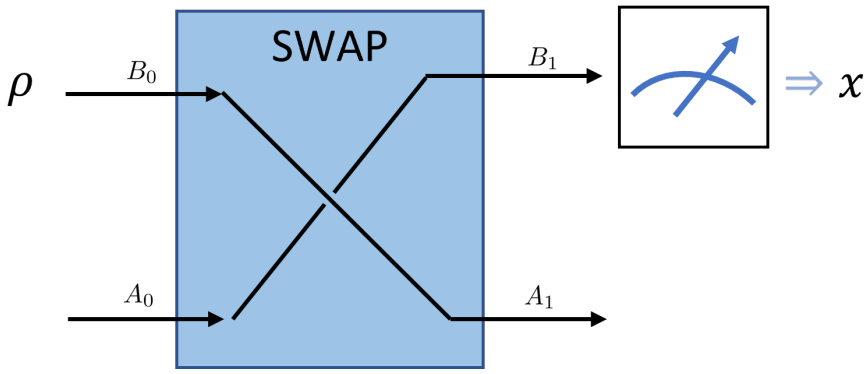

Figure 6.5: Example of probabilistic universal quantum programming with quantum processor $\mathcal{E} = \text{SWAP}$. Here, the PPOVM Θ_x is chosen as $\{\rho_{B_0}, M_x = |x\rangle\langle x|\}$. In this case, the probability distribution associated with measurement outcome x depends on the input state on system A_0.

We remark that, generally, the probability p_x not only depends on the quantum processor $\mathcal{E}$ and the manipulation process Θ_x, but also on the input state $\sigma_{A_0} \in \mathfrak{D}(A_0)$. For instance, consider the following example: the quantum processor $\mathcal{E}$ is chosen as the SWAP gate between systems A and B (see Fig. 6.5). Meanwhile, we set the PPOVM $\{\Theta_x\}_x$ as $\{\rho_{B_0}, |x\rangle\langle x|\}_x$. In this case, it is straightforward to check that the probability distribution p_x associated with the measurement outcome x depends on the input state of the system A_0. Moreover, p_x can be any real numbers between 0 and 1. In particular, by preparing the quantum state $\sigma_{A_0} = \sum_x \lambda_x |x\rangle\langle x|$, we see that $p_x = \lambda_x$ holds for any probability distribution $\{\lambda_x\}_x$. Hence, even though the quantum processor $\mathcal{E}$ and PPOVM $\{\Theta_x\}_x$ are already given, it is generally impossible to determine the value of p_x when the input state on system A_0 is absent. Let us assume that the input state

on A_0 system is σ_{A_0}, then the probability distribution p_x is given by

$$p_x = \mathrm{Tr}[\mathcal{E}_{A_0B_0\to A_1B_1}(\sigma_{A_0}\otimes\rho_{B_0R})\cdot I_{A_1}\otimes M_x]. \tag{6.16}$$

Similarly, the conditional channel $\mathcal{F}_x$ not only depends on the quantum processor $\mathcal{E}$ and the manipulation process Θ_x, but also on the input state in system A_0. Assume the input state is $\sigma_{A_0}\in\mathfrak{D}(A_0)$, it then follows that

$$\mathcal{F}_x(\sigma_{A_0}) = \frac{\mathrm{Tr}_{B_1R}[\mathcal{E}_{A_0B_0\to A_1B_1}(\sigma_{A_0}\otimes\rho_{B_0R})\cdot I_{A_1}\otimes M_x]}{\mathrm{Tr}[\mathcal{E}_{A_0B_0\to A_1B_1}(\sigma_{A_0}\otimes\rho_{B_0R})\cdot I_{A_1}\otimes M_x]} = \frac{\Theta_x[\mathcal{E}](\sigma_{A_0})}{p_x}. \tag{6.17}$$

Then the channel fidelity between $\mathcal{F}_x$ and $\mathcal{U}$ is given by

$$F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U}) = \frac{1}{|A_0|^2}\mathrm{Tr}[J^{\mathcal{F}_x}\cdot J^{\mathcal{U}}]. \tag{6.18}$$

Using this, we get a bound on the quantity $p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})$ as

$$\begin{aligned} p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U}) &= \frac{1}{|A_0|^2}\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}\cdot J^{\mathcal{U}}] \\ &= \frac{1}{|A_0|^2}\mathrm{Tr}[(J^{\Theta_x}_{B_0B_1})^{\mathbf{T}_{B_0B_1}}\cdot J^{\mathcal{E}}_{A_0A_1B_0B_1}\cdot J^{\mathcal{U}}_{A_0A_1}] \\ &\leqslant \max_{\Theta_x\in\mathrm{PPOVM}}\frac{1}{|A_0|^2}\mathrm{Tr}[(J^{\Theta_x}_{B_0B_1})^{\mathbf{T}_{B_0B_1}}\cdot J^{\mathcal{E}}_{A_0A_1B_0B_1}\cdot J^{\mathcal{U}}_{A_0A_1}]. \end{aligned} \tag{6.19}$$

Denote the operator $N(\mathcal{E},\mathcal{U})$, or briefly N, as

$$N(\mathcal{E},\mathcal{U}) := \frac{1}{|A_0|^2}\mathrm{Tr}_{A_0A_1}[J^{\mathcal{E}}_{A_0A_1B_0B_1}\cdot J^{\mathcal{U}}_{A_0A_1}]. \tag{6.20}$$

Now the upper bound of $p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})$ becomes

$$\begin{aligned} \max\quad & \mathrm{Tr}[J^{\Theta_x}\cdot N] \\ \text{s.t.}\quad & J^{\Theta_x}\geqslant 0,\ J^{\Theta_x}\leqslant\rho_{B_0}\otimes I_{B_1},\ \mathrm{Tr}[\rho_{B_0}]=1, \end{aligned} \tag{6.21}$$

which equals to

$$\begin{aligned} \min\quad & \lambda \\ \text{s.t.}\quad & X\geqslant N,\ \mathrm{Tr}_{B_1}[X]=\lambda\ I_{B_0}. \end{aligned} \tag{6.22}$$

Without loss of generality, we assume that the optimal value of Eq. (6.22) is achieved by $\lambda^*(N)$; that is

$$\begin{aligned} \lambda^*(N) := \min \quad & \lambda \\ \text{s.t.} \quad & X \geqslant N,\ \mathrm{Tr}_{B_1}[X] = \lambda\ I_{B_0}, \end{aligned} \tag{6.23}$$

which depends only on the operator N. Consider another optimization problem

$$\begin{aligned} \mu^*(N) := \min \quad & \mu \\ \text{s.t.} \quad & Y \geqslant N,\ \mathrm{Tr}_{B_1}[Y] \leqslant \mu\ I_{B_0}, \end{aligned} \tag{6.24}$$

with the optimal value being denoted as $\mu^*(N)$. Now it is straightforward to check that

$$\lambda^*(N) \geqslant \mu^*(N). \tag{6.25}$$

On the other hand, suppose that $\mu^*(N)$ in Eq. (6.24) is achieved by some solution Y^*, then we can construct the following operator

$$X^* := Y^* + (\mu^*(N)\ I_{B_0} - \mathrm{Tr}_{B_1}[Y^*]) \otimes \tau_{B_1}, \tag{6.26}$$

with τ_{B_1} being an arbitrary quantum state acting on system B_1. And the operator X^* satisfies the following properties

$$X^* \geqslant N, \tag{6.27}$$

$$\mathrm{Tr}_{B_1}[X^*] = \mu^*(N)\ I_{B_0}, \tag{6.28}$$

which are exactly the restrictions of Eq. (6.23). Thus, we have

$$\lambda^*(N) \leqslant \mu^*(N). \tag{6.29}$$

Combining Eq. (6.25) with 6.29, we obtain the following lemma,

Lemma 6.5. *For the operator N defined in Eq. (6.20), the optimization problems Eq. (6.23) and (6.24) are equivalent; namely*

$$\lambda^*(N) = \mu^*(N). \tag{6.30}$$

It is now straightforward to check that $\{Y = N, \mu = \lambda_1(\mathrm{Tr}_{B_1}[N])\}$ forms a dual feasible solution of the optimization problem Eq. (6.21), and $\{J^{\Theta_x} = \rho_{B_0} \otimes I_{B_1}, \rho_{B_0} = |\phi\rangle\langle\phi|\}$, with $\mathrm{Tr}_{B_1}[N]|\phi\rangle = \lambda_1(\mathrm{Tr}_{B_1}[N])|\phi\rangle$, forms a feasible solution of the primal formulation of Eq. (6.21). This implies that

$$\lambda_1(\mathrm{Tr}_{B_1}[N]) \geqslant \mu^*(N) \geqslant \lambda_1(\mathrm{Tr}_{B_1}[N]). \tag{6.31}$$

Namely,

$$\mu^*(N) = \lambda_1(\mathrm{Tr}_{B_1}[N]). \tag{6.32}$$

The previous proof depends heavily on the duality of SDPs. Alternatively, Eq. (6.21) can be directly solved as follows. We can assume ρ_{B_i} is a full rank density matrix, because otherwise, one can always add an arbitrary small perturbation to make it full rank. Then, by introducing the new variable $\tilde{J} \coloneqq (\rho_{B_i}^{-1/2} \otimes I)\, J^{\Theta_x}\, (\rho_{B_i}^{-1/2} \otimes I)$, the constraint of the optimization program can equivalently be written as

$$0 \leqslant \tilde{J} \leqslant I_{B_i B_o},\ \mathrm{Tr}[\rho_{B_i}] = 1,\ \rho_{B_i} \geqslant 0. \tag{6.33}$$

With this new variable, the objective function becomes

$$\mathrm{Tr}[J^{\Theta_x} \cdot N] = \mathrm{Tr}[(\rho_{B_i}^{1/2} \otimes I)\, \tilde{J}\, (\rho_{B_i}^{1/2} \otimes I) N] \tag{6.34}$$

$$= \mathrm{Tr}[\tilde{J}\, (\rho_{B_i}^{1/2} \otimes I) N\, (\rho_{B_i}^{1/2} \otimes I)]. \tag{6.35}$$

Since $(\rho_{B_i}^{1/2} \otimes I) N\, (\rho_{B_i}^{1/2} \otimes I) \geqslant 0$, this is clearly maximized at $\tilde{J} = I_{B_i B_o}$, which turns the objective function into

$$\mathrm{Tr}[(\rho_{B_i}^{1/2} \otimes I) N\, (\rho_{B_i}^{1/2} \otimes I)] = \mathrm{Tr}[(\rho_{B_i} \otimes I)\, N] = \mathrm{Tr}_{B_i}[\rho_{B_i} \mathrm{Tr}_{B_o} N]. \tag{6.36}$$

Maximizing this over ρ_{B_i} under the conditions in (6.33) then results in $\lambda_1(\mathrm{Tr}_{B_o}[N])$.

Thus, for any probabilistic physical process Θ_x, representing an element of PPOVM, when applied on a fixed processor, we get the following proposition between the highest success probability and the channel fidelity between the target unitary and resultant channel.

Proposition 7. Given a bipartite quantum processor $\mathcal{E} \in \mathrm{CPTP}(A_0 B_0 \to A_1 B_1)$ and a target quantum channel $\mathcal{U} \in \mathrm{CPTP}(A_0 \to A_1)$, denote the output quantum operation under PPOVM Θ_x as $p_x \mathcal{F}_x$, where p_x stands for the success probability and $\mathcal{F}_x$ represents the output channel. Then, for any input state on

system A_0, the success probability p_x and channel fidelity $F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})$ in probabilistic programming scenario satisfy the following relation

$$p_x F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U}) \leqslant \lambda_1(\mathrm{Tr}_{B_1}[N]), \tag{6.37}$$

where N is defined in Eq. (6.20).

It now follows immediately that

Corollary 6.6. *Given a bipartite quantum processor $\mathcal{E} \in \mathrm{CPTP}(A_0B_0 \to A_1B_1)$ and a target quantum channel $\mathcal{U}$, if there exist a PPOVM $\Theta_x = \{\rho_{B_0R}, M_x\}$ and a input state σ_{A_0} such that $F_C(\mathcal{F}_x,\mathcal{U}) \geqslant 1-\epsilon$ for target gate $\mathcal{U}$, then*

$$\epsilon \geqslant 1 - \frac{\lambda_1(\mathrm{Tr}_{B_1}[N])}{p_x}, \tag{6.38}$$

where p_x is given by Eq. (6.16), and N is defined in Eq. (6.20).

We remark that Eq. (6.31) implies that, to achieve the optimal bound for successful probability and performance, Bob should select $|\phi\rangle$ satisfying

$$\mathrm{Tr}_{B_1}[N]|\phi\rangle = \lambda_1(\mathrm{Tr}_{B_1}[N])|\phi\rangle, \tag{6.39}$$

as the program state. Let us move on to investigating the relation between the operator M defined in Eq. (6.8) and the operator N defined in Eq. (6.20), we thus see

$$\mathrm{Tr}_{B_1}[N] = M. \tag{6.40}$$

Hence, the optimal bound for deterministic protocols obtained in Thm. 6.3 coincides with the optimal bound of the product of successful probability and channel fidelity in probabilistic protocols derived in proposition 7. In other words, one of the optimal strategies for Bob to achieve the highest $p_x F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})$ is a deterministic protocol consisted of $\{|\psi\rangle\langle\psi|_{B_0}, \mathrm{Tr}_{B_1}\}$.

6.2.4 Average performance of a quantum processor

To quantify the performance of a processor $\mathcal{E}$ with respect to a desired unitary channel $\mathcal{U}$, we choose the average performance of $\mathcal{E}$ as our figure of merit. We consider the average performance because finding the fidelity between two channels by optimizing over all quantum states is very hard to compute. The average

performance is a good indicator of how well a unitary channel can be simulated by a fixed processor. It can also be used to compare how well two processors can simulate a target unitary. The performance of $\mathcal{E} \in \mathrm{CPTP}(A_0B_0 \to A_1B_1)$ averaged over all input states $\psi \in \mathfrak{D}(A_0)$ is given as follows

$$\overline{P_x(\mathcal{E},\mathcal{U})} := \int d\psi \, p_x^\psi F(\mathcal{F}_x(\psi),\mathcal{U}(\psi)), \tag{6.41}$$

where the resultant channel $\mathcal{F}_x(\psi)$ is given by

$$\mathcal{F}_x(\psi) = \frac{\mathrm{Tr}_{B_1}[M_x \cdot \mathcal{E}(\psi \otimes \rho)]}{p_x^\psi}, \tag{6.42}$$

and outcome x happens with probability p_x^ψ, which is characterized by

$$p_x^\psi = \mathrm{Tr}[M_x \cdot \mathcal{E}(\psi \otimes \rho)]. \tag{6.43}$$

Here, F stands for the Uhlmann's fidelity for quantum states, ρ is the program state, and $d\psi$ denotes the Haar measure, i.e. $\int d\psi = 1$. Another useful quantity is the so-called average fidelity $F_{\mathrm{A}}(\mathcal{E},\mathcal{U})$, which is defined as

Definition 6.7 (Average Fidelity)**.** Given a quantum channel $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ and a unitary channel $\mathcal{U} \in \mathrm{CPTP}(A_0 \to A_1)$, we define the average fidelity between them as

$$F_{\mathrm{A}}(\mathcal{E},\mathcal{U}) := \int d\psi \langle\psi|\mathcal{U}^\dagger \circ \mathcal{E}(\psi)|\psi\rangle, \tag{6.44}$$

where $\mathcal{U}^\dagger(\cdot) = U^\dagger(\cdot)U$.

Here, the average fidelity quantifies how well the channel $\mathcal{E}$ simulates a desired gate $\mathcal{U}$, and is closely related with the channel fidelity defined in Def. 6.2. More precisely, they satisfy the following relation:

Lemma 6.8. *Given a quantum channel $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$ and a unitary channel $\mathcal{U} \in \mathrm{CPTP}(A_0 \to A_1)$, their average fidelity $F_A(\mathcal{E},\mathcal{U})$ and channel fidelity $F_C(\mathcal{E},\mathcal{U})$ are connected through the following equation*

$$F_A(\mathcal{E},\mathcal{U}) = \frac{|A_0|F_C(\mathcal{E},\mathcal{U}) + 1}{|A_0| + 1}. \tag{6.45}$$

Note that, above Lem. 6.8 was first proved by M. Horodecki, P. Horodecki, and R. Horodecki in Ref. [328]. (See also Ref. [329] for a simpler proof.)

To simplify the formula of Eq. (6.41), we use the fact that $\overline{P_x(\mathcal{E},\mathcal{U})}$ is invariant under channel twirling operation $\mathcal{T}$:

Definition 6.9 (Channel Twirling)**.** Given a quantum channel $\mathcal{E} \in \mathrm{CPTP}(A_0 \to A_1)$, with $|A_0| = |A_1|$, we define the channel twirling $\mathcal{T}$ acting on $\mathcal{E}$ as

$$\mathcal{T}(\mathcal{E}) := \int dU\, \mathcal{U}^\dagger \circ \mathcal{E} \circ \mathcal{U}, \tag{6.46}$$

where $\mathcal{U}^\dagger(\cdot) = U^\dagger(\cdot)U$, and dU stands for the Haar measure over special unitary group $\mathrm{SU}(|A_0|)$.

