2025-09-16

Maps and Higher-Order Maps for the Manipulation of Quantum Resources

Zanoni, Elia

Zanoni, E. (2025). Maps and higher-order maps for the manipulation of quantum resources (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca.

https://hdl.handle.net/1880/122862

UNIVERSITY OF CALGARY

Maps and Higher-Order Maps for the Manipulation of Quantum Resources

by

Elia Zanoni

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

GRADUATE PROGRAM IN MATHEMATICS AND STATISTICS

CALGARY, ALBERTA
SEPTEMBER, 2025

Abstract

In this Thesis, we present the framework of quantum resource theories, a versatile and valuable mathematical framework for describing quantum phenomena and evaluating quantum advantages. We first focus on the resource theory of entanglement, useful for investigating the quantum advantages in communication protocols between distant parties. We derive a closed-form expression for computing how well distant parties can approximate a desired target state using a given initial state and performing only local operations and classical communication. We then utilize this expression to characterize a phenomenon known as entanglement embezzlement.

Next, we focus on the resource theory of quantum thermodynamics. We provide a formal definition of cooling and heating and determine how much an agent, allowed to perform only closed thermal operations, can heat and cool a target system using another system that is out of thermal equilibrium. Moreover, we demonstrate that the ability to heat and cool qubits completely characterizes the conversion of quasi-classical states in the resource theory of thermodynamics.

Then, we identify a property common in many resource theories: An operation is free if and only if its Choi matrix (properly renormalized) is a free state. This property simplifies the numerical and analytical solution of many problems in these resource theories. We refer to such resource theories as 'Choi-defined' resource theories and investigate their properties.

Lastly, we study higher-order quantum maps, a necessary step in defining resource theories at higher orders, where the resources are higher-order maps. We develop a system based on types to label higher-order maps. We define the sequential and parallel composition of higher-order maps, we generalize the concepts of Hermitian-preserving and completely positive-preserving maps, and we show that the Choi isomorphism, properly extended to higher-order maps, preserves Hermitian-preserving and completely positive-preserving higher-order maps.

Preface

This Thesis is based on the following published works

[1] E. Zanoni, T. Theurer, and G. Gour, Complete characterization of entanglement embezzlement, Quantum **8**, 1368 (2024),

[2] T. Theurer, E. Zanoni, C. M. Scandolo, and G. Gour, Thermodynamic state convertibility is determined by qubit cooling and heating, New J. Phys. **25**, 123017 (2023),

[3] E. Zanoni and C. M. Scandolo, Choi-defined resource theories, Phys. Rev. A **111**, 062407 (2025),

and on the unpublished work 'Higher-order quantum maps' done in collaboration with S. Steakley and C. M. Scandolo. Specifically,

- Section 2.3 presents the results about closed thermal operations that appear in Ref. [2] (Proposition 2.3.2 and Proposition 2.3.3). Since I was not the first author of the paper, the presentation in this Thesis differs from the one in Ref. [2]. However, to preserve the validity of the results, the proofs are mostly unchanged.
- Chapter 3 presents the results of Ref. [1], with minor changes to adapt the paper to this Thesis.
- Chapter 4 presents the results of Ref. [2], excluding the ones already presented in Section 2.3. The same considerations made for Section 2.3 apply to this Chapter as well. The only exception is subsection 4.1.2, which presents the results of Marcy Orr's undergraduate thesis [4], of which I was one of the supervisors.
- Chapter 5 presents the results of Ref. [3], with minor changes to adapt the paper to this Thesis.
- Chapter 6 presents the unpublished results concerning higher-order quantum maps developed with S. Steakley and C. M. Scandolo.

Chapter 4, Chapter 5, Chapter 5, and Chapter 6 contain only original results unless explicitly stated otherwise. Chapter 2 contains a very brief review of quantum resource theories based on existing literature, except for Proposition 2.3.2 and Proposition 2.3.3, which are original results. Chapter 1 and Chapter 7 serve as the introduction and conclusion, respectively.

The works listed above were done in collaboration with other members of the Quantum Information Group at the University of Calgary. My contributions and the contributions of each coauthor are as follows.

- Gilad Gour and Carlo Maria Scandolo acted as group leaders and supervisors, providing ideas, directions, and reviewing the work at every stage.
- In Ref. [1], I derived all the results and wrote the manuscript. Thomas Theurer supervised my work, helping with the simplification of some proofs, fixing minor gaps, and reviewing the manuscript.
- In Ref. [2], Thomas Theurer wrote the manuscript. He derived the results about the minimum and maximum temperatures for cooling and heating (Section 4.1) and the results about closed thermal operations (Proposition 2.3.2 and Proposition 2.3.3). I derived the results about the alternative characterization of state conversion (Section 4.2).
- In Ref. [3], I derived all the results and wrote the manuscript.
- The work ‘Higher-order quantum maps’ started from ideas that Carlo Maria Scandolo and I had while working on Ref. [3]. I derived many of the results regarding types and the results concerning sets of higher-order linear maps and higher-order Hermitian maps. Later, Samuel Steakley joined the project. I worked with him on all the results concerning the product of types, the product of maps, completely positive-preserving maps, and the Choi isomorphism.

Acknowledgments

Many of, if not all, the people I met in these four years, and more generally in my life, contributed to a greater or lesser extent to this achievement. For this reason, I want to start these acknowledgments with a great thank you to all the friends, relatives, professors, colleagues, and coaches whom I will not mention explicitly below. Some of them fostered my passion for physics and mathematics, others helped and supported me along the way, and some others improved me as a person. I am grateful to all of them.

The first (explicit) and most important thank you goes to my supervisor, Prof. Carlo Maria Scandolo. When we started, four years ago, I was a complete novice in quantum information, and I am grateful for everything he taught me in these years. Every time I stepped into his office, I had no idea when the meeting would end. We spent hours discussing the projects we were working on and various other topics in quantum information. Most of all, I am thankful to him because he cared about me. Like many Ph.D. students, I had low moments during my Ph.D., and Prof. Carlo Maria Scandolo helped me overcome them, listening and sharing his personal experience.

Prof. Gilad Gour deserves a great thank you as well. As a co-supervisor, he stepped in during the COVID years, when Prof. Carlo Maria Scandolo was blocked in Italy, and his ideas, experience, and support have been essential in these four years.

I want to thank the other two members of the supervisory committee, Prof. Barry Sanders and Prof. David Feder, as well. I thank them for their commitment: not many other Ph.D. students had as many supervisory committee meetings as I did. In all of them, they listened carefully to the presentation of my work and, with their questions, they helped me understand where I could improve and gave me new research ideas.

A huge thank you goes to Thomas Theurer. He was a postdoc in the Quantum Information Group, and he taught me how to write a scientific paper properly. I am thankful for all the hours we spent working side by side on the paper about entanglement embezzlement. His attention to detail is truly German, and I have learned a lot from the way he works.

Another big thank you goes to Gaurav Saxena, my first and only office mate and another of the members of the Quantum Information Group. He gave me all the help I needed to get started in the department, and every time I had a question (whether a scientific conundrum or practical issue), I knew he would have an answer. In addition, he had extensive knowledge of Calgary's restaurants, and I am very grateful that he shared it with us.

All the other members of the Quantum Information Group deserve a big thank you as well. Starting with Samuel Steakley, with whom I worked for more than a year on the project on higher-order quantum maps. Both of our antipodal approaches to the topic were essential for solving the challenging problems we faced.

Our interactions were not limited to scientific projects: I can still remember the amazing lunch he prepared for me and other members of the group this Spring. Next, I thank Takla Nateeboon and Doyeong Kim. They shared the office next to mine, which led to many hours spent discussing our projects and almost as many hours chatting about anything but mathematics or physics. I thank Marcy Orr and Eve Taplin for the passion and commitment they showed with their undergraduate theses. Working with them on those projects has been a pleasure. Lastly, I thank the members of the Quantum Information Group with whom I interacted less, Trace Harms, Durgesh Kumar, Aaron Lengyel, and Ibrahim Sultan. Their questions and points of view have been a key ingredient for the success of our group meetings.

Moving to people outside the academic world, the greatest thank you goes to my mum, my dad, and my sisters, to whom this Thesis is dedicated. Their love made this achievement possible. They have always supported me. Even if I was last in a swimming race or I was not able to put two notes together on the piano, they always had encouraging words. Again, huge thanks go to them.

Speaking of love and support, a huge thank you goes to my partner, Negar Dehghan Noudeh. She experienced firsthand everything I went through in these four years, both the good and the bad moments. We celebrated the firsts, and she was there for the latter. Writing this Thesis would have been extremely harder without her. She continuously supported me, and she patiently helped me every time I got stuck and could not write a single word for days.

I thank all my friends, in particular, Pietro Fiorini, Paola Fazzioni, Don Luca Bernardello, Enrico Visconti, Marco Gasparini, Federico Zanella, Federico Pomari, Federica Lorenzi, and Marco Zonescuti. Whenever I needed a break from a world full of physicists and mathematicians, they were just a phone call away (if I was in Canada) or ready to meet (if I was in Italy).

If you have read these acknowledgments, expecting to find your name mentioned somewhere, but you have arrived here without finding it, this paragraph is for you. I apologize if I have forgotten you in this list, and I am sure you deserve a heartfelt thank you. Since you know me well enough, you are aware that I easily forget many things, and you will laugh reading these lines.

Lastly, I acknowledge the support of the Eric Milner's Graduate Scholarship, the Alberta Graduate Excellence Scholarship (AGES), and the Eyes High International Doctoral Scholarship.

To my parents, Paolo and Tamara,
and my sisters, Giuditta and Rebecca.

Contents

Symbols

Notation	**Description**
$A, B, C, \ldots$	Quantum systems and corresponding Hilbert spaces.
$X, Y, Z, \ldots$	Classical systems and corresponding Hilbert spaces.
$\|A\|$	Dimension of system A or, equivalently, dimension of the associated Hilbert space.
$\mathfrak{L}(\mathcal{H}_A, \mathcal{H}_B)$	The set of all the linear maps from $\mathcal{H}_A$ to $\mathcal{H}_B$.
$\mathfrak{L}(A)$	The set of all the endomorphisms of the Hilbert space A.
$\mathrm{Herm}(A)$	The set of all the Hermitian operators in $\mathfrak{L}(A)$.
$\mathrm{Pos}(A)$	The set of all the positive semidefinite operators in $\mathrm{Herm}(A)$.
$\mathfrak{D}(A)$	The set of all the quantum states of system A, i.e., the set of all operators with unit trace in $\mathrm{Pos}(A)$.
$\rho_A, \sigma_A, \tau_A, \ldots$	Quantum states in $\mathfrak{D}(A)$.
$\mathrm{PURE}(A)$	The set of all the pure quantum states of system A, i.e., the set of all rank-one operators in $\mathfrak{D}(A)$.
$\psi_A, \varphi_A, \chi_A, \ldots$	Pure quantum states in $\mathfrak{D}(A)$.
$\mathrm{Prob}(n)$	The set of probability vectors of dimension n.
$\mathrm{Prob}^{\downarrow}(n)$	The set of probability vectors of dimension n with components in non-increasing order.
$\mathbf{p}, \mathbf{q}, \mathbf{r}, \ldots$	Probability vectors.
$[n]$	The set $\{ 1, \ldots, n \}$.
$\mathfrak{L}(A \to B)$	The set of all the linear maps from $\mathfrak{L}(A)$ to $\mathfrak{L}(B)$.
$\mathrm{CP}(A \to B)$	The set of all the completely positive-preserving maps in $\mathfrak{L}(A \to B)$.
$\mathfrak{D}(A \to B)$	The set of all the quantum channels from system A to system B, i.e., the set of all trace-preserving operators in $\mathrm{CP}(A)$.
$\mathcal{M}^{A \to B}, \mathcal{N}^{A \to B}, \ldots$	Quantum channels in $\mathfrak{D}(A \to B)$.
$\mathbb{1}^A$	The identity operator in $\mathfrak{L}(A)$.
$\mathcal{I}^A$	The identity channel in $\mathfrak{D}(A \to A)$.
$\mathcal{S}^{AB \to BA}$	The swap channel in $\mathfrak{D}(AB \to BA)$.

Notation	**Description**
$\mathfrak{F}(A \to B)$	The set of free quantum channels from system A to system B in a resource theory.
$\mathfrak{F}(A)$	The set of free quantum states of system A in a resource theory.

Acronyms

Notation	**Description**
CP	Completely positive.
LO	Local operations.
LOCC	Local operations and classical communication.
SR	Schmidt rank.
TO	Thermal operations.
CTO	Closed thermal operations.
CRNG	Completely resource-non-generating.
CD	Choi-defined.
CDRT	Choi-defined resource theory.

Chapter 1

Introduction

Quantum mechanics revolutionized our understanding of the laws of nature. We live in a predominantly classical world. We can easily come to terms with Newton's laws, Maxwell's laws, and even partially with special and general relativity. However, quantum mechanics introduced many new, counterintuitive phenomena with no classical counterpart, which not only proved to be true but were also fundamental to understanding molecules, atoms, electrons, and all the other subatomic and fundamental particles. Among these phenomena, entanglement emerges, in Schrödinger's words [5], as the defining trait of quantum mechanics, the one that significantly separates quantum from classical mechanics.

Two quantum systems are entangled if their combined state cannot be described with one state of the first system and one state of the second system (this definition applies to pure states, i.e., states of which we have maximal knowledge; we will later define mixed entangled states as well). This clashes with our classical intuition. If we consider a system of two coins, the possible states are 'Tail-Tail', 'Head-Head', 'Tail-Head', and 'Head-Tail'. In each of these cases, the global state is fully determined by the states of the first and second coins. If we perform the same exercise with a two-dimensional quantum state, e.g., the spin of an electron, we have states that can be decomposed into local states, e.g., $|\uparrow\uparrow\rangle$, and states that cannot, e.g., $|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle$. The latter are called entangled states. In entangled states, the knowledge of the global system is more than the sum of the knowledge of its parts. In certain cases, one can possess zero local knowledge and yet have perfect global knowledge. Moreover, entanglement persists even if the systems are separated, which leads to a violation of the principles of locality and realism [6]. For this reason, quantum theory has been considered incomplete for almost thirty years: There was hope that another theory, which respects locality and realism, was more fundamental than quantum mechanics, and phenomena such as entanglement were just a manifestation of our ignorance of one or more variables. This line of thought came to an abrupt end with Bell's inequality [7–9], which shows that no theory that obeys locality and realism is consistent with quantum mechanics.

Once it was determined that quantum mechanics, however counterintuitive it may be from our classical perspective, is a fundamental theory of nature, the key question became: can we achieve something with quantum mechanics that we cannot achieve with classical mechanics? The first results in this direction are significantly influenced by the seminal work on classical information theory by Shannon [10] and opened the field of quantum information theory. Since then, this field has produced many foundational results and technological applications, for example, in the field of quantum computing [11]. Indeed, since classical information theory is concerned with the transmission of information between distant parties, and quantum mechanics is non-local, it was natural to attempt to exploit the features of quantum mechanics to create communication protocols that outperform any classical ones. For example, Refs. [12–15] propose a public key distribution protocol based on quantum systems that has an arbitrarily high probability of detecting an eavesdropper. Instead, if only classical systems are used, a very cunning and resourceful eavesdropper can intercept the key, read it, copy it, and transmit it again without errors, leaving the parties unaware that the key has been compromised. Of a similar flavour are the results of quantum teleportation [16] and superdense coding [17]. In the first protocol, a sender can teleport an unknown quantum state to a receiver using only a classical channel if they share a pair of entangled systems. Similarly, in the second protocol, a sender can transmit two bits of information by sending a single qubit to the receiver if they share a pair of entangled systems. When the sender and the receiver use protocols involving quantum systems, they can achieve something that would be impossible if they were only using classical systems. The key ingredient of these protocols, at the origin of the quantum advantage, is entanglement.

One may think that entangled states are the key to any quantum advantage. However, this is not the case. For example, considering the computational advantages of quantum algorithms [18–20], one may want to build a quantum computer. However, to truly exploit the advantages of these algorithms, the quantum computer should be superior to a classical one, i.e., it should not be possible to simulate it efficiently with a classical one. The Gottesman-Knill theorem [21] states that quantum circuits composed of Clifford group gates, measurements of Pauli group operators and classically conditioned Clifford group operations acting on a finite number of qubits initially in their ground states can be efficiently simulated classically. A quantum computer performing only these operations, which are called *stabilizer* operations is not universal. However, if one allows for the preparation of 'magic' states, universality is achieved [22]. In this context, magic states are the key for universal quantum computation, not entangled states, which can be prepared from non-entangled states with stabilizer operations (we assume that everything is local and that the computer is not spread among many distant laboratories). Which quantum states provide a quantum advantage is highly situational; a quantum state useful in a particular application is completely useless in another.

Why is entanglement useful in communication protocols? Why are magic states useful for quantum computation? To answer that, we consider the communication protocols in which entanglement does provide an advantage. They all share the same operational setup. There are two agents, typically referred to as Alice and Bob. These agents live in distant laboratories. For simplicity, it is assumed that these two agents have perfect control over any system in their laboratories. However, given their distance, they have no reliable way of exchanging quantum information. Any quantum system that is transmitted from Alice to Bob (or vice versa) would inevitably interact with the environment, and its state would change. Over long distances, environmental noise is so strong that any information encoded initially in the state is lost during transmission. Therefore, Alice and Bob are limited to classical communication only. Indeed, with modern technologies, it is safe to assume that a classical message can be carried over long distances without any errors (or, in the event of an error, it can be corrected). These considerations limit the global operations that can be performed on a composite system, of which Alice holds one subsystem and Bob holds another. The two agents can only perform local operations and classical communication (LOCC). It turns out that with these operations, Alice and Bob can only prepare separable states, while entangled states are impossible to prepare. If Alice and Bob share one (maybe because a third party with better technology shared it with them), they gain access to something they cannot prepare by themselves and can now exploit it. From this operational perspective, entangled states emerge as crucial resources in communication protocols. The example concerning quantum computation is analogous. A circuit composed of stabilizer operations can be efficiently simulated, but with these operations, one cannot prepare all the quantum states. The states that cannot be prepared are called 'magic' states. However, one can simulate non-stabilizer operations by performing stabilizer operations on a magic state. Similarly to what happens in entanglement theory, the agent cannot prepare magic states if restricted to stabilizer operations. Still, if he is given a magic state, he can simulate a non-stabilizer operation.

This operational approach, which is based on identifying which operations are allowed or easy to implement, is at the core of the framework of quantum resource theories. Indeed, in every resource theory, a subset of all quantum operations is identified as free, e.g., LOCC in entanglement theory and stabilizer operations in computation. Moreover, states that can be prepared with free operations are called free, while states that cannot are resources. Once again, entangled states are resources when free operations are LOCC, and non-stabilizer states are resources when free operations are stabilizer operations. This framework was so successful that it has been applied to a variety of different quantum phenomena. A non-exhaustive list includes entanglement [23–27], superselection rules and reference frames [28–32], athermality [33–35], magic [36–40], imaginarity [41–43], asymmetry [44, 45], and coherence [46–51]. We dedicate Chapter 2 to an introduction to quantum resource theories, where we provide details about two resource theories relevant to this Thesis: the resource theory of entanglement and the resource theory of quantum thermodynamics.

In a resource theory, the central question is what can be achieved with a given resource under the constraints imposed by the free operations: If Alice and Bob have access to an entangled state, what can they do with it? Can they convert it to another entangled state using only LOCC? This question is known as the conversion problem. For example, it has been shown that if Alice and Bob share enough entangled qubits, they can concentrate their entanglement into a pair of maximally entangled qubits with only LOCC [23], which can later be used in one of the protocols discussed above. The conversion problem for pure states entanglement has been solved by Nielsen [26]. However, sometimes, it may not be possible to convert the initial state exactly into the target state, but rather into a state that is very close to it. Indeed, it may be so close that it is indistinguishable from the target state by any physical apparatus. In this case, we talk about approximate conversion. In Chapter 3, we present the results concerning approximate conversion previously published in [1]. We derive a closed-form expression to compute how close to the desired pure target state Alice and Bob can get using a given input state. In the second part of Chapter 3, we use this result to characterize a phenomenon known as entanglement embezzlement [52, 53].

In Chapter 4, which is based on Ref. [2], we turn our attention to the resource theory of quantum thermodynamics. Following the example of classical thermodynamics, in which a bath cannot be used to produce work if it is at thermal equilibrium with its environment, in the resource theory of quantum thermodynamics, resources are states that are out of thermal equilibrium with the environment, while free states are states at equilibrium with the environment. Free operations consist of 1) preparing equilibrium states, 2) performing energy-preserving unitary operations, and 3) discarding subsystems (or more precisely, free operations consist of the closure of such operations). In this setup, we focus on cooling and heating.

We first explain what it means to cool and heat a quantum system and then derive formulas to compute the highest and lowest temperatures to which a quantum system can be heated or cooled with a given resource, using only free operations. In addition, we demonstrate that the ability to heat and cool qubits completely characterizes the conversion problem for quasi-classical states, i.e., states that exhibit no coherence between different eigenspaces of the system's Hamiltonian. In other words, a quasi-classical state can be converted into another if the first cools and heats qubits to higher and lower temperatures than the second does.

While Chapter 3 and Chapter 4 focus on specific resource theories, in Chapter 5, we investigate a property common to many resource theories. Indeed, having a common framework for different phenomena allows for the development of rules and tools that can be easily applied to many resource theories. The starting point is the realization that in many important resource theories, for example, the resource theories of separable and non-positive partial transpose entanglement [54–64], magic states [36–40], a channel is free if and only if its Choi matrix (properly renormalized) is a free state. Moreover, these resource theories have been so successful because the one-to-one correspondence between free states and free channels simplifies many problems, allowing for both numerical and analytical solutions. We refer to resource theories that exhibit this one-to-one correspondence as 'Choi-defined' resource theories. We find the conditions under which it is possible to construct such resource theories from a given set of free states (which determines the free operations as well, given the one-to-one correspondence). We then prove properties that are shared by all Choi-defined resource theories. The results presented in this Chapter were published in Ref. [3].