Now it is straightforward to check that

$$\begin{aligned}\int d\psi \langle\psi|\mathcal{T} \circ \mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}](\psi)|\psi\rangle &= \int d\psi \int dV \langle\psi|V^\dagger \mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}](V\psi V^\dagger)V|\psi\rangle \\ &= \int d\psi \langle\psi|\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}](\psi)|\psi\rangle \\ &= \overline{P_x(\mathcal{E},\mathcal{U})}. \end{aligned} \tag{6.47}$$

On the other hand, the Choi operator of $\mathcal{T}$ is given by

$$J^{\mathcal{T}} = \Phi^+ \otimes \Phi^+ + \frac{1}{|A_0|^2 - 1}(I - \Phi^+) \otimes (I - \Phi^+), \tag{6.48}$$

where Φ^+ represents the maximally entangled state. Thus, the average performance $\overline{P_x(\mathcal{E},\mathcal{U})}$ can be rewritten as

$$\overline{P_x(\mathcal{E},\mathcal{U})} = \int d\psi \langle\psi|\mathcal{T} \circ \mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}](\psi)|\psi\rangle \tag{6.49}$$

$$= \int d\psi \langle\psi|J^{\mathcal{T}} \star J^{\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}]} \star \psi|\psi\rangle \tag{6.50}$$

$$= \int d\psi \langle\psi|[(\Phi^+ \star J^{\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}]})\Phi^+ + ((I - \Phi^+) \star J^{\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}]})\frac{I - \Phi^+}{|A_0|^2 - 1}] \star \psi|\psi\rangle. \tag{6.51}$$

For the sake of convenience, we define a_x and b_x as

$$a_x := \Phi^+ \star J^{\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}]} = \frac{1}{|A_0|}\mathrm{Tr}[J^{\mathcal{U}} \cdot J^{\Theta_x[\mathcal{E}]}] = |A_0| p_x F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U}), \tag{6.52}$$

$$b_x := \mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]. \tag{6.53}$$

Equipped with this notation, Eq. (6.51) can be further simplified as

$$\overline{P_x(\mathcal{E},\mathcal{U})} = \int d\psi \langle\psi|(\frac{b_x - a_x}{|A_0|^2 - 1} I + \frac{a_x|A_0|^2 - b_x}{|A_0|^2 - 1}\Phi^+) \star \psi|\psi\rangle \tag{6.54}$$

$$= \frac{a_x}{|A_0| + 1} + \frac{b_x}{|A_0|(|A_0| + 1)}. \tag{6.55}$$

Hence, for each measurement outcome x, the average performance is upper bounded by the channel fidelity and the trace of $J^{\Theta_x\circ\mathcal{E}}$; that is

$$\overline{P_x(\mathcal{E},\mathcal{U})} = \frac{|A_0|p_x F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})}{|A_0|+1} + \frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)} \tag{6.56}$$

$$\leqslant \max_{\Theta_x\in\mathrm{PPOVM}} \left(\frac{|A_0|p_x F_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})}{|A_0|+1}\right) + \max_{\Theta_x\in\mathrm{PPOVM}} \left(\frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)}\right) \tag{6.57}$$

$$= \frac{|A_0|\lambda_1(\mathrm{Tr}_{B_1}[N])}{|A_0|+1} + \frac{|A_0|}{|A_0|(|A_0|+1)} \tag{6.58}$$

$$= \frac{|A_0|\lambda_1(\mathrm{Tr}_{B_1}[N])+1}{|A_0|+1}, \tag{6.59}$$

where Eq. (6.58) is obtained by applying our proposition 7. To conclude, we have the following theorem.

Theorem 6.10. *Given a bipartite quantum processor $\mathcal{E}\in\mathrm{CPTP}(A_0B_0\to A_1B_1)$ and a target quantum channel $\mathcal{U}\in\mathrm{CPTP}(A_0\to A_1)$. Denote the output quantum operation under PPOVM Θ_x as $p_x^\psi\mathcal{F}_x$, where $\psi\in\mathfrak{D}(A_0)$, p_x^ψ stands for the success probability and $\mathcal{F}_x$ represents the output channel. Then, the average performance $\overline{P_x(\mathcal{E},\mathcal{U})}$ has the following upper bound*

$$\int d\psi\, p_x^\psi F(\mathcal{F}_x(\psi),\mathcal{U}(\psi)) \leqslant \frac{|A_0|\lambda_1(\mathrm{Tr}_{B_1}[N])+1}{|A_0|+1}, \tag{6.60}$$

where N is defined in Eq. (6.20).

Note that here the equality can be achieved by taking deterministic protocols. Moreover, Eq. (6.56) is equivalent to

$$\int d\psi\langle\psi|\mathcal{U}^\dagger\circ\Theta_x[\mathcal{E}](\psi)|\psi\rangle = \frac{|A_0|\langle\Phi^+|\mathcal{U}^\dagger\circ\Theta_x[\mathcal{E}](\Phi^+)|\Phi^+\rangle}{|A_0|+1} + \frac{\mathrm{Tr}[J^{\mathcal{U}^\dagger\circ\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)}. \tag{6.61}$$

It is straightforward to check that the same equation holds for any completely positive (CP) and trace-non-increasing (TNI) map Λ_x,

$$\int d\psi\langle\psi|\Lambda_x(\psi)|\psi\rangle = \frac{|A_0|\langle\Phi^+|\Lambda_x(\Phi^+)|\Phi^+\rangle}{|A_0|+1} + \frac{\mathrm{Tr}[J^{\Lambda_x}]}{|A_0|(|A_0|+1)}, \tag{6.62}$$

which generalizes Lem. 6.8 investigated in Ref. [328].

Besides the average performance, we are also interested in the average probability $\overline{p_x} := \int d\psi\, p_x^\psi$. Written explicitly, we have

$$\overline{p_x} := \int d\psi\, p_x^\psi = \int d\psi \mathrm{Tr}[\Theta_x[\mathcal{E}](\psi)], \tag{6.63}$$

which is also invariant under channel twirling as

$$\int d\psi \mathrm{Tr}[\mathcal{T} \circ \Theta_x[\mathcal{E}](\psi)] = \int d\psi \int dU \mathrm{Tr}[\mathcal{U}^\dagger \circ \Theta_x[\mathcal{E}] \circ \mathcal{U}(\psi)] \tag{6.64}$$

$$= \int dU \int d\psi \mathrm{Tr}[\Theta_x[\mathcal{E}] \circ \mathcal{U}(\psi)] \tag{6.65}$$

$$= \int d\psi \mathrm{Tr}[\Theta_x \circ \mathcal{E}(\psi)]. \tag{6.66}$$

It now follows immediately that

$$\int d\psi \, p_x^\psi = \int d\psi \mathrm{Tr}[\mathcal{T} \circ \Theta_x[\mathcal{E}](\psi)] \tag{6.67}$$

$$= \int d\psi \, \left(I \star J^{\mathcal{T}} \star J^{\Theta_x[\mathcal{E}]} \star \psi \right) \tag{6.68}$$

$$= \int d\psi \, \left(I \star \left[(\Phi^+ \star J^{\Theta_x[\mathcal{E}]})\Phi^+ + ((I - \Phi^+) \star J^{\Theta_x[\mathcal{E}]}) \frac{I - \Phi^+}{|A_0|^2 - 1} \right] \star \psi \right). \tag{6.69}$$

Similar to the definitions of Eqs. 6.52 and 6.53, we define the quantity c_x as

$$c_x := \Phi^+ \star J^{\Theta_x[\mathcal{E}]}. \tag{6.70}$$

Note that in general $a_x \neq c_x$. Now $\int d\psi \, p_x^\psi$ can be further simplified as

$$\int d\psi \, p_x^\psi = \int d\psi \, \left(I \star \left[\frac{c_x |A_0|^2 - b_x}{|A_0|^2 - 1} \Phi^+ + \frac{b_x - c_x}{|A_0|^2 - 1} I \right] \star \psi \right) \tag{6.71}$$

$$= \int d\psi \, \left(\left[\frac{c_x |A_0|^2 - b_x}{|A_0|(|A_0|^2 - 1)} I + \frac{|A_0|(b_x - c_x)}{|A_0|^2 - 1} I \right] \star \psi \right) \tag{6.72}$$

$$= \frac{b_x}{|A_0|}. \tag{6.73}$$

Then we have the following lemma.

Lemma 6.11. *Given a quantum channel $\mathcal{E}$ and a protocol Θ_x, the average probability of obtaining x over all input states $\psi \in \mathcal{D}(A_0)$ is characterized by the following equation*

$$\int d\psi \, \mathrm{Tr}[\Theta_x[\mathcal{E}](\psi)] = \frac{b_x}{|A_0|} = \frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|}. \tag{6.74}$$

This implies the following corollary

Corollary 6.12. *Given a bipartite quantum processor $\mathcal{E} \in \mathrm{CPTP}(A_0 B_0 \to A_1 B_1)$ and a target quantum channel $\mathcal{U}$, if there exists a PPOVM $\Theta_x = \{\rho_{B_0 R}, M_x\}$ such that $F(\mathcal{F}_x^\psi(\psi), \mathcal{U}(\psi)) \geqslant 1 - \epsilon$ for any input state*

ψ, then

$$\epsilon \geqslant 1 - \frac{|A_0|(|A_0|\lambda_1(\mathrm{Tr}_{B_1}[N])+1)}{b_x(|A_0|+1)} = 1 - \frac{|A_0|(|A_0|\lambda_1(\mathrm{Tr}_{B_1}[N])+1)}{\mathrm{Tr}[J^{\Theta_x\circ\mathcal{E}}](|A_0|+1)}, \tag{6.75}$$

where N is defined in Eq. (6.20).

Here the proof is a simple combination of Thm. 6.10 and Lem. 6.11.

Note that the above limitation on the error in Eq. (6.75) depends on Θ_x and thus can be improved. In finding the upper bound of the average performance $\overline{P_x(\mathcal{E},\mathcal{U})}$, rather than taking the maximization over all of RHS of Eq. (6.56), if we only take the maximization over either of the terms on RHS, we can get a tighter bound. That is, let

$$m = \min\left(\max_{\Theta_x\in\mathrm{PPOVM}}\left(\frac{|A_0|p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})}{|A_0|+1}\right) + \frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)}, \frac{|A_0|p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})}{|A_0|+1} + \max_{\Theta_x\in\mathrm{PPOVM}}\left(\frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)}\right)\right) \tag{6.76}$$

$$= \min\left(\frac{|A_0|\lambda_1(\mathrm{Tr}_{B_o}[N])}{|A_0|+1} + \frac{\mathrm{Tr}[J^{\Theta_x[\mathcal{E}]}]}{|A_0|(|A_0|+1)}, \frac{|A_0|p_xF_{\mathrm{C}}(\mathcal{F}_x,\mathcal{U})}{|A_0|+1} + \frac{|A_0|}{|A_0|(|A_0|+1)}\right). \tag{6.77}$$

Then the error is lower bounded by

$$\epsilon \geqslant 1 - \frac{|A_0|m}{\mathrm{Tr}[J^{\Theta_x\circ\mathcal{E}}]}.$$

6.2.5 Trade-off between success probability and average fidelity error

Another advantage of using the average fidelity is that we can express the trade-off between success probability and average fidelity error as a semidefinite program. Since we want to maximize the success probability p_x of simulating a unitary channel $\mathcal{U}$, keeping the average fidelity between the resultant channel $\mathcal{F}_x$ and $\mathcal{U}$ lower bounded by $1-\epsilon$, we can express this maximization as an optimization problem as

$$\max \quad p_x \tag{6.78}$$

$$\text{s.t.} \quad \Theta_x[\mathcal{E}] = p_x\mathcal{F}_x \tag{6.79}$$

$$F_A(\mathcal{F}_x,\mathcal{U}) \geqslant 1-\epsilon \tag{6.80}$$

Now using the relation between average and channel fidelity, we can write the maximum success probability as

$$\max \quad p_x \tag{6.81}$$

$$\text{s.t.} \quad J^{\Theta_x[\mathcal{E}]} = p_x \mathcal{F}_x, \tag{6.82}$$

$$\frac{\mathrm{Tr}\left[J^{\Theta_x[\mathcal{E}]} J^{\mathcal{U}}\right] + |A_i|}{|A_i|(|A_i|+1)} \geqslant 1-\epsilon \tag{6.83}$$

If we define $\omega_{A_i A_o} := p_x J^{\mathcal{F}_x}_{A_i A_o}$, then the above optimization problem can be expresed as

$$\max \quad \frac{\mathrm{Tr}[\omega]}{|A_0|} \tag{6.84}$$

$$\text{s.t.} \quad J^{\Theta_x} \star J^{\varepsilon} = \omega, \tag{6.85}$$

$$\frac{\frac{|A_i|}{\mathrm{Tr}[\omega]} \mathrm{Tr}[\omega J^{\mathcal{U}}] + |A_i|}{|A_i||A_i+1|} \geqslant 1-\epsilon, \tag{6.86}$$

$$J^{\Theta_x}_{B_i B_o} \leqslant \rho_{B_i} \otimes I_{B_o}, \omega_{A_i A_o} \geqslant 0, \omega_{A_i} = \frac{I_{A_i}}{|A_i|} \mathrm{Tr}[\omega] \tag{6.87}$$

which becomes the following semidefinite program

$$\max \quad \frac{\mathrm{Tr}[\omega]}{|A_0|} \tag{6.88}$$

$$\text{s.t.} \quad J^{\Theta_x}_B \star J^{\mathcal{E}}_{AB} = \omega, \tag{6.89}$$

$$J^{\Theta_x} \leqslant \rho_{B_i} \otimes I_{B_o}, \tag{6.90}$$

$$\frac{1}{\mathrm{Tr}[\omega]} \mathrm{Tr}_r\left[\omega_A J^{\mathcal{U}}_A\right] \geqslant |A_i| - \epsilon(|A_i|+1) \tag{6.91}$$

$$\omega_A \geqslant 0, \ \omega_{A_i} = \frac{I_{A_i}}{|A_i|} \mathrm{Tr}[\omega] \tag{6.92}$$

6.3 Resource-theoretic bounds for quantum programming

Practically, the agent Bob might not have access to all possible quantum operations and measurements on his side due to the restrictions on his laboratory. It is thus quite natural to apply a quantum resource-theoretic framework to the study of universal quantum programming. In this subsection, we investigate the fundamental trade-off between successful probability and performance from a perspective of dynamical coherence, where all classical communications, i.e. classical channels, are 'free' [37].

6.3.1 Limitations in static coherence

In this subsection, we investigate the case where all the program states $\{\rho_{\mathcal{U}}\}$ are incoherent, see Fig. 6.3. The channel fidelity between $J^{\mathcal{E}} \star (\rho_{\mathcal{U}} \otimes I_{B_1})$ and the target gate $\mathcal{U}$ is upper bounded by

$$\begin{aligned} \max \quad & \mathrm{Tr}[M(\mathcal{E},\mathcal{U}) \cdot \rho] \\ \text{s.t.} \quad & \rho \geqslant 0,\ \mathrm{Tr}[\rho] = 1,\ \mathcal{D}_{B_0}(\rho) = \rho, \end{aligned} \tag{6.93}$$

whose dual is given by

$$\begin{aligned} \iota^*(M) := \min \quad & \iota \\ \text{s.t.} \quad & \iota I \geqslant M + Z,\ \mathcal{D}_{B_0}(Z) = \mathbf{0},\ Z^\dagger = Z. \end{aligned} \tag{6.94}$$

Written in full, that is

Theorem 6.13 (Optimal Performance with Incoherent Program States)**.** *Given a bipartite quantum processor $\mathcal{E} \in \mathrm{CPTP}(A_0B_0 \to A_1B_1)$ and a target quantum gate $\mathcal{U}$, then the optimal performance of universal quantum programming with incoherent program states $\rho_{\mathcal{U}}$ is upper bounded by*

$$F_C(\mathcal{F},\mathcal{U}) \leqslant \iota^*(M), \tag{6.95}$$

where $\mathcal{F} := \mathrm{Tr}_{B_1} \circ \mathcal{E}(\rho_{\mathcal{U}})$ and $\iota^*(M)$ is defined in Eq. (6.94).

Then we get the following limitation on the performance with incoherence program states.

Corollary 6.14. *If there exists an incoherent program state $\rho_{\mathcal{U}}$ such that $F_C(\mathrm{Tr}_{B_1} \circ \mathcal{E}(\rho_{\mathcal{U}}),\mathcal{U}) \geqslant 1 - \epsilon$ for target gate $\mathcal{U}$, then*

$$\epsilon \geqslant 1 - \iota^*(M). \tag{6.96}$$

6.3.2 Limitations in dynamical coherence

For probabilistic universal quantum programming in dynamical coherence, we start by introducing some relevant concepts in probabilistic causal networks and dynamical coherence. First of all, to manipulating quantum channels probabilistically, we need to implement the so-called *sub-superchannel*. For more information, refer to Refs. [122, 125, 330]. A completely positive map Θ is called a sub-superchannel, if there

exists a superchannel Ξ such that $\Xi \geqslant \Theta$. Here the greater-than-equal notation means $\Xi - \Theta$ is a CP map.

In chapter 3, we considered a set of free superchannels called MISC. Using this set of superchannels, we can now define a set of free sub-superchannels and call it maximally incoherent sub-superchannels (MIS_{sub}). Like the definition of MISC, we can define MIS_{sub} as follows

Definition 6.15 (Maximal Incoherent Sub-Superchannel (MIS_{sub}))**.** A sub-superchannel Θ taking quantum channels in $\text{CPTP}(B_0^{'} \to B_0)$ to completely positive trace-preserving maps in $\text{CPTNI}(B_1 \to B_1^{'})$ is said to be maximal incoherent sub-superchannel (MIS_{sub}) if it satisfies

$$\Delta_{B'} \circ \Theta \circ \Delta_B = \Theta \circ \Delta_B, \tag{6.97}$$

where $\Delta_{B'} := \mathcal{D}_{B_1^{'}} \otimes \mathcal{D}_{B_0^{'}}$, and $\Delta_B := \mathcal{D}_{B_1} \otimes \mathcal{D}_{B_0}$.

Denoting the set of all possible sub-superchannels from dynamical system $(B_0^{'} \to B_0)$ to dynamical system $(B_1 \to B_1^{'})$ by $\mathcal{S}_{\text{sub}}(B_0^{'} \to B_0, B_1 \to B_1^{'})$, and all probabilistic protocols for Bob – PPOVMs on B – by $\text{PPOVM}(\mathbb{C} \to B_0, B_1 \to \mathbb{C})$, it now follows immediately that

$$\text{PPOVM}(\mathbb{C} \to B_0, B_1 \to \mathbb{C}) \subseteq \mathcal{S}_{\text{sub}}(B_0^{'} \to B_0, B_1 \to B_1^{'}). \tag{6.98}$$

In the case of dynamical coherence, the free maximal incoherent process POVM (MIPPOVM) is then defined as

Definition 6.16 (Maximal Incoherent Process POVM (MIPPOVM))**.** A process POVM $\{\Theta_x\}_x$ from $(1 \to B_0)$ to $(B_1 \to \mathbb{C})$ is said to be maximal incoherent process POVM (MIPPOVM) if it satisfies

$$\Theta_x \in \text{PPOVM}(\mathbb{C} \to B_0, B_1 \to \mathbb{C}) \cap \text{MIS}_{\text{sub}}, \tag{6.99}$$

for all x.

As the earliest input system $B_0^{'}$ and the latest output system $B_1^{'}$ of PPOVMs from $(\mathbb{C} \to B_0)$ to $(B_1 \to \mathbb{C})$ are trivialized, i.e. $B_0^{'} = B_1^{'} = \mathbb{C}$, the set of all MIPPOVMs coincides with the set of all PPOVMs; that is

$$\text{MIPPOVM} = \text{PPOVM}. \tag{6.100}$$

Thus, the average performance of the processor under MIS_{sub} is same as using a PPOVM that we discussed earlier.

6.4 Numerical examples

By using the port-based teleportation processor as considered in [50], we first plot the lower bound of error using Equation (6.77) (which is tighter than Equation (6.75)) for a range of single-parameter family of unitary gates as target. A unitary gate belonging to this family has the form

$$U(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2)i \\ -\sin(\theta/2)i & \cos(\theta/2) \end{pmatrix}.$$

By fixing our PPOVM Θ_x, i.e., by fixing the input program state and by fixing the POVM element, the relation between the minimum error and the angle of the target unitary can be seen in Fig.6.6.

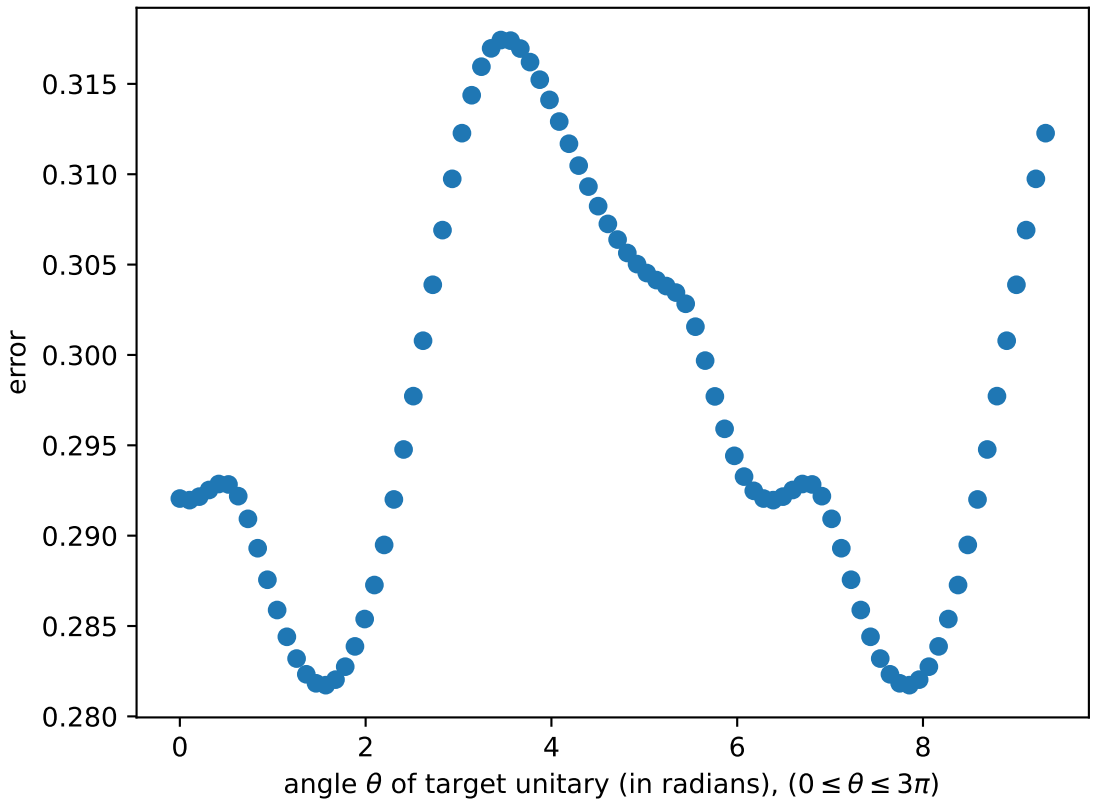

Figure 6.6: Variation of minimum error with varying target unitary keeping the program state and measurement fixed

Next, using the results of Section 6.2.5, we also plot the trade-off between probability and average error for a fixed target unitary gate for the case when the processor is the dephasing map and the case when the processor is the depolarizing map. The action of the dephasing map Δ_p on a state ρ_{A_0} is defined as

$$\Delta_p(\rho) := (1-p)\text{diag}(\rho) + p\rho \tag{6.101}$$

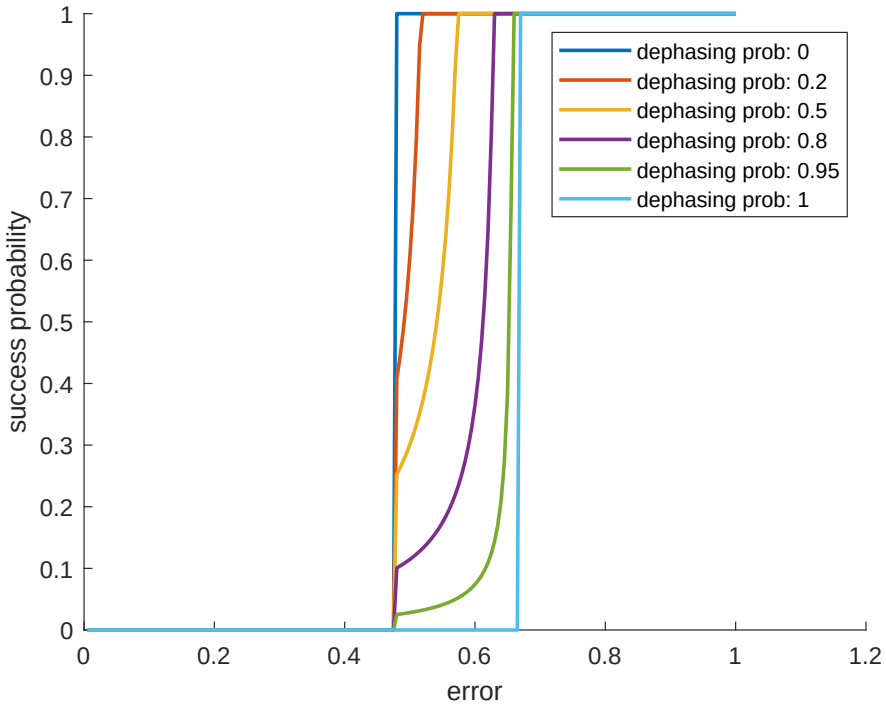

Figure 6.7: Trade-off between success probability and average fidelity error for different parameters of the dephasing channel

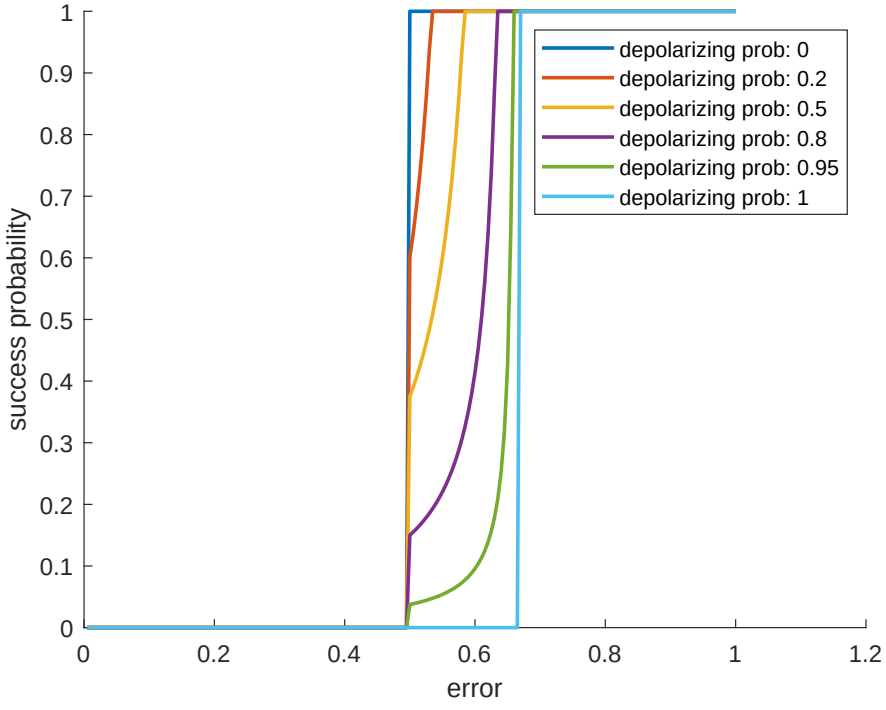

Figure 6.8: Trade-off between success probability and average fidelity error for different parameters of the depolarizing channel

where $p \in [0,1]$. If $p = 0$, we get a completely dephasing map $\mathcal{D}$ defined as $\mathcal{D}(\cdot) = \sum_i |i\rangle\langle i|(\cdot)|i\rangle\langle i| = \text{diag}(\cdot)$ and for $p = 1$, we get the identity channel. The action of the depolarizing map Δ'_p on a state ρ_{A_0} is defined as

$$\Delta'_p := (1-p)\frac{I}{|A_i|} + p\rho \tag{6.102}$$

where $p \in [0,1]$. Using these two channels as processors and a fixed unitary channel $\mathcal{U}$ whose Choi matrix $J^{\mathcal{U}}$ is given by

$$J^{\mathcal{U}} = \begin{pmatrix} 0.5636 & 0.4959 & 0.4959 & -0.5636 \\ 0.4959 & 0.4364 & 0.4364 & -0.4959 \\ 0.4959 & 0.4364 & 0.4364 & -0.4959 \\ -0.5636 & -0.4959 & -0.4959 & 0.5636 \end{pmatrix}, \tag{6.103}$$

we get the trade-off plots between the average fidelity error and the success probability for varying parameter p of the dephasing channel in Fig. 6.7 and that of the depolarizing channel in Fig. 6.8. Moreover, with the same target unitary channel we also plot the trade-off between probability and average fidelity error for a random unital channel with real entries, which is given in Fig. 6.9. The entries of the Choi matrix of the unital channel used as a processor are given in Fig. 6.10

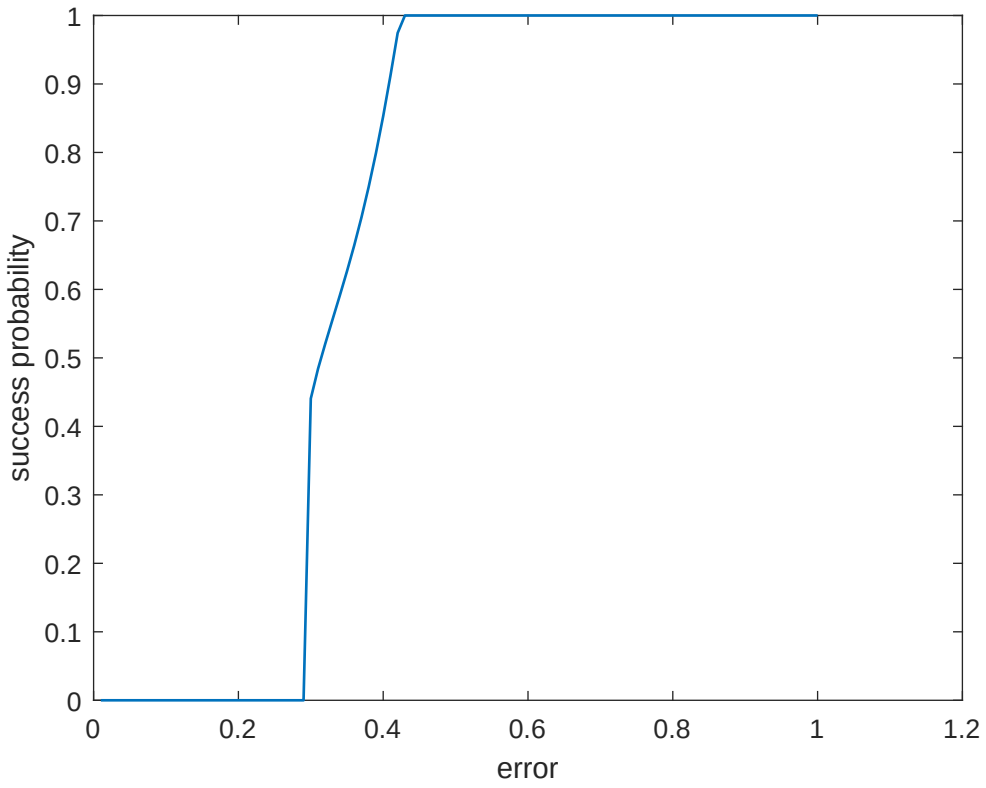

Figure 6.9: Trade-off between success probability and average fidelity error when a random unital channel is used as a processor

0.2367	-0.0479	-0.0074	0.023	0.0392	-0.0275	-0.0099	0.0064	0.0411	0.0686	0.0501	0.0256	0.0053	-0.0072	-0.0413	0.1088
-0.0479	0.2866	-0.0766	-0.0042	-0.0675	-0.089	-0.0397	0.0389	-0.111	-0.1085	-0.0105	0.1591	-0.0775	-0.0215	0.1017	0.0054
-0.0074	-0.0766	0.2813	-0.0339	0.0276	0.1947	0.0223	0.0858	0.0438	-0.0011	-0.01	-0.0885	-0.0292	-0.0126	0.0902	-0.0909
0.023	-0.0042	-0.0339	0.1954	-0.0115	0.0477	0.0667	0.0275	-0.0395	0.0479	0.0277	0.0773	0.0048	0.0764	-0.1692	-0.074
0.0392	-0.0675	0.0276	-0.0115	0.3047	-0.0106	-0.0268	0.0685	0.0322	-0.1053	0.0443	0.0646	-0.018	-0.0004	-0.1031	0.061
-0.0275	-0.089	0.1947	0.0477	-0.0106	0.2563	0.0841	0.0828	0.1161	0.0186	-0.0031	-0.12	-0.0109	0.091	-0.006	-0.1136
-0.0099	-0.0397	0.0223	0.0667	-0.0268	0.0841	0.2285	0.001	0.0048	-0.0162	-0.0682	0.0066	-0.0018	0.0538	-0.0191	-0.1318
0.0064	0.0389	0.0858	0.0275	0.0685	0.0828	0.001	0.2106	-0.0719	-0.1034	-0.0008	0.0174	0.0507	0.0355	-0.0293	-0.054
0.0411	-0.111	0.0438	-0.0395	0.0322	0.1161	0.0048	-0.0719	0.2624	0.0824	0.071	-0.1018	-0.0377	0.0774	0.0206	0.0819
0.0686	-0.1085	-0.0011	0.0479	-0.1053	0.0186	-0.0162	-0.1034	0.0824	0.3043	0.0627	-0.0149	0.0762	0.0293	-0.0055	0.0341
0.0501	-0.0105	-0.01	0.0277	0.0443	-0.0031	-0.0682	-0.0008	0.071	0.0627	0.1265	0.0598	-0.0163	0.0275	-0.0537	0.0622
0.0256	0.1591	-0.0885	0.0773	0.0646	-0.12	0.0066	0.0174	-0.1018	-0.0149	0.0598	0.3067	-0.0703	-0.0105	-0.048	0.0621
0.0053	-0.0775	-0.0292	0.0048	-0.018	-0.0109	-0.0018	0.0507	-0.0377	0.0762	-0.0163	-0.0703	0.1962	-0.0238	-0.0369	0.0103
-0.0072	-0.0215	-0.0126	0.0764	-0.0004	0.091	0.0538	0.0355	0.0774	0.0293	0.0275	-0.0105	-0.0238	0.1528	-0.0702	-0.0637
-0.0413	0.1017	0.0902	-0.1692	-0.1031	-0.006	-0.0191	-0.0293	0.0206	-0.0055	-0.0537	-0.048	-0.0369	-0.0702	0.3637	-0.027
0.1088	0.0054	-0.0909	-0.074	0.061	-0.1136	-0.1318	-0.054	0.0819	0.0341	0.0622	0.0621	0.0103	-0.0637	-0.027	0.2873

Figure 6.10: Entries of the Choi matrix of the Unital channel

Chapter 7

Summary and future outlook

In this chapter, I present a summary of my results followed by some open problems.

7.1 Summary of results

In chapter 3, we developed the resource theory of coherence present in quantum channels. In this resource theory, the underlying physical principle we followed was that the free dynamical objects cannot preserve or manipulate coherence. This led us to identify classical channels as the free elements in this theory. Thus, all other coherent quantum channels were classified as resources, including the identity channel as it preserves the coherence of quantum systems. To manipulate dynamical resources, we considered two free sets of superchannels, maximally incoherent superchannels (MISC) and dephasing covariant incoherent superchannels (DISC), and characterized these superchannels in terms of their Choi matrices in Sec. 3.3. After the partition of the set of channels as free and resources, we formulated various types of resource measures to quantify the coherence of quantum channels in Sec. 3.4. Our first resource measure was a complete family of monotones. If the value of all the monotones of this family for a channel is greater than that of another channel, then the former channel can be converted to the latter using the free superchannels. Next, in Sec. 3.4.2, we showed that there are only three out of the six relative entropies defined for channels that are dynamical coherence monotones under MISC. For DISC, we defined a new divergence-based monotone as the channel divergence between a given channel and the channel evolved under the dephasing superchannel. Our third measure was the max-relative entropy of coherence or the log-robustness of coherence for channels defined in Sec. 3.4.3. We showed that log-robustness of coherence is an operational monotone which was used to express the coherence cost of simulating a channel from the maximally coherent state. After formulating these various types of dynamical coherence measures, we solved some resource interconversion problems in

Sec. 3.5. First, in Sec. 3.5.1, we defined the conversion distance from a channel to another, and showed that it is an SDP and, hence, can be efficiently computed. If the conversion distance from one channel to another is zero, it implies that the former can be deterministically converted to the latter using free superchannels. After defining the conversion distance, we defined the coherence cost and the coherence distillation of a channel. The coherence cost of a channel is defined as the log of the minimum dimension of the maximally coherent state that can be converted to the channel using free superchannels. We found that the exact coherence cost of a channel can be expressed using the log-robustness of coherence. We also formulated expressions for approximate and liberal coherence costs of a channel. Lastly, we defined the single-shot exact distillable coherence as the log of the maximum dimension of the maximally coherent state that can be obtained from a given quantum channel. We showed that single-shot coherence distillation can be expressed as a non-linear optimization problem. However, when we allow for probabilistic distillation, the optimization problem can be expressed as an SDP.

In chapter 4, we developed a resource theory of multi-qubit magic channels. Unlike the odd-dimensional case where the negativity of the Wigner representation quantifies the magic of quantum channels, no such function has been formulated for the multi-qubit domain. Therefore, alternative approaches need to be taken to quantify multi-qubit magic resources. In our work, we identified the set of completely stabilizer preserving operations (CSPOs) as the set of free channels because CSPOs strictly contain the set of stabilizer operations and any circuit formed using CSPOs is efficiently simulable classically. The resources then, are all the other multi-qubit magic channels. For resource manipulation, we defined two sets of free superchannels in Sec. 4.3, namely, the CSPO preserving and the completely CSPO preserving superchannels. We characterized these sets of superchannels in terms of their Choi matrices. We found that a superchannel is a completely CSPO preserving superchannel if and only if its normalized Choi matrix is a stabilizer state. In Sec. 4.4, we generalized the key magic monotones from the state to the channel domain, namely, the generalized robustness of magic and the min relative entropy of magic. Using these monotones, we formulated the bounds on the cost and distillation of magic resources in Sec. 4.5. Unlike the resource theory of coherence, the resource theory of magic does not admit a unique maximal resource, and any pure magic state can be used as a resource to achieve universal quantum computation. So, we define the cost of simulating a magic channel using a pure magic state as the minimum copies of a given state that are required to convert it to the desired channel using free superchannels. Similarly, distillation of a pure magic state is defined as the maximum number of copies of the target state that can be obtained from a given channel. Apart from cost and distillation, we also found the conditions for interconversion among qubits under CSPOs in Sec. 4.5.1. We showed that it is a linear programming feasibility problem and hence can be efficiently solved. Lastly, we devised a classical simulation algorithm to estimate the expectation values of an observable in Sec. 4.6.