The resource theories presented so far are resource theories of states, in which the resources are states, and agents act on these resources with free channels. However, quantum operations can also be seen as resources. For example, in LOCC, a noiseless quantum channel from Alice to Bob would be a great resource if available. With this perspective, quantum teleportation can be described as a sequence of LOCC 'super'-operations [65, 66] that convert an entangled state into a noiseless quantum channel, which is later used to transmit quantum information. These considerations lead to resource theories of channels [64, 67–72], in which channels are viewed as resources and agents act on them with free super-operations (or superchannels). A critical result concerning superchannels is that they can all be decomposed into a circuit composed of two channels with memory. One can then go one step further and consider circuits with three, four, or n channels with memories. If states are of order zero, channels of order one, and superchannels of order two, these are maps of order higher than two. Interestingly, higher-order maps that cannot be decomposed into channels with memories provide advantages over maps that can [73–79]. Thus, it is natural to approach these higher-order maps with the framework of quantum resource theories. However, there is yet no good framework to describe them. The best attempt is the work by Bisio and Perinotti [80], which is based on *types*. In Chapter 6, we improve the type system proposed in [80]. One of the most significant improvements concerns the parallel composition of higher-order maps, which was missing in the original work. We also generalize the notions of Hermitian-preserving and completely positive-preserving to higher-order maps. We introduce the Choi isomorphism for maps of any order, and we show that it preserves Hermitian-preserving maps and completely positive-preserving maps, a generalization of the results about the Choi isomorphism between states and channels [66]. These results are based on unpublished work done in collaboration with S. Steakley and C. M. Scandolo.

1.1 Notations and Preliminaries

The author of this Thesis assumes that the reader is already familiar with the basic concepts of quantum information. For a proper introduction to the topic, we refer the reader to the great books by Nielsen and Chuang [11], Wilde [81], Watrous [82], and Gour [83]. Here, we provide a very brief summary of the fundamental objects of quantum information, i.e., systems, states, and channels, and introduce relevant notations.

Each quantum system A is associated with a Hilbert space $\mathcal{H}_A$. We will often use A to denote both the system and the Hilbert space. In this Thesis, we consider solely finite-dimensional systems, and we denote the finite dimension of the system (or equivalently, the dimension of the Hilbert space) as $|A|$. If a system is composed of two subsystems, A and B, the global Hilbert space is $\mathcal{H}_A \otimes \mathcal{H}_B$. Such a system and its corresponding Hilbert spaces are denoted as AB.

A quantum state of a system A is a positive semidefinite map in $\mathfrak{L}(A)$ (the set of endomorphisms on $\mathcal{H}_A$) with unit trace. We denote the set of all quantum states of the system A with $\mathfrak{D}(A)$. Quantum states are usually labelled with lowercase Greek letters, e.g., ρ^A, σ^A, τ^A. Of particular interest are rank-one states, also known as pure states. The set of pure states of system A is denoted with $\mathrm{PURE}(A)$, and the states are labelled with the Greek letters $\psi^A, \varphi^A, \chi^A$. The normalized eigenvector of a pure state ψ^A is denoted with $|\psi\rangle^A$, and we have $\psi^A = |\psi\rangle\langle\psi|^A$, where $\langle\psi|^A$ is the dual of $|\psi\rangle^A$.

Mathematically, classical systems are a special case of quantum systems. Each classical system X comes equipped with an orthonormal basis $\left\{ |x\rangle^X \right\}_x$, such that every state of X has the form $\sum_x p_x |x\rangle\langle x|^X$. Moreover, there is a one-to-one correspondence between states of a classical system X and probability vectors in $\mathrm{Prob}(|X|)$:

$$\sum_x p_x |x\rangle\langle x|^X \leftrightarrow \begin{pmatrix} p_1 \\ \vdots \\ p_{|X|} \end{pmatrix}. \tag{1.1}$$

Note that for each state ρ^A of a quantum system A there exists an orthonormal basis $\left\{ \left|x_\rho\right\rangle^A \right\}$ such that $\rho^A = \sum_x p_x |x_\rho\rangle\langle x_\rho|^A$. However, A is not a classical system because not *all* the possible states of A are diagonal in the same basis.

Quantum channels from system A to system B are linear, completely positive (CP), trace-preserving maps from $\mathfrak{L}(A)$ to $\mathfrak{L}(B)$. The set of quantum channels from A to B is denoted with $\mathfrak{D}(A \to B)$. For each channel $\mathcal{M}^{A\to B} \in \mathfrak{D}(A \to B)$ there exists a set of linear operators $\{ K_x \}_x$ from $\mathcal{H}_A$ to $\mathcal{H}_B$, called Kraus operators, such that $\sum_x K_x^\dagger K_x = \mathbb{1}^A$ and for all $\rho^A \in \mathfrak{D}(A)$

$$\mathcal{M}^{A\to B}(\rho^A) = \sum_x K_x \rho^A K_x^\dagger. \tag{1.2}$$

Eq. (1.2) is called the operator-sum representation of the channel $\mathcal{M}^{A\to B}$.

A quantum instrument is a quantum channel with classical-quantum output. For each quantum instrument $\mathcal{E}^{A\to XB} \in \mathfrak{D}(A \to BX)$, there exists a set of completely positive operators $\left\{ \mathcal{E}_x^{A\to B} \right\}$ such that $\mathcal{E}^{A\to XB}(\rho^A) = \sum_x |x\rangle\langle x|^X \otimes \mathcal{E}_x^{A\to B}(\rho^A)$ for all $\rho^A \in \mathfrak{D}(A)$ and $\sum_x \mathcal{E}_x^{A\to B}$ is trace preserving. Alternatively, one can define a quantum instrument as a set $\left\{ \mathcal{E}_x^{A\to B} \right\}$ of CP operators that sum to a trace-preserving operator. The two definitions are equivalent.

A positive operator-valued measure on A (POVM) is a collection $\left\{ F_x^A \right\}_x$ of positive semidefinite matrices on A such that $\sum_x F_x^A = \mathbb{1}^A$.

We conclude this very brief introduction with the definition of the two norms that we will use throughout this Thesis. For any $M \in \mathfrak{L}(A)$, and $1 \leq p < \infty$ the Schatten p-norm is defined as

$$\|M\|_p = \left(\operatorname{Tr}\left(\sqrt{M^\dagger M}\right)^p\right)^{\frac{1}{p}}, \tag{1.3}$$

and $\|M\|_\infty$ is equal to largest eigenvalue of $\sqrt{M^\dagger M}$. For any $k \in [|A|]$ the k-th Ky Fan norm of $M \in \mathfrak{L}(A)$ is

$$\|M\|_{(k)} = m_1 + m_2 + \cdots + m_k,$$

where $m_1 \geq m_2 \geq \cdots \geq m_k$ are the k largest eigenvalues of $\sqrt{M^\dagger M}$. From the Schatten 1-norm, one defines the trace distance between two quantum states $\rho^A, \sigma^A \in \mathfrak{Q}(A)$ as

$$T(\rho^A, \sigma^A) = \frac{1}{2}\left\|\rho^A - \sigma^A\right\|_1, \tag{1.4}$$

the fidelity as

$$F(\rho^A, \sigma^A) = \left\|\sqrt{\sigma^A}\sqrt{\rho^A}\right\|_1, \tag{1.5}$$

and the purified distance as

$$P(\rho^A, \sigma^A) = \sqrt{1 - F^2(\rho^A, \sigma^A)}. \tag{1.6}$$

The Schatten and Ky-Fan norms of vectors follow from the corresponding matrix norms:

$$\begin{aligned} \|\mathbf{v}\|_p &:= \|\operatorname{diag}(\mathbf{v})\|_p, \\ \|\mathbf{v}\|_{(k)} &:= \|\operatorname{diag}(\mathbf{v})\|_{(k)}, \end{aligned} \tag{1.7}$$

where $\operatorname{diag}(\mathbf{v})$ is the diagonal matrix with diagonal entries v_x. From these, one defines the trace distance $T(\mathbf{v}, \mathbf{w})$, the fidelity $F(\mathbf{v}, \mathbf{w})$, and the purified distance $P(\mathbf{v}, \mathbf{w})$ between vectors.

Chapter 2

Quantum Resource Theories

In the last two decades of the twentieth century, it became clear that entangled states provide an advantage in many communication tasks. Examples in cryptography [12–15] and about the transmission of classical [17] and quantum [16] bits show that if the two distant parties involved in the protocols share a pair of entangled states, they can achieve results that would be unattainable with any purely classical protocol.

After these works, quantum entanglement is recognized as a valuable resource, and the works in the 90s focus on how two distant parties can manipulate quantum entanglement [23–26]. These articles share a common framework. The problem is described with an operational approach, i.e., all these works focus on what operations the two distant parties can or cannot perform. Since there is no reliable way of sending quantum systems over long distances while preserving their state, the two distant parties cannot exchange quantum information. As a consequence, the actions that the two parties can perform are only local quantum operations and classical communications (LOCC). Interestingly, entangled states are the ones that the two parties cannot create by using only LOCC. Indeed, they can only create separable states. Once again, entanglement is considered a valuable and scarce resource because it cannot be created freely.

The operational approach described here has been applied in the last 20 years to a multitude of quantum phenomena and constitutes the core of all quantum resource theories. A non-exhaustive list of quantum resource theories includes entanglement [23–27], superselection rules and reference frames [28–32], athermality [33–35], magic [36–40], imaginarity [41–43], asymmetry [44, 45], and coherence [46–51]. Given the variety of the results and methods applied in these resource theories, it is beyond the scope of this thesis to give a complete account of the framework of quantum resource theories. For that, we refer the reader to Ref. [84], an excellent review paper, and Ref. [83], a very recent book about quantum resource theories. However, in the next Section, we will introduce the framework of quantum resource theories, together with those aspects of the framework that are relevant to this thesis. Later on, in Section 2.2 and Section 2.3, we will present entanglement and athermality as examples of resource theories. The results presented in these two sections will be the starting points for the works about entanglement embezzlement [1] and qubits cooling and heating [2] presented in Chapter 3 and Chapter 4, respectively.

2.1 The Mathematical Framework of Quantum Resource Theories

Following the example of quantum entanglement, the first step towards defining a quantum resource theory is the identification of the allowed quantum operations. Formally, one has to identify for every ordered pair of finite dimensional systems A and B a subset $\mathfrak{F}(A \to B)$ of all the possible quantum channels $\mathfrak{Q}(A \to B)$. If a quantum channel is in $\mathfrak{F}(A \to B)$, we say that it is free. These sets of free channels have to satisfy the following constraints to give origin to a resource theory:

1. For every system A, the identity channel $\mathcal{I}^A$ is free. This property is equivalent to requiring that not acting on a system is always allowed.
2. For every triple (A, B, C) of quantum systems, if $\mathcal{M}^{A \to B} \in \mathfrak{F}(A \to B)$ and $\mathcal{N}^{B \to C} \in \mathfrak{F}(B \to C)$, then $\mathcal{N}^{B \to C} \circ \mathcal{M}^{A \to B} \in \mathfrak{F}(A \to C)$. In other words, if an agent can freely perform two operations in sequence, then the resulting operation is free.
3. For every quadruplet (A, B, C, D) of quantum systems, if $\mathcal{M}^{A \to B} \in \mathfrak{F}(A \to B)$ and $\mathcal{N}^{C \to D} \in \mathfrak{F}(C \to D)$, then $\mathcal{M}^{A \to B} \otimes \mathcal{N}^{C \to D} \in \mathfrak{F}(AC \to BD)$. In other words, if an operator performs two free operations in parallel on two different systems, then the resulting operation on the joint system is free.
4. For every pair (A, B) of quantum systems, the swap channel $\mathcal{S}^{AB \to BA}$ is free. That is, an operator can always describe joint systems in the order that they prefer.
5. For every quantum system A, the trace operator Tr_A is free. Equivalently, an operator can always forget the information about a quantum system.

These conditions, which are reasonable operational constraints, constitute the core of every quantum resource theory.

Definition 2.1.1: Quantum resource theories

A map $\mathfrak{F}$ that associates with every pair (A, B) of finite dimensional quantum systems a subset $\mathfrak{F}(A \to B) \subseteq \mathfrak{Q}(A \to B)$ is a quantum resource theory if $\mathfrak{F}$ satisfies the conditions *1-5* listed above.

Continuing with the operational approach, one defines the free states of a quantum system A as those states that can be created with free operations starting from the trivial quantum system. With an abuse of notation, we denote the set of free states of the system A with $\mathfrak{F}(A) := \mathfrak{F}(I \to A)$, where I denotes the trivial system.

Definition 2.1.2: Free states and resources

For every finite-dimensional system A, the set of free states of system A is $\mathfrak{F}(A) := \mathfrak{F}(I \to A)$. Quantum states in $\mathfrak{Q}(A) \setminus \mathfrak{F}(A)$ are resources.

Every set of free states must satisfy the following properties as a consequence of the definition of quantum resource theories.

Proposition 2.1.3: Properties of free states

Let A, B be quantum systems. Then,

1. $\mathfrak{F}(A) \otimes \mathfrak{F}(B) \subseteq \mathfrak{F}(AB)$,
2. $\mathcal{S}^{AB\to BA}(\mathfrak{F}(AB)) = \mathfrak{F}(BA)$,
3. $\mathrm{Tr}_B\, \mathfrak{F}(AB) \subseteq \mathfrak{F}(A)$.

Proof. All the conditions follow trivially from the definition of resource theories and free states. *1*) Let $\mu^A \in \mathfrak{F}(A)$, $\nu^B \in \mathfrak{F}(B)$, and let $\mathcal{M}^{I\to A} \in \mathfrak{F}(I \to A)$ and $\mathcal{N}^{I\to B} \in \mathfrak{F}(I \to B)$ be their preparation channels. Since the set of free operations is closed under parallel composition, $\mathcal{M}^{I\to A} \otimes \mathcal{N}^{I\to B} \in \mathfrak{F}(I \to AB)$, and therefore, $\mu^A \otimes \nu^B \in \mathfrak{F}(AB)$. *2*) Let $\mu^{AB} \in \mathfrak{F}(AB)$ and let $\mathcal{M}^{I\to AB}$ be the free channel associated with it. Since the swap channel is free and the sequential composition of free channels is free, we have that $\mathcal{S}^{AB\to BA} \circ \mathcal{M}^{I\to AB} \in \mathfrak{F}(BA)$. Therefore, $\mathcal{S}^{AB\to BA}(\mu^{BA}) \in \mathfrak{F}(BA)$. This proves that $\mathcal{S}^{AB\to BA}(\mathfrak{F}(AB)) \subseteq \mathfrak{F}(BA)$. The proof of the opposite inclusion is the analogous. *3*) Using the same techniques detailed in the previous points, condition *3* follows from the fact that the trace and the identity are free channels and parallel composition of free channels is free. □

We note that different sources provide different lists of minimal properties for a quantum resource theory. For example, in Ref. [83], the definition of a quantum resource theory includes only conditions *1*, *2*, and *5*, reserving conditions *3* and *4* to a special class of resource theories that they refer to as resource theory with tensor product structure. We decided to include conditions *3* and *4* in the definition of resource theory because every time more than one system is considered, it is reasonable to assume that the agent can act on one system while doing nothing on the other, therefore operations such as $\mathcal{M}^{A\to B} \otimes \mathcal{I}^C$, where $\mathcal{I}^C$ is the identity channel on C, should be free if $\mathcal{M}^{A\to B}$ is free. Similarly, $\mathcal{I}^B \otimes \mathcal{N}^{C\to D}$ should be free, and therefore, by sequential composition, $\mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$ is free as well. Another example of a different minimal list of properties of a resource theory is Ref. [85]. In this case, the authors include only properties *1-4*. However, we decided to include condition *5* as well because agents can always freely discard information.

There is a sixth property that most, but not all, resource theories share: Convexity. From a physical perspective, if an agent can perform $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{A\to B}$, then they could toss a coin or use any other random source to perform the channel $\mathcal{M}^{A\to B}$ with probability $0 \le p \le 1$, and the channel $\mathcal{N}^{A\to B}$ with probability $1-p$. This is equivalent to performing the channel $p\mathcal{M}^{A\to B} + (1-p)\mathcal{N}^{A\to B}$. In other words, in convex resource theories, if $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{A\to B}$ are free, so is the channel $p\mathcal{M}^{A\to B} + (1-p)\mathcal{N}^{A\to B}$.

Definition 2.1.4: Convex quantum resource theories

A quantum resource theory $\mathfrak{F}$ is convex if for all pairs (A, B) of finite dimensional quantum systems, the set $\mathfrak{F}(A \to B)$ is convex.

As we mentioned, not all quantum resource theories are convex. One of which is the resource theory of local operations, which is similar to the resource theory of entanglement, with the exception that the two parties cannot even communicate classically. In this case, the two parties do not share a common random source, and therefore, they cannot freely prepare convex combinations of free channels.

The reader may have realized that in all the definitions above, we have only considered finite-dimensional quantum systems. From this point on, every quantum system that we introduce will be finite-dimensional, even if we do not explicitly state that. Infinite dimensional resource theories exist (see, e.g., Refs. [86–90]), but are beyond the scope of this thesis.

2.1.1 Resource Manipulation

Once we understand free operations, we turn our attention to what we can achieve with them. It follows from the definition of quantum resource theories that if $\rho^A \in \mathfrak{F}(A)$, and $\mathcal{M}^{A\to B} \in \mathfrak{F}(A \to B)$, then $\mathcal{M}^{A\to B}(\rho^A) \in \mathfrak{F}(B)$. Indeed, let $\mathcal{R}^{I\to A}$ be the free channel that prepares ρ^A, then $\mathcal{M}^{A\to B}(\rho^A) = \mathcal{M}^{A\to B} \circ \mathcal{R}^{I\to A}(1)$, where the resulting preparation channel on the right is free because it is the sequential composition of free channels. In other words, if an agent acts on free states with free channels, one can only obtain free states. This is sometimes referred to as the *golden rule* of quantum resource theories.

The golden rule, as it is stated, applies to quantum instruments in $\mathfrak{F}(A \to BX)$, where X is a classical state. If one starts with a free state ρ^A and applies a free instrument $\sum_x \mathcal{E}_x^{A\to B} \otimes |x\rangle\langle x|^X$, the resulting state $\sum_x \mathcal{E}_x^{A\to B}(\rho^A) \otimes |x\rangle\langle x|$ is free. However, if one measures the classical system, one obtains the state $\sigma_x^B = \frac{\mathcal{E}_x^{A\to B}(\rho^A)}{\mathrm{Tr}_B[\mathcal{E}_x^{A\to B}(\rho^A)]}$ with probability $p_x = \mathrm{Tr}_B[\mathcal{E}_x^{A\to B}(\rho^A)]$. If σ_x^B were not free, one would be able to generate a resource from free objects with non-zero probability. For this reason, every time that we have free instruments, we generalize the golden rule to the axiom of free instruments [83].

Definition 2.1.5: Axiom of free instruments

Let A, B be quantum systems and let X be a classical system. For every instrument $\sum_x \mathcal{E}_x^{A\to B}(\rho^A) \otimes |x\rangle\langle x|$ in $\mathfrak{F}(A \to BX)$ and every free state $\rho^A \in \mathfrak{F}(A)$, the states $\sigma_x^B = \frac{\mathcal{E}_x^{A\to B}(\rho^A)}{\mathrm{Tr}_B[\mathcal{E}_x^{A\to B}(\rho^A)]}$ are free.

Thanks to the golden rule, one can define resource theories using a bottom-up approach, that is, starting with free states and then defining free operations as all those operations that preserve the set of free states.

Definition 2.1.6: Completely resource non-generating operations

Let A, B be finite-dimensional quantum systems. The set of completely resource-non-generating (CRNG) operations from A to B is

$$\begin{aligned}\mathrm{CRNG}(A \to B) = \{\mathcal{M}^{A\to B} \in \mathfrak{Q}(A \to B) \mid \\ (\mathcal{M}^{A\to B} \otimes \mathcal{I}^C)(\rho^{AC}) \in \mathfrak{F}(BC), \\ \forall \rho^{AC} \in \mathfrak{F}(AC), \forall C \text{ s.t. } |C| < \infty\}.\end{aligned} \tag{2.1}$$

where $\mathfrak{F}(\cdot)$ denotes the set of free states, and it satisfies the properties in Proposition 2.1.3.

Note that we require completely resource-non-generating operations to preserve free states even when tensored with the identity. This is a requirement because of our definition of resource theory, specifically because the identity channel and the parallel composition of free maps are free.

In case an agent wants to obtain a state σ^B that is not free and can only perform free operations, the agent must start with another resource ρ^A. However, not all resources ρ^A can be freely converted into a state σ^B. The key question in every resource theory is whether, given two resources ρ^A, and σ^B, there exists a free channel $\mathcal{M}^{A\to B}$ such that $\mathcal{M}^{A\to B}(\rho^A) = \sigma^B$. If such a channel exists, we say that ρ^A can be freely converted into σ^B, and we write $\rho^A \xrightarrow{\mathfrak{F}} \sigma^B$.