The runtime of this algorithm depends on the allowed error and the generalized robustness of the circuit elements.

In chapter 5, we worked on the problem of interconversion among pairs of channels. We investigated how such an interconversion is related to the testing region of the channels. The motivation behind this stems from the equivalence of relative and Lorenz majorization among pairs of probability distributions established by Blackwell in the mid of the twentieth century. A pair of probability distributions is said to Lorenz majorize another pair if the testing region, a convex symmetric region, of the former pair contains that of the latter. This equivalence was also shown to hold for a pair of qubits in Ref. [313]. The definition of testing region and Lorenz majorization was generalized for quantum states in Ref. [308], which were used as a tool to find conditions for interconversion among pairs of quantum states in various resource theories. In our work, in Sec. 5.3, we generalized these definitions from the state to the channel domain, and characterized Lorenz majorization among channels using diamond norm and Hilbert α divergences. We also addressed interconversion among pairs of classical channels in Sec. 5.4. We showed the equivalence of Lorenz majorization and relative majorization in two cases. In the first case, we showed that if a pair of classical states Lorenz majorizes the pair of classical channels, then the former pair can be converted to the latter. In the second case, we showed that any pair of classical channels can be converted to a pair of classical channels (with the latter pair's output being two-dimensional) if and only if the former pair Lorenz majorizes the other.

In chapter 6, we considered the problem of programming a bipartite processor to simulate a unitary channel. In the settings usually considered in the literature, the processor is scalable, thus allowing for universal computation. However, given the current scenario where it is hard to prevent errors upon scaling, we consider fixed processors, i.e., processors with fixed input and output dimensions. Further, suppose the bipartite processor is shared between Alice and Bob, where Bob needs to apply some unitary gate on his quantum system by using the fixed processor. To do that, Alice prepares a bipartite program state and inputs one of her systems into the processor. She uses the other system of her program state to make a joint measurement on the output state she got from the processor. Given this setting, we quantified the average performance of a processor to simulate a target unitary and found limitations on the performance of deterministic and probabilistic protocols in Sec. 6.2. The average performance helps in comparing two processors to simulate a target unitary, and the limitations help in identifying the optimal resources to achieve the best performance. We also formulated a trade-off relation between the success probability and the average fidelity error in Sec. 6.2.5. We showed that this trade-off can be expressed in terms of an SDP. In the end, we found limitations on simulating a unitary channel using a fixed processor for two cases when Alice has limited computational power. In one case, the restriction we imposed was that Alice can only

input classical states, and in the other case, Alice can only perform maximally incoherent operations. In the end, we numerically plotted the trade-off between success probability and average fidelity error for different processors.

7.2 Conclusion and open problems

In our work on the coherence of quantum channels in Chapter 3, we classified classical channels as free and all quantum channels as resources. To manipulate dynamical resources, we defined the free set of superchannels as the set of linear maps that preserve classical channels. It is not yet clear if there is an operational way to define another set of free superchannels. To quantify the resources we defined three relative entropies of coherence and showed that the max relative entropy of coherence has an operational interpretation in terms of exact coherence cost. It is an interesting problem to define other coherence measures not based on relative entropies and to find their operational meaning. Further, we worked on single-shot resource interconversion, specifically the exact, approximate, and liberal coherence costs of channels, and the exact and probabilistic distillation of coherence. Finding expressions or bounds on other types of resource interconversion problems like probabilistic coherence cost or catalytic resource interconversion problems are still open.

In our work on the resource theory of multi-qubit magic channels in Chapter 4, the criterion we used to identify the free channels was an efficient classical simulation of circuits. It is an open problem for the multi-qubit case to formulate a function (like the Wigner function for the odd case) to characterize the magic resources useful for achieving universal quantum computation in the magic state/channel model of quantum computation. Since finding such a function is a very hard problem, finding alternate approaches to characterize and quantify magic resources is an interesting problem. Further, given that the set of CSPOs can be classically simulated, it is not yet known whether there exists a larger set of operations from which pure magic states cannot be distilled. Apart from identifying such channels, it is also not clear if there is an operational way to define the set of free superchannels. Moreover, using the free sets of superchannels that we defined, we were unable to find the upper bound on the cost and the lower bound on distillation under completely CSPO preserving superchannels, and have left it as open. Lastly, we devised a classical simulation algorithm to estimate the expectation value of an observable given a general quantum circuit. An interesting future direction would be to generalize simulation algorithms not based on quasiprobability simulators, and figure out if there is a way to substantially improve the runtime.

In chapter 5, we aimed at finding interconversion conditions among pairs of quantum channels. Using our generalization, we were able to find conditions for interconversion among pairs of classical channels when the resultant channels' output dimension is two. It is an open problem to find out if the same conditions

hold when the dimension of the resultant pair is greater than two. If proven, it would be an elegant result for classical channels, implying that the problem of interconversion among pairs of classical channels is a completely geometric problem, similar to the classical state case. Apart from that, how Lorenz and relative majorization are related to channel interconversion for different dynamical resource theories is open for investigation. Another direction for further research would be to extend our results to interconversion among families of channels. To solve this, one needs to generalize and characterize the testing region and find if the equivalence between Lorenz and relative majorization still holds for families of classical channels with the same constraints.

In chapter 6, we quantified the performance of a fixed processor using average fidelity. An open problem is to figure out other ways to quantify and compare the performance of fixed processors. The limitations on various protocols also indicate that there is an underlying uncertainty principle guiding the programmability of quantum processors. Other possible directions that are still unexplored include finding limitations by restricting the program states and measurements to certain types of states and measurements. Besides, we only considered single use of the fixed processor. It would be interesting to investigate how several uses of the processor can be used to better approximate a desired unitary channel.

Bibliography

[1] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, Jun 1982.

[2] P. W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, SFCS '94, page 124–134, USA, 1994. IEEE Computer Society.

[3] D. R. Simon. On the power of quantum computation. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, SFCS '94, page 116–123, USA, 1994. IEEE Computer Society.

[4] David P. DiVincenzo. Quantum computation. *Science*, 270(5234):255–261, 1995.

[5] Artur Ekert and Richard Jozsa. Quantum computation and shor's factoring algorithm. *Rev. Mod. Phys.*, 68:733–753, Jul 1996.

[6] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[7] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences*, 114(29):7555–7560, 2017.

[8] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin-Lic Chan. Low-depth quantum simulation of materials. *Phys. Rev. X*, 8:011044, Mar 2018.

[9] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, Jun 2018.

[10] Eli Biham, Gilles Brassard, Dan Kenigsberg, and Tal Mor. Quantum computing without entanglement. *Theoretical Computer Science*, 320(1):15–33, 2004.

[11] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Rev. Mod. Phys.*, 92:015003, Mar 2020.

[12] A. P. Lund, Michael J. Bremner, and T. C. Ralph. Quantum sampling problems, bosonsampling and quantum supremacy. *npj Quantum Information*, 3(1):15, Apr 2017.

[13] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, Oct 2019.

[14] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, Feb 2019.

[15] Dominik Hangleiter and Jens Eisert. Computational advantage of quantum random sampling, 2022.

[16] John Preskill. Quantum computing and the entanglement frontier, 2012.

[17] Aram W. Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549(7671):203–209, Sep 2017.

[18] John Preskill. Quantum computing 40 years later, 2021.

[19] Diana Franklin and Frederic T. Chong. *Challenges in Reliable Quantum Computing*, pages 247–266. Springer US, Boston, MA, 2004.

[20] P. Zoller, Th. Beth, D. Binosi, R. Blatt, H. Briegel, D. Bruss, T. Calarco, J. I. Cirac, D. Deutsch, J. Eisert, A. Ekert, C. Fabre, N. Gisin, P. Grangiere, M. Grassl, S. Haroche, A. Imamoglu, A. Karlson,

J. Kempe, L. Kouwenhoven, S. Kröll, G. Leuchs, M. Lewenstein, D. Loss, N. Lütkenhaus, S. Massar, J. E. Mooij, M. B. Plenio, E. Polzik, S. Popescu, G. Rempe, A. Sergienko, D. Suter, J. Twamley, G. Wendin, R. Werner, A. Winter, J. Wrachtrup, and A. Zeilinger. Quantum information processing and communication. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics*, 36(2):203–228, Nov 2005.

[21] Won-Young Hwang. Quantum key distribution with high loss: Toward global secure communication. *Phys. Rev. Lett.*, 91:057901, Aug 2003.

[22] C. G. Almudever, L. Lao, X. Fu, N. Khammassi, I. Ashraf, D. Iorga, S. Varsamopoulos, C. Eichler, A. Wallraff, L. Geck, A. Kruth, J. Knoch, H. Bluhm, and K Bertels. The engineering challenges in quantum computing. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*, pages 836–845, 2017.

[23] Rupak Biswas, Zhang Jiang, Kostya Kechezhi, Sergey Knysh, Salvatore Mandrà, Bryan O'Gorman, Alejandro Perdomo-Ortiz, Andre Petukhov, John Realpe-Gómez, Eleanor Rieffel, Davide Venturelli, Fedir Vasko, and Zhihui Wang. A nasa perspective on quantum computing: Opportunities and challenges. *Parallel Computing*, 64:81–98, 2017. High-End Computing for Next-Generation Scientific Discovery.

[24] Antonio D. Córcoles, Abhinav Kandala, Ali Javadi-Abhari, Douglas T. McClure, Andrew W. Cross, Kristan Temme, Paul D. Nation, Matthias Steffen, and Jay M. Gambetta. Challenges and opportunities of near-term quantum computing systems. *Proceedings of the IEEE*, 108(8):1338–1352, 2020.

[25] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[26] P.W. Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65, 1996.

[27] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation. *Science*, 279(5349):342–345, 1998.

[28] Andrew Steane. Multiple-particle interference and quantum error correction. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 452(1954):2551–2577, 1996.

[29] Hans Aschauer. *Quantum communication in noisy environments.* Ludwig-Maximilians-Universität München, April 2005.

[30] A.-S.F. Obada, N.M. Arafa, and N. Metwally. Quantum communication in the presence of noise local environments. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 12(1):55–60, 2012.

[31] Yu. I. Bogdanov, A. Yu. Chernyavskiy, Alexander Holevo, V. F. Lukichev, and A. A. Orlikovsky. Modeling of quantum noise and the quality of hardware components of quantum computers. In Alexander A. Orlikovsky, editor, *International Conference Micro- and Nano-Electronics 2012*, volume 8700, page 87001A. International Society for Optics and Photonics, SPIE, 2013.

[32] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio. Colloquium: Quantum coherence as a resource. *Rev. Mod. Phys.*, 89:041003, Oct 2017.

[33] Gilles Brassard, Ashwin Nayak, Alain Tapp, Dave Touchette, and Falk Unger. Noisy interactive quantum communication, 2019.

[34] Matthew Otten and Stephen K. Gray. Accounting for errors in quantum algorithms via individual error reduction. *npj Quantum Information*, 5(1):11, Jan 2019.

[35] Konstantinos Georgopoulos, Clive Emary, and Paolo Zuliani. Modeling and simulating the noisy behavior of near-term quantum computers. *Phys. Rev. A*, 104:062432, Dec 2021.

[36] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, 94:015004, Feb 2022.

[37] Gaurav Saxena, Eric Chitambar, and Gilad Gour. Dynamical resource theory of quantum coherence. *Phys. Rev. Research*, 2:023298, Jun 2020.

[38] Johan Aberg. Quantifying superposition. *arXiv:quant-ph/0612146*, 2006.

[39] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying coherence. *Phys. Rev. Lett.*, 113:140401, Sep 2014.

[40] Eric Chitambar and Gilad Gour. Critical examination of incoherent operations and a physically consistent resource theory of quantum coherence. *Phys. Rev. Lett.*, 117:030401, Jul 2016.

[41] Iman Marvian and Robert W. Spekkens. How to quantify coherence: Distinguishing speakable and unspeakable notions. *Phys. Rev. A*, 94:052324, Nov 2016.

[42] G. Vidal, L. Masanes, and J. I. Cirac. Storing quantum dynamics in quantum states: A stochastic programmable gate. *Phys. Rev. Lett.*, 88:047905, Jan 2002.

[43] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Phys. Rev. A*, 81:062348, Jun 2010.

[44] John Preskill. Reliable quantum computers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):385–410, 1998.

[45] Adam Brazier, Vladimír Bužek, and Peter L. Knight. Probabilistic programmable quantum processors with multiple copies of program states. *Phys. Rev. A*, 71:032306, Mar 2005.

[46] Bartosz Regula. Probabilistic transformations of quantum resources. *Phys. Rev. Lett.*, 128:110505, Mar 2022.

[47] M. A. Nielsen and Isaac L. Chuang. Programmable quantum gate arrays. *Phys. Rev. Lett.*, 79:321–324, Jul 1997.

[48] Paweł Horodecki, Michał Horodecki, and Ryszard Horodecki. Binding entanglement channels. *Journal of Modern Optics*, 47(2-3):347–354, 2000.

[49] Chung-Yun Hsieh. Resource Preservability. *Quantum*, 4:244, March 2020.

[50] Leonardo Banchi, Jason Pereira, Seth Lloyd, and Stefano Pirandola. Convex optimization of programmable quantum computers. *npj Quantum Information*, 6(1):42, May 2020.

[51] Maximilian Schlosshauer, Johannes Kofler, and Anton Zeilinger. A snapshot of foundational attitudes toward quantum mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):222–230, 2013.

[52] Max Jammer. *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective*. New York: Wiley, 1974.

[53] Kent A. Peacock. *The Quantum Revolution: A Historical Perspective*. Greenwood Press, 2008.

[54] Claude Elwood Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 1948.

[55] SIMON KOCHEN and E. P. SPECKER. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.

[56] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131:2766–2788, Sep 1963.

[57] E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys. Rev. Lett.*, 10:277–279, Apr 1963.

[58] L Mandel. Non-classical states of the electromagnetic field. *Physica Scripta*, T12:34–42, jan 1986.

[59] Alexander Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems Inform. Transmission*, 9, 1973.

[60] A. Holevo. Some statistical problems for quantum gaussian states. *IEEE Transactions on Information Theory*, 21(5):533–543, 1975.

[61] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, Jun 1969.

[62] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, jan 1983.

[63] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[64] Roman S. Ingarden. Quantum information theory. *Reports on Mathematical Physics*, 10(1):43–72, August 1976.

[65] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical hamiltonian model of computers as represented by turing machines. *Journal of Statistical Physics*, 22(5):563–591, May 1980.

[66] Yu. I. Manin. Computable and uncomputable. *Sovetskoe Radio, Moscow*, 1980.

[67] David Deutsch and Roger Penrose. Quantum theory, the church–turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.

[68] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[69] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[70] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[71] Francesco Bova, Avi Goldfarb, and Roger G. Melko. Commercial applications of quantum computing. *EPJ Quantum Technology*, 8(1):2, Jan 2021.

[72] Gilad Gour and Mark M. Wilde. Entropy of a quantum channel. *Phys. Rev. Research*, 3:023096, May 2021.

[73] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[74] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2 edition, 2017.

[75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[76] A S Holevo. Quantum coding theorems. *Russian Mathematical Surveys*, 53(6):1295–1331, dec 1998.

[77] Richard P. Feynman. *The Feynman lectures on physics*, volume 3. Reading, Mass. : Addison-Wesley Pub. Co., 1965.

[78] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 2 edition, 2017.

[79] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[80] Dik Bouwmeester, Jian-Wei Pan, Klaus Mattle, Manfred Eibl, Harald Weinfurter, and Anton Zeilinger. Experimental quantum teleportation. *Nature*, 390(6660):575–579, Dec 1997.

[81] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu. Experimental realization of teleporting an unknown pure quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 80:1121–1125, Feb 1998.

[82] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, Apr 2019.

[83] A. Yu. Kitaev. Quantum measurements and the abelian stabilizer problem, 1995.

[84] Lov K. Grover. Quantum computers can search rapidly by using almost any transformation. *Phys. Rev. Lett.*, 80:4329–4332, May 1998.

[85] Andris Ambainis. Quantum walk algorithm for element distinctness. *arXiv*, 2003.

[86] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[87] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74:145–195, Mar 2002.

[88] Gilad Gour and Carlo Maria Scandolo. Dynamical entanglement. *Phys. Rev. Lett.*, 125:180505, Oct 2020.

[89] Gilad Gour and Carlo Maria Scandolo. Entanglement of a bipartite channel. *Phys. Rev. A*, 103:062422, Jun 2021.

[90] Xin Wang, Mark M. Wilde, and Yuan Su. Quantifying the magic of quantum channels. *New Journal of Physics*, 21(10):103002, oct 2019.

[91] Gaurav Saxena and Gilad Gour. Quantifying multiqubit magic channels with completely stabilizer-preserving operations. *Phys. Rev. A*, 106:042422, Oct 2022.

[92] Miguel Navascués and Luis Pedro García-Pintos. Nonthermal quantum channels as a thermodynamical resource. *Phys. Rev. Lett.*, 115:010405, Jul 2015.

[93] Stefan Bäuml, Siddhartha Das, and Mark M. Wang, Xinand Wilde. Resource theory of entanglement for bipartite quantum channels. *arXiv:1907.04181*, 2019.

[94] Thomas Theurer, Saipriya Satyajit, and Martin B. Plenio. Quantifying dynamical coherence with dynamical entanglement. *Phys. Rev. Lett.*, 125:130401, Sep 2020.

[95] Eric Chitambar, Gilad Gour, Kuntal Sengupta, and Rana Zibakhsh. Quantum bell nonlocality as a form of entanglement. *Phys. Rev. A*, 104:052208, Nov 2021.

[96] Quntao Zhuang, Peter W. Shor, and Jeffrey H. Shapiro. Resource theory of non-gaussian operations. *Phys. Rev. A*, 97:052317, May 2018.

[97] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, jan 2014.

[98] David Blackwell. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, 24(2):265 – 272, 1953.

[99] Wim van Dam and Gadiel Seroussi. Efficient quantum algorithms for estimating gauss sums, 2002.

[100] John Watrous. Quantum computational complexity, 2008.

[101] Hefeng Wang, Sixia Yu, and Hua Xiang. Efficient quantum algorithm for solving structured problems via multi-step quantum computation, 2019.

[102] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Mip*=re. *arXiv*, 2020.

[103] Ashley Montanaro. Quantum algorithms: an overview. *npj Quantum Information*, 2(1):15023, Jan 2016.

[104] William K. Wootters and W. S. Leng. Quantum entanglement as a quantifiable resource. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 356(1743):1717–1731, 1998.

[105] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[106] Julio I de Vicente. On nonlocality as a resource theory and nonlocality measures. *Journal of Physics A: Mathematical and Theoretical*, 47(42):424017, oct 2014.

[107] Paul Hausladen, Richard Jozsa, Benjamin Schumacher, Michael Westmoreland, and William K. Wootters. Classical information capacity of a quantum channel. *Phys. Rev. A*, 54:1869–1876, Sep 1996.

[108] Howard Barnum, M. A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57:4153–4175, Jun 1998.

[109] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.*, 83:3081–3084, Oct 1999.

[110] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem, 2001.

[111] A. S. Holevo. On entanglement-assisted classical capacity. *Journal of Mathematical Physics*, 43(9):4326–4333, 2002.

[112] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, Sep 2017.

[113] Rajendra Bhatia. *Matrix Analysis*, volume 169. Springer, 1997.

[114] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge; New York, 1994.

[115] David Bohm. *Quantum Theory*. Dover Books, 1989.

[116] Paul Adrien Maurice Dirac. *The Principles of Quantum Mechanics*. Clarendon Press, 1930.

[117] Mário Ziman. Process positive-operator-valued measure: A mathematical framework for the description of process tomography experiments. *Phys. Rev. A*, 77:062112, Jun 2008.

[118] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[119] Gilad Gour. Comparison of quantum channels by superchannels. *IEEE Transactions on Information Theory*, 65(9):5880–5904, 2019.

[120] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Transforming quantum operations: Quantum supermaps. *EPL (Europhysics Letters)*, 83(3):30004, jul 2008.

[121] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Quantum circuit architecture. *Phys. Rev. Lett.*, 101:060401, Aug 2008.

[122] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Phys. Rev. A*, 80:022339, Aug 2009.

[123] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Phys. Rev. A*, 88:022318, Aug 2013.

[124] Paolo Perinotti. *Causal Structures and the Classification of Higher Order Quantum Computations*, pages 103–127. Springer International Publishing, Cham, 2017.

[125] John Burniston, Michael Grabowecky, Carlo Maria Scandolo, Giulio Chiribella, and Gilad Gour. Necessary and sufficient conditions on measurements of quantum channels. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 476(2236):20190832, 2020.

[126] Alessandro Bisio and Paolo Perinotti. Theoretical framework for higher-order quantum theory. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475(2225):20180706, 2019.

[127] Marco Tomamichel. *Quantum Information Processing with Finite Resources*. Springer International Publishing, 2016.

[128] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[129] John Watrous. Semidefinite programs for completely bounded norms. *Theory of Computing*, 5(11):217–238, 2009.

[130] Gilad Gour. Uniqueness and optimality of dynamical extensions of divergences. *PRX Quantum*, 2:010313, Jan 2021.

[131] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, Jun 1975.

[132] Alfréd Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 547–561, Berkeley, Calif., 1961. University of California Press.

[133] Hisaharu Umegaki. Conditional expectation in an operator al gebra. IV. Entropy and information. *Kodai Mathematical Seminar Reports*, 14(2):59 – 85, 1962.

[134] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23(1):57–65, 1986.

[135] Mark M. Wilde, Andreas Winter, and Dong 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, Oct 2014.

[136] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013.

[137] Nilanjana Datta and Felix Leditzky. A limit of the quantum rényi divergence. *Journal of Physics A: Mathematical and Theoretical*, 47(4):045304, jan 2014.

[138] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative rényi entropy. *Journal of Mathematical Physics*, 54(12):122201, 2013.

[139] Fumio Hiai, Milan Mosonyi, Denes Petz, and Cedric Beny. Quantum f-divergences and error correction. *Reviews in Mathematical Physics*, 23(07):691–747, 2011.

[140] Fumio Hiai and Milán Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, 2017.

[141] Mark M Wilde. Optimized quantum f-divergences and data processing. *Journal of Physics A: Mathematical and Theoretical*, 51(37):374002, aug 2018.

[142] Koenraad M. R. Audenaert and Nilanjana Datta. α-z-rényi relative entropies. *Journal of Mathematical Physics*, 56(2):022202, 2015.

[143] Yunchao Liu and Xiao Yuan. Operational resource theory of quantum channels. *Phys. Rev. Research*, 2:012035, Feb 2020.

[144] Zi-Wen Liu and Andreas Winter. Resource theories of quantum channels and the universal role of resource erasure. *arXiv:1904.04201*, 2019.

[145] Gilad Gour and Andreas Winter. How to quantify a dynamical quantum resource. *Phys. Rev. Lett.*, 123:150401, Oct 2019.

[146] Tom Cooney, Milán Mosonyi, and Mark M. Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. *Communications in Mathematical Physics*, 344(3):797–829, Jun 2016.

[147] Felix Leditzky, Eneet Kaur, Nilanjana Datta, and Mark M. Wilde. Approaches for approximate additivity of the holevo information of quantum channels. *Phys. Rev. A*, 97:012332, Jan 2018.

[148] Mark M. Wilde, Mario Berta, Christoph Hirche, and Eneet Kaur. Amortized channel divergence for asymptotic quantum channel discrimination. *Letters in Mathematical Physics*, 110(8):2277–2336, Aug 2020.

[149] Alexander Barvinok. *A Course in Convexity*, volume 54. Graduate Studies in Mathematics, University of Michigan, Ann Arbor, MI, 2002.

[150] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[151] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series. Princeton University Press, Princeton, N. J., 1970.

[152] Yu. Nesterov and A. Nemirovsky. Conic formulation of a convex programming problem and duality. *Optimization Methods and Software*, 1(2):95–115, 1992.

[153] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[154] Haotian Jiang, Tarun Kathuria, Yin Tat Lee, Swati Padmanabhan, and Zhao Song. A faster interior point method for semidefinite programming. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 910–918, 2020.

[155] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*, STOC '84, page 302–311, New York, NY, USA, 1984. Association for Computing Machinery.

[156] Kurt M. Anstreicher. The volumetric barrier for semidefinite programming. *Mathematics of Operations Research*, 25(3):365–380, 2000.

[157] Lieven Vandenberghe, V. Ragu Balakrishnan, Ragnar Wallin, Anders Hansson, and Tae Roh. *Interior-Point Algorithms for Semidefinite Programming Problems Derived from the KYP Lemma*, pages 195–238. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[158] Florian A. Potra and Stephen J. Wright. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(1):281–302, 2000. Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations.

[159] Jos F. Sturm. Implementation of interior point methods for mixed semidefinite and second order cone optimization problems. *Optimization Methods and Software*, 17(6):1105–1154, 2002.

[160] R. H. Tütüncü, K. C. Toh, and M. J. Todd. Solving semidefinite-quadratic-linear programs using sdpt3. *Mathematical Programming*, 95(2):189–217, Feb 2003.

[161] Michal Kočvara and Michael Stingl. Pennon: A code for convex nonlinear and semidefinite programming. *Optimization Methods and Software*, 18(3):317–333, 2003.

[162] Samuel Burer and Changhui Choi. Computational enhancements in low-rank semidefinite programming. *Optimization Methods and Software*, 21(3):493–512, 2006.

[163] Brian Borchers and Joseph G. Young. Implementation of a primal–dual method for sdp on a shared memory parallel architecture. *Computational Optimization and Applications*, 37(3):355–369, Jul 2007.

[164] Zaiwen Wen, Donald Goldfarb, and Wotao Yin. Alternating direction augmented lagrangian methods for semidefinite programming. *Mathematical Programming Computation*, 2(3):203–230, Dec 2010.

[165] Eftychios Pnevmatikakis and Liam Paninski. Fast interior-point inference in high-dimensional sparse, penalized state-space models. In Neil D. Lawrence and Mark Girolami, editors, *Proceedings of the*

Fifteenth International Conference on Artificial Intelligence and Statistics, volume 22 of *Proceedings of Machine Learning Research*, pages 895–904, La Palma, Canary Islands, 21–23 Apr 2012. PMLR.

[166] Brendan O'Donoghue, Eric Chu, Neal Parikh, and Stephen Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068, Jun 2016.

[167] Lijun Ding, Alp Yurtsever, Volkan Cevher, Joel A. Tropp, and Madeleine Udell. An optimal-storage approach to semidefinite programming using approximate complementarity, 2019.

[168] Baihe Huang, Shunhua Jiang, Zhao Song, Runzhou Tao, and Ruizhe Zhang. Solving sdp faster: A robust ipm framework and efficient implementation, 2021.

[169] Brian Borchers. Csdp, a c library for semidefinite programming. *Optimization Methods and Software*, 11(1-4):613–623, Jan 1999.

[170] Jos F. Sturm. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[171] K. C. Toh, M. J. Todd, and R. H. Tütüncü. Sdpt3 — a matlab software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, 11(1-4):545–581, 1999.

[172] Makoto Yamashita, Katsuki Fujisawa, and Masakazu Kojima. Implementation and evaluation of sdpa 6.0 (semidefinite programming algorithm 6.0). *Optimization Methods and Software*, 18(4):491–505, 2003.

[173] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.

[174] CVX Research Inc. CVX: Matlab software for disciplined convex programming, version 2.0. http://cvxr.com/cvx, Aug 2012.

[175] Alp Yurtsever, Joel A. Tropp, Olivier Fercoq, Madeleine Udell, and Volkan Cevher. Scalable semidefinite programming. *SIAM Journal on Mathematics of Data Science*, 3(1):171–200, 2021.

[176] Wojciech Hubert Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715–775, May 2003.

[177] Maximilian Schlosshauer. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Rev. Mod. Phys.*, 76:1267–1305, Feb 2005.

[178] Daniel A. Lidar and K. Birgitta Whaley. *Decoherence-Free Subspaces and Subsystems*, pages 83–120. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[179] Dave Morris Bacon. *Decoherence, Control, and Symmetry in Quantum Computers*. PhD thesis, University of California, Berkeley, 2001. AAI3044374.

[180] Martin B. Plenio and Shashank Virmani. An introduction to entanglement measures. *Quantum Info. Comput.*, 7(1):1–51, January 2007.

[181] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.*, 79:555–609, Apr 2007.

[182] John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14):143001, feb 2016.

[183] Gilad Gour, Markus P. Muller, Varun Narasimhachar, Robert W. Spekkens, and Nicole Yunger Halpern. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports*, 583:1 – 58, 2015. The resource theory of informational nonequilibrium in thermodynamics.

[184] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[185] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84:621–669, May 2012.

[186] Ángel Rivas, Susana F. Huelga, and Martin B. Plenio. Quantum non-markovianity: characterization, quantification and detection. *Reports on Progress in Physics*, 77(9):094001, aug 2014.

[187] Gerardo Adesso, Thomas R. Bromley, and Marco Cianciaruso. Measures and applications of quantum correlations. *Journal of Physics A: Mathematical and Theoretical*, 49(47):473001, nov 2016.

[188] Kavan Modi, Aharon Brodutch, Hugo Cable, Tomasz Paterek, and Vlatko Vedral. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.*, 84:1655–1707, Nov 2012.

[189] Wojciech H. Zurek. Decoherence and the transition from quantum to classical. *Phys. Today*, 44N10:36–44, 1991.

[190] Wojciech Hubert Zurek. *Decoherence and the Transition from Quantum to Classical — Revisited*, pages 1–31. Birkhäuser Basel, Basel, 2007.

[191] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76:722–725, Jan 1996.

[192] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996.

[193] Hoi-Kwong Lo and Sandu Popescu. Concentrating entanglement by local actions: Beyond mean values. *Phys. Rev. A*, 63:022301, Jan 2001.

[194] Eric Chitambar and Gilad Gour. Comparison of incoherent operations and measures of coherence. *Phys. Rev. A*, 94:052336, Nov 2016.

[195] Eric Chitambar and Gilad Gour. Erratum: Comparison of incoherent operations and measures of coherence [phys. rev. a 94, 052336 (2016)]. *Phys. Rev. A*, 95:019902, Jan 2017.

[196] Thomas Theurer, Dario Egloff, Lijian Zhang, and Martin B. Plenio. Quantifying operations with an application to coherence. *Phys. Rev. Lett.*, 122:190405, May 2019.

[197] Khaled Ben Dana, María García Díaz, Mohamed Mejatty, and Andreas Winter. Resource theory of coherence: Beyond states. *Phys. Rev. A*, 95:062327, Jun 2017.

[198] Maria Garcia Diaz, Kun Fang, Xin Wang, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Andreas Winter. Using and reusing coherence to realize quantum processes. *Quantum 2, 100 (2018)*, 2018.

[199] Azam Mani and Vahid Karimipour. Cohering and decohering power of quantum channels. *Phys. Rev. A*, 92:032331, Sep 2015.

[200] María García-Díaz, Dario Egloff, and Martin B. Plenio. A note on coherence power of n-dimensional unitary operators. *Quantum Info. Comput.*, 16(15–16):1282–1294, November 2016.

[201] Paolo Zanardi, Georgios Styliaris, and Lorenzo Campos Venuti. Measures of coherence-generating power for quantum unital operations. *Phys. Rev. A*, 95:052307, May 2017.

[202] M Takahashi and E Chitambar. Comparing coherence and entanglement under resource non-generating unitary transformations. *Journal of Physics A: Mathematical and Theoretical*, 51(41):414003, sep 2018.

[203] Zi-Wen Liu, Xueyuan Hu, and Seth Lloyd. Resource destroying maps. *Phys. Rev. Lett.*, 118:060502, Feb 2017.

[204] Denis Rosset, Francesco Buscemi, and Yeong-Cherng Liang. Resource theory of quantum memories and their faithful verification with minimal assumptions. *Phys. Rev. X*, 8:021033, May 2018.

[205] Timo Simnacher, Nikolai Wyderka, Cornelia Spee, Xiao-Dong Yu, and Otfried Gühne. Certifying quantum memories with coherence. *Phys. Rev. A*, 99:062319, Jun 2019.

[206] Xiao Yuan, Yunchao Liu, Qi Zhao, Bartosz Regula, Jayne Thompson, and Mile Gu. Universal and operational benchmarking of quantum memories. *npj Quantum Information*, 7(1):108, Jul 2021.

[207] Stephanie Wehner, David Elkouss, and Ronald Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412), 2018.

[208] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. Fundamental limits of repeaterless quantum communications. *Nature Communications*, 8(1):15043, Apr 2017.

[209] Eric Chitambar and Min-Hsiu Hsieh. Relating the resource theories of entanglement and quantum coherence. *Phys. Rev. Lett.*, 117:020402, Jul 2016.

[210] Lu Li, Kaifeng Bu, and Zi-Wen Liu. Quantifying the resource content of quantum channels: An operational approach. *Phys. Rev. A*, 101:022335, Feb 2020.

[211] Xin Wang and Mark M. Wilde. Exact entanglement cost of quantum states and channels under ppt-preserving operations. *arXiv:1809.09592*, 2018.

[212] Philippe Faist, Mario Berta, and Fernando Brandão. Thermodynamic capacity of quantum processes. *Phys. Rev. Lett.*, 122:200601, May 2019.

[213] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. A mathematical theory of resources. *Information and Computation*, 250:59–86, 2016. Quantum Physics and Logic.

[214] Eneet Kaur and Mark M Wilde. Amortized entanglement of a quantum channel and approximately teleportation-simulable channels. *Journal of Physics A: Mathematical and Theoretical*, 51(3):035303, dec 2017.

[215] Thomas R. Bromley, Marco Cianciaruso, and Gerardo Adesso. Frozen quantum coherence. *Phys. Rev. Lett.*, 114:210401, May 2015.

[216] Alexander Streltsov, Uttam Singh, Himadri Shekhar Dhar, Manabendra Nath Bera, and Gerardo Adesso. Measuring quantum coherence with entanglement. *Phys. Rev. Lett.*, 115:020403, Jul 2015.

[217] Andreas Winter and Dong Yang. Operational resource theory of coherence. *Phys. Rev. Lett.*, 116:120404, Mar 2016.

[218] Benjamin Yadin, Jiajun Ma, Davide Girolami, Mile Gu, and Vlatko Vedral. Quantum processes which do not use coherence. *Phys. Rev. X*, 6:041028, Nov 2016.

[219] Giulio Chiribella and Yuxiang Yang. Optimal quantum operations at zero energy cost. *Phys. Rev. A*, 96:022327, Aug 2017.

[220] Julio I de Vicente and Alexander Streltsov. Genuine quantum coherence. *Journal of Physics A: Mathematical and Theoretical*, 50(4):045301, dec 2016.

[221] Kun Fang, Omar Fawzi, Renato Renner, and David Sutter. Chain rule for the quantum relative entropy. *Phys. Rev. Lett.*, 124:100501, Mar 2020.

[222] Kun Fang, Xin Wang, Marco Tomamichel, and Mario Berta. Quantum channel simulation and the channel's smooth max-information. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 2326–2330, 2018.

[223] Stefan Bauml, Siddhartha Das, Xin Wang, and Mark M. Wilde. Resource theory of entanglement for bipartite quantum channels. *arXiv:1907.04181*, 2019.

[224] Leonid Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 10–19, New York, NY, USA, 2003. ACM.

[225] Nilanjana Datta. Max-relative entropy of entanglement, alias log robustness. *International Journal of Quantum Information*, 07(02):475–491, 2009.

[226] N. Datta. Min- and max-relative entropies and a new entanglement monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, June 2009.

[227] F. Buscemi and N. Datta. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Transactions on Information Theory*, 56(3):1447–1460, March 2010.

[228] 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, March 2011.

[229] Kaifeng Bu, Uttam Singh, Shao-Ming Fei, Arun Kumar Pati, and Junde Wu. Maximum relative entropy of coherence: An operational coherence measure. *Phys. Rev. Lett.*, 119:150405, Oct 2017.

[230] Gilad Gour, Iman Marvian, and Robert W. Spekkens. Measuring the quality of a quantum reference frame: The relative entropy of frameness. *Phys. Rev. A*, 80:012307, Jul 2009.

[231] Bartosz Regula, Kun Fang, Xin Wang, and Gerardo Adesso. One-shot coherence distillation. *Phys. Rev. Lett.*, 121:010401, Jul 2018.

[232] Q. Zhao, Y. Liu, X. Yuan, E. Chitambar, and A. Winter. One-shot coherence distillation: Towards completing the picture. *IEEE Transactions on Information Theory*, 65(10):6441–6453, Oct 2019.

[233] Daniel Gottesman. Stabilizer codes and quantum error correction. *arXiv:quant-ph/9705052*, 1997.

[234] John Preskill. Fault-tolerant quantum computation. *arXiv:quant-ph/9712048*, 1997.

[235] Daniel Gottesman. An introduction to quantum error correction and fault-tolerant quantum computation. *arXiv:0904.2557*, 2009.

[236] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, Feb 2005.

[237] Earl T. Campbell and Dan E. Browne. Bound states for magic state distillation in fault-tolerant quantum computation. *Phys. Rev. Lett.*, 104:030503, Jan 2010.

[238] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A*, 86:032324, Sep 2012.

[239] Adam M. Meier, Bryan Eastin, and Emanuel Knill. Magic-state distillation with the four-qubit code. *arXiv:1204.4221*, 2012.

[240] Cody Jones. Multilevel distillation of magic states for quantum computing. *Phys. Rev. A*, 87:042305, Apr 2013.