We observe that if $\rho^A \xrightarrow{\mathfrak{F}} \sigma^B$, then ρ^A is at least as resourceful as σ^B in any task. Indeed, suppose one can manipulate σ^B with free operations to obtain τ^C, then one can convert ρ^A into τ^C with free operations as well, since the sequential composition of free channels is free. In this sense, ρ^A is at least as resourceful as σ^B. To characterize the relative resourcefulness of two quantum states, we introduce the relation $\succeq_{\mathfrak{F}}$, defined as $\rho^A \succeq_{\mathfrak{F}} \sigma^B$ if $\rho^A \xrightarrow{\mathfrak{F}} \sigma^B$. This relation is reflexive and transitive because the set of free operations contains the identity channel and is closed under sequential composition. Therefore, '$\succeq_{\mathfrak{F}}$' is a preorder. Note that, '$\succeq_{\mathfrak{F}}$' is not a partial order because $\rho^A \succeq_{\mathfrak{F}} \sigma^B$ and $\sigma^B \succeq_{\mathfrak{F}} \rho^A$, denoted as $\rho^A \sim_{\mathfrak{F}} \sigma^B$, do not imply $\rho^A = \sigma^B$. As a counterexample, choose $\rho^A = 1$ and σ^B any free state of a system with dimension greater than one. Since they are free, one can freely convert one into the other; however, they are not equal.

2.1.2 Resource Monotones

The next task in a quantum resource theory is to quantify a resource. Given that we are dealing with a preorder, a convenient resource quantifier is a function $M : \bigcup_A \mathfrak{D}(A) \to \mathbb{R}$ that is compatible with the preorder.

Definition 2.1.7: Resource measure

Let $\mathfrak{F}$ be a quantum resource theory. A resource measure M is a function that maps every quantum state into $\mathbb{R}$ and satisfies

1. $\rho^A \succeq_{\mathfrak{F}} \sigma^B \Rightarrow M(\rho^A) \geq M(\sigma^B)$,
2. $M(1) = 0$,

where $1 \in \mathbb{C}$ is the only state of the trivial quantum system.

Note that this definition implies that $M(\rho^A) = 0$ for any free state ρ^A, and $M(\sigma^A) \geq 0$ for any quantum state σ^A. Indeed, if $\rho^A \in \mathfrak{F}(A)$, then $1 \sim_{\mathfrak{F}} \rho^A$, and therefore $M(\rho^A) = M(1) = 0$. Similarly, $\sigma^A \succ_{\mathfrak{F}} (1)$, and therefore $M(\sigma^A) \geq M(1) = 0$. We have decided to define $M(1) = 0$ rather than any other real number to align with the intuition that the value of free states is zero.

In the upcoming chapters, we will introduce various resource measures tailored to the needs of the resource theory under consideration. However, there is a general way to construct a resource measure based on any quantum divergence $\mathbb{D}(\cdot\|\cdot)$, that is a function from a pair of quantum states of the same system to the real numbers that satisfies the data processing inequality

$$\mathbb{D}(\mathcal{M}^{A\to B}(\rho^A)\|\mathcal{M}^{A\to B}(\sigma^A)) \leq \mathbb{D}(\rho^A\|\sigma^A), \tag{2.2}$$

and the normalization $\mathbb{D}(1\|1) = 0$. The $\mathbb{D}$-divergence of a resource, defined as

$$\mathbb{D}(\rho^A) = \inf_{\sigma^A \in \mathfrak{F}(A)} \mathbb{D}(\rho^A \| \sigma^A), \tag{2.3}$$

is a resource measure [83]. An example of resource measure is the measure based on the Umegaki relative entropy and constructed as in Eq. (2.3):

$$D(\rho^A) = \inf_{\sigma^A \in \mathfrak{F}(A)} D(\rho^A \| \sigma^A), \tag{2.4}$$

and $D(\rho^A \| \sigma^A) = \mathrm{Tr}_A[\rho^A \log \rho^A] - \mathrm{Tr}_A[\rho^A \log \sigma^A]$.

The definition of resource measure is not very strict. For example, the function $M(\rho^A) = 0$ for all $\rho^A \in \mathfrak{Q}(A)$ is a resource measure. However, it is not very helpful in characterizing the preorder '$\succ_{\mathfrak{F}}$'. It is often useful that a measure is at least able to distinguish between free states and resources. That is, $M(\rho^A) = 0 \Leftrightarrow \rho^A \in \mathfrak{F}(A)$. This property is called *faithfulness*.

Another useful property for a resource measure is *convexity*:

$$M\left(\sum_x p_x \rho_x^A\right) \leq \sum_x p_x M\left(\rho_x^A\right). \tag{2.5}$$

The physical intuition behind this is as follows. Assume one has an ensemble of states $\left\{ p_x, \rho_x^A \right\}_x$. The average of the resource measure on the ensemble is $\sum_x p_x M\left(\rho_x^A\right)$. The ensemble is described by the classical-quantum state $\sum_x p_x \rho_x^A \otimes |x\rangle\langle x|^X$. If one discards the classical flag, one obtains the state $\sum_x p_x \rho_x^A$. Convexity states that the resource measure of the mixed state obtained by forgetting the classical flag of the ensemble is less than or equal to the average resource measure of the ensemble. Note that convexity is a useful mathematical property even in those resource theories in which the physical intuition is not valid.

If the resource theory admits a free instrument $\mathcal{M}^{A \to BX}$, one could start from a state ρ^A and produce an ensemble of states $\{ p_x, \sigma^x \}_x$, defined as in Definition 2.1.5. We say that a resource measure is *strongly monotone* if $M(\rho^A) \geq \sum_x p_x M(\sigma_x^B)$.

Since convexity and strong monotonicity (when a resource theory allows for free instruments) are very desirable properties, we name convex and strongly monotone resource measures as resource monotones.

Definition 2.1.8: Resource monotones

A resource monotone is a convex and strongly monotone (if the resource theory contains free instruments) resource measure.

For example, the resource monotone based on the Umegaki relative entropy introduced above is also a resource monotone.

Whenever the resource monotone satisfies $M(\rho^A \otimes \sigma^B) \leq M(\rho^A) + M(\sigma^B)$ we say that it is subadditive. It is additive if the equality holds for all states ρ^A, σ^B of any pair of systems A, B.

2.1.3 Approximate Conversion

So far, we have only considered exact conversions $\rho^A \xrightarrow{\mathfrak{F}} \sigma^B$. However, every physical setup is affected by noise, and a finite number of measurements will only lead to finite accuracy. It is, therefore, practically impossible to distinguish the desired target state from a close approximation. This is the physical motivation to investigate whether a state can be *approximately* converted to the target using free operations [91]. Once we choose a metric D on $\mathfrak{L}(B)$ (among the many metrics that formalize the notion of closeness between quantum states [92–96]), and fix an error $\epsilon > 0$, we say that ρ^A can be approximately converted into σ^B up to an error ϵ if there exists $\mathcal{M}^{A\to B} \in \mathfrak{F}(A \to B)$ such that $D(\mathcal{M}^{A\to B}(\rho^A), \sigma^B) < \epsilon$. If that is the case, we denote it with $\rho^A \xrightarrow{\mathfrak{F}}_{\epsilon} \sigma^B$. Here, we use the letter D to denote a metric and not the Umegaki relative entropy introduced below Eq. (2.3). The letter D is commonly used in literature to denote both these quantities. Still, one can easily tell the difference because $D(\cdot\|\cdot)$ is the Umegaki relative entropy while $D(\cdot,\cdot)$ is a metric.

When dealing with approximate conversion, a natural question is how close to σ^B it is possible to convert ρ^A using only free operations. To answer that, one has to solve the following optimization problem:

$$\inf_{\mathcal{M}^{A\to B}\in\mathfrak{F}(A\to B)} D(\mathcal{M}^{A\to B}(\rho^A), \sigma^B), \tag{2.6}$$

where D is the fixed metric. We call such a quantity conversion distance [69, 70, 97, 98], and we denote it with $D(\rho^A \xrightarrow{\mathfrak{F}} \sigma^B)$.

2.2 The Resource Theory of Entanglement

In this Section, we present the resource theory of entanglement, which has a central role both in the history of resource theories and in the works presented in this Thesis. We have already provided the intuition behind the resource theory of (bipartite) entanglement at the beginning of this Chapter. There are two distant parties, Alice and Bob, who can only communicate classically. Every LOCC protocol consists of one of multiple rounds in which a party performs an operation in their lab and sends classical information to the other party, who in turn performs an operation conditioned on the classical information received. The number of rounds can be arbitrarily large. The assumptions that Alice and Bob have perfect control over the systems in their laboratories and can perform an arbitrarily large number of communication rounds are operationally impractical. Still, they are useful in finding bounds to what Alice and Bob can achieve when quantum communication is forbidden. With LOCC, Alice and Bob can prepare all and only separable states, i.e., states that can be written as convex combination of elementary tensors of states: $\rho^{AB} = \sum_x p_x \rho_x^A \otimes \rho_x^B$, where each ρ_x^A and ρ_x^B is a quantum state, $p_x \geq 0$, and $\sum_x p_x = 1$. States that are not separable are called entangled.

Before we present the results of the resource theory of entanglement that are relevant to this Thesis, we introduce notations and assumptions that we will use throughout this and the following Chapters. First, we denote Alice's systems with the letter A and Bob's systems with the letter B. The spatial separation is always between Alice's and Bob's systems. Second, the systems in the resource theory of entanglement are always bipartite, e.g., AB. Specifically, when in the conditions for a resource theory, we read that the identity channel is free, it means, in the case of entanglement, that for every bipartite system AB, the identity channel $\mathcal{I}^{AB}$ is free. There is no free quantum identity channel from Alice to Bob that would result in quantum communication, which is not allowed. Similarly, when we say that the swap channel is free, we do not mean that Alice and Bob can freely swap systems between them, but the bipartite system A_1B_1 can be freely swapped with A_2B_2. Third, we assume that all systems have the same dimension unless we specify otherwise. Indeed, both Alice and Bob can embed their local system into larger systems with local isometries. For any bipartite system AB, let $A'B'$ be such that $|A'| = |B'| \geq \{ |A|, |B| \}$, and let $\mathcal{U}^{A\to A'}, \mathcal{V}^{B\to B'}$ be isometries. Then, $\mathcal{U}^{A\to A'} \otimes \mathcal{V}^{B\to B'} \in \mathfrak{F}(AB \to A'B')$, and for any ρ^{AB}, we have that $\rho^{AB} \succ_{\mathfrak{F}} \rho^{A'B'} = \mathcal{U}^{A\to A'} \otimes \mathcal{V}^{B\to B'}(\rho^{AB})$. However, since the adjoint of an isometry is its left inverse and since $(\mathcal{U}^{A\to A'})^\dagger \otimes (\mathcal{V}^{B\to B'})^\dagger \in \mathfrak{F}(A'B' \to AB)$, we also have that $\rho^{A'B'} \succ_{\mathfrak{F}} \rho^{AB}$. Therefore, $\rho^{AB} \sim_{\mathfrak{F}} \rho^{A'B'}$, which means that they are equivalent as resources. This argument can be generalized to cases involving multiple systems. Therefore, without loss of generality, we can assume that all the systems have the same dimension.

2.2.1 Pure State Conversion

The definition of LOCC, which allows for an arbitrarily large number of rounds of communication and local operations, makes checking whether $\rho^{AB} \xrightarrow{\text{LOCC}} \sigma^{AB}$, for arbitrary $\rho^{AB}, \sigma^{AB} \in \mathfrak{D}(AB)$, a daunting challenge. However, this task is much simpler if one starts with a pure state.

Theorem 2.2.1: Lo-Popescu's theorem [25]

Any deterministic LOCC pure-state manipulation can be simulated with the following protocol: 1) Alice performs a generalized measurement, 2) Alice sends the measurement outcome to Bob, and 3) Bob performs a unitary operation conditioned on the measurement outcome. That is, if $\psi^{AB} \xrightarrow{\text{LOCC}} \sigma^{AB}$, then there exists a generalized measurement $\{ M_A^x \}_x$ (where each M_A^x is a complex matrix on A and $\sum_x (M_A^x)^\dagger M_A^x = \mathbb{1}^A$) and a family of unitary matrices $\{ U_B^x \}_x$ such that

$$\sigma^{AB} = \sum_x (M_A^x \otimes U_B^x) \psi^{AB} (M_A^x \otimes U_B^x)^\dagger. \tag{2.7}$$

Lo-Popescu's Theorem is remarkable because it states that any LOCC deterministic protocol that starts with a pure state, regardless of the round of communication, can be exactly simulated with a protocol involving at most one round of communication.

When dealing with pure states, one can use another important result from linear algebra: the Schmidt decomposition of a bipartite vector. Indeed, for every bipartite vector $|\psi^{AB}\rangle$ there exist orthonormal bases $\{ |x_A\rangle \}_x$ and $\{ |x_B\rangle \}_x$ for A and B, and real non-negative coefficients $\{ \alpha_x \}_x$ such that

$$|\psi^{AB}\rangle = \sum_x \alpha_x |xx\rangle^{AB} . \tag{2.8}$$

Moreover, by assuming that $|\psi^{AB}\rangle$ is a vector of norm one, we obtain that $\sum_x \alpha_x^2 = 1$. Therefore, the elements of $\{ \alpha_x^2 \}_x$ can be seen as the entries of a probability vector in $\text{Prob}(|A|)$. We say that $\mathbf{p} \in \text{Prob}^{\downarrow}(|A|)$ is the vector containing the Schmidt coefficients of ψ^{AB} when the entries of $\mathbf{p}$ are the elements of $\{ \alpha_x^2 \}_x$, organized in non-increasing order. The number of non-zero entries of $\mathbf{p}$ is known as the Schmidt rank of ψ^{AB}, and we denote it with $\text{SR}(\psi^{AB})$. Since local unitary are free operations, one can fix orthonormal bases $\{ |\tilde{x}_A\rangle \}_x$ and $\{ |\tilde{x}_B\rangle \}_x$, and every state $|\psi^{AB}\rangle = \sum_x \alpha_x |xx\rangle^{AB}$ is equivalent to a state $|\tilde{\psi}^{AB}\rangle = \sum_x \sqrt{p_x} |\tilde{x}x\rangle^{AB}$, where p_x are the entries of $\mathbf{p}$, the vector containing the Schmidt coefficients of ψ^{AB}. We say that the vector $|\tilde{\psi}^{AB}\rangle$ is in *standard form*, and thanks to the LOCC-equivalence, we can from now on assume, w.l.o.g., that all pure states are in standard form.

With these notations at hand, we are ready to present the most important result about pure-to-pure state conversion in the resource theory of entanglement.

Theorem 2.2.2: Nielsen's theorem [26]

Let $\psi^{AB}, \varphi^{AB} \in \mathrm{PURE}(AB)$, and let $\mathbf{p}, \mathbf{q} \in \mathrm{Prob}^{\downarrow}(|A|)$ be their corresponding Schmidt probability vectors. Then $\psi^{AB} \xrightarrow{\text{LOCC}} \varphi^{AB}$ if and only if $\sum_{x=1}^{k} q_x \geq \sum_{x=1}^{k} p_x$ for all $k \in [|A|]$.

Nielsen's Theorem completely characterizes pure state conversion. We notice that $\sum_{x=1}^{k} p_x$ is the k-th Ky Fan norm of $\mathbf{p}$, denoted as $\|\mathbf{p}\|_{(k)}$, then we can define a family of pure-state entanglement monotones as $E_k(\psi^{AB}) = 1 - \|\mathbf{p}\|_{(k)}$. Indeed, Nielsen's Theorem can be equivalently formulated as $\psi^{AB} \xrightarrow{\text{LOCC}} \varphi^{AB}$ if and only if $E_k(\psi^{AB}) \geq E_k(\varphi^{AB})$. Moreover, we observe that the condition $\sum_{x=1}^{k} q_x \geq \sum_{x=1}^{k} p_x$ for all $k \in [|A|]$ is nothing else than the definition of vector majorization [99], which we denote as $\mathbf{q} \succ \mathbf{p}$. Therefore, another way of stating Nielsen theorem is $\psi^{AB} \xrightarrow{\text{LOCC}} \varphi^{AB}$ if and only if $\mathbf{q} \succ \mathbf{p}$.

We conclude this overview of pure-state entanglement conversion by mentioning that Nielsen's Theorem was generalized in Ref. [100] to the case in which the target state is an ensemble of pure states:

$$\psi^{AB} \xrightarrow{\text{LOCC}} \left\{ t_z, \varphi_z^{AB} \right\}_z \Leftrightarrow E_k(\psi^{AB}) \geq \sum_z t_z E_k(\varphi_z^{AB}), \forall k \in [|A|]. \tag{2.9}$$

2.2.2 SEP and NPT Entanglement

As mentioned before, LOCC maps are difficult to characterize. Historically, there have been two generalizations of the resource theory of LOCC entanglement with simpler sets of free operations. The first is obtained by considering the complete set of resource-non-generating operations that are compatible with the free states of the resource theory of entanglement, i.e., with separable states. The CRNG operations coincide with the separable channels, that is, those channels that admit an operator sum representation in which all Kraus operators are tensor products. That is, $\mathcal{N}^{AB\to AB}$ is separable if and only if there exists a set of Kraus operators $\left\{ M_A^x \otimes N_B^x \right\}_x$ such that $\mathcal{N}^{AB\to AB}(\rho^{AB}) = \sum_x (M_A^x \otimes N_B^x)\rho^{AB}(M_A^x \otimes N_B^x)^\dagger$. In Chapter 5, we will present an easier characterization based on the Choi isomorphism. The resource theory with separable operations as free operations is often called the resource theory of separable entanglement (SEP) [54–56].

The second generalization is a consequence of the Peres-Horodecki criterion for separable states [57, 58]: Every separable state has positive partial transpose. Therefore, one can generalize entanglement theory and consider as free states all those states that have positive partial transpose (PPT). The completely resource-non-generating operations, in this case, are all those operations that preserve PPT in a complete sense, i.e., even when tensored with the identity [59–63]. The resource theory with these operations as free operations is known as the resource theory of non-positive partial transpose entanglement (NPT).

2.3 The Resource Theory of Quantum Thermodynamics

When classical thermodynamics was first developed, the goal was to optimize work production through cyclic processes. Carnot's theorem tells us that this is only possible if one has access to two baths at different temperatures, one of which is typically the environment. The greater the temperature difference, the higher the efficiency of the engine. In resource theoretic terms, a resource is a bath not at thermal equilibrium with the environment. We will see in this Section that the intuition behind the resource theory of thermodynamics or athermality is the same.

2.3.1 Free States in the Resource Theory of Quantum Thermodynamics

In defining a resource theory, we usually start with free operations. However, in this case, it is more enlightening first to define which states are free and then construct free operations. We assume that the existence of a large heath bath (or environment) at a fixed temperature T, or equivalently at an inverse temperature $\beta = (k_B T)^{-1}$, where k_B is Boltzmann's constant, the only free state of a system S with Hamiltonian H^S is the state at thermal equilibrium with the bath:

$$\gamma^S = \frac{e^{-\beta H^S}}{\mathrm{Tr}_S[e^{-\beta H^S}]}. \tag{2.10}$$

This state is the quantum canonical ensemble and is known as the Gibbs state. In the remainder of this Subsection, we will present different motivations that lead to singling out the Gibbs state as the only free state.

Principle of Maximum Entropy

Suppose one is given a system S with Hamiltonian H^S, and the only information that is known is that the energy of the system is E. What is the best description of the state? Jaynes [101] suggests answering this question with principles of information theory. In particular, one should apply the principle of maximum entropy and choose among all the states with energy E the one that maximizes the entropy. That is, we are looking for a state γ^S, such that

- $\mathrm{Tr}_S[\gamma^S H^S] = E$,
- $-\mathrm{Tr}_S[\gamma^S \ln \gamma^S] \geq -\mathrm{Tr}_S[\rho^S \ln \rho^S]$ for all $\rho^S \in \mathfrak{Q}(S)$.

Using the Lagrange multiplier method, one finds that the only quantum states that satisfy these conditions are:

$$\frac{e^{-\lambda H^S}}{\mathrm{Tr}_S[e^{-\lambda H^S}]}, \tag{2.11}$$

where λ is the real parameter. We will discuss later the meaning of λ.

Principle of Minimum Energy

With a symmetric argument, one may fix the entropy of a system and look for the state that minimizes the energy. That is, we want to find a state γ^S such that

- $-\operatorname{Tr}_S[\gamma^S \ln \gamma^S] = S$,
- $\operatorname{Tr}_S[\gamma^S H^S] \leq \operatorname{Tr}_S[\rho^S H^S]$ for all $\rho^S \in \mathfrak{Q}(S)$.

Once again, one finds the family of states

$$\frac{e^{-\lambda H^S}}{\operatorname{Tr}_S[e^{-\lambda H^S}]}. \tag{2.12}$$

Passivity, Structural Stability, and Consistency

Lenard [102] derives the formula for the Gibbs state from three assumptions: passivity, structural stability, and consistency.

First, a state ρ^S of a system S with Hamiltonian H^S is passive if no work can be extracted from it. This is equivalent to requiring that $\operatorname{Tr}_S[H^S \rho^S] \leq \operatorname{Tr}_S[H^S U \rho^S U^\dagger]$ for any unitary operator U. Indeed, if some work can be extracted from ρ^S it means that there exists a unitary process U (the evolution of a quantum system in quantum mechanics is unitary) such that the energy of the system S at the end of the process is less than the energy at the beginning: $\operatorname{Tr}_S[H^S U \rho^S U^\dagger] < \operatorname{Tr}_S[H^S \rho^S]$. By requesting that the mean energy of S at the end of the process be always greater than or equal to the energy at the beginning, we guarantee that no work is extracted from the system.

Second, a passive state ρ^S is structurally stable if a small change in the Hamiltonian H^S results in a small change of the passive state. That is, for every neighbourhood $\mathfrak{U}$ of ρ^S there exists a neighbourhood $\mathfrak{V}$ of H^S such that for every $\tilde{H} \in \mathfrak{V}$ there exists at least a state in $\mathfrak{U}$ that is passive for $\tilde{H}$.