[241] Earl T. Campbell and Mark Howard. Unified framework for magic state distillation and multiqubit gate synthesis with reduced resource cost. *Phys. Rev. A*, 95:022316, Feb 2017.

[242] Matthew B. Hastings and Jeongwan Haah. Distillation with sublogarithmic overhead. *Phys. Rev. Lett.*, 120:050504, Jan 2018.

[243] Christopher Chamberland and Andrew W. Cross. Fault-tolerant magic state preparation with flag qubits. *Quantum*, 3:143, May 2019.

[244] Daniel Litinski. Magic State Distillation: Not as Costly as You Think. *Quantum*, 3:205, Dec 2019.

[245] Xin Wang, Mark M. Wilde, and Yuan Su. Efficiently computable bounds for magic state distillation. *Phys. Rev. Lett.*, 124:090505, Mar 2020.

[246] Ernesto F. Galvão. Discrete wigner functions and quantum computational speedup. *Phys. Rev. A*, 71:042302, Apr 2005.

[247] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. Negative quasi-probability as a resource for quantum computation. *New Journal of Physics*, 14(11):113011, nov 2012.

[248] Dan Stahlke. Quantum interference as a resource for quantum speedup. *Phys. Rev. A*, 90:022302, Aug 2014.

[249] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, Jun 2014.

[250] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. Estimating outcome probabilities of quantum circuits using quasiprobabilities. *Phys. Rev. Lett.*, 115:070501, Aug 2015.

[251] Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf. Wigner function negativity and contextuality in quantum computation on rebits. *Phys. Rev. X*, 5:021003, Apr 2015.

[252] Sergey Bravyi and David Gosset. Improved classical simulation of quantum circuits dominated by clifford gates. *Phys. Rev. Lett.*, 116:250501, Jun 2016.

[253] Mark Howard and Earl Campbell. Application of a resource theory for magic states to fault-tolerant quantum computing. *Phys. Rev. Lett.*, 118:090501, Mar 2017.

[254] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. Contextuality as a resource for models of quantum computation with qubits. *Phys. Rev. Lett.*, 119:120505, Sep 2017.

[255] Robert Raussendorf, Dan E. Browne, Nicolas Delfosse, Cihan Okay, and Juan Bermejo-Vega. Contextuality and wigner-function negativity in qubit quantum computation. *Phys. Rev. A*, 95:052334, May 2017.

[256] Ryan S. Bennink, Erik M. Ferragut, Travis S. Humble, Jason A. Laska, James J. Nutaro, Mark G. Pleszkoch, and Raphael C. Pooser. Unbiased simulation of near-clifford quantum circuits. *Phys. Rev. A*, 95:062337, Jun 2017.

[257] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, September 2019.

[258] James R. Seddon and Earl T. Campbell. Quantifying magic for multi-qubit operations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475(2227):20190251, 2019.

[259] Patrick Rall, Daniel Liang, Jeremy Cook, and William Kretschmer. Simulation of qubit quantum circuits via pauli propagation. *Phys. Rev. A*, 99:062337, Jun 2019.

[260] Ryuji Takagi, Bartosz Regula, and Mark M. Wilde. One-shot yield-cost relations in general quantum resource theories. *arXiv:2110.02212*, 2021.

[261] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. Stabilizer rényi entropy. *Phys. Rev. Lett.*, 128:050402, Feb 2022.

[262] Ben W. Reichardt. Quantum universality from magic states distillation applied to css codes. *Quantum Information Processing*, 4(3):251–264, Aug 2005.

[263] Sergey Bravyi and Jeongwan Haah. Magic-state distillation with low overhead. *Phys. Rev. A*, 86:052329, Nov 2012.

[264] Hussain Anwar, Earl T Campbell, and Dan E Browne. Qutrit magic state distillation. *New Journal of Physics*, 14(6):063006, jun 2012.

[265] Earl T. Campbell, Hussain Anwar, and Dan E. Browne. Magic-state distillation in all prime dimensions using quantum reed-muller codes. *Phys. Rev. X*, 2:041021, Dec 2012.

[266] Bryan Eastin. Distilling one-qubit magic states into toffoli states. *Phys. Rev. A*, 87:032321, Mar 2013.

[267] Jeongwan Haah, Matthew B. Hastings, D. Poulin, and D. Wecker. Magic state distillation with low space overhead and optimal asymptotic input count. *Quantum*, 1:31, October 2017.

[268] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.*, 119:180509, Nov 2017.

[269] Anirudh Krishna and Jean-Pierre Tillich. Towards low overhead magic state distillation. *Phys. Rev. Lett.*, 123:070507, Aug 2019.

[270] Markus Heinrich and David Gross. Robustness of Magic and Symmetries of the Stabiliser Polytope. *Quantum*, 3:132, April 2019.

[271] James R. Seddon, Bartosz Regula, Hakop Pashayan, Yingkai Ouyang, and Earl T. Campbell. Quantifying quantum speedups: Improved classical simulation from tighter magic monotones. *PRX Quantum*, 2:010345, Mar 2021.

[272] Daniel Gottesman. Theory of fault-tolerant quantum computation. *Phys. Rev. A*, 57:127–137, Jan 1998.

[273] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[274] Barbara Amaral. Resource theory of contextuality. *arXiv:1904.04182*, 2019.

[275] Tamal Guha, Mir Alimuddin, Sumit Rout, Amit Mukherjee, Some Sankar Bhattacharya, and Manik Banik. Quantum Advantage for Shared Randomness Generation. *Quantum*, 5:569, October 2021.

[276] Mehdi Ahmadi, Hoan Bui Dang, Gilad Gour, and Barry C. Sanders. Quantification and manipulation of magic states. *Phys. Rev. A*, 97:062332, Jun 2018.

[277] Zi-Wen Liu and Andreas Winter. Many-body quantum magic. *arXiv:2010.13817*, 2020.

[278] Kun Fang and Zi-Wen Liu. No-go theorems for quantum resource purification. *Phys. Rev. Lett.*, 125:060405, Aug 2020.

[279] Oliver Hahn, Alessandro Ferraro, Lina Hultquist, Giulia Ferrini, and Laura García-Álvarez. Quantifying qubit magic with gottesman-kitaev-preskill encoding. *arXiv:2109.13018*, 2021.

[280] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, 2006.

[281] D. Gross. Non-negative wigner functions in prime dimensions. *Applied Physics B*, 86(3):367–370, Feb 2007.

[282] A. Mari and J. Eisert. Positive wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.*, 109:230503, Dec 2012.

[283] Sergey Bravyi, Graeme Smith, and John A. Smolin. Trading classical and quantum computational resources. *Phys. Rev. X*, 6:021043, Jun 2016.

[284] Nicolas Delfosse, Cihan Okay, Juan Bermejo-Vega, Dan E Browne, and Robert Raussendorf. Equivalence between contextuality and negativity of the wigner function for qudits. *New Journal of Physics*, 19(12):123024, dec 2017.

[285] Nikolaos Koukoulekidis and David Jennings. Constraints on magic state protocols from the statistical mechanics of wigner negativity. *arXiv:2106.15527*, 2021.

[286] Cecilia Cormick, Ernesto F. Galvão, Daniel Gottesman, Juan Pablo Paz, and Arthur O. Pittenger. Classicality in discrete wigner functions. *Phys. Rev. A*, 73:012301, Jan 2006.

[287] Lucas Kocia and Peter Love. Discrete wigner formalism for qubits and noncontextuality of clifford gates on qubit stabilizer states. *Phys. Rev. A*, 96:062134, Dec 2017.

[288] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. Phase-space-simulation method for quantum computation with magic states on qubits. *Phys. Rev. A*, 101:012350, Jan 2020.

[289] Michael Zurel, Cihan Okay, and Robert Raussendorf. Hidden variable model for universal quantum computation with magic states on qubits. *Phys. Rev. Lett.*, 125:260404, Dec 2020.

[290] Robert Raussendorf, Cihan Okay, Michael Zurel, and Polina Feldmann. Clifford covariance of wigner functions, positive representation of pauli measurement, and cohomology. *arXiv:2110.11631*, 2021.

[291] Hammam Qassim, Joel J. Wallman, and Joseph Emerson. Clifford recompilation for faster classical simulation of quantum circuits. *Quantum*, 3:170, August 2019.

[292] Cihan Okay, Michael Zurel, and Robert Raussendorf. On the extremal points of the λ-polytopes and classical simulation of quantum computation with magic states. *arXiv:2104.05822*, 2021.

[293] Guifré Vidal and Rolf Tarrach. Robustness of entanglement. *Phys. Rev. A*, 59:141–155, Jan 1999.

[294] Héctor J. García, Igor L. Markov, and Andrew W. Cross. On the geometry of stabilizer states. *Quantum Info. Comput.*, 14(7 & 8):683–720, May 2014.

[295] Yunchao Liu and Xiao Yuan. Operational resource theory of quantum channels. *Phys. Rev. Research*, 2:012035, Feb 2020.

[296] Mark M. Wilde, Mario Berta, Christoph Hirche, and Eneet Kaur. Amortized channel divergence for asymptotic quantum channel discrimination. *Letters in Mathematical Physics*, 110(8):2277–2336, Aug 2020.

[297] Kun Fang, Omar Fawzi, Renato Renner, and David Sutter. Chain rule for the quantum relative entropy. *Phys. Rev. Lett.*, 124:100501, Mar 2020.

[298] Gilad Gour and Mark M. Wilde. Entropy of a quantum channel. *Phys. Rev. Research*, 3:023096, May 2021.

[299] Tzu-Chieh Wei and Paul M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, 68:042307, Oct 2003.

[300] Arne Heimendahl, Markus Heinrich, and David Gross. The axiomatic and the operational approaches to resource theories of magic do not coincide, 2020.

[301] Bartosz Regula and Ryuji Takagi. One-shot manipulation of dynamical quantum resources. *Phys. Rev. Lett.*, 127:060402, Aug 2021.

[302] Xiao Yuan, Pei Zeng, Minbo Gao, and Qi Zhao. One-shot dynamical resource theory. *arXiv:2012.02781*, 2020.

[303] G. H. Hardy, John E. Littlewood, and George Pólya. *Inequalities*. Cambridge University Press, Cambridge, 1988.

[304] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83:436–439, Jul 1999.

[305] Michael A. Nielsen and Guifré Vidal. Majorization and the interconversion of bipartite states. *Quantum Inf. Comput.*, 1:76–93, 2001.

[306] Gilad Gour, David Jennings, Francesco Buscemi, Runyao Duan, and Iman Marvian. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics. *Nature Communications*, 9(1):5352, Dec 2018.

[307] Joseph M. Renes. Relative submajorization and its use in quantum resource theories. *Journal of Mathematical Physics*, 57(12):122202, 2016.

[308] Francesco Buscemi and Gilad Gour. Quantum relative lorenz curves. *Phys. Rev. A*, 95:012110, Jan 2017.

[309] M. O. Lorenz. Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219, 1905.

[310] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and its Applications*, volume 143. Springer, second edition, 2011.

[311] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4(1):2059, Jun 2013.

[312] Fernando G. S. L. Brandão, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. Resource theory of quantum states out of thermal equilibrium. *Phys. Rev. Lett.*, 111:250404, Dec 2013.

[313] P.M. Alberti and A. Uhlmann. A problem relating to positive linear maps on matrix algebras. *Reports on Mathematical Physics*, 18(2):163–176, 1980.

[314] Michele Dall'Arno, Francesco Buscemi, and Valerio Scarani. Extension of the Alberti-Ulhmann criterion beyond qubit dichotomies. *Quantum*, 4:233, February 2020.

[315] Nikolaos Koukoulekidis and David Jennings. Constraints on magic state protocols from the statistical mechanics of wigner negativity. *npj Quantum Information*, 8(1):42, Apr 2022.

[316] MICHAL HORODECKI and JONATHAN OPPENHEIM. (quantumness in the context of) resource theories. *International Journal of Modern Physics B*, 27(01n03):1345019, 2013.

[317] Xin Wang and Mark M. Wilde. Resource theory of asymmetric distinguishability for quantum channels. *Phys. Rev. Research*, 1:033169, Dec 2019.

[318] Soorya Rethinasamy and Mark M. Wilde. Relative entropy and catalytic relative majorization. *Phys. Rev. Research*, 2:033455, Sep 2020.

[319] Ligong Wang and Renato Renner. One-shot classical-quantum capacity and hypothesis testing. *Phys. Rev. Lett.*, 108:200501, May 2012.

[320] F. Dupuis, L Kramer, P. Faist, J. M. Renese, and R. Renner. GENERALIZED ENTROPIES. In *XVIIth International Congress on Mathematical Physics*, pages 134–153. WORLD SCIENTIFIC, oct 2013.

[321] Nilanjana Datta and Felix Leditzky. A limit of the quantum rényi divergence. *Journal of Physics A: Mathematical and Theoretical*, 47(4):045304, jan 2014.

[322] Satoshi Ishizaka and Tohya Hiroshima. Asymptotic teleportation scheme as a universal programmable quantum processor. *Phys. Rev. Lett.*, 101:240501, Dec 2008.

[323] Aleksander M. Kubicki, Carlos Palazuelos, and David Pérez-García. Resource quantification for the no-programing theorem. *Phys. Rev. Lett.*, 122:080505, Feb 2019.

[324] Yuxiang Yang, Renato Renner, and Giulio Chiribella. Optimal universal programming of unitary gates. *Phys. Rev. Lett.*, 125:210501, Nov 2020.

[325] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, Nov 1999.

[326] Jaehyun Kim, Yongwook Cheong, Jae-Seung Lee, and Soonchil Lee. Storing unitary operators in quantum states. *Phys. Rev. A*, 65:012302, Dec 2001.

[327] Mark Hillery, Vladimir Bužek, and Mario Ziman. Programmable quantum gate arrays. *Fortschritte der Physik*, 49(10-11):987–992, 2001.

[328] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, 60:1888–1898, Sep 1999.

[329] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 2002.

[330] Bartosz Regula and Ryuji Takagi. Fundamental limitations on distillation of quantum channel resources. *Nature Communications*, 12(1):4411, Jul 2021.

[331] Francesco Buscemi and Nilanjana Datta. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Transactions on Information Theory*, 56(3):1447–1460, 2010.

Appendix A

Appendix for chapter 3

A.1 Proof of dual of the log-robustness

Finding the dual of the log-robustness($LR_{\mathfrak{C}}(\mathcal{N}_A)$) is equivalent to finding the dual of $2^{LR_{\mathfrak{C}}(\mathcal{N}_A)}$. From (3.88), we can write $2^{LR_{\mathfrak{C}}(\mathcal{N}_A)}$ as

$$\min\Big\{\frac{1}{|A_0|}\mathrm{Tr}[\omega_A] \ : \ \omega_A \geqslant J_A^{\mathcal{N}} \ , \ \mathcal{D}_A[\omega_A] = \omega_A \ , \ \omega_{A_0} = \mathrm{Tr}[\omega_A]u_{A_0} \ , \ \omega_A \geqslant 0\Big\} \tag{A.1}$$

where $u_{A_0} = \frac{I_{A_0}}{|A_0|}$. The primal problem of the above conic linear program can be stated as

$$\min\Big\{\frac{1}{|A_0|}\mathrm{Tr}[\omega_A I_A] \ : \ \Gamma(\omega_A) - H_2 \in \mathfrak{K}_2 \ , \ \omega \geqslant 0\Big\} \tag{A.2}$$

where $\Gamma(\omega_A)$ is a linear map and is expressed as a 3-tuple such that $\Gamma(\omega_A) = (\omega_{A_0} - \mathrm{Tr}[\omega_{A_0}]u_{A_0} \ , \ \omega_A \ , \ \omega_A - \mathcal{D}(\omega_A))$. The separation of elements in the tuple can be understood as a direct sum between the subspaces in a larger vector space. Likewise, H_2 is also expressed as a 3-tuple such that $H_2 = (0_{A_0}, J_A^{\mathcal{N}}, 0_A)$. The cone $\mathfrak{K}_2$ can be expressed as a 3-tuple as $\mathfrak{K}_2 = \{(0_{A_0} \ , \ \zeta_A \ , \ 0_A) : \zeta_A \geqslant 0\}$. Hence, the dual cone $\mathfrak{K}_2^* = \{(Z_{A_0}, \beta_A, W_A) : Z_{A_0} \in \mathrm{Herm}(A_0), \ \beta_A \geqslant 0, \ W_A \in \mathrm{Herm}(A)\}$.

Therefore, it is easy to see that the dual to the above primal problem is

$$\max\Big\{\frac{1}{|A_0|}\mathrm{Tr}[\beta_A J_A^{\mathcal{N}}] \ : \ I_A - \Gamma^*(Z_{A_0}, \beta_A, W_A) \geqslant 0 \ , \ Z_{A_0} \in \mathrm{Herm}(A_0) \ , \ W_A \in \mathrm{Herm}(A) \ , \ \beta_A \geqslant 0\Big\} \tag{A.3}$$

In order to find $\Gamma^*(Z_{A_0}, \beta_A, W_A)$, we need to equate

$$\mathrm{Tr}[(Z_{A_0}, \beta_A, W_A)\Gamma(\omega_A)] = \mathrm{Tr}[\Gamma^*(Z_{A_0}, \beta_A, W_A)\omega_A] \tag{A.4}$$

From the LHS of (A.4), we find

$$\mathrm{Tr}[(Z_{A_0},\beta_A,W_A)\Gamma(\omega_A)] = \mathrm{Tr}[Z_{A_0}(\omega_{A_0}-\mathrm{Tr}[\omega_{A_0}]u_{A_0})] + \mathrm{Tr}[\beta_A\,\omega_A] + \mathrm{Tr}[W_A\,(\omega_A-\mathcal{D}(\omega_A))] \tag{A.5}$$

Therefore,

$$\Gamma^*(Z_{A_0},\beta_A,W_A) = Z_{A_0}\otimes I_{A_1} - \mathrm{Tr}[Z_{A_0}]u_{A_0}\otimes I_{A_1} + \beta_A + W_A - \mathcal{D}(W_A) \tag{A.6}$$

So, we can rewrite the first constraint in the dual problem as

$$\frac{I_A}{|A_0|} - Z_{A_0}\otimes I_{A_1} + \mathrm{Tr}[Z_{A_0}]u_{A_0}\otimes I_{A_1} - \beta_A - W_A + \mathcal{D}(W_A) \geqslant 0 \tag{A.7}$$

Now let $\eta_A \geqslant 0$ obey the following conditions

$$\mathcal{D}_A(\eta_A) = \mathcal{D}_{A_0}\left(\eta_{A_0}\right)\otimes u_{A_1} \;\;,\;\; \mathcal{D}_{A_1}[\eta_{A_1}] = I_{A_1} \tag{A.8}$$

Any such matrix can be expressed as $(u_{A_0} - Z_{A_0} + \mathrm{Tr}[Z_{A_0}]u_{A_0})\otimes I_{A_1} - W_A + \mathcal{D}(W_A)$. Hence, we can express (A.7) as

$$\eta_A \geqslant \beta_A \geqslant 0 \tag{A.9}$$

Since, $J_A^{\mathcal{N}} \geqslant 0$, therefore from the above equation we get

$$\mathrm{Tr}[\,\eta_A\,J_A^{\mathcal{N}}\,] \;\geqslant\; \mathrm{Tr}[\,\beta_A\,J_A^{\mathcal{N}}\,] \tag{A.10}$$

Hence, we can recast the dual problem in the following form

$$\max\Big\{\mathrm{Tr}[\,\eta_A\,J_A^{\mathcal{N}}\,] \;:\; \mathcal{D}_A(\eta_A) = \mathcal{D}_{A_0}(\eta_{A_0})\otimes u_{A_1}\;,\; \mathcal{D}_{A_1}[\eta_{A_1}] = I_{A_1}\;,\; \eta_A\geqslant 0\Big\} \tag{A.11}$$

Therefore,

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\max\Big\{\mathrm{Tr}[\,\eta_A\,J_A^{\mathcal{N}}\,] \;:\; \mathcal{D}_A(\eta_A) = \mathcal{D}_{A_0}(\eta_{A_0})\otimes u_{A_1}\;,\; \mathcal{D}_{A_1}[\eta_{A_1}] = I_{A_1}\;,\; \eta_A\geqslant 0\Big\} \tag{A.12}$$

which is Eq.(3.89).

A.2 Proof of Theorem 3.13 and the dual of the conversion distance for MISC and DISC

In [129], it was shown that the diamond norm can be expressed as the following SDP

$$\frac{1}{2}\|\mathcal{E}_B - \mathcal{F}_B\|_\diamond = \min_{\omega\geqslant 0;\omega\geqslant J_B^{\mathcal{E}-\mathcal{F}}} \|\omega_{B_0}\|_\infty \;\forall\, \mathcal{E},\mathcal{F}\in \mathrm{CPTP}(B_0\to B_1)\,. \tag{A.13}$$

Note that (A.13) can be rewritten as [145]

$$\frac{1}{2}\|\mathcal{E}_B - \mathcal{F}_B\|_\diamond = \min\{\lambda\;:\;\lambda\mathcal{Q}_B\geqslant \mathcal{E}_B-\mathcal{F}_B\,;\mathcal{Q}_B\in\mathrm{CPTP}(B_0\to B_1)\}\,. \tag{A.14}$$

Taking $\mathcal{E}_B=\Theta_{A\to B}[\mathcal{N}_A]$ and $\mathcal{F}_B=\mathcal{M}_B$, $d_{\mathfrak{F}}(\mathcal{N}_A\to\mathcal{M}_B)$ in (3.115) becomes

$$d_{\mathfrak{F}}(\mathcal{N}_A\to\mathcal{M}_B)=\min\{\lambda\;:\;\lambda\mathcal{Q}_B\geqslant\Theta_{A\to B}[\mathcal{N}_A]-\mathcal{M}_B\,,\mathcal{Q}_B\in\mathrm{CPTP}(B_0\to B_1)\,,\;\Theta\in\mathfrak{F}(A\to B)\}\,. \tag{A.15}$$

For the case $\mathfrak{F}=$ MISC, let us start by denoting ω_B as the Choi matrix of $\lambda\mathcal{Q}_B$ and α_{AB} as the Choi matrix of Θ, we can express $d_{\mathfrak{F}}(\mathcal{N}_A\to\mathcal{M}_B)$ as

$$\begin{aligned}
&d_{\mathfrak{F}}\left(\mathcal{N}_A\to\mathcal{M}_B\right)=\min\lambda\\
&\text{subject to}: (1)\;\lambda I_{B_0}\geqslant\omega_{B_0}\,,\;(2)\;\omega_B\geqslant 0\,,\;(3)\;\omega_B\geqslant \mathrm{Tr}_A\left[\alpha_{AB}\left((J_A^{\mathcal{N}})^T\otimes I_B\right)\right]-J_B^{\mathcal{M}}\;,\\
&\qquad\qquad (4)\;\alpha_{AB}\geqslant 0\,,\;(5)\;\alpha_{AB_0}=\alpha_{A_0B_0}\otimes u_{A_1}\,,\;(6)\;\alpha_{A_1B_0}=I_{A_1B_0}\,,\\
&\qquad\qquad (7)\;\mathrm{Tr}[\alpha_{AB}X_{AB}^i]=0\;\forall\, i=1,\ldots,n
\end{aligned} \tag{A.16}$$

where $n\equiv|AB|(|B|-1)$ and $\{X_{AB}^i\}_{i=1}^n$ are the bases of the subspace $\mathfrak{K}_{\mathfrak{F}}$ defined in (3.63). Here, constraints (1-3) are due to diamond norm, constraints (4-6) follow from the requirement of Θ to be a superchannel and constraint (7) is due to the requirement that $\Theta\in\mathfrak{F}$.

Now consider a linear map $\mathcal{L}:\mathbb{R}\oplus\mathrm{Herm}(B)\oplus\mathrm{Herm}(AB)\to\mathrm{Herm}(B_0)\oplus\mathrm{Herm}(B)\oplus\mathrm{Herm}(AB_0)\oplus\mathrm{Herm}(A_1B_0)\oplus^n\mathbb{R}$ where $\oplus^n\mathbb{R}$ denotes $\underbrace{\mathbb{R}\oplus\ldots\oplus\mathbb{R}}_{n}$.

Its action on a generic element $\mu=(\lambda\,,\,\omega_B\,,\,\alpha_{AB})$ of $\mathbb{R}\oplus\mathrm{Herm}(B)\oplus\mathrm{Herm}(AB)$ such that $\lambda\in\mathbb{R}_+\,,\;\omega_B\geqslant 0\,,\;\alpha_{AB}\geqslant 0$ is

$$\begin{aligned}
\mathcal{L}(\mu)\coloneqq\Big(&\lambda I_{B_0}-\omega_{B_0}\,,\omega_B-\mathrm{Tr}\left[\alpha_{AB}\left((J_A^{\mathcal{N}})^T\otimes I_B\right)\right]\;,\;\alpha_{AB_0}-\alpha_{A_0B_0}\otimes u_{A_1}\;,\;\alpha_{A_1B_0}\;,\\
&\mathrm{Tr}[\alpha_{AB}X_{AB}^1]\;,\;\ldots\;,\mathrm{Tr}[\alpha_{AB}X_{AB}^n]\Big)
\end{aligned} \tag{A.17}$$

Taking a generic element $\nu = (\beta_{B_0}\ ,\ \gamma_B\ ,\ \tau_{AB_0}\ ,\ \zeta_{A_1B_0}\ , t_1, \ldots, t_n)$ of $\mathrm{Herm}(B_0) \oplus \mathrm{Herm}(B) \oplus \mathrm{Herm}(AB_0) \oplus \mathrm{Herm}(A_1B_0) \oplus^n \mathbb{R}$ such that $\beta_{B_0} \geqslant 0\ ,\ \gamma_B \geqslant 0\ ,$ we have

$$\mathcal{L}^*(\nu) = \Big(\mathrm{Tr}[\beta_{B_0}]\ ,\ \gamma_B - \beta_{B_0} \otimes I_{B_1}\ ,\ \tau_{AB_0} \otimes I_{B_1} - (J_A^{\mathcal{N}})^T \otimes \gamma_B - \tau_{A_0B_0} \otimes u_{A_1} \otimes I_{B_1} + \tau_{A_1B_0} \otimes I_{A_0B_1} + \sum_i t_i X_{AB}^i \Big)\ . \tag{A.18}$$

Following [149], the dual is given by

$$d_{\mathfrak{F}}\left(\mathcal{N}_A \to \mathcal{M}_B\right) = \max\big\{ -\mathrm{Tr}\left[J_B^{\mathcal{M}} \gamma_B\right] + \mathrm{Tr}\left[\zeta_{A_1B_0}\right] \big\} \tag{A.19}$$

where the maximum is subject to

$$\begin{aligned} &\beta_{B_0} \otimes I_{B_1} \geqslant \gamma_B \geqslant 0\ ,\ 1 \geqslant \mathrm{Tr}[\beta_{B_0}]\ , \\ &\zeta_{A_1B_0} \in \mathrm{Herm}(A_1B_0)\ ,\ \tau_{AB_0} \in \mathrm{Herm}(AB_0)\ ,\ t_1, \ldots, t_n \in \mathbb{R}\ , \\ &J_A^{\mathcal{N}} \otimes \gamma_B + \tau_{A_0B_0} \otimes u_{A_1} \otimes I_{B_1} - \tau_{AB_0} \otimes I_{B_1} - \tau_{A_1B_0} \otimes I_{A_0B_1} - \sum_i t_i X_{AB}^i \geqslant 0\ . \end{aligned} \tag{A.20}$$

For the case of $\mathfrak{F} = \mathrm{DISC}$, note that the only distinction is in the choice of basis of the subspace $\mathfrak{K}_{\mathfrak{F}}$. So, in this case, the dual is given by

$$d_{\mathfrak{F}}\left(\mathcal{N}_A \to \mathcal{M}_B\right) = \max\big\{ -\mathrm{Tr}\left[J_B^{\mathcal{M}} \gamma_B\right] + \mathrm{Tr}\left[\zeta_{A_1B_0}\right] \big\} \tag{A.21}$$

where the maximum is subject to

$$\begin{aligned} &\beta_{B_0} \otimes I_{B_1} \geqslant \gamma_B \geqslant 0\ ,\ 1 \geqslant \mathrm{Tr}[\beta_{B_0}]\ , \\ &\zeta_{A_1B_0} \in \mathrm{Herm}(A_1B_0)\ ,\ \tau_{AB_0} \in \mathrm{Herm}(AB_0)\ ,\ t_1, \ldots, t_n \in \mathbb{R}\ , \\ &J_A^{\mathcal{N}} \otimes \gamma_B + \tau_{A_0B_0} \otimes u_{A_1} \otimes I_{B_1} - \tau_{AB_0} \otimes I_{B_1} - \tau_{A_1B_0} \otimes I_{A_0B_1} - \sum_i t_i Y_{AB}^i \geqslant 0\ . \end{aligned} \tag{A.22}$$

Therefore, we see that $d_{\mathfrak{F}}(\mathcal{N}_A \to \mathcal{M}_B)$ is an SDP in the dynamical resource theory of quantum coherence if the free superchannels belong to MISC or DISC.

A.3 Upper bound on the log-robustness of coherence and the log-robustness of quantum Fourier transform channel and the maximally coherent replacement channel

The log-robustness of a channel $\mathcal{N}_A$ can be expressed as

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\max\left\{\mathrm{Tr}\left[\eta_A J_A^{\mathcal{N}}\right] \ : \ \mathcal{D}_A\left(\eta_A\right) = \mathcal{D}_{A_0}\left(\eta_{A_0}\right)\otimes u_{A_1} \ , \ \mathcal{D}_{A_1}\left[\eta_{A_1}\right] = I_{A_1} \ , \ \eta_A \geqslant 0\right\} \tag{A.23}$$

$$= \log\max\left\{|A_0A_1|\,\mathrm{Tr}\left[\frac{\eta_A}{|A_1|}\frac{J_A^{\mathcal{N}}}{|A_0|}\right] \ : \ \mathcal{D}_A\left(\eta_A\right) = \mathcal{D}_{A_0}\left(\eta_{A_0}\right)\otimes u_{A_1} \ , \ \mathcal{D}_{A_1}\left[\eta_{A_1}\right] = I_{A_1} \ , \ \eta_A \geqslant 0\right\} \tag{A.24}$$

Let $\rho_A := \dfrac{\eta_A}{|A_1|}$, $\sigma_A := \dfrac{J_{\mathcal{N}}}{|A_0|}$ and observing that $\rho_A, \sigma_A \in \mathfrak{D}(A)$, we can rewrite the above expression as

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log\max\left\{|A_0A_1|\mathrm{Tr}\left[\rho_A\sigma_A\right] \ : \ \mathcal{D}_A(\rho_A) = \mathcal{D}_{A_0}\left(\rho_{A_0}\right)\otimes\frac{u_{A_1}}{|A_1|} \ , \ \mathcal{D}_{A_1}[\rho_{A_1}] = u_{A_1} \ , \ \rho_A \geqslant 0\right\} \tag{A.25}$$

Recall that the maximum of the trace of the product of two density matrices is 1 and can be obtained if the two density matrices are same and pure, i.e., $\rho_A = \sigma_A$ and $\mathrm{Tr}[\sigma_A^2] = \mathrm{Tr}[\sigma_A] = 1$. Therefore, for any channel $\mathcal{N}_A$

$$LR_{\mathfrak{C}}(\mathcal{N}_A) \leqslant \log|A_0A_1| \tag{A.26}$$

with $|A_0| = |A_1| = d$, the upper bound on the log-robustness of coherence becomes

$$LR_{\mathfrak{C}}(\mathcal{N}_A) \leqslant \log d^2 \ . \tag{A.27}$$

To achieve this upper bound of the log-robustness of coherence, we require

$$\frac{\eta_A}{|A_1|} = \frac{J_A^{\mathcal{N}}}{|A_0|} \tag{A.28}$$

and $\sigma_A = \dfrac{J_A^{\mathcal{N}}}{|A_0|}$ to be pure. Thus, for σ_A to be pure, $\mathcal{N}_A$ has to be a unitary channel, i.e.,

$$J_A^{\mathcal{N}} = \sum_{x,y}|x\rangle\langle y|\otimes U|x\rangle\langle y|U^\dagger \ , \tag{A.29}$$

Since for a unitary channel $J_A^{\mathcal{N}} = \eta_A$ ($|A_0| = |A_1|$ for a unitary channel), $J_A^{\mathcal{N}}$ has to follow the constraints

in (A.23) which can be expressed as

$$\mathcal{D}_A(J_A^{\mathcal{N}}) = \mathcal{D}_{A_0}\left(J_{A_0}^{\mathcal{N}}\right) \otimes u_{A_1} \ , \tag{A.30}$$

$$\mathcal{D}_{A_1}[J_{A_1}^{\mathcal{N}}] = I_{A_1} \ , \tag{A.31}$$

$$J_A^{\mathcal{N}} \geqslant 0 \ . \tag{A.32}$$

Constraint in (A.32) follows from the definition of a Choi matrix of a channel and (A.31) follows trivially for the Choi matrix of any unitary channel U. From constraint in (A.30), we can find the condition on the unitary matrix to achieve the upper bound of the log-robustness of coherence of channels in the following way. First, we can write the lhs of (A.30) as

$$\mathcal{D}_A(J_A^{\mathcal{N}}) = \mathcal{D}_A\left(\sum_{x,y} |x\rangle\langle y| \otimes U|x\rangle\langle y|U^\dagger\right) \tag{A.33}$$

$$= \sum_{x,y,i,j} \delta_{x,i}\delta_{y,j}|i\rangle\langle i| \otimes \langle j|U|x\rangle\langle y|U^\dagger|j\rangle|j\rangle\langle j| \tag{A.34}$$

$$= \sum_{i,j} |i\rangle\langle i| \otimes \langle j|U|i\rangle\langle i|U^\dagger|j\rangle|j\rangle\langle j| \tag{A.35}$$

$$= \sum_{i,j} |u_{ij}|^2 \ |i\rangle\langle i| \otimes |j\rangle\langle j| \tag{A.36}$$

where $u_{ij} = \langle j|U|i\rangle$. Simplifying the rhs of constraint in (A.30) we get

$$\mathcal{D}_{A_0}\left(J_{A_0}^{\mathcal{N}}\right) \otimes u_{A_1} = I_{A_0} \otimes u_{A_1} \ . \tag{A.37}$$

Now equating (A.36) and (A.37) we get

$$\sum_{i,j} |u_{ij}|^2 \ |i\rangle\langle i| \otimes |j\rangle\langle j| = I_{A_0} \otimes u_{A_1} \ , \tag{A.38}$$

which implies that

$$|u_{ij}|^2 = \frac{1}{|A_1|} = \frac{1}{d} \ \ \forall \ \ i, j \tag{A.39}$$

of a unitary channel that achieves the upper bound of the log-robustness of coherence.

A.3.1 Log-robustness of coherence of QFT channel

The action of a Quantum Fourier Transform (QFT) on a basis vector is given by

$$F_d|i\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \omega^{ik} |k\rangle \tag{A.40}$$

where d denotes the dimension of the system the QFT is acting on and ω is the complex root of unity. Therefore, a general element of the QFT matrix can be written as

$$\langle j|F_d|i\rangle = \frac{\omega^{ij}}{\sqrt{d}} \,. \tag{A.41}$$

It is trivial to check that it follows the condition required to achieve the upper bound of the log-robustness of coherence, i.e.,

$$|\langle j|F_d|i\rangle|^2 = \left|\frac{\omega^{ij}}{\sqrt{d}}\right|^2 \tag{A.42}$$

$$= \frac{1}{d} \,. \tag{A.43}$$

Therefore, the log-robustness of quantum Fourier transform channel ($\mathcal{N}_A^{F_d}$) is

$$LR_{\mathfrak{C}}(\mathcal{N}_A^{F_d}) = \log d^2 = 2 \log d \,. \tag{A.44}$$

A.3.2 Log-robustness of maximal replacement channels

The Choi matrix of a maximal replacement channels $\mathcal{N}_A$ is given by

$$J_A^{\mathcal{N}} = \mathsf{id} \otimes \phi_{A_1}^{+} \tag{A.45}$$

where the density matrix of the maximally coherent state in dimension d is given by

$$\phi_{A_1}^{+} = \frac{1}{d} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}_{d \times d} \tag{A.46}$$

Using (A.23), we can find the log-robustness of $J_A^{\mathcal{N}}$ as follows

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log \max \mathrm{Tr}[\eta_A J_A^{\mathcal{N}}] \tag{A.47}$$

$$= \log \max \mathrm{Tr}[\eta_A (\mathsf{id}_{A_0} \otimes \phi^+_{A_1})] \tag{A.48}$$

$$= \log_2 \max \mathrm{Tr}[\eta_{A_1} \phi^+_{A_1}] \tag{A.49}$$

where η_A follows the following conditions

$$\mathcal{D}_A(\eta_A) = \mathcal{D}_{A_0}(\eta_{A_0}) \otimes u_{A_1} \ , \ \mathcal{D}_{A_1}[\eta_{A_1}] = I_{A_1} \ , \ \eta_A \geqslant 0 \tag{A.50}$$

(A.49) suggests that the Choi matrix of the maximal replacement channel is the log of the max of the sum of all the elements of η_{A_1} divided by d. And from the constraints on η_A, we know that the diagonal elements of η_{A_1} are all 1's. Hence, the maximum value would be obtained when all the elements of η_{A_1} are equal to 1. If any off-diagonal element (and so its diagonally opposite element) are greater than 1 then the determinant of the leading principal minor will be negative which contradicts the positive semi-definite constraint imposed on η_A. Hence, the log-robustness of a maximally coherent replacement channel, i.e., a channel $\mathcal{N}_A$ that replaces any input by the maximally coherent state of dimension d is given by

$$LR_{\mathfrak{C}}(\mathcal{N}_A) = \log d \, . \tag{A.51}$$

In conclusion, we find that the ratio between the log-robustness of coherence of QFT channel and the maximal coherent replacement channels is always 2 implying that 2 maximal replacement channels are required to simulate a QFT channel. One interpretation of this finding can be given by combining the resources of entanglement and coherence. As noted above, a distinguishing feature between the QFT and the maximal replacement channel is that the latter is entanglement-breaking. Hence even though it can generate maximal coherence, in the process of doing so it will destroy any entanglement the primary system may have with an external one. For example, when acting on the first subsystem in the entangled state $\sqrt{1/2}(|00\rangle + |11\rangle)$, the qubit replacement channel will yield $\rho_1 = \phi_2^+ \otimes I/2$, while the QFT will yield $|\psi_2\rangle = \sqrt{1/2}(|0+\rangle + |1-\rangle)$, where $|\pm\rangle = \sqrt{1/2}(|0\rangle \pm |1\rangle)$.