Lastly, a consistent family $\mathfrak{F}$ of structurally stable passive states is a function that associate with every system S with Hamiltonian H^S a passive state ρ^S and such that if ρ^S and σ^R are passive state with Hamiltonian H^S and H^R, then $\rho^S \otimes \sigma^R$ is the passive state associated with the system SR with Hamiltonian

$$H^{SR} = H^S \otimes \mathbb{1}^R + \mathbb{1}^S \otimes H^R. \tag{2.13}$$

A thermal equilibrium state ρ^S for the system S with Hamiltonian H^S is a state for which a consistent family of structurally stable passive states exists and such that ρ^S is the state associated with the pair (S, H^S).

Lenard [102] shows that thermal equilibrium states associated with a system S with Hamiltonian H^S can only be of three kinds:

- multiple of the projector associated with the smallest eigenvalues of H^S,
- $e^{-\lambda H^S}/\operatorname{Tr}_S[e^{-\lambda H^S}]$ for some $\lambda \in \mathbb{R}$,
- $\frac{1}{|S|}\mathbb{1}^S$.

Note that the first and third cases are limit cases obtained from the second by taking the limits for $\lambda \longrightarrow \infty$ or $\lambda \longrightarrow 0$, respectively.

The Meaning of λ

The last detail is to understand the meaning of λ. Here, we show that it is equal to the inverse temperature of the system. If we consider $e^{-\lambda H^S}/\mathrm{Tr}_S[e^{-\lambda H^S}]$ as a statistical ensemble, we can compute the temperature associated with it as $\frac{1}{k_B T_S} = \frac{\partial S(\gamma^S)}{\partial E}$.

E is the average energy of the systems

$$E = \left\langle H^S \right\rangle_{\gamma^S} = \frac{\mathrm{Tr}_S[H^S \gamma^S]}{\mathrm{Tr}_S[e^{-\lambda H^S}]} = \frac{\sum_i E_i e^{-\lambda E_i}}{\sum_j e^{-\lambda E_j}}, \tag{2.14}$$

and S is its entropy:

$$\begin{aligned} S(\gamma^S) &= -\mathrm{Tr}_S[\gamma^S \ln \gamma^S] \\ &= \frac{\sum_i e^{-\lambda E_i}(\lambda E_i)}{\sum_j e^{-\lambda E_j}} + \ln \mathrm{Tr}_S[e^{-\lambda H^S}] \\ &= \lambda E + \ln \mathrm{Tr}_S[e^{-\lambda H^S}]. \end{aligned} \tag{2.15}$$

Therefore,

$$\begin{aligned} \frac{\partial S(\gamma^S)}{\partial E} &= \lambda + \frac{\partial \lambda}{\partial E} E + \frac{\partial \ln \mathrm{Tr}_S[e^{-\lambda H^S}]}{\partial \lambda} \frac{\partial \lambda}{\partial E} \\ &= \lambda + \frac{\partial \lambda}{\partial E} E - \frac{\mathrm{Tr}_S[H^S e^{-\lambda H^S}]}{\mathrm{Tr}_S[e^{-\lambda H^S}]} \frac{\partial \lambda}{\partial E} \\ &= \lambda + \frac{\partial \lambda}{\partial E} E - \left\langle H^S \right\rangle_{\gamma^S} \frac{\partial \lambda}{\partial E} \\ &= \lambda + \frac{\partial \lambda}{\partial E} E - E \frac{\partial \lambda}{\partial E} \\ &= \lambda. \end{aligned} \tag{2.16}$$

This shows that $\lambda = \frac{1}{k_B T_S}$ is the inverse temperature of the ensemble. For γ^S to be a free state, this inverse temperature must be equal to β, the inverse temperature of the background environment. If not, the global state associated with S and the environment would no longer be passive.

Complete Passivity

All the arguments exposed above indicate that Gibbs states should be free in the resource theory of quantum thermodynamics. However, one may wonder if we should include other free states. Once again, Lenard [102] comes to our help. He argues that passivity is not enough. One should require complete passivity, i.e., a state is completely passive if no work can be extracted from arbitrarily many copies of it. He showed that completely passive states are either thermal states or *ground* states, defined as ρ^S is ground if $\mathrm{Tr}_S[H^S \rho^0] \leq \mathrm{Tr}_S[H^S \rho^S]$ for all $\rho^S \in \mathfrak{Q}(S)$.

If we added states that are not completely passive to the set of free states, our resource theory would be trivial. Indeed, we could create as many of them as we want for free and extract an arbitrarily large amount of work from them, which we could then reuse. Therefore, if we fix the background temperature β, for every system S with Hamiltonian H^S there is only one free state, the Gibbs state

$$\gamma^S = \frac{e^{-\beta H^S}}{\mathrm{Tr}[e^{-\beta H^S}]}. \tag{2.17}$$

The Minus First Law

Lastly, the Gibbs state is the state to which any quantum system evolves under weak coupling with a bath [103]. This mirrors the well-known fact in classical thermodynamics that if two or more systems, isolated from the rest of the world, are allowed to exchange heat, they will eventually reach a state of equilibrium. The fact that quantum systems move towards their equilibrium states has been called the minus first law of thermodynamics [104, 105].

2.3.2 Free Operations

We identified the free states in the previous Subsection. Therefore, for each free state γ^B, the preparation channel $1 \longrightarrow \gamma^B$ is free. Since the parallel composition of free operations is free, for every state ρ^S the channel $\rho^S \longrightarrow \rho^S \otimes \gamma^B$ is free as well for every free state γ^B.

Now, if we want to act on the composite system $R = SB$ relying only on the resourcefulness of ρ^S, we cannot introduce new energy into the system. That is, we can only perform energy-preserving unitary transformations, i.e., $[U, H^R] = 0$, where $H^R = \mathbb{1}^S \otimes H^B + H^S \otimes \mathbb{1}^B$. While the request $[U, H^R] = 0$ seems very restrictive, one should notice that the choice of the system B is arbitrary, and one can choose a system B in a way that H^R has many degenerate energy eigenvalues. In this way, U can act non-trivially on the eigenspaces associated with those eigenvalues. If one applies an energy-preserving unitary transformation to $\rho^S \otimes \gamma^B$ one obtains $\rho^S \longrightarrow U(\rho^S \otimes \gamma^B)U^\dagger$. Lastly, we are free to discard unnecessary systems. For every S' and B' such that $R = S'B'$ and $H^R = \mathbb{1}^{S'} \otimes H^{B'} + H^{S'} \otimes \mathbb{1}^{B'}$ (where S' and B' may or may not be equal to S and B), the operation

$$\mathcal{M}^{S \to S'} = \mathrm{Tr}_{B'}[U\rho^S \otimes \gamma^B U^\dagger] \tag{2.18}$$

is free. Such operations, called thermal operations, are the free operations in the resource theory of thermodynamics. As one notices in Eq. (2.18), the final system S' may be different from the initial system and could have different energy. Only the unitary transformation is energy-preserving, the overall thermal operation may not be. Indeed, one could start with a qubit S in the excited state, tensor it with a qubit B at equilibrium with the environment, pick as U the identity operator, which is energy-preserving, and finally trace over S. At the end of the process one is left with a qubit at equilibrium with the environment which has less energy than a qubit in the excited state.

Observe that $[U, H^R] = 0$ implies $[U, \gamma^S \otimes \gamma^B] = 0$, therefore thermal operations are Gibbs-preserving operations. This guarantees that the only states that can be prepared for free are the Gibbs states, as requested in the previous Section.

To denote that a state ρ^S can be converted to a state $\sigma^{S'}$ with thermal operations (TO), we may be tempted to write $\rho^S \xrightarrow{\text{TO}} \sigma^{S'}$, as done in entanglement theory. However, the free operations are not fixed as in the case of LOCC, but they depend on the temperature of the heat bath and the Hamiltonian of the system. The Gibbs state $\gamma^S = \frac{e^{-\beta H^S}}{\mathrm{Tr}_S[e^{-\beta H^S}]}$ encodes both of these pieces of information. The state γ^S is the only free state of the system S, once β and H^S are fixed. Any state of S that is not γ^S is a resource. Therefore, we will denote a resource in the resource theory of athermality as (ρ^S, γ^S), where γ^S is used as the reference free state of the system S, and if the conversion with thermal operation is possible, we write $(\rho^S, \gamma^S) \xrightarrow{\text{TO}} (\sigma^{S'}, \gamma^{S'})$.

Closed Thermal Operations

The set of thermal operations is not topologically closed with respect to any topology induced by a metric on the set of channels (an example of such metrics is the metric induced by the diamond norm introduced below in Eq. (2.20)). From an operational point of view, it makes sense to consider an operation as free if it belongs to the closure of the set of thermal operations. Indeed, if one can approximate a channel arbitrarily well with thermal operations, then there is no operational way to distinguish it from a thermal operation.

Definition 2.3.1: Closed thermal operations (CTO)

Let S, S' be two systems with Gibbs states γ^S, $\gamma^{S'}$. Then $\mathcal{N}^{S\to S'}$ is a closed thermal operation if there exists a sequence of thermal operations $\left\{ \mathcal{N}_k^{S\to S'} \right\}_k$ such that

$$\lim_{k\to\infty} \left\| \mathcal{N}^{S\to S'} - \mathcal{N}_k^{S\to S'} \right\|_\diamond = 0, \tag{2.19}$$

where $\|\cdot\|_\diamond$ is the diamond norm [82, 106], defined as

$$\left\| \mathcal{M}^{S\to S'} \right\|_\diamond = \max \left\{ \left\| (\mathcal{M}^{S\to S'} \otimes \mathcal{I}^S)(X) \right\|_1 \mid X \in \mathfrak{L}(S), \|X\|_1 \leq 1 \right\}, \tag{2.20}$$

for a $\mathcal{M}^{S\to S'} \in \mathfrak{L}(S \to S')$.

To ensure that the set of closed thermal operations can be used as the set of free operations in a resource theory, we must verify that they are closed under sequential and parallel composition. The results presented here about closed thermal operations are taken from our work about cooling and heating in the resource theory of athermality (Ref. [2]).

Proposition 2.3.2: Parallel and sequential composition of CTOs

If $\mathcal{M}^{A\to B}$, $\mathcal{N}^{B\to C}$ and $\mathcal{P}^{C\to D}$ are closed thermal operations then $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$ and $\mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$ are closed thermal operations.

Proof. Let $\left\{ \mathcal{M}_k^{A\to B} \right\}_k$, $\left\{ \mathcal{N}_k^{B\to C} \right\}_k$, $\left\{ \mathcal{P}_k^{C\to D} \right\}_k$ be sequences of thermal operations converging to $\mathcal{M}^{A\to B}$, $\mathcal{N}^{B\to C}$ and $\mathcal{P}^{C\to D}$ in the diamond norm. We show now that $\left\{ \mathcal{N}_k^{B\to C} \circ \mathcal{M}_k^{A\to B} \right\}_k$ and $\left\{ \mathcal{M}_k^{A\to B} \otimes \mathcal{P}_k^{C\to D} \right\}_k$ converge to $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$ and $\mathcal{M}^{A\to B} \otimes \mathcal{P}^{C\to D}$.

$$\begin{aligned} &\lim_{k\to\infty} \left\| \mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B} - \mathcal{N}_k^{B\to C} \circ \mathcal{M}_k^{A\to B} \right\|_\diamond \\ &\quad \leq \lim_{k\to\infty} \left\| \mathcal{N}^{B\to C} - \mathcal{N}_k^{B\to C} \right\|_\diamond + \lim_{k\to\infty} \left\| \mathcal{M}^{A\to B} - \mathcal{M}_k^{A\to B} \right\|_\diamond \\ &\quad = 0. \end{aligned} \tag{2.21}$$

For the properties of the diamond norm, see Ref. [82].

$$\begin{aligned}
&\lim_{k\to\infty} \left\| \mathcal{M}_k^{A\to B} \otimes \mathcal{P}_k^{C\to D} - \mathcal{M}^{A\to B} \otimes \mathcal{P}^{C\to D} \right\|_\diamond \\
&\quad = \lim_{k\to\infty} \left\| \mathcal{M}_k^{A\to B} \otimes \mathcal{P}_k^{C\to D} - \mathcal{M}^{A\to B} \otimes \mathcal{P}_k^{C\to D} + \mathcal{M}^{A\to B} \otimes \mathcal{P}_k^{C\to D} - \mathcal{M}^{A\to B} \otimes \mathcal{P}^{C\to D} \right\|_\diamond \\
&\quad \le \lim_{k\to\infty} \left\| (\mathcal{M}_k^{A\to B} - \mathcal{M}^{A\to B}) \otimes \mathcal{P}_k^{C\to D} \right\|_\diamond + \lim_{k\to\infty} \left\| \mathcal{M}^{A\to B} \otimes (\mathcal{P}_k^{C\to D} - \mathcal{P}^{C\to D}) \right\|_\diamond \\
&\quad = \lim_{k\to\infty} \left\| (\mathcal{M}_k^{A\to B} - \mathcal{M}^{A\to B}) \right\|_\diamond \left\| \mathcal{P}_k^{C\to D} \right\|_\diamond + \lim_{k\to\infty} \left\| \mathcal{M}^{A\to B} \right\|_\diamond \left\| (\mathcal{P}_k^{C\to D} - \mathcal{P}^{C\to D}) \right\|_\diamond \\
&\quad = \lim_{k\to\infty} \left\| (\mathcal{M}_k^{A\to B} - \mathcal{M}^{A\to B}) \right\|_\diamond + \lim_{k\to\infty} \left\| (\mathcal{P}_k^{C\to D} - \mathcal{P}^{C\to D}) \right\|_\diamond \\
&\quad = 0.
\end{aligned} \tag{2.22}$$

Here, we have used the fact that the diamond norm of the tensor product of two linear maps is the product of their diamond norms and the fact that the diamond norm of a channel is always 1.

Since thermal operations are closed under sequential and parallel composition, $\left\{ \mathcal{N}_k^{B\to C} \circ \mathcal{M}_k^{A\to B} \right\}_k$ and $\left\{ \mathcal{M}_k^{A\to B} \otimes \mathcal{P}_k^{C\to D} \right\}_k$ are sequences of thermal operations and therefore $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$ and $\mathcal{M}^{A\to B} \otimes \mathcal{P}^{C\to D}$ are closed thermal operations. □

As requested, CTOs are approximated arbitrarily well by TOs. In the following, we will use the trace distance, see Eq. (1.4), to evaluate distances between quantum states.

Proposition 2.3.3: CTO vs TO

The following two statements are equivalent.

1. The athermality state (ρ^A, γ^A) can be converted to $(\sigma^{A'}, \gamma^{A'})$ by CTO, i.e.,

$$(\rho^A, \gamma^A) \xrightarrow{\text{CTO}} (\sigma^{A'}, \gamma^{A'}). \tag{2.23}$$

2. For every $\epsilon > 0$, there exist athermality states $(\tilde{\rho}^A, \gamma^A)$ and $(\tilde{\sigma}^{A'}, \gamma^{A'})$ ϵ-close to (ρ^A, γ^A) and $(\sigma^{A'}, \gamma^{A'})$, respectively, such that

$$(\tilde{\rho}^A, \gamma^A) \xrightarrow{\text{TO}} (\tilde{\sigma}^{A'}, \gamma^{A'}). \tag{2.24}$$

Proof. We begin by showing that *2* follows from *1*. Let $(\mathcal{N}^{A\to A'}, \gamma^A)$ be a closed thermal operation that converts (ρ^A, γ^A) to $(\sigma^{A'}, \gamma^{A'})$ and let $\{(\mathcal{N}_n^{A\to A'}, \gamma^A)\}_{n\in\mathbb{N}}$ be a sequence of thermal operations that has $(\mathcal{N}^{A\to A'}, \gamma^A)$ as its limit.

Next, introduce

$$\sigma_n^{A'} := \mathcal{N}_n^{A\to A'}\left(\rho^A\right). \tag{2.25}$$

By definition, the thermal operation $(\mathcal{N}_n^{A\to A'}, \gamma^A)$ then converts (ρ^A, γ^A) to $(\sigma_n^{A'}, \gamma^{A'})$ and

$$\sigma^{A'} = \lim_{n\to\infty} \sigma_n^{A'}. \tag{2.26}$$

This further implies that for every $\epsilon > 0$, there exists an m such that

$$\frac{1}{2}\left\| \sigma_m^{A'} - \sigma^{A'} \right\|_1 \le \epsilon, \tag{2.27}$$

which finishes the first direction by choosing

$$(\tilde{\rho}^A, \gamma^A) = (\rho^{A'}, \gamma^A),$$
$$(\tilde{\sigma}^{A'}, \gamma^{A'}) = (\sigma_m^{A'}, \gamma^{A'}).$$

Now let $\{\epsilon_k\}_{k\in\mathbb{N}}$ be a sequence of non-negative numbers such that

$$\lim_{k\to\infty} \epsilon_k = 0. \tag{2.28}$$

To show the reverse, assume that *2* holds. This implies that for every $k \in \mathbb{N}$, there exists a $(\tilde{\rho}_k^A, \gamma^A)$ ϵ_k-close to (ρ^A, γ^A) and a thermal operation $(\mathcal{N}_k^{A\to A'}, \gamma^A)$ that converts $(\tilde{\rho}_k^A, \gamma^A)$ to a state $(\tilde{\sigma}_k^{A'}, \gamma^{A'})$ ϵ_k-close to $(\sigma^{A'}, \gamma^{A'})$. Since the set of quantum channels from A to A' is compact, there exists a converging sub-sequence $\{\mathcal{N}_{k_l}^{A\to A'}\}_l$. We can thus define

$$\mathcal{N}^{A\to A'} := \lim_{l\to\infty} \mathcal{N}_{k_l}^{A\to A'}. \tag{2.29}$$

By definition, $(\mathcal{N}^{A\to A'}, \gamma^A) \in \mathrm{CTO}(\gamma^{A'} \leftarrow \gamma^A)$. Moreover,

$$\mathcal{N}^{A\to A'}(\rho^A) = \lim_{l\to\infty} \mathcal{N}_{k_l}^{A\to A'}(\tilde{\rho}_{k_l}^A) = \lim_{l\to\infty} \tilde{\sigma}_{k_l}^{A'} = \sigma^{A'}, \tag{2.30}$$

which finishes the proof. □

Gibbs-Preserving Operations

From a set of free states, it is possible to construct a resource theory in which the free operations are the operations that preserve the free states in a complete sense (see Definition 2.1.6). In this case, $\mathcal{M}^{A\to A'}$ is a completely resource-non-generating operation if

$$(\mathcal{M}^{A\to A'} \otimes \mathcal{I}^R)(\gamma^{AR}) = \gamma^{A'R} \tag{2.31}$$

is a free state. However, this condition can be simplified by noticing that $\gamma^{AR} = \gamma^A \otimes \gamma^R$ and $\gamma^{A'R} = \gamma^{A'} \otimes \gamma^R$. Therefore, $\mathcal{M}^{A\to A'}$ is completely resource-non-generating if and only if

$$\mathcal{M}^{A\to A'}(\gamma^A) = \gamma^{A'}. \tag{2.32}$$

Such operations are referred to as Gibbs-preserving operations (GPOs). It is straightforward to prove that $\mathrm{TO} \subseteq \mathrm{CTO} \subseteq \mathrm{GPO}$.

2.3.3 Quasi-Classical States and Relative Majorization

While the conversion problem in the resource theory of athermality is so far unsolved in full generality, it has a solution for *quasi-classical* states, i.e., states that do not show coherence between eigenspaces associated with different eigenvalues of the Hamiltonian. These states are all and only the states $\rho^S \in \mathfrak{D}(S)$ such that $[\rho^S, H^S] = 0$. As a consequence $[\rho^S, \gamma^S] = 0$ as well. This implies that there exists an orthonormal basis $\{\,|i\rangle\,\}_{i=1}^{|S|}$ in which both ρ^S and γ^S are diagonal. That is,

$$\rho^S = \sum_i p_i^S |i\rangle\langle i|^S, \quad \gamma^S = \sum_i g_i^S |i\rangle\langle i|^S, \tag{2.33}$$

where $p_i^S, g_i^S \geq 0$ and $\sum_{i=1}^{|S|} p_i = \sum_{i=1}^{|S|} g_i = 1$. Therefore, $\left\{ p_i^S \right\}_{i=1}^{|S|}$ and $\left\{ g_i^S \right\}_{i=1}^{|S|}$ are the entries of probability vectors $\mathbf{p}^S, \mathbf{g}^S \in \mathrm{Prob}(|S|)$.

Once the basis is fixed, ρ^S and γ^S are completely characterized by $\mathbf{p}^S$ and $\mathbf{g}^S$. As mentioned, the conversion problem for quasi-classical states has been fully solved.

Theorem 2.3.4: Conversion problem for quasi-classical states [107]

Let (ρ^S, γ^S) and (σ^R, γ^R) be quasi classical states. The following are equivalent

1. $(\rho^S, \gamma^S) \xrightarrow{\text{CTO}} (\sigma^R, \gamma^R)$,
2. $(\rho^S, \gamma^S) \xrightarrow{\text{GPO}} (\sigma^R, \gamma^R)$,
3. There exists a column stochastic matrix S such that $\mathbf{p}^R = S\mathbf{p}^S$ and $\mathbf{g}^R = Sg^S$,

where $(\mathbf{p}^S, \mathbf{g}^S)$ and $(\mathbf{p}^R, \mathbf{g}^R)$ are the pairs of probability vectors associated with (ρ^S, γ^S) and (σ^R, γ^R).

The last condition is known as relative majorization[108–115].