A combined resource theory of entanglement and coherence was studied in Ref. [209]. In particular, the asymptotic convertibility of states using local incoherent operations was considered, and for a given state ρ^{AB}, one can define the optimal rate sum $R_C + R_E$ of coherent bits (R_C) and entangled bits (R_E) needed to asymptotically prepare the state ρ^{AB}. It turns out that state $|\psi_2\rangle$ has twice the resource cost as state

ρ_1, when resource cost is measured in terms of the smallest rate sum $R_C + R_E$. Here, R_C and R_E are the asymptotic rates of coherence and entanglement used to build the state. Hence the greater resource power of the QFT versus the replacement channel becomes operationally manifest in this way.

A.4 Proof of proposition 1 (Maximal Probability of Success in Distillation of Dynamical Coherence under MISC)

For a given error tolerance ϵ and quantum channel $\mathcal{N}$, the maximal probability of achieving $F(\sigma, \phi_d^+) \geqslant 1-\epsilon$ with $\Theta(\mathcal{N}) = p\sigma$ is equal to

$$P^{\epsilon}_{\mathfrak{F}_{\text{prob}}}(\mathcal{N}; d) = \max\{p \,|\, \Theta[\mathcal{N}] = p\phi^+_{d,\epsilon}, \Theta \in \mathfrak{S}_{\text{prob}}\}. \tag{A.52}$$

with $\phi^+_{d,\epsilon} = (1-\epsilon)\phi^+_d + \epsilon(I - \phi^+_d)/(d-1)$.

Similar to ϕ^+_d, state $\phi^+_{d,\epsilon}$ is also invariant under twirling operation, which is a free channel in the theory of coherence; that is $\mathcal{T}(\phi^+_{d,\epsilon}) = \phi^+_{d,\epsilon}$. Then for any superchannel Θ satisfying $\Theta[\mathcal{N}] = p\phi^+_{d,\epsilon}$, we also have $\mathcal{T} \circ \Theta[\mathcal{N}] = p\phi^+_{d,\epsilon}$. Mathematically, the Choi operator of $\mathcal{T} \circ \Theta$ is given by

$$J^{\mathcal{T}\circ\Theta}_{A_0A_1B_1} = \text{Tr}_{B_1}[\phi^+_{d,B_1} \cdot J^{\Theta}] \otimes \phi^+_{d,B_1} + \frac{\text{Tr}_{B_1}[(I_{B_1} - \phi^+_{d,B_1}) \cdot J^{\Theta}]}{d-1} \otimes (I_{B_1} - \phi^+_{d,B_1}). \tag{A.53}$$

Define X and Y as

$$X^{\text{T}_{A_0A_1}} := \text{Tr}_{B_1}[\phi^+_{d,B_1} \cdot J^{\Theta}], \tag{A.54}$$

$$Y^{\text{T}_{A_0A_1}} := \text{Tr}_{B_1}[(\mathbb{1}_{B_1} - \phi^+_{d,B_1}) \cdot J^{\Theta}]/(d-1), \tag{A.55}$$

which are positive semidefinite operators on systems $A = A_0A_1$. With these notations, the Choi operator of $\mathcal{T} \circ \Theta$ can be simplified as

$$J^{\mathcal{T}\circ\Theta} = X^{\text{T}} \otimes \phi^+_d + Y^{\text{T}} \otimes (\mathbb{1} - \phi^+_d). \tag{A.56}$$

Hence the maximal probability of distilling ϕ_d^+ with error tolerance ϵ under $\mathrm{MIS}_{\mathrm{sub}}$ can be characterized as

$$\begin{aligned} P^{\epsilon}_{\mathrm{MISC}_{\mathrm{prob}}}(\mathcal{N};d) = \max \quad & \mathrm{Tr}[Z_A \cdot J_A^{\mathcal{N}}] \\ \text{s.t.} \quad & \mathrm{Tr}[X_A \cdot J_A^{\mathcal{N}}] \geqslant (1-\epsilon)\mathrm{Tr}[Z_A \cdot J_A^{\mathcal{N}}], \\ & 0 \leqslant X_A \leqslant Z_A \leqslant \rho_{A_0} \otimes I_{A_1}, \\ & \mathcal{D}(Z) = d\mathcal{D}(X),\ \mathrm{Tr}[\rho_{A_0}] = 1, \end{aligned} \tag{A.57}$$

which is theorem 1.

Appendix B

Appendix for chapter 4

B.1 Interconversion Distance

We define the interconversion distance from a state $\rho \in \mathfrak{D}(A_0)$ to another state $\sigma \in \mathfrak{D}(A_1)$ as

$$d(\rho_{A_0} \to \sigma_{A_1}) = \frac{1}{2} \min_{\mathcal{E} \in \mathrm{CSPO}(A_0 \to A_1)} \|\mathcal{E}(\rho) - \sigma\|_1 \tag{B.1}$$

$$= \min_{\mathcal{E} \in \mathrm{CSPO}} \left(\max_{0 \leqslant P \leqslant I} \mathrm{Tr}\left[(\mathcal{E}(\rho) - \sigma) P \right] \right) \tag{B.2}$$

Using the dual of trace norm, we can express the above interconversion distance as follows

$$d(\rho \to \sigma) = \min \mathrm{Tr}[X + Y] \tag{B.3}$$

$$\text{s.t.} \begin{pmatrix} X & \mathcal{E}(\rho) - \sigma \\ \mathcal{E}(\rho) - \sigma & Y \end{pmatrix} \geqslant 0 \,, \tag{B.4}$$

$$X \geqslant 0 \,,\ Y \geqslant 0 \,, \tag{B.5}$$

$$J^{\mathcal{E}}_{A_0 A_1} \geqslant 0 \,,\ J^{\mathcal{E}}_{A_0} = I_{A_0} \,, \tag{B.6}$$

$$\frac{J^{\mathcal{E}}_{A_0 A_1}}{|A_0|} \in \mathrm{STAB} \tag{B.7}$$

B.2 Proof of additivity of min-relative entropy of magic for qubits

To prove the additivity of min-relative entropy of magic for qubits, first note that the projector onto the support of a qubit state is identity if the state is mixed, else it is the state itself if it is pure. For the proof, we construct the following four possible cases for qubits ρ_1 or ρ_2

1. For $\rho_1, \rho_2 > 0$, we get

$$D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) = -\log_2 \max_{\psi \in \overline{\mathrm{STAB}}} \mathrm{Tr}[(P_{\rho_1} \otimes P_{\rho_2})\psi] \tag{B.8}$$

$$= -\log_2 \max \mathrm{Tr}[(I \otimes I)\psi] \tag{B.9}$$

$$= 0 \tag{B.10}$$

$$= D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2) \tag{B.11}$$

2. For $\rho_1 > 0$ and $\rho_2 = |\chi\rangle\langle\chi|$, we get

$$D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) = -\log_2 \max_{\psi \in \overline{\mathrm{STAB}}(A_1 A_2)} \mathrm{Tr}[(P_{\rho_1} \otimes P_{\rho_2})\psi] \tag{B.12}$$

$$= -\log_2 \max_{\phi \in \overline{\mathrm{STAB}}(A_2)} \mathrm{Tr}[|\chi\rangle\langle\chi|\phi] \tag{B.13}$$

$$= D_{\min}^{\mathrm{STAB}}(\rho_2) \tag{B.14}$$

$$= D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2) \tag{B.15}$$

3. For $\rho_1 = |\chi\rangle\langle\chi|$ and $\rho_2 > 0$, we get the same result as obtained in 2, i.e.,

$$D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) = D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2) \tag{B.16}$$

4. For the case when both ρ_1 and ρ_2 are pure and let $\rho_1 = |\chi\rangle\langle\chi|$ and $\rho_2 = |\omega\rangle\langle\omega|$, we get

$$D_{\min}^{\mathrm{STAB}}(\rho_1 \otimes \rho_2) = -\log_2 \max_{\psi \in \overline{\mathrm{STAB}}} \mathrm{Tr}[(|\chi\rangle\langle\chi| \otimes |\omega\rangle\langle\omega|)\psi] \tag{B.17}$$

$$= -\log_2 F(|\chi\rangle\langle\chi| \otimes |\omega\rangle\langle\omega|) \tag{B.18}$$

$$= -\log_2 \left(F(|\chi\rangle\langle\chi|)F(|\omega\rangle\langle\omega|)\right) \tag{B.19}$$

$$= D_{\min}^{\mathrm{STAB}}(\rho_1) + D_{\min}^{\mathrm{STAB}}(\rho_2) \tag{B.20}$$

where the second equality follows from the definition of stabilizer fidelity as defined in [257]. The third equality follows from Theorem 5 and Corollary 3 of [257].

Therefore, for single-qubit states we find that the min-relative entropy of magic is additive.

B.3 Robustness of magic

We define the robustness of magic of a quantum state as

$$\mathrm{R}(\rho) = \min\left\{\lambda \geqslant 0 : \frac{\rho + \lambda\sigma}{\lambda + 1} \in \mathrm{STAB},\ \sigma \in \mathrm{STAB}\right\} \tag{B.21}$$

which is slightly different from how it was originally defined in [253]. We use this definition because any resource monotone must be zero for free elements. Likewise, we define channel robustness of magic of a quantum channel $\mathcal{N}$ as

$$\mathrm{R}(\mathcal{N}_A) = \min\left\{\lambda \geqslant 0 : \frac{\mathcal{N} + \lambda\mathcal{E}}{\lambda + 1} \in \mathrm{CSPO},\ \mathcal{E} \in \mathrm{CSPO}\right\} \tag{B.22}$$

which agains differs slightly from the definition of channel robustness of magic in [258].

Both these quantities are magic monotones and are sub-multiplicative under tensor products. Therefore, the log of the robustness of magic (denoted as LR) is sub-additive i.e.,

$$\mathrm{LR}(\rho^{\otimes m}) \leqslant m\mathrm{LR}(\rho)\,, \tag{B.23}$$

$$\mathrm{LR}(\mathcal{N}^{\otimes m}) \leqslant m\mathrm{LR}(\mathcal{N})\,. \tag{B.24}$$

where $\mathrm{LR}(\rho) = \log(1 + \mathrm{R}(\rho))$ and $\mathrm{LR}(\mathcal{N}) = \log(1 + \mathrm{R}(\mathcal{N}))$.

B.4 Hypothesis testing relative entropy of magic

The hypothesis testing relative entropy of magic or the operator smoothed min-relative entropy of magic is defined as

$$D_{\min}^{\epsilon,\,\mathrm{STAB}}(\rho) = \min_{\sigma\in\mathrm{STAB}} D_{\min}^{\epsilon}(\rho\|\sigma) \tag{B.25}$$

$$= \min_{\sigma\in\mathrm{STAB}}(-\log\min \mathrm{Tr}[E\sigma] \tag{B.26}$$

$$\text{s.t. } 0 \leqslant E \leqslant I, \tag{B.27}$$

$$\mathrm{Tr}[E\rho] \geqslant 1 - \epsilon\ ) \tag{B.28}$$

For $\epsilon = 0$, the hypothesis testing relative entropy of magic becomes equal to the min-relative entropy of magic, i.e., $D_{\min}^{\epsilon=0,\,\mathrm{STAB}}(\rho) = D_{\min}^{\mathrm{STAB}}(\rho)$.

B.5 Proof of proposition 5

First we note that for any $\mathcal{E}_A \in \mathrm{CPTP}(A)$ and any $\Theta \in \mathrm{CSPSC}(A \to B_1)$ we have

$$D_{\min}\left(\psi^{\otimes k} \| \Theta[\mathcal{E}]\right) = -\log_2 \mathrm{Tr}\left[\psi^{\otimes k} \Theta[\mathcal{E}]\right] \tag{B.29}$$

$$\geqslant D_{\min}^{\mathrm{STAB}}(\psi^{\otimes k}) \tag{B.30}$$

$$= k\, D_{\min}^{\mathrm{STAB}}(\psi) \tag{B.31}$$

where the inequality follows from the definition of min-relative entropy of magic for states and the last equality follows from its additivity for single-qubit states.

The hypothesis testing relative entropy [319, 331] between two states ρ_1 and ρ_2 is given by

$$D_{\mathrm{Hyp}}^{\epsilon}(\rho_1 \| \rho_2) \coloneqq -\log_2 \min\{\mathrm{Tr}[M\rho_2] : 0 \leqslant M \leqslant I\,,\ \mathrm{Tr}[M\rho_1] \geqslant 1-\epsilon\}\,. \tag{B.32}$$

and its channel counterpart can be given as

$$D_{\mathrm{Hyp}}^{\epsilon}\left(\mathcal{N}_A \| \mathcal{M}_A\right) \coloneqq \sup_{\psi \in \mathfrak{D}(R_0 A_0)} D_{\mathrm{Hyp}}^{\epsilon}\left(\mathcal{N}(\psi_{R_0 A_0}) \| \mathcal{M}(\psi_{R_0 A_0})\right) \tag{B.33}$$

Using this definition, we then have

$$k\, D_{\min}^{\mathrm{STAB}}(\psi) \leqslant \min_{\mathcal{E} \in CSPO} D_{\min}(\psi^{\otimes k} \| \Theta[\mathcal{E}]) \tag{B.34}$$

$$\leqslant \min_{\mathcal{E} \in CSPO} D_{\mathrm{Hyp}}^{\epsilon}(\Theta[\mathcal{N}] \| \Theta[\mathcal{E}]) \tag{B.35}$$

$$\leqslant \min_{\mathcal{E} \in CSPO} D_{\mathrm{Hyp}}^{\epsilon}(\mathcal{N} \| \mathcal{E}) \tag{B.36}$$

where the second inequality follows from the definition of hypothesis testing relative entropy and the last inequality follows from the data-processing inequality. And therefore, we get

$$\mathrm{DISTILL}_{\psi}^{\epsilon}(\mathcal{N}_A) \leqslant \frac{\min_{\mathcal{E} \in \mathrm{CSPO}(A_0 \to A_1)} D_{\mathrm{Hyp}}^{\epsilon}(\mathcal{N} \| \mathcal{E})}{D_{\min}^{\mathrm{STAB}}(\psi)} \tag{B.37}$$

which for exact distillation process (i.e., $\epsilon = 0$) will become

$$\mathrm{DISTILL}_{\psi}(\mathcal{N}_A) \leqslant \frac{D_{\min}^{\mathrm{CSPO}}(\mathcal{N}_A)}{D_{\min}^{\mathrm{STAB}}(\psi)} \tag{B.38}$$

B.6 Single qubit Unitary CSPOs

Table B.1 lists the set of 24 unitary gates which are completely stabilizer preserving along with corresponding (unnormalized) Choi matrices. Table B.2 gives an account of the states generated by these unitary CSPOs. Since a single qubit state can be represented as a vector $(r_1, r_2, r_3)^T$ in the Bloch sphere, we will give below the vectors to which this vector transforms on the application of the above unitaries.

Unitary gate	state corresponding to associated Choi matrix
I	$\|00\rangle + \|11\rangle$
X	$\|01\rangle + \|10\rangle$
Z	$\|00\rangle - \|11\rangle$
XZ	$\|01\rangle - \|10\rangle$
H	$\|0+\rangle + \|1-\rangle$
HX	$\|0-\rangle + \|1+\rangle$
HZ	$\|0+\rangle - \|1-\rangle$
HXZ	$\|0-\rangle - \|1+\rangle$
S	$\|00\rangle + i\|11\rangle$
XS	$\|01\rangle + i\|10\rangle$
ZS	$\|00\rangle - i\|11\rangle$
XZS	$\|01\rangle - i\|10\rangle$
HS	$\|0+\rangle + i\|1-\rangle$
HSZ	$\|0+\rangle - i\|1-\rangle$
HXS	$\|0-\rangle + i\|1+\rangle$
$HXSZ$	$\|0-\rangle - i\|1+\rangle$
SH	$\|0 + i\rangle + \|1 - i\rangle$
SHZ	$\|0 + i\rangle - \|1 - i\rangle$
SHX	$\|0 - i\rangle + \|1 + i\rangle$
$SHXZ$	$\|0 - i\rangle - \|1 + i\rangle$
SHS	$\|0 + i\rangle + i\|1 - i\rangle$
$SHSZ$	$\|0 + i\rangle - i\|1 - i\rangle$
$SHSX$	$i\|0 - i\rangle + \|1 + i\rangle$
$SHSXZ$	$i\|0 - i\rangle - \|1 + i\rangle$

Table B.1: Unitary CSPOs and their Choi matrices.

Unitary gate	Transformed vector
I	r_1, r_2, r_3
SH	r_2, r_3, r_1
HSZ	r_3, r_1, r_2
X	$r_1, -r_2, -r_3$
$SHXZ$	$r_2, -r_3, -r_1$
HS	$r_3, -r_1, -r_2$
Z	$-r_1, -r_2, r_3$
SHX	$-r_2, -r_3, r_1$
$HXSZ$	$-r_3, -r_1, r_2$
Y	$-r_1, r_2, -r_3$
SHZ	$-r_2, r_3, -r_1$
HXS	$-r_3, r_1, -r_2$
SHS	$r_1, r_3, -r_2$
HZ	$r_3, r_2, -r_1$
XZS	$r_2, r_1, -r_3$
$SHSX$	$r_1, -r_3, r_2$
H	$r_3, -r_2, r_1$
ZS	$r_2, -r_1, r_3$
$SHSZ$	$-r_1, r_3, r_2$
HX	$-r_3, r_2, r_1$
S	$-r_2, r_1, r_3$
$SHSXZ$	$-r_1, -r_3, -r_2$
HY	$-r_3, -r_2, -r_1$
XS	$-r_2, -r_1, -r_3$

Table B.2: Possible transformations of a Bloch vector using unitary CSPOs.

Appendix C

Copyright and content reuse statements

In this appendix, we reproduce the statements we received from the sources of our published works [37, 91], allowing us to reuse the content of these articles in this thesis.