Definition 2.3.5: Relative majorization

Let $\mathbf{p}_1, \mathbf{q}_1 \in \mathrm{Prob}(n)$ and let $\mathbf{p}_2, \mathbf{q}_2 \in \mathrm{Prob}(m)$. then $(\mathbf{p}_1, \mathbf{q}_1)$ relative majorizes $(\mathbf{p}_2, \mathbf{q}_2)$, denoted as $(\mathbf{p}_1, \mathbf{q}_1) \succ (\mathbf{p}_2, \mathbf{q}_2)$, if there exists a column stochastic matrix S such that $\mathbf{p}_2 = S\mathbf{p}_1$ and $\mathbf{q}_2 = S\mathbf{q}_1$.

Finding a stochastic matrix that satisfies the conditions in Definition 2.3.5 or proving that such a matrix does not exist is equivalent to solving a system of linear equations, which becomes quite difficult when m and n are large. However, there is an easier way based on Lorenz curves [116–119]. The construction of a Lorenz curve associated with a pair of probability vectors $\mathbf{p}, \mathbf{q} \in \mathrm{Prob}(n)$ is the following:

1. Find a permutation $\pi(j)$ such that $\left\{ p_{\pi(k)}/q_{\pi_j} \right\}_j$ are in non-increasing order.
2. For all $k \in [n]$, define the points $(x_k^{(\mathbf{p},\mathbf{q})}, y_k^{(\mathbf{p},\mathbf{q})})$ in the (x, y)-plane as

$$(x_k^{(\mathbf{p},\mathbf{q})}, y_k^{(\mathbf{p},\mathbf{q})}) = \left(\sum_{x=1}^{k} p_{\pi(j)}, \sum_{x=1}^{k} q_{\pi(j)} \right) \tag{2.34}$$

3. The lower Lorenz curve associate with $(\mathbf{p}, \mathbf{q})$ is the curve that connects $(0, 0)$ and the ordered points $\left\{ x_k^{(\mathbf{p},\mathbf{q})}, y_k^{(\mathbf{p},\mathbf{q})} \right\}_k$ with straight lines. This curve can be expressed both as a function of x or as a function of y. In Chapter 4, we will extensively use Lorenz curves as functions of y, and we denote them as $\alpha^{(\mathbf{p},\mathbf{q})}(y)$.

As an example, consider the pair of probability vectors $\mathbf{p} = (1/4, 1/3, 5/12)$ and $\mathbf{q} = (1/5, 2/3, 2/15)$. If we take the ratio of each component, we obtain $(5/4, 1/2, 25/8)$, which are ordered as $25/8 > 5/4 > 1/2$. The permutation that orders the entries in non-increasing order consists of a circular right shift of the triples. Applying this permutation to $\mathbf{p}$ and $\mathbf{q}$ we obtain $(5/12, 1/4, 1/3)$ and $(2/15, 1/5, 2/3)$. Computing the partial sums as in point *2)*, we get the points $\{ (5/12, 2/15), (2/3, 1/3), (1, 1) \}$. Note that the last point is always

$(1,1)$ because we are dealing with probability vectors. These are the 'elbows' of the Lorenz curve associated with $(\mathbf{p},\mathbf{q})$, as depicted in Figure 2.1.

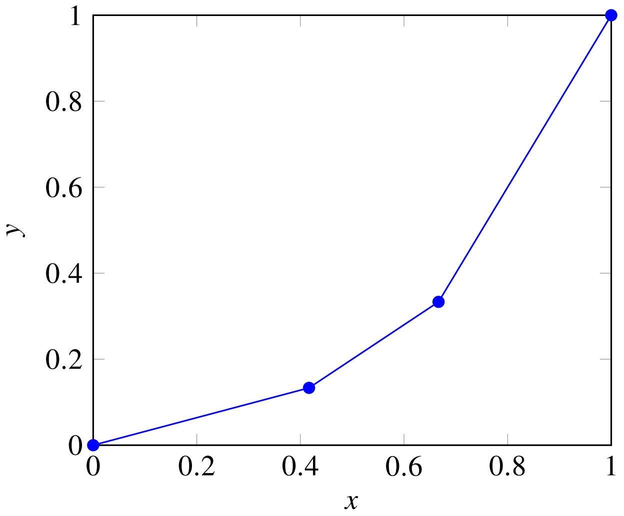


Figure 2.1: Lorenz curve associated with $\mathbf{p} = (1/4, 1/3, 5/12)$ and $\mathbf{q} = (1/5, 2/3, 2/15)$

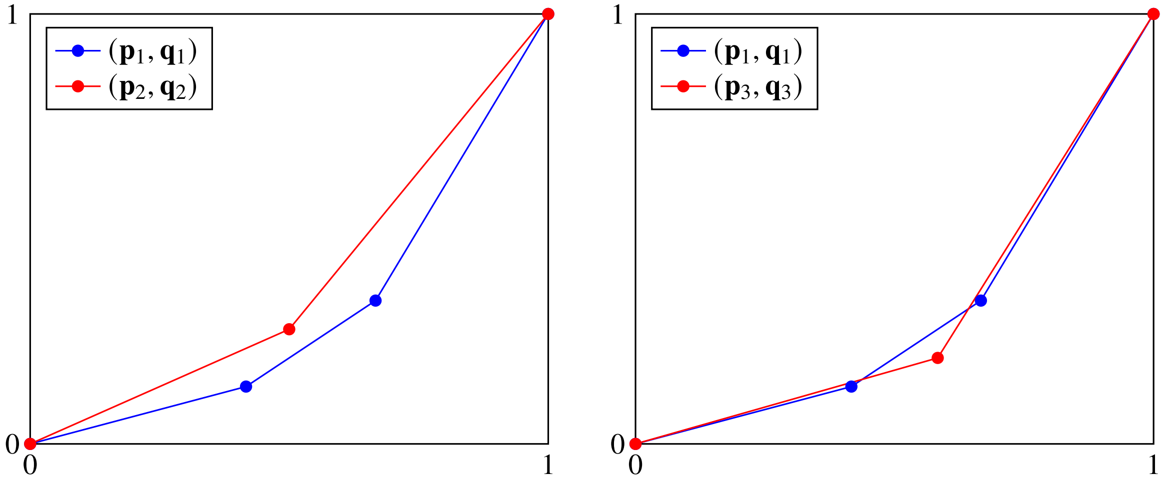
With Lorenz curves, we can state an equivalent condition for relative majorization.

Theorem 2.3.6: Relative majorization and Lorenz curves [83, 116, 118, 120, 121]

Let $\mathbf{p}_1, \mathbf{q}_1 \in \text{Prob}(n)$ and let $\mathbf{p}_2, \mathbf{q}_2 \in \text{Prob}(m)$. The following are equivalent:

1. $(\mathbf{p}_1, \mathbf{q}_1) \succ (\mathbf{p}_2, \mathbf{q}_2)$,
2. $\alpha^{(\mathbf{p}_1,\mathbf{q}_1)}(y) \geq \alpha^{(\mathbf{p}_2,\mathbf{q}_2)}(y)$ for all $y \in [0,1]$.

However, since $\alpha^{(\mathbf{p},\mathbf{q})}(y)$ is always a concave function, it is enough to check that all the elbows of $(\mathbf{p}_3, \mathbf{q}_3)$ are on the left of the Lorenz curve associated with $(\mathbf{p}_1, \mathbf{q}_1)$ [119] (see Figure 2.2).



(a) $(\mathbf{p}_1, \mathbf{q}_1) \succ (\mathbf{p}_2, \mathbf{q}_2)$

(b) $(\mathbf{p}_1, \mathbf{q}_1) \not\succ (\mathbf{p}_3, \mathbf{q}_3)$

Figure 2.2: Relative majorization versus Lorenz curves. In the first plot, one observes that $(\mathbf{p}_1, \mathbf{q}_1) \succ (\mathbf{p}_2, \mathbf{q}_2)$ because the Lorenz curve associated with $(\mathbf{p}_1, \mathbf{q}_1)$ is on the right of the Lorenz curve associated with $(\mathbf{p}_2, \mathbf{q}_2)$. In the second plot, instead, one notices that $(\mathbf{p}_1, \mathbf{q}_1) \not\succ (\mathbf{p}_3, \mathbf{q}_3)$ because one of the elbows of the Lorenz curve associated with $(\mathbf{p}_3, \mathbf{q}_3)$ is on the right of the Lorenz curve associated with $(\mathbf{p}_1, \mathbf{q}_1)$.

Chapter 3

Approximate LOCC Conversion and Entanglement Embezzlement

Quantum entanglement [23, 26, 27, 122–124] describes correlations between different particles with no classical counterpart and has both deep foundational implications [6] and numerous applications in quantum information science [12, 13, 16, 17, 125–131]. As described in Section 2.2, if two or more parties are far apart, in practice, they are restricted to local operations and classical communication (LOCC), and entanglement is a resource and thus studied within the framework of quantum resource theories (QRTs) [84, 85]. Consuming entangled states or, more generally, resourceful states can lead to operational advantages, e.g., in communication scenarios [12, 13, 16, 17, 125–128]. Importantly, suppose a state can be converted with a free operation into another one. In that case, the former is at least as valuable as the latter in any application that only allows for free operations. Answering the question of which states can be converted into each other is thus a central question in any QRT. In entanglement theory, Nielsen's Theorem [26] provides a characterization of the deterministic conversion between pure states that was later extended to probabilistic conversions [100]. The first result that we will present in this Chapter is a generalization of these results that characterizes the conversion of a pure state into a mixed state.

The results on state conversions mentioned so far require that an initial state is exactly converted into a target state. However, as shown in Section 2.1, approximate conversion is often enough, considering that every physical apparatus comes with finite accuracy. Conversion distances measure the error in approximate conversions, i.e., given two states, they are defined as the smallest distance between the second state and the output of any free operation on the first one. We introduce a new conversion distance under LOCC defined on pure states. Using our result on exact conversions, we prove that it is topologically equivalent to the other conversion distances and derive a closed formula for it.

This, in turn, will allow us to characterize entanglement embezzlement [52, 53]: Whilst it is impossible to create entanglement with LOCC [84, 122, 124, 132], it is possible to embezzle it in the sense that one converts a given entangled state approximately to itself *and* a copy of another entangled target state. A family of pure states such that one can do this for any target state with arbitrary accuracy is called a universal embezzling family. Embezzlement is thus a generalization of quantum catalysis, a phenomenon discovered in the early years of the resource theory of entanglement [133]: While in catalysis, the catalyst must be preserved exactly, embezzlement allows for it to be changed by an arbitrarily small amount. Very recently, we have witnessed a renewed interest in the catalysis of various quantum resources [134–150], with applications far beyond state conversion. In Ref. [144], for example, it was shown that a catalytic quantum teleportation protocol outperforms the standard teleportation protocol [16, 129, 130] in terms of teleportation fidelity. For further details and examples, see the recent review articles, Refs. [151, 152].

Universal embezzling families are valuable resources in various applications. They are, for example, necessary resources for the efficient simulation of noisy quantum channels with noiseless channels [128, 131], a result known in quantum information theory as 'Quantum Reverse Shannon Theorem'. They are also an important component in the elementary proofs of Grothendiek theorems [153, 154], which are of fundamental importance in the theories of Banach spaces and C^*- algebras (see, e.g., Ref. [155]). Furthermore, embezzling families are necessary to win various quantum guessing games with certainty [156–159]. Lastly, due to their close relation to catalysis, we expect that they will prove useful in applications such as teleportation, where catalysis provides advantages.

The first universal embezzling family was introduced in Ref. [52], and we refer to it as the van Dam and Hayden embezzling family. More recently, additional families have been proposed in Ref. [53]. In these works, embezzlement is considered using only local operations. In this Chapter, we extend the framework and additionally allow for classical communication, provide a complete characterization of universal embezzling families under LOCC, and discuss in what sense the van Dam and Hayden family is unique.

3.1 Approximate LOCC Conversion

As mentioned in Section 2.1.3 and at the beginning of this Chapter, when exact conversion is not possible, conversion distances are used to characterize how well Alice and Bob can *approximate* $\sigma \in \mathfrak{D}(A'B')$ with $\rho \in \mathfrak{D}(AB)$ and LOCC. As a note, from now on, we will drop the superscripts with the labels of the quantum system whenever they are obvious from the context. Moreover, as detailed in Section 2.2, we will assume all the systems to have the same dimension. In LOCC, the conversion distance associated with the trace distance is

$$T(\rho \to \sigma) = \inf_{\mathcal{N} \in \mathrm{LOCC}(AB \to A'B')} \frac{1}{2} \left\| \mathcal{N}(\rho) - \sigma \right\|_1 . \tag{3.1}$$

Note that this conversion distance has an operational interpretation in terms of a result in state discrimination known as Holevo-Helstrom Theorem [160, 161] (see Ref. [82, Theorem 3.4] for a review). Indeed, if a single copy of either $\sigma \in \mathfrak{D}(AB)$ or $\mathcal{N}(\rho) \in \mathfrak{D}(AB)$, with $\mathcal{N} \in$ LOCC, is given with equal probability, then the maximum probability $p_{\max}$ of correctly identifying the given state is bounded by

$$p_{\max} \geq \frac{1}{2}(1 + T(\rho \to \sigma)). \tag{3.2}$$

Moreover, for every fixed ρ and σ, there always exists an $\mathcal{N} \in$ LOCC such that $p_{\max}$ is arbitrarily close to this lower bound. In this sense, $T(\rho \to \sigma)$ describes how well we can approximate σ given access to ρ and LOCC.

Similarly, the conversion distance associated with the purified distance is

$$P(\rho \to \sigma) = \inf_{\mathcal{N} \in \mathrm{LOCC}(AB \to A'B')} \sqrt{1 - F^2(\mathcal{N}(\rho), \sigma)}, \tag{3.3}$$

where $F(\mathcal{N}(\rho), \sigma) = \left\| \sqrt{\mathcal{N}(\rho)}\sqrt{\sigma} \right\|_1$ is the fidelity.

In this Chapter, we restrict our analysis to pure initial and target states. Inspired by the fact that Schmidt coefficients play a crucial role in pure state LOCC, we define the following quantities.

Definition 3.1.1: Star conversion distances

Let ψ and $\varphi \in \mathrm{PURE}(AB)$, and let $\mathbf{p}$ and $\mathbf{q} \in \mathrm{Prob}^{\downarrow}(|A|)$ be their Schmidt coefficients. The trace star conversion distance is defined as

$$T_{\star}(\psi \to \varphi) = \min_{\mathbf{r} \succ \mathbf{p}} T(\mathbf{r}, \mathbf{q}), \tag{3.4}$$

and the purified star conversion distance as

$$P_{\star}(\psi \to \varphi) = \min_{\mathbf{r} \succ \mathbf{p}} P(\mathbf{r}, \mathbf{q}), \tag{3.5}$$

where $T(\mathbf{r}, \mathbf{q})$ and $P(\mathbf{r}, \mathbf{q})$ are the trace and purified distances defined in Section 1.1.

In the following two Subsections, we will present the properties of $P_{\star}$ and $T_{\star}$. Before doing that, we observe that even when dealing with pure initial and target states, we cannot use Nielsen's Theorem. Indeed, close to φ_{AB}, there are mixed states as well, while Nielsen's theorem only applies to pure-to-pure exact state conversions. Our first step is to generalize Nielsen's theorem to pure-to-mixed state conversions.

Proposition 3.1.2: Pure to mixed state conversion

For $\psi \in \mathrm{PURE}(AB)$ and $\sigma \in \mathfrak{D}(AB)$, $\psi \xrightarrow{\text{LOCC}} \sigma$ if and only if there exists a pure state decomposition $\{\, p_z, \chi_z \,\}$ of σ, that is, $\sigma = \sum_z p_z \chi_z$, such that

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z p_z E_k(\chi_z) \right\} \geq 0, \tag{3.6}$$

where $E_k(\cdot) = 1 - \|\cdot\|_k$.

Proof. The sufficient condition follows from Refs. [100, 162]. Indeed, the authors show that

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z p_z E_k(\chi_z) \right\} \geq 0 \tag{3.7}$$

if and only if the probabilistic conversion $\psi \xrightarrow{\text{LOCC}} \{\, p_z, \chi_z^{AB} \,\} = \sum_z p_z |z\rangle\langle z|^X \otimes \chi_z^{AB}$ is achievable. By discarding the classical system X, which is a free LOCC operation, one obtains the state $\sum_z p_z \chi_z = \sigma$. This proves $\psi \xrightarrow{\text{LOCC}} \sigma$.

For the converse, if $\psi \xrightarrow{\text{LOCC}} \sigma$, then there exist a generalized measurement $\{\, M_z^A \,\}$ and a family of unitary matrices $\{\, U_z^B \,\}$ such that [25, 163]

$$\sigma^{AB} = \sum_z (M_z^A \otimes U_z^B) \psi^{AB} (M_z^A \otimes U_z^B)^{\dagger}. \tag{3.8}$$

If Alice performs the generalized measurement M_z^A, records the outcome of the measurement in a classical system X, and sends the outcome to Bob, who performs the unitary U_z^B associated with the outcome of the measurement, they obtain the state

$$\begin{aligned}\sum_z p_z |z\rangle\langle z|^X \otimes (M_z^A \otimes U_z^B) \psi^{AB} (M_z^A \otimes U_z^B)^\dagger &= \sum_z p_z |z\rangle\langle z|^X \otimes \chi_z^{AB} \\ &= \left\{ p_z, \chi_z^{AB} \right\},\end{aligned} \tag{3.9}$$

where $|\chi_z\rangle^{AB} = \frac{(M_z^A \otimes U_z^B)|\psi\rangle^{AB}}{\|(M_z^A \otimes U_z^B)|\psi\rangle^{AB}\|}$. Therefore, $\psi^{AB} \xrightarrow{\text{LOCC}} \left\{ p_z, \chi_z^{AB} \right\}$, and this is equivalent to

$$\min_{k \in [|A|]} \left\{ E_k(\psi) - \sum_z p_z E_k(\chi_z) \right\} \geq 0, \tag{3.10}$$

as shown in Refs. [100, 162]. □

3.1.1 Properties of the Purified Star Conversion Distance

The main result of this Subsection is that two conversion distances based on the purified distance coincide on pure states, i.e., $P(\psi \to \varphi) = P_\star(\psi \to \varphi)$. To this end, we need the following Lemma.

Lemma 3.1.3: Concavity of $f(\mathbf{v}) = \left(\sum_{x=1}^d \sqrt{q_x v_x}\right)^2$

Let $\mathbf{q} \in \mathrm{Prob}(d)$. The function $f\colon \mathrm{Prob}(d) \to [0,1]$ defined as $f(\mathbf{v}) = \left(\sum_{x=1}^d \sqrt{q_x v_x}\right)^2$ is concave.

Proof. A twice differentiable function is concave if and only if its Hessian matrix is negative semi-definite (see Ref. [164] for properties of concave functions). The functions $f_{x,y}(v_x, v_y) := \sqrt{q_x v_x}\sqrt{q_y v_y}$ are twice differentiable in v_x, v_y for $v_x, v_y \in (0,1]$ and their Hessian matrices are given by

$$H\left(v_x, v_y\right) = \frac{1}{4}\begin{bmatrix} -\sqrt{\frac{q_x q_y v_y}{v_x^3}} & \sqrt{\frac{q_x q_y}{v_x v_y}} \\ \sqrt{\frac{q_x q_y}{v_x v_y}} & -\sqrt{\frac{q_x q_y v_x}{v_y^3}} \end{bmatrix}. \tag{3.11}$$

For all $v_x,\ v_y \in (0,1]$, the determinant of these Hessian matrices is zero. Therefore, the eigenvalues of $H(v_x, v_y)$ are 0 and $\operatorname{Tr} H(v_x, v_y) \leq 0$. This implies that the functions $f_{x,y}(v_x, v_y)$ are concave for $v_x, v_y \in (0,1]$. Since $0 = f_{x,y}(0, v_y) = f_{x,y}(v_x, 0) = f_{x,y}(0,0)$, we obtain that $f_{x,y}(v_x, v_y)$ is in fact concave for $v_x, v_y \in [0,1]$. As a consequence f is concave:

$$f\left(\mathbf{v}\right) = \sum_x v_x q_x + \sum_y \sum_{x \neq y} \sqrt{q_x v_x}\sqrt{q_y v_y} \tag{3.12}$$

and the sum of concave functions is a concave function. □

We are now ready to prove the promised theorem.

Theorem 3.1.4: Equality of P and $P_\star$

Let $\psi, \varphi \in \mathrm{PURE}(AB)$, then $P\left(\psi \to \varphi\right) = P_\star\left(\psi \to \varphi\right)$.

Proof. We assume w.l.o.g. that all pure states are in standard form, that is, $|\psi\rangle = \sum_x \sqrt{p_x}\,|xx\rangle_{AB}$, where $\mathbf{p} \in \mathrm{Prob}^{\downarrow}(d)$ and $\{\,|x\rangle_A\,\}$ and $\{\,|x\rangle_B\,\}$ are fixed bases for A and B, respectively. Let $\mathbf{p}, \mathbf{q} \in \mathrm{Prob}^{\downarrow}(d)$ be the Schmidt coefficients or ψ and φ, respectively. With $\mathbf{r} \in \mathrm{Prob}^{\downarrow}(d)$, as a consequence of Nielsen's Theorem [26], $\mathbf{r} \succ \mathbf{p}$ if and only if there exists an $\mathcal{N} \in \mathrm{LOCC}$ such that $\mathcal{N}(\psi) \in \mathrm{PURE}(AB)$ has Schmidt coefficients $\mathbf{r}$. We notice that

$$\begin{aligned} P(\mathcal{N}(\psi), \varphi) &= \sqrt{1 - F^2(\mathcal{N}(\psi), \varphi)} = \sqrt{1 - |\langle \mathcal{N}(\psi) \,|\, \varphi \rangle|^2} \\ &= \sqrt{1 - \left(\sum_x \sqrt{r_x q_x} \right)^2} = \sqrt{1 - F^2(\mathbf{r}, \mathbf{q})} \\ &= P(\mathbf{r}, \mathbf{q}). \end{aligned} \tag{3.13}$$

This implies that

$$\begin{aligned} P_\star\left(\psi \to \varphi\right) &= \min_{\mathbf{r} \succ \mathbf{p}} P\left(\mathbf{r}, \mathbf{q}\right) \\ &= \min_{\substack{\mathcal{N} \in \mathrm{LOCC} \\ \mathcal{N}(\psi) \in \mathrm{PURE}(AB)}} P\left(\mathcal{N}\left(\psi\right), \varphi\right) \\ &\geq \inf_{\mathcal{N} \in \mathrm{LOCC}} P\left(\mathcal{N}\left(\psi\right), \varphi\right) \\ &= P\left(\psi \to \varphi\right). \end{aligned} \tag{3.14}$$

The non-trivial part is to show that the opposite inequality also holds. To this end, we want to show that for every mixed state σ such that $\psi \xrightarrow{\mathrm{LOCC}} \sigma$, there exists a pure state χ such that $\psi \xrightarrow{\mathrm{LOCC}} \chi$ and $P(\sigma, \varphi) \geq P(\chi, \varphi)$. It is then sufficient to consider only pure output states for the computation of the purified conversion distance, which implies the reverse inequality.

Let $\sigma = \mathcal{M}(\psi) \in \mathfrak{Q}(AB)$, where $\mathcal{M} \in \mathrm{LOCC}$. Also, let $\{\, t_z, \chi_z \,\}$ be a pure state decomposition of σ that satisfies Eq. (3.6) (where the χ_z are not necessarily in standard form), and $\mathbf{s}^{(z)}$ be the Schmidt coefficient of χ_z for every z. Furthermore, for every χ_z, let $\tilde{\chi}_z$ be the pure state in standard form that is equal to χ_z up to local unitary transformations, let $\tilde{\sigma} = \sum_z t_z \tilde{\chi}_z$, and define $\chi \in \mathrm{PURE}(AB)$ as the pure bipartite state (in standard form) with Schmidt coefficients $\mathbf{s} = \sum_z t_z \mathbf{s}^{(z)}$. We notice that for all $k \in [d]$

$$\begin{aligned} E_k(\chi) &= 1 - \sum_{x=1}^{k} s_x = 1 - \sum_{x=1}^{k} \sum_z t_z s_x^{(z)} \\ &= 1 - \sum_z t_z (1 - E_k(\chi_z)) = \sum_z t_z E_k(\chi_z) \\ &\leq E_k(\psi), \end{aligned} \tag{3.15}$$

where the last inequality follows from the fact that $\{\, t_z, \chi_z \,\}$ satisfies Eq. (3.6). This implies that $\mathbf{s} \succ \mathbf{p}$.

The next step is to show that $P(\sigma, \varphi) \geq P(\mathbf{s}, \mathbf{q})$. First, we note that due to the von Neumann trace inequality [165, 166],

$$F^2\left(\tilde{\sigma}, \varphi\right) = \sum_z t_z \left(\sum_x \sqrt{q_x s_x^{(z)}} \right)^2 \geq F^2\left(\sigma, \varphi\right), \tag{3.16}$$

and

$$F^2(\chi, \varphi) = F^2\left(\mathbf{s}, \mathbf{q}\right) = \left(\sum_x \sqrt{q_x \sum_z t_z s_x^{(z)}} \right)^2. \tag{3.17}$$

Second, we introduce the concave function $f(\mathbf{v}) = \left(\sum_x \sqrt{q_x v_x}\right)^2$, for $\mathbf{v} \in \text{Prob}(d)$ (see Lemma 3.1.3), and rewrite Eq. (3.16) and Eq. (3.17) as

$$F^2(\sigma,\varphi) \leq F^2(\tilde{\sigma},\varphi) = \sum_z t_z f\left(\mathbf{s}^{(z)}\right), \quad F^2(\mathbf{s},\mathbf{q}) = f\left(\sum_z t_z \mathbf{s}^{(z)}\right). \tag{3.18}$$

Finally, the concavity of f and the multidimensional Jensen inequality [167] imply

$$F^2(\sigma,\varphi) \leq F^2(\tilde{\sigma},\varphi) = \sum_z t_z f\left(\mathbf{s}^{(z)}\right) \leq f\left(\sum_z t_z \mathbf{s}^{(z)}\right) = F^2(\mathbf{s},\mathbf{q}), \tag{3.19}$$

which is equivalent to $P(\sigma,\varphi) \geq P(\mathbf{s},\mathbf{q})$. So far, we have shown that for every $\sigma = \mathcal{M}(\psi)$, with $\mathcal{M} \in \text{LOCC}$, there exists an $\mathbf{s}^{(\mathcal{M}(\psi))} \in \text{Prob}^{\downarrow}(d)$ such that $\mathbf{s}^{(\mathcal{M}(\psi))} \succ \mathbf{p}$ and $P(\sigma,\varphi) \geq P(\mathbf{s}^{(\mathcal{M}(\psi))},\mathbf{q})$, thus

$$\begin{aligned} P(\psi \to \varphi) &= \inf_{\mathcal{M}\in\text{LOCC}} P(\mathcal{M}(\psi),\varphi) \\ &\geq \inf_{\mathcal{M}\in\text{LOCC}} P\left(\mathbf{s}^{(\mathcal{M}(\psi))},\mathbf{q}\right) \\ &\geq \min_{\mathbf{r}\succ\mathbf{p}} P(\mathbf{r},\mathbf{q}) \\ &= P_\star(\psi \to \varphi). \end{aligned} \tag{3.20}$$

Eq. (3.14) and Eq. (3.20) imply that $P(\psi \to \varphi) = P_\star(\psi \to \varphi)$, and this concludes the proof. □

3.1.2 Properties of the Trace Star Conversion Distance

In this Subsection, we present properties of the star conversion distance based on the trace distance. We begin by proving that the trace star conversion distance is topologically equivalent to the standard conversion distance defined via the trace distance.

Lemma 3.1.5: Topological equivalence of the conversion distances T and $T_\star$

Let $\psi, \varphi \in \text{PURE}(AB)$. Then

$$\frac{1}{2}\left[T_\star(\psi \to \varphi)\right]^2 \leq T(\psi \to \varphi) \leq \sqrt{2T_\star(\psi \to \varphi)}. \tag{3.21}$$

Proof. First, we observe that the trace distance and the purified distance are equivalent [95], i.e., for all $\rho, \sigma \in \mathfrak{D}(AB)$:

$$T(\rho,\sigma) \leq P(\rho,\sigma) \leq \sqrt{2T(\rho,\sigma)}. \tag{3.22}$$

Then, thanks to Theorem 3.1.4, we obtain

$$\begin{aligned} T(\psi \to \varphi) &= \inf_{\mathcal{N}\in\mathrm{LOCC}} T(\mathcal{N}(\psi), \varphi) \\ &\le \inf_{\mathcal{N}\in\mathrm{LOCC}} P(\mathcal{N}(\psi), \varphi) \\ &= P(\psi \to \varphi) = P_\star(\psi \to \varphi) \\ &= \min_{\mathbf{r}\succ\mathbf{p}} P(\mathbf{r}, \mathbf{q}) \\ &\le \min_{\mathbf{r}\succ\mathbf{p}} \sqrt{2T(\mathbf{r}, \mathbf{q})} \\ &= \sqrt{2T_\star(\psi \to \varphi)}, \end{aligned} \tag{3.23}$$

Analogously,

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\mathbf{r}\succ\mathbf{p}} T(\mathbf{r}, \mathbf{q}) \\ &\le \min_{\mathbf{r}\succ\mathbf{p}} P(\mathbf{r}, \mathbf{q}) \\ &= P_\star(\psi \to \varphi) \\ &= P(\psi \to \varphi) \\ &= \inf_{\mathcal{N}\in\mathrm{LOCC}} P(\mathcal{N}(\psi), \varphi) \\ &\le \inf_{\mathcal{N}\in\mathrm{LOCC}} \sqrt{2T(\mathcal{N}(\psi), \varphi)} \\ &= \sqrt{2T(\psi \to \varphi)}. \end{aligned} \tag{3.24}$$

□

This proves that $T_\star(\psi \to \varphi)$ is topologically equivalent to $T(\psi \to \varphi)$.

The next theorem is the key result of this part about conversion distances. It provides a closed-form expression for the trace star conversion distance, and it shows that any non-separable state is more valuable than any separable state for approximating a non-separable state.

Theorem 3.1.6: Closed-form expression for $T_\star$

Let $\psi, \varphi \in \mathrm{PURE}(AB)$ and let $\mathbf{p}, \mathbf{q} \in \mathrm{Prob}^{\downarrow}(|A|)$ be their corresponding Schmidt coefficients. Then,

$$T_\star(\psi \to \varphi) = \max_{k\in[\mathrm{SR}(\psi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\}. \tag{3.25}$$

If $\xi \in \mathrm{PURE}(AB)$ is separable and ψ, φ are not, then

$$T_\star(\psi \to \varphi) < T_\star(\xi \to \varphi). \tag{3.26}$$

Proof. Let $\varepsilon \in [0, 1]$ and $\mathfrak{B}^{\varepsilon}_{\mathbf{q}} = \left\{ \mathbf{q}' : \frac{1}{2} \|\mathbf{q} - \mathbf{q}'\|_1 \le \varepsilon \right\}$. As shown in Ref. [111], there exist probability vectors $\bar{\mathbf{q}}^{(\varepsilon)} \in \mathfrak{B}^{\varepsilon}_{\mathbf{q}}$ called steepest ε-approximations of $\mathbf{q}$ such that $\bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{q}'$ for all $\mathbf{q}' \in \mathfrak{B}^{\varepsilon}_{\mathbf{q}}$. Moreover, these steepest ε-approximations can be constructed explicitly: If $\frac{1}{2} \|\mathbf{q} - \mathbf{e}_1\|_1 \le \varepsilon$, then $\bar{\mathbf{q}}^{(\varepsilon)} = \mathbf{e}_1$, otherwise let

$k_\varepsilon \in [d]$ be the index satisfying

$$\sum_{x=1}^{k_\varepsilon} q_x \leq 1-\varepsilon \quad \text{and} \quad \sum_{x=1}^{k_\varepsilon+1} q_x > 1-\varepsilon. \tag{3.27}$$

The components of $\bar{\mathbf{q}}^{(\varepsilon)}$ are then given by

$$\bar{q}_x^{(\varepsilon)} = \begin{cases} q_1+\varepsilon & \text{if } x=1, \\ q_x & \text{if } x \in \{2,\ldots,k_\varepsilon\}, \\ 1-\varepsilon-\sum_{x=1}^{k_\varepsilon} q_x & \text{if } x = k_\varepsilon+1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.28}$$

We now show that

$$\min_{\mathbf{r}\succ\mathbf{p}} \left\{ \frac{1}{2}\|\mathbf{q}-\mathbf{r}\|_1 \right\} = \min\left\{ \varepsilon\in[0,1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \tag{3.29}$$

First, we notice that if $\mathbf{r}_\star$ is an optimizer of $\min_{\mathbf{r}\succ\mathbf{p}}\left\{\frac{1}{2}\|\mathbf{q}-\mathbf{r}\|_1\right\}$ and $\tilde{\varepsilon} = \frac{1}{2}\|\mathbf{q}-\mathbf{r}_\star\|_1$, then $\mathbf{r}_\star \in \mathfrak{B}_{\mathbf{q}}^{\tilde{\varepsilon}}$ and therefore $\bar{\mathbf{q}}^{(\tilde{\varepsilon})} \succ \mathbf{r}_\star$. By transitivity, we also have $\bar{\mathbf{q}}^{(\tilde{\varepsilon})} \succ \mathbf{p}$, which implies that

$$\min\left\{ \varepsilon\in[0,1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\} \leq \tilde{\varepsilon} = \min_{\mathbf{r}\succ\mathbf{p}} \left\{ \frac{1}{2}\|\mathbf{q}-\mathbf{r}\|_1 \right\}. \tag{3.30}$$

For the reverse inequality, let $\varepsilon_\star$ be the optimizer of $\min\left\{ \varepsilon\in[0,1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}$. By definition, $\bar{\mathbf{q}}^{(\varepsilon_\star)} \succ \mathbf{p}$ and $\frac{1}{2}\left\|\mathbf{q}-\bar{\mathbf{q}}^{(\varepsilon_\star)}\right\|_1 \leq \varepsilon_\star$, thus

$$\min_{\mathbf{r}\succ\mathbf{p}} \left\{ \frac{1}{2}\|\mathbf{q}-\mathbf{r}\|_1 \right\} \leq \varepsilon_\star = \min\left\{ \varepsilon\in[0,1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \tag{3.31}$$

This shows that

$$\begin{aligned} T_\star(\psi\longrightarrow\varphi) &= \min_{\mathbf{r}\in\mathrm{Prob}^{\downarrow}(d)} \left\{ \frac{1}{2}\|\mathbf{q}-\mathbf{r}\|_1 : \mathbf{r}\succ\mathbf{p} \right\} \\ &= \min\left\{ \varepsilon\in[0,1] : \bar{\mathbf{q}}^{(\varepsilon)} \succ \mathbf{p} \right\}. \end{aligned} \tag{3.32}$$

We observe that $\left\|\bar{\mathbf{q}}^{(\varepsilon)}\right\|_{(k)} = \min\left\{1, \|\mathbf{q}\|_{(k)}+\varepsilon\right\}$, thus the condition $\bar{\mathbf{q}}^{\varepsilon} \succ \mathbf{p}$ is equivalent to

$$\|\mathbf{p}\|_{(k)} \leq \min\left\{1, \|\mathbf{q}\|_{(k)}+\varepsilon\right\}, \quad \forall k\in[d]. \tag{3.33}$$

This expression is further simplified by noticing that $\mathbf{p}$ is a probability vector, and therefore $\|\mathbf{p}\|_{(k)} \leq 1$ for all $k\in[d]$. Consequently,

$$\begin{aligned} \bar{\mathbf{q}}^{\varepsilon} \succ \mathbf{p} \quad &\Leftrightarrow \quad \|\mathbf{p}\|_{(k)} \leq \|\mathbf{q}\|_{(k)}+\varepsilon,\ \forall k\in[d] \\ &\Leftrightarrow \quad \max_{k\in[d]}\left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\} \leq \varepsilon. \end{aligned} \tag{3.34}$$

In combination with the minimization in Eq. (3.32) follows that

$$T_\star(\psi\longrightarrow\varphi) = \max_{k\in[d]}\left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\}. \tag{3.35}$$

To conclude the proof of the first part of the theorem, we observe that for $k \geq \mathrm{SR}(\psi^{AB})$, $\|\mathbf{p}\|_{(k)} = 1$ and $\|\mathbf{q}\|_{(k)}$ is non-decreasing with k. Thus, we can restrict the maximization to $k \leq \mathrm{SR}(\psi^{AB})$.

For the second part, we notice that according to Eq. (3.32),

$$\begin{aligned} T_\star(\xi \to \varphi) &= \min_{\mathbf{r}\in\mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2}\left\|\mathbf{q}-\mathbf{r}\right\|_1 : \mathbf{r} \succ \mathbf{e}_1 \right\} \\ &= \frac{1}{2}\left\|\mathbf{q}-\mathbf{e}_1\right\|_1 \\ &= 1 - q_1 . \end{aligned} \tag{3.36}$$

Furthermore, from $\mathbf{e}_1 \succ \mathbf{p}$ and the transitivity of the majorization-relation, it follows that

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\mathbf{r}\in\mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2}\left\|\mathbf{q}-\mathbf{r}\right\|_1 : \mathbf{r} \succ \mathbf{p} \right\} \\ &\le \min_{\mathbf{r}\in\mathrm{Prob}^\downarrow(d)} \left\{ \frac{1}{2}\left\|\mathbf{q}-\mathbf{r}\right\|_1 : \mathbf{r} \succ \mathbf{e}_1 \right\} \\ &= T_\star(\xi \to \varphi). \end{aligned} \tag{3.37}$$

To rule out equality, suppose that $T_\star(\psi \to \varphi) = \max_{k\in[d]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{q}\|_{(k)} \right\} = 1 - q_1$, and denote with $k_\star$ an index that achieves this maximum. Then

$$\|\mathbf{p}\|_{(k_\star)} = 1 + \|\mathbf{q}\|_{(k_\star)} - q_1 . \tag{3.38}$$

From this expression follows that $\|\mathbf{p}\|_{(k_\star)} \le 1$ only if either $\mathbf{q} = \mathbf{e}_1$ or if $k_\star = 1$, and therefore $\mathbf{p} = \mathbf{e}_1$. These conditions are in contrast with the assumption that ψ and φ are not separable. As a consequence, equality in Eq. (3.37) is unachievable, which proves the second part of the theorem. □

Next, we show that the star conversion distance satisfies a triangle inequality. Let ψ, φ, and $\chi \in \mathrm{PURE}(AB)$ and let $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ be their Schmidt coefficients. With the help of Eq. (3.35), this implies that

$$\begin{aligned} T_\star(\psi \to \chi) &= \max_{n\in[d]} \left(\|\mathbf{p}\|_{(n)} - \|\mathbf{r}\|_{(n)}\right) \\ &= \max_{n\in[d]} \left(\|\mathbf{p}\|_{(n)} - \|\mathbf{q}\|_{(n)} + \|\mathbf{q}\|_{(n)} - \|\mathbf{r}\|_{(n)}\right) \\ &\le \max_{n\in[d]} \left(\|\mathbf{p}\|_{(n)} - \|\mathbf{q}\|_{(n)}\right) + \max_{n\in[d]} \left(\|\mathbf{q}\|_{(n)} - \|\mathbf{r}\|_{(n)}\right) \\ &= T_\star(\psi \to \varphi) + T_\star(\varphi \to \chi). \end{aligned} \tag{3.39}$$

Lastly, we generalize the definition of the conversion distance to initial and target systems of different dimensions. We notice that an equivalent definition of $P_\star$ and $T_\star$ is:

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\substack{\mathcal{N}\in\mathrm{LOCC} \\ \mathcal{N}(\psi)\in\mathrm{PURE}(AB)}} T(\mathcal{N}(\psi), \varphi), \\ P_\star(\psi \to \varphi) &= \min_{\substack{\mathcal{N}\in\mathrm{LOCC} \\ \mathcal{N}(\psi)\in\mathrm{PURE}(AB)}} P(\mathcal{N}(\psi), \varphi). \end{aligned} \tag{3.40}$$

This definition can be easily generalized to encompass systems of different dimensions. Let $\psi \in \mathfrak{Q}(AB)$ and $\varphi \in \mathfrak{Q}(A'B')$, where $|A| = |B| = d$ and $|A'| = |B'| = d'$. Then,

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\substack{\mathcal{N}\in\mathrm{LOCC}(AB\to A'B') \\ \mathcal{N}(\psi)\in\mathrm{PURE}(A'B')}} T(\mathcal{N}(\psi), \varphi), \\ P_\star(\psi \to \varphi) &= \min_{\substack{\mathcal{N}\in\mathrm{LOCC}(AB\to A'B') \\ \mathcal{N}(\psi)\in\mathrm{PURE}(A'B')}} P(\mathcal{N}(\psi), \varphi). \end{aligned} \tag{3.41}$$

Now, let D denote either the trace or the purified distance. Let both $\bar{d} = \max\{ d, d' \}$, $|\bar{A}| = \bar{d} = |\bar{B}|$ and let $\mathcal{U}^{AB\to\bar{A}\bar{B}}$ and $\mathcal{V}^{A'B'\to\bar{A}\bar{B}}$ be the tensor product of two local isometries. Since D is invariant under local isometries, then

$$\begin{aligned} D(\mathcal{N}(\psi), \varphi) &= D(\mathcal{V} \circ \mathcal{N}(\psi), \mathcal{V}(\varphi)) \\ &= D(\mathcal{V} \circ \mathcal{N} \circ \mathcal{U}^\dagger(\mathcal{U}(\psi)), \mathcal{V}(\varphi)). \end{aligned} \tag{3.42}$$

Moreover, since $\mathcal{U}^{AB\to\bar{A}\bar{B}}$ and $\mathcal{V}^{A'B'\to\bar{A}\bar{B}}$ are in LOCC, $\mathcal{V} \circ \mathcal{N} \circ \mathcal{U}^\dagger \in \text{LOCC} \Leftrightarrow \mathcal{N} \in \text{LOCC}$, and $\mathcal{V} \circ \mathcal{N}(\psi) \in \text{PURE}(\bar{A}\bar{B}) \Leftrightarrow \mathcal{N}(\psi) \in \text{PURE}(A'B')$. Let us denote $\bar{\mathcal{N}} = \mathcal{V} \circ \mathcal{N} \circ \mathcal{U}^\dagger$, $\bar{\psi} = \mathcal{U}(\psi)$, $\bar{\varphi} = \mathcal{V}(\varphi)$. We can write the trace and purified conversion distances as follows:

$$\begin{aligned} D_\star(\psi \to \varphi) &= \min_{\substack{\mathcal{N} \in \text{LOCC}(AB\to A'B') \\ \mathcal{N}(\psi) \in \text{PURE}(A'B')}} D(\mathcal{N}(\psi), \varphi) \\ &= \min_{\substack{\mathcal{N} \in \text{LOCC}(AB\to A'B') \\ \mathcal{N}(\psi) \in \text{PURE}(A'B')}} D(\mathcal{V} \circ \mathcal{N} \circ \mathcal{U}^\dagger(\mathcal{U}(\psi)), \mathcal{V}(\varphi)) \\ &= \min_{\substack{\bar{\mathcal{N}} \in \text{LOCC}(\bar{A}\bar{B}\to\bar{A}\bar{B}) \\ \bar{\mathcal{N}}(\bar{\psi}) \in \text{PURE}(\bar{A}\bar{B})}} D(\bar{\mathcal{N}}(\bar{\psi}), \bar{\varphi}) \\ &= \min_{\mathbf{r} \succ \bar{\mathbf{p}}} D(\mathbf{r}, \bar{\mathbf{q}}), \end{aligned} \tag{3.43}$$

where $\bar{\mathbf{p}}$ and $\bar{\mathbf{q}}$ are the Schmidt coefficients of $\bar{\psi}$ and $\bar{\varphi}$, respectively. Since $\bar{\psi}$ is locally isometric to ψ, the non-zero entries of $\bar{\mathbf{p}}$ are the Schmidt coefficients of ψ. Analogously, the non-zero entries of $\bar{\mathbf{q}}$ coincide with the Schmidt coefficients of φ. Therefore, the star conversion distances between systems of different dimensions are:

$$\begin{aligned} T_\star(\psi \to \varphi) &= \min_{\mathbf{r} \succ \bar{\mathbf{p}}} T(\mathbf{r}, \bar{\mathbf{q}}), \\ P_\star(\psi \to \varphi) &= \min_{\mathbf{r} \succ \bar{\mathbf{p}}} P(\mathbf{r}, \bar{\mathbf{q}}), \end{aligned} \tag{3.44}$$

where $\bar{\mathbf{p}}, \bar{\mathbf{q}} \in \text{Prob}(\bar{d})$ are probability vectors whose non-zero entries are the non-zero Schmidt coefficients of ψ and φ, respectively.

3.2 Entanglement Embezzlement

It is impossible to create additional entanglement with LOCC alone [84, 122, 124, 132]: If $\rho \in \mathfrak{Q}(A'B')$ is an entangled state, this implies that there cannot exist a $\psi \in \text{PURE}(AB)$ and a channel $\mathcal{N} \in \text{LOCC}(AB \to ABA'B')$ such that $\mathcal{N}(\psi) = \psi \otimes \rho$, because this would increase the total amount of entanglement between systems AA' and BB' with respect to any additive entanglement measure. It might, however, be possible to *approximate* $\psi \otimes \rho$ in the sense that $T(\psi \to \psi \otimes \rho) \leq \varepsilon$ for a small ε. In this case, it is hard to distinguish $\psi \otimes \rho$ from the approximation. By keeping the systems $A'B'$, one would thus embezzle entanglement from the owner of ψ - and if one would be able to make ε arbitrarily small, it would be impossible to detect.

In fact, in Ref. [52], van Dam and Hayden showed that it is possible to embezzle *any* bipartite state $\sigma \in \mathfrak{Q}(A'B')$ arbitrarily well from a family of pure states $\{ \chi_n^{AB} \}_{n\in\mathbb{N}}$ in the sense that $\lim_{n\to\infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$. This implies that an arbitrarily good approximation of any σ can be embezzled from χ_n whilst changing χ_n arbitrarily little, as long as n is large enough. This motivates the following definition.

Definition 3.2.1: Universal embezzling family

A family of pure bipartite states $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ is called a *universal embezzling family* if $\lim_{n\to\infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$ for every bipartite finite dimensional state σ.

This definition is very similar to the one provided in Refs. [52, 53]. The main difference is that these works only consider protocols using local operations (LO), whereas this work also allows for classical communication. The set of operations that we consider is, therefore, larger, and as a result, if a family of states is not an embezzling family according to our definition, then it is not an embezzling family in the sense of Refs. [52, 53]. Surprisingly, the original embezzling family proposed in Ref. [52] is unique in the sense that we will specify later, even when we allow for classical communication. Another difference with Refs. [52, 53] is that in those works, the fidelity was used to quantify the conversion error. Since the fidelity, or more accurately $1 - F$, and the trace distance are topologically equivalent, in the sense that there exist two positive constants C, and D such that $CT(\rho, \sigma) \leq 1 - F(\rho, \sigma) \leq DT(\rho, \sigma)$ for all states $\rho^A, \sigma^A \in \mathfrak{D}(A)$ and for all systems A, the choice of metric is irrelevant.

Evaluating whether a family of states is an embezzling family with the definition above is quite challenging. However, we show in the following lemma that it is enough to evaluate the limit of star conversion distance $T_\star(\chi_n \to \chi_n \otimes \Phi_2)$, where Φ_m is the maximally entangled state with m Schmidt coefficients.

Lemma 3.2.2: Equivalent definitions of universal embezzling families (cf. Ref. [53, Lemma 2 (LO)])

Let $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ be a family of pure bipartite states. The following three statements are equivalent

1. $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ is a universal embezzling family.
2. $\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_2) = 0$.
3. $\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 0$ for every $m \geq 2$.

Proof. Clearly, 2. and 3. follow from 1. due to the definition of universal embezzling families and the topological equivalence of T and $T_\star$. Moreover, 3. follows from 2. because $\Phi_2^{\otimes\lceil \log_2 m\rceil} \xrightarrow{\text{LOCC}} \Phi_m$, and therefore

$$\begin{aligned}
T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil - 1}) + T_\star(\chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil - 1} \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq T_\star(\chi_n \to \chi_n \otimes \Phi_2) + T_\star(\chi_n \otimes \Phi_2 \to \chi_n \otimes \Phi_2^{\otimes 2}) + \dots \\
&\qquad + T_\star(\chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil - 1} \to \chi_n \otimes \Phi_2^{\otimes\lceil \log_2 m\rceil}) \\
&\leq \lceil \log_2 m\rceil T_\star(\chi_n \to \chi_n \otimes \Phi_2).
\end{aligned} \tag{3.45}$$

By taking the limit $n \to \infty$ on both sides, we obtain the desired result. To conclude, we note that 3. implies 1., since for all σ^{AB}, $\Phi_{|A|} \xrightarrow{\text{LOCC}} \sigma^{AB}$. □

It is also important to note that technically, one could have required that $\liminf_{n\to\infty} T(\chi_n \to \chi_n \otimes \sigma) = 0$, since this would also allow to embezzle any state arbitrarily well. However, since one can always choose a subfamily, we decided to keep the definition in line with Ref. [53]. In the following, we find an equivalent characterization of embezzling families based on the closed formula for the star conversion distance derived in Theorem 3.1.6. To achieve that, first, we write an explicit expression for $T_\star(\chi \to \chi \otimes \Phi_m)$, for all $m \in \mathbb{N}$.

Lemma 3.2.3: An explicit expression for $T_\star(\chi \to \chi \otimes \Phi_m)$

Let $\mathbf{p} \in \mathrm{Prob}^{\downarrow}(d)$ be the Schmidt coefficients of χ. Then

$$T_\star(\chi \to \chi \otimes \Phi_m) = \max_{k \in [\mathrm{SR}(\chi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{p}\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1} \right\}, \tag{3.46}$$

where $a_k = \lfloor k/m \rfloor$, $\lfloor \cdot \rfloor$ denotes the floor, and $b_k = k - a_k m$.

Proof. The Schmidt coefficients of the input and target states are

$$\begin{aligned} \mathbf{p} \otimes \mathbf{e}_1 &= (p_1, \ldots, p_d, \underbrace{0, \ldots, 0}_{d\cdot(m-1)\text{ times}}), \\ \mathbf{p} \otimes \mathbf{u}^{(m)} &= \frac{1}{m}(\underbrace{p_1, \ldots, p_1}_{m\text{ times}}, \ldots, \underbrace{p_d, \ldots, p_d}_{m\text{ times}}), \end{aligned} \tag{3.47}$$

where $\mathbf{u}^{(m)} = (1/m, \ldots, 1/m)$. It is straightforward to see that

$$\|\mathbf{p} \otimes \mathbf{e}_1\|_{(k)} = \begin{cases} \|\mathbf{p}\|_{(k)} & \text{if } k \in [d], \\ 1 & \text{if } d < k \leq d \cdot m, \end{cases} \tag{3.48}$$

and by writing $k = a_k m + b_k$, where $a_k = \lfloor k/m \rfloor$,

$$\left\| \mathbf{p} \otimes \mathbf{u}^{(m)} \right\|_{(k)} = \sum_{x=1}^{a_k} p_x + b_k \frac{p_{a_k+1}}{m} = \|\mathbf{p}\|_{(a_k)} + b_k \frac{p_{a_k+1}}{m}. \tag{3.49}$$

Using the closed formula for the star conversion distance given in Theorem 3.1.6, we obtain

$$T_\star(\chi \to \chi \otimes \Phi_m) = \max_{k \in [\mathrm{SR}(\chi)]} \left\{ \|\mathbf{p}\|_{(k)} - \|\mathbf{p}\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\}. \tag{3.50}$$

□

As promised, here we derive a necessary and sufficient condition for embezzling families based on the star conversion distance.

Theorem 3.2.4: Characterization of universal embezzling families

A family of pure bipartite states $\{ \chi_n \}_{n\in\mathbb{N}}$ with corresponding Schmidt coefficients $\{ \mathbf{p}^{(n)} \}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if

$$\lim_{n\to\infty} \max_{l\in[A_n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} = 0, \tag{3.51}$$

where $A_n = \lceil \mathrm{SR}(\chi_n)/2 \rceil$.

Proof. Due to Lemma 3.2.2 and Eq. (3.46), $\{ \chi_n \}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if

$$\begin{aligned} 0 &= \lim_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_2) \\ &= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p^{(n)}_{a_k+1} \right\}, \end{aligned} \tag{3.52}$$

where $a_k = \lfloor k/2 \rfloor$ and $k = 2a_k + b_k$. First, we prove the necessary condition. Let $\{ \chi_n \}_{n\in\mathbb{N}}$ be a universal embezzling family. We observe that

$$\begin{aligned} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p^{(n)}_{a_k+1} \right\} &\geq \left\|\mathbf{p}^{(n)}\right\|_{(1)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_1)} - \frac{b_1}{2} p^{(n)}_{a_1+1} \\ &= p^{(n)}_1 - 0 - \frac{1}{2} p^{(n)}_1 > 0. \end{aligned} \tag{3.53}$$

Taking the limit for $n \longrightarrow \infty$ in the expression above, we obtain $\lim_{n\to\infty} p^{(n)}_1 = 0$. Since $p^{(n)}_1$ is the largest Schmidt coefficient, $p^{(n)}_{a_k+1}$ converges to zero too. Since $0 \leq \frac{b_k}{2} \leq \frac{1}{2}$,

$$\lim_{n\to\infty} \frac{b_k}{2} p^{(n)}_{a_k+1} = 0 \tag{3.54}$$

and thus

$$\begin{aligned} 0 &= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p^{(n)}_{a_k+1} \right\} \\ &= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\} \\ &\geq \lim_{n\to\infty} \max_{l\in[\lceil \mathrm{SR}(\chi_n)/2 \rceil]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} \geq 0. \end{aligned} \tag{3.55}$$

For the sufficient condition, we observe that

$$0 = \lim_{n\to\infty} \max_{l\in[A_n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(2l-1)} - \left\|\mathbf{p}^{(n)}\right\|_{(l-1)} \right\} \geq \lim_{n\to\infty} p^{(n)}_1, \tag{3.56}$$

which implies that $\lim_{n\to\infty} p^{(n)}_1 = 0$. Furthermore,

$$\begin{aligned} \lim_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_2) &= \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{2} p^{(n)}_{a_k+1} \right\} \\ &\leq \lim_{n\to\infty} \max_{k\in[\mathrm{SR}(\chi_n)]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \right\} \end{aligned} \tag{3.57}$$

At this point, we have a closer look at

$$\max_{k\in[\mathrm{SR}(\chi_n)]}\left\{\left\|\mathbf{p}^{(n)}\right\|_{(k)}-\left\|\mathbf{p}^{(n)}\right\|_{(a_k)}\right\}. \tag{3.58}$$

If k is even, then $a_k = k/2 = a_{k+1}$, and thus

$$\left\|\mathbf{p}^{(n)}\right\|_{(k)}-\left\|\mathbf{p}^{(n)}\right\|_{(a_k)} \leq \left\|\mathbf{p}^{(n)}\right\|_{(k+1)}-\left\|\mathbf{p}^{(n)}\right\|_{(a_{k+1})}. \tag{3.59}$$

If $\mathrm{SR}(\chi_n)$ is odd, we can thus, without loss of generality, restrict the maximization to run over odd $k \in [\mathrm{SR}(\chi_n)]$. If $\mathrm{SR}(\chi_n)$ is even, we must additionally consider $k = \mathrm{SR}(\chi_n)$. Assume that this is the case: It then holds that

$$\begin{aligned}&\left|\left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n))}-\left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n)/2)}-\left(\left\|\mathbf{p}^{(n)}\right\|_{(\mathrm{SR}(\chi_n)-1)}-\left\|\mathbf{p}^{(n)}\right\|_{\left(\frac{\mathrm{SR}(\chi_n)}{2}-1\right)}\right)\right|\\ &\quad=\left|p^{(n)}_{\mathrm{SR}(\chi_n)}-p^{(n)}_{\mathrm{SR}(\chi_n)/2}\right|\\ &\quad\leq p^{(n)}_1,\end{aligned} \tag{3.60}$$

which vanishes in the limit $n \longrightarrow \infty$. As a consequence,

$$\begin{aligned}\lim_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_2) &\leq \lim_{n\to\infty}\max_{k\in[\mathrm{SR}(\chi_n)]}\left\{\left\|\mathbf{p}^{(n)}\right\|_{(k)}-\left\|\mathbf{p}^{(n)}\right\|_{(a_k)}\right\}\\ &= \lim_{n\to\infty}\max_{k\in[\mathrm{SR}(\chi_n)],k\text{ odd}}\left\{\left\|\mathbf{p}^{(n)}\right\|_{(k)}-\left\|\mathbf{p}^{(n)}\right\|_{(a_k)}\right\}\\ &= \lim_{n\to\infty}\max_{l\in\lceil\mathrm{SR}(\chi_n)/2\rceil}\left\{\left\|\mathbf{p}^{(n)}\right\|_{(2l-1)}-\left\|\mathbf{p}^{(n)}\right\|_{(l-1)}\right\}\\ &= 0.\end{aligned} \tag{3.61}$$

Eq. (3.61) shows that $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is an embezzling family and concludes the proof. □

An important and easy-to-check necessary condition for a universal embezzling family is given in the following Corollary.

Corollary 3.2.5: Necessary condition for universal embezzling family (cf. Ref. [53, Lemma 3 (LO)])

If a family of pure bipartite states $\{\,\chi_n\,\}_{n\in\mathbb{N}}$ is a universal embezzling family, then $\lim_{n\to\infty} p_1^{(n)} = 0$, where $\mathbf{p}^{(n)} \in \mathrm{Prob}^{\downarrow}(d_n)$ are the Schmidt coefficients of χ_n.

3.2.1 Regular Embezzling Families of States

Often, families of states are defined in terms of a positive, monotonic, and continuous function. This motivates the following definition (compare to Ref. [53]).

Definition 3.2.6: Regular family

A family of states $\{ \chi_n \}_{n\in\mathbb{N}}$ is a regular family if there exists a monotonic function $f\colon \mathbb{N} \to (0,\infty)$ such that

$$|\chi_n\rangle = \frac{1}{\sqrt{F_n}} \sum_{x=1}^{n} \sqrt{f(x)}\,|xx\rangle \quad \text{for all} \quad n \in \mathbb{N}, \tag{3.62}$$

where $F_n = \sum_{x=1}^{n} f(x)$.

Note that if $\{ \chi_n \}_{n\in\mathbb{N}}$ is a regular family, then there exists a sequence $\left\{ n_j \right\}_{j\in\mathbb{N}}$ such that the family $\left\{ \chi_{n_j} \right\}_{j\in\mathbb{N}}$ is a universal embezzling family if and only if $\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 0$. However, according to our definition, the family $\left\{ \chi_{n_j} \right\}_{j\in\mathbb{N}}$ is then no longer regular, unless $n_j = j$ for all $j \in \mathbb{N}$.

For the following proofs, it is beneficial to extend f to a monotonic continuous function on $[1,\infty)$. This can often be done trivially by simply extending the domain of f. An example is the van Dam and Hayden family where $f(x) = x^{-1}$. Otherwise, we can extend f by connecting two consecutive points with straight lines. If the function f is multiplied by a constant factor, this does not change the corresponding family of states. Therefore, from now on, we assume for simplicity that $f(1) = 1$.

Let $\{ \chi_n \}_{n\in\mathbb{N}}$ be a regular family of states, and let f be the function associated with it. We define the non-increasing functions f_n, $n \in \mathbb{N}$, by

$$f_n(x) = \begin{cases} f(x) & \text{if } f \text{ is non-increasing,} \\ \frac{f(n+1-x)}{f(n)} & \text{if } f \text{ is non-decreasing.} \end{cases} \tag{3.63}$$

By our extension of f, f_n is also naturally extended to a continuous function $g(x,y)$ such that $g(x,n) = f_n(x)$ via the definition

$$g(x,y) = \begin{cases} f(x) & \text{if } f \text{ is non-increasing,} \\ \frac{f(y+1-x)}{f(y)} & \text{if } f \text{ is non-decreasing.} \end{cases} \tag{3.64}$$

The (by definition non-increasing) Schmidt coefficients of χ_n are therefore given by

$$p_x^{(n)} = \frac{f_n(x)}{F_n}. \tag{3.65}$$

Since $p_1^{(n)} = \frac{1}{F_n}$, we can restate Corollary 3.2.5 for regular families.

Corollary 3.2.7: Necessary condition for regular universal embezzling family (cf. Ref. [53, Lemma 3 (LO)])

If $\{ \chi_n \}_{n\in\mathbb{N}}$ is a regular universal embezzling family, then $\lim_{n\to\infty} F_n = +\infty$, where F_n is defined in Definition 3.2.6.

In the following proposition, we present bounds on the limit of the conversion distance $T_\star(\chi_n \to \chi_n \otimes \Phi_m)$ in terms of the function $g(x,y)$ introduced in Eq. (3.64), which will be of use later.

Proposition 3.2.8: Upper and lower bounds for $T_\star$ in the limit $n \to \infty$

Let $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ be a regular family of states, f be the function associated with it, and g be defined as in Eq. (3.64). If $\lim_{n\to\infty} F_n = \infty$, then

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \geq \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{\, \int_a^{am} g(x,y)\,\mathrm{d}x \,\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}, \tag{3.66}$$

and

$$\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \leq \limsup_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{\, \int_a^{am} g(x,y)\,\mathrm{d}x \,\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}. \tag{3.67}$$

If $\lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m}\{\, \int_a^{am} g(x,y)\,\mathrm{d}x \,\}}{\int_1^y g(x,y)\,\mathrm{d}x}$ exists, this implies that

$$\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{\, \int_a^{am} g(x,y)\,\mathrm{d}x \,\right\}}{\int_1^y g(x,y)\,\mathrm{d}x}. \tag{3.68}$$

Proof. For better readability, we divide the proof into steps and use the notation $G(x,y) = \int_1^x g(t,y)\,\mathrm{d}t$.

Step 1: Starting from Eq. (3.46), we find

$$\begin{aligned}
&\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\
&\quad = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&\quad = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(a_k m)} + \sum_{x=a_k m+1}^{k} p^{(n)}_x - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&\quad = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(a_k m)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} + \frac{\sum_{x=a_k m+1}^{a_k m+b_k} f_n(x) - \frac{b_k}{m} f_n(a_k+1)}{F_n} \right\}
\end{aligned} \tag{3.69}$$

Now we observe that the last contribution in the expression above vanishes because by construction $f_n(x) \leq 1$ and

$$\begin{aligned}
0 &\leq \frac{\sum_{x=a_k m+1}^{a_k m+b_k} f_n(x)}{F_n} \leq \frac{m}{F_n}, \\
0 &\leq \frac{\frac{b_k}{m} f_n(a_k+1)}{F_n} \leq \frac{1}{F_n}.
\end{aligned} \tag{3.70}$$

Moreover, by assumption, the right-hand sides converge to zero in the limit $n \to \infty$, and this ensures that the limit inferior is additive. As a result,

$$\begin{aligned}
\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &= \liminf_{n\to\infty} \max_{a\in\{\,0,\ldots,\lfloor n/m\rfloor\,\}} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\} \\
&= \liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\},
\end{aligned} \tag{3.71}$$

where we used that $\left\|\mathbf{p}^{(n)}\right\|_{(0)} = 0$.

Step 2: We start with finding bounds for $\sum_{x=1}^{n} f_n(x) - \int_1^n f_n(x)\,\mathrm{d}x$. To this end we observe that since $f_n(x)$ is by construction non-increasing,

$$\int_1^n f_n(x)\,\mathrm{d}x \le \int_1^n f_n(x)\,\mathrm{d}x + f_n(n) \le \sum_{x=1}^{n} f_n(x) \le f_n(1) + \int_1^n f_n(x)\,\mathrm{d}x. \tag{3.72}$$

By subtracting $\int_1^n f_n(x)\,\mathrm{d}x$ from all terms, the desired bounds follow:

$$0 \le \sum_{x=1}^{n} f_n(x) - \int_1^n f_n(x)\,\mathrm{d}x \le f_n(1) = 1. \tag{3.73}$$

Dividing by $F_n = \sum_{x=1}^{n} f_n(x)$ and taking the limit $n \to \infty$, we obtain $\lim_{n\to\infty} 1 - \frac{\int_1^n f_n(x)\,\mathrm{d}x}{\sum_{x=1}^{n} f_n(x)} = 0$. Since $f_n(x) = g(x,n)$, the limit can be rewritten as $\lim_{n\to\infty} \frac{G(n,n)}{F_n} = 1$.

Step 3: In this step, we want to show that

$$\begin{aligned}\liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\} &= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} \\ &= \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{\int_1^n f_n(x)\,\mathrm{d}x}.\end{aligned} \tag{3.74}$$

Analogously to the previous step, one can obtain the following bounds for $\frac{\sum_{x=a+1}^{am} f_n(x)}{F_n}$:

$$\frac{\int_{a+1}^{am} f_n(x)\,\mathrm{d}x}{F_n} \le \frac{\sum_{x=a+1}^{am} f_n(x)}{F_n} \le \frac{1 + \int_{a+1}^{am} f_n(x)\,\mathrm{d}x}{F_n}, \tag{3.75}$$

and after taking the maximum over $a \in [\lfloor n/m\rfloor]$ and the limit inferior the bounds become

$$\begin{aligned}\liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n} &\le \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} \\ &\le \liminf_{n\to\infty} \frac{1 + \max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}.\end{aligned} \tag{3.76}$$

We can rewrite the last expression as

$$\begin{aligned}&\liminf_{n\to\infty} \left(\frac{1}{F_n} + \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n} \right) \\ &\quad = \liminf_{n\to\infty} \left(\frac{1}{F_n} \right) + \liminf_{n\to\infty} \left(\frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n} \right) \\ &\quad = \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}.\end{aligned} \tag{3.77}$$

By combining Eq. (3.76) with Eq. (3.77), one obtains

$$\liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \sum_{x=a+1}^{am} f_n(x) \right\}}{F_n} = \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \left\{ \int_{a+1}^{am} f_n(x)\,\mathrm{d}x \right\}}{F_n}. \tag{3.78}$$

This is equivalent to

$$\liminf_{n\to\infty} \max_{a\in[\lfloor n/m\rfloor]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(am)} - \left\|\mathbf{p}^{(n)}\right\|_{(a)} \right\} = \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \{\, G(am,n) - G(a+1,n) \,\}}{G(n,n)}, \tag{3.79}$$

where we replaced F_n with $G(n,n)$, which we can do according to step 2.

Step 4: Let c_n be the value that maximizes $\max_{1\le a\le n/m} \{\, G(am,n) - G(a+1,n) \,\}$ and let c'_n be the natural number satisfying $c'_n \le c_n < c'_n + 1$. This implies

$$\begin{aligned} G(c_n m,n) - G(c_n+1,n) &= \int_{c_n+1}^{c_n m} g(t,n)\,\mathrm{d}t \\ &\le \int_{c'_n+1}^{c'_n m} g(t,n)\,\mathrm{d}t + \int_{c'_n m}^{(c'_n+1)m} g(t,n)\,\mathrm{d}t \\ &\le \int_{c'_n+1}^{c'_n m} g(t,n)\,\mathrm{d}t + m. \end{aligned} \tag{3.80}$$

Dividing by $G(n,n)$, taking the limit inferior, and arguing as in the previous step, we obtain

$$\begin{aligned} &\liminf_{n\to\infty} \frac{\max_{1\le a\le n/m} \{\, G(am,n) - G(a+1,n) \,\}}{G(n,n)} \\ &\quad= \liminf_{n\to\infty} \frac{G(c_n m,n) - G(c_n+1,n)}{G(n,n)} \\ &\quad\le \liminf_{n\to\infty} \frac{G(c'_n m,n) - G(c'_n+1,n) + m}{G(n,n)} \\ &\quad= \liminf_{n\to\infty} \frac{G(c'_n m,n) - G(c'_n+1,n)}{G(n,n)} \\ &\quad\le \liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \{\, G(am,n) - G(a+1,n) \,\}}{G(n,n)}. \end{aligned} \tag{3.81}$$

The reverse inequality follows from the fact that we maximize over a larger set. This proves that

$$\begin{aligned} &\liminf_{n\to\infty} \frac{\max_{a\in[\lfloor n/m\rfloor]} \{\, G(am,n) - G(a+1,n) \,\}}{G(n,n)} \\ &\quad= \liminf_{n\to\infty} \frac{\max_{1\le a\le n/m} \{\, G(am,n) - G(a+1,n) \,\}}{G(n,n)}. \end{aligned} \tag{3.82}$$

Step 5: Combining the results of the previous steps, and in particular Eq. (3.71), Eq. (3.79), and Eq. (3.82), we have proven that

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n\to\infty} \frac{\max_{1\le a\le n/m} \left\{ \int_{a+1}^{am} g(x,n)\,\mathrm{d}x \right\}}{\int_1^n g(x,n)\,\mathrm{d}x}. \tag{3.83}$$

In the above equation, we can replace $\int_{a+1}^{am} g(x,y)\,\mathrm{d}x$ with $\int_a^{am} g(x,y)\,\mathrm{d}x$ because the difference of the two integrals is finite and divided by a diverging term. Thus

$$\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = \liminf_{n\to\infty} \frac{\max_{1\le a\le n/m} \left\{ \int_a^{am} g(x,y)\,\mathrm{d}x \right\}}{\int_1^n g(x,y)\,\mathrm{d}x}. \tag{3.84}$$

The expression inside the lim inf on the right-hand side of Eq. (3.84) is a function of $n \in \mathbb{N} \setminus \{ 1 \}$ and we call it $M(n)$,

$$M(n) = \frac{\max_{1 \le a \le n/m} \left\{ \int_a^{am} g(x,n)\, \mathrm{d}x \right\}}{\int_1^n g(x,n)\, \mathrm{d}x}. \tag{3.85}$$

We extend this function to real numbers $y > 1$ as follows:

$$M(y) = \frac{\max_{1 \le a \le y/m} \left\{ \int_a^{am} g(x,y)\, \mathrm{d}x \right\}}{\int_1^y g(x,y)\, \mathrm{d}x}. \tag{3.86}$$

Since the natural numbers are a subset of the real numbers,

$$\liminf_{n\to\infty} M(n) \ge \liminf_{y\to\infty} M(y), \tag{3.87}$$

which proves Eq. (3.66). Furthermore, Steps 1-4 can be repeated with the same arguments for the lim sup, and they imply

$$\limsup_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) = \limsup_{n\to\infty} M(n) \le \limsup_{y\to\infty} M(y). \tag{3.88}$$

This proves Eq. (3.67).

In the following, we assume that $\lim_{y\to\infty} M(y)$ exists. This implies

$$\lim_{y\to\infty} M(y) = \liminf_{y\to\infty} M(y) = \limsup_{y\to\infty} M(y). \tag{3.89}$$

From Eq. (3.87), Eq. (3.88), and Eq. (3.89) we obtain

$$\lim_{y\to\infty} M(y) \le \liminf_{n\to\infty} M(n) \le \limsup_{n\to\infty} M(n) \le \lim_{y\to\infty} M(y). \tag{3.90}$$

Therefore, $\lim_{n\to\infty} M(n)$ exists and is equal to $\lim_{y\to\infty} M(y)$. From Eq. (3.84) and Eq. (3.88) we derive

$$\lim_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) = \lim_{n\to\infty} M(n) = \lim_{y\to\infty} M(y). \tag{3.91}$$

This proves Eq. (3.68) and concludes the proof.

□

Corollary 3.2.7 combined with Proposition 3.2.8 leads to the following characterization of embezzling families.

Corollary 3.2.9: Necessary and sufficient conditions for regular universal embezzling families

Let $\{ \chi_n \}_{n\in\mathbb{N}}$ be a regular family of states, f be the function associated with it, and g be defined as in Eq. (3.64). If

$$\lim_{y\to\infty} \frac{\max_{1 \le a \le y/m} \left\{ \int_a^{am} g(x,y)\, \mathrm{d}x \right\}}{\int_1^y g(x,y)\, \mathrm{d}x} \tag{3.92}$$

exists, then the family of states $\{ \chi_n \}_{n\in\mathbb{N}}$ is an universal embezzling family if and only if $\lim_{n\to\infty} F_n = \infty$ and

$$\lim_{y\to\infty} \frac{\max_{1 \le a \le y/m} \left\{ \int_a^{am} g(x,y)\, \mathrm{d}x \right\}}{\int_1^y g(x,y)\, \mathrm{d}x} = 0. \tag{3.93}$$

In addition, thanks to Proposition 3.2.8, the task of determining if a family of states is a universal embezzling family is converted into an optimization problem. It is enough to find the maximum of the differentiable function $G_y(a) = \int_a^{am} g(x, y)\,\mathrm{d}x$ on $1 \le a \le y/m$, which is easily done with the help of its derivative

$$mg(am, y) - g(a, y), \quad 1 \le a \le y/m. \tag{3.94}$$

3.2.2 Generalization of the van Dam and Hayden Family

The universal embezzling family $\{ \chi_n \}_{n\in\mathbb{N}}$ introduced by van Dam and Hayden consists of the states

$$|\chi_n\rangle = \frac{1}{\sqrt{H_n}} \sum_{x=1}^{n} \sqrt{x^{-1}}\, |xx\rangle\,, \tag{3.95}$$

where $H_n = \sum_{x=1}^{n} x^{-1}$ is the n-th harmonic number. We generalize this family of states as follows: For every $\alpha \in \mathbb{R}$, we introduce the family $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$, where

$$\left|\chi_n^{(\alpha)}\right\rangle = \frac{1}{\sqrt{H_n^{(-\alpha)}}} \sum_{x=1}^{n} \sqrt{x^{\alpha}}\, |xx\rangle \tag{3.96}$$

and $H_n^{(-\alpha)} = \sum_{x=1}^{n} x^{\alpha}$ is the n-th generalized harmonic number. Note that the families of states $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$ are regular families of states, and the van Dam and Hayden family is recovered for $\alpha = -1$. The corresponding Schmidt coefficients, as per our convention, arranged in non-increasing order, are

$$\mathbf{p}^{(n|\alpha)} = \begin{cases} \frac{1}{H_n^{(-\alpha)}}(1, 2^{\alpha}, \dots, n^{\alpha}) & \text{if } \alpha \le 0, \\ \frac{1}{H_n^{(-\alpha)}}(n^{\alpha}, \dots, 2^{\alpha}, 1) & \text{if } \alpha > 0. \end{cases} \tag{3.97}$$

In the remaining part of this Subsection, we show that the family of states $\left\{ \chi_n^{(\alpha)} \right\}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $\alpha = -1$ and derive bounds on the star conversion distance. For a visual representation of the results of this Subsection see Figure 3.1.

Case $\alpha < -1$. For $\alpha < -1$, $\lim_{n\to\infty} H_n^{(-\alpha)}$ is finite, thus the regular family of states $\{ \chi_n^{\alpha} \}_{n\in\mathbb{N}}$ is not a universal embezzling family (see Corollary 3.2.7). The largest Schmidt coefficient provides a lower bound on the star conversion distance: Using Eq. (3.35), we obtain

$$\begin{aligned} T_\star(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) &= \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n|\alpha)}\right\|_{(k)} - \left\|\mathbf{p}^{(n|\alpha)}\right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n|\alpha)} \right\} \\ &\ge p_1^{(n|\alpha)} \left(1 - \frac{1}{m}\right) \\ &= \frac{1}{H_n^{(-\alpha)}} \left(1 - \frac{1}{m}\right), \end{aligned} \tag{3.98}$$

where the inequality follows from choosing $k = 1$. By taking the limit, we obtain

$$\liminf_{n\to\infty} T_\star(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) \ge \liminf_{n\to\infty} \frac{1}{H_n^{(-\alpha)}} \left(1 - \frac{1}{m}\right) = \frac{1}{\zeta(-\alpha)} \left(1 - \frac{1}{m}\right), \tag{3.99}$$

where $\zeta(-\alpha) = \sum_{x=1}^{\infty} x^{\alpha}$ is the Riemann Zeta function. To obtain an upper bound, we observe that, since $\alpha < -1$,

$$\sum_{x=a+1}^{(a+1)m} x^{\alpha} \leq (a+1)^{\alpha} + \int_{a+1}^{(a+1)m} x^{\alpha}\,\mathrm{d}x = (a+1)^{\alpha} + \frac{(a+1)^{\alpha+1}}{\alpha+1}\left(m^{\alpha+1}-1\right). \tag{3.100}$$

The right-hand side is decreasing in a, and thus

$$\begin{aligned}
&\max_{k\in[n]}\left\{\left\|\mathbf{p}^{(n|\alpha)}\right\|_{(k)} - \left\|\mathbf{p}^{(n|\alpha)}\right\|_{(a_k)} - \frac{b_k}{m}p^{(n|\alpha)}_{a_k+1}\right\} \\
&\quad\leq \frac{1}{H_n^{(-\alpha)}}\max_{k\in[n]}\left\{\sum_{x=a_k+1}^{k} x^{\alpha}\right\} \\
&\quad\leq \frac{1}{H_n^{(-\alpha)}}\max_{a\in\{0,\dots,\lfloor n/m\rfloor\}}\left\{\sum_{x=a+1}^{(a+1)m} x^{\alpha}\right\} \\
&\quad\leq \frac{1}{H_n^{(-\alpha)}}\max_{a\in\{0,\dots,\lfloor n/m\rfloor\}}\left\{(a+1)^{\alpha} + \frac{(a+1)^{\alpha+1}}{\alpha+1}\left(m^{\alpha+1}-1\right)\right\} \\
&\quad= \frac{1}{H_n^{(-\alpha)}}\left(1 + \frac{m^{\alpha+1}-1}{\alpha+1}\right).
\end{aligned} \tag{3.101}$$

Taking the limit $n \to \infty$, and considering that the conversion distance is by definition smaller than one, we obtain

$$\limsup_{n\to\infty} T_{\star}(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) \leq \min\left\{1, \frac{1}{\zeta(-\alpha)}\left(1 + \frac{m^{\alpha+1}-1}{\alpha+1}\right)\right\}. \tag{3.102}$$

Case $\alpha = -1$. This is the van Dam and Hayden family. Analogously to the previous case, we observe that

$$\begin{aligned}
\sum_{x=a+1}^{(a+1)m} x^{-1} &\leq (a+1)^{-1} + \int_{a+1}^{(a+1)m} x^{-1}\,\mathrm{d}x \\
&= (a+1)^{-1} + \ln m \\
&\leq 1 + \ln m.
\end{aligned} \tag{3.103}$$

Since $H_n^{(1)} = H_n$ diverges, by following the same steps, we have

$$\lim_{n\to\infty} T_{\star}(\chi_n^{(-1)} \to \chi_n^{(-1)} \otimes \Phi_m) \leq \lim_{n\to\infty}\frac{1+\ln m}{H_n} = 0. \tag{3.104}$$

Thus, the limit for the conversion distance is zero, as expected for the van Dam and Hayden family.

Case $-1 < \alpha < 0$. In this scenario $H_n^{(-\alpha)}$ diverges when $n \to \infty$, and we can use the results of Section 3.2.1. Here $g(x,y) = x^{\alpha}$, and the derivative with respect to a of the function $\int_a^{am} g(x,y)\,\mathrm{d}x$ is

$$mg(am,y) - g(a,y) = a^{\alpha}(m^{\alpha+1}-1), \tag{3.105}$$

which is positive in the domain $1 \leq a \leq y/m$. This implies that the function $\int_a^{am} g(x,y)\,\mathrm{d}x$ is non-decreasing and the maximum

$$\max_{1\leq a\leq y/m}\left\{\int_a^{am} g(x,y)\,\mathrm{d}x\right\} \tag{3.106}$$

is obtained for $a = y/m$. Due to Proposition 3.2.8,

$$\lim_{n\to\infty} T_\star(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) = \lim_{y\to\infty} \frac{\int_{y/m}^{y} x^\alpha \,\mathrm{d}x}{\int_1^y x^\alpha \,\mathrm{d}x} = 1 - \frac{1}{m^{\alpha+1}}. \tag{3.107}$$

Case $\alpha = 0$. Here we have

$$p_x^{(n)} = \frac{1}{n}, \quad (\mathbf{p}^{(n)} \otimes \mathbf{u}^{(m)})_x = \frac{1}{nm}. \tag{3.108}$$

We can compute the star conversion distance directly using Eq. (3.35) and obtain

$$T_\star(\chi_n^{(0)} \to \chi_n^{(0)} \otimes \Phi_m) = \max_{k\in[n]} \left\{ \frac{k}{n} - \frac{k}{nm} \right\} = 1 - \frac{1}{m}. \tag{3.109}$$

Case $\alpha > 0$. Also here $H_n^{(-\alpha)}$ diverges for $n \to \infty$ and we can use the results of Section 3.2.1. The function $g(x, y)$ is defined as $g(x, y) = (y + 1 - x)^\alpha / y^\alpha$, because x^α is increasing. From this follows

$$\begin{aligned} mg(am, y) - g(a, y) \geq 0 &\Leftrightarrow m(y + 1 - am)^\alpha \geq (y + 1 - a)^\alpha \\ &\Leftrightarrow a \leq \frac{(y+1)\left(m^{\frac{1}{\alpha}} - 1\right)}{m^{1+\frac{1}{\alpha}} - 1} := a_{\max}. \end{aligned} \tag{3.110}$$

Since $a_{\max}$ increases linearly with y and $a_{\max} \leq \frac{y}{m}$ is equivalent to $y \geq \frac{m(m^{1/\alpha}-1)}{m-1}$, for large enough y, the value $a_{\max}$ belongs to the interval $[1, y/m]$ and is a global maximum. Substituting $a_{\max}$ into Proposition 3.2.8, we obtain

$$\begin{aligned} \lim_{n\to\infty} T_\star(\chi_n^{(\alpha)} \to \chi_n^{(\alpha)} \otimes \Phi_m) &= \lim_{y\to\infty} \frac{\int_{a_{\max}}^{a_{\max} m} \frac{(y+1-x)^\alpha}{y^\alpha} \,\mathrm{d}x}{\int_1^y \frac{(y+1-x)^\alpha}{y^\alpha} \,\mathrm{d}x} \\ &= (m-1) \left(\frac{m-1}{m^{1+\frac{1}{\alpha}} - 1} \right)^\alpha. \end{aligned} \tag{3.111}$$

This completes the study of regular families of states defined by $f(x) = x^\alpha$. The star conversion distance vanishes only for $\alpha = -1$. Thus, the only universal embezzling family of this form is the one introduced by van Dam and Hayden. Furthermore, exact values for the limit of the star conversion distance for $\alpha \geq -1$ and lower and upper bounds for $\alpha < -1$ were provided.

3.2.3 Uniqueness of the van Dam and Hayden Embezzling Family

In this Subsection, we provide further results on the uniqueness of the van Dam and Hayden embezzling family. Given a regular family of states $\{ \chi_n \}_{n\in\mathbb{N}}$, the asymptotic behaviour of the function f associated with it is relevant to determine whether $\{ \chi_n \}_{n\in\mathbb{N}}$ is a universal embezzling family or not (e.g., see Corollary 3.2.7). We thus use the notations little-ω and little-o (see, e.g., Ref. [168]) to describe asymptotic relations between two functions $g, h\colon [1, \infty) \to (0, \infty)$,

$$h \in \omega(g) \Leftrightarrow \lim_{x\to\infty} \frac{h(x)}{g(x)} = +\infty, \quad h \in o(g) \Leftrightarrow \lim_{x\to\infty} \frac{h(x)}{g(x)} = 0. \tag{3.112}$$

Before we state the results about the uniqueness of the van Dam and Hayden embezzling family, we prove the following Lemma, which is a direct consequence of Theorem 2 in Ref. [169], which we restate here to

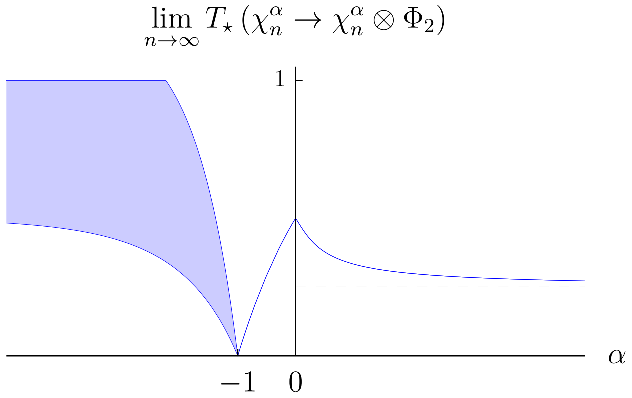


Figure 3.1: The family of states $\left\{ \chi_n^\alpha \right\}_{n\in\mathbb{N}}$ is an embezzling family if and only if $\alpha = -1$. This can be seen in the above plot showing the analytically derived exact value of $\lim_{n\to\infty} T_\star(\chi_n^\alpha \to \chi_n^\alpha \otimes \Phi_2)$ for $\alpha \geq -1$ and lower and upper bounds for $\alpha < -1$.

improve readability: Let μ be a measure on the real line $\mathbb{R}$, and let f_i, g_i $(i = 1, 2)$ be four Borel-measurable functions: $\mathbb{R} \to \mathbb{R}$ such that $f_2 \geq 0$ and $g_2 \geq 0$, and $\int \left|f_i g_j\right| \mathrm{d}\mu < \infty$ $(i, j = 1, 2)$. If f_1/f_2 and g_1/g_2 are monotonic in the same direction, then

$$\int f_1 g_1 \,\mathrm{d}\mu \int f_2 g_2 \,\mathrm{d}\mu \geq \int f_1 g_2 \,\mathrm{d}\mu \int f_2 g_1 \,\mathrm{d}\mu. \tag{3.113}$$

Lemma 3.2.10: An inequality about integrals of continuous functions with non-increasing ratio

Let f, g be continuous, positive functions on $[a, b]$ such that $f(x)/g(x)$ is a non-decreasing function on $[a, b]$. Then, for every x_1 such that $a < x_1 < b$,

$$\int_a^b f(x)\,\mathrm{d}x \int_a^{x_1} g(x)\,\mathrm{d}x \geq \int_a^{x_1} f(x)\,\mathrm{d}x \int_a^b g(x)\,\mathrm{d}x. \tag{3.114}$$

Proof. Let $\tilde{h}(x) = 1$ and $h(x) = \chi_{[a,x_1]}(x)$, where $\chi_S(x)$ is the characteristic function of the set S. Furthermore, let

$$h_k(x) = \begin{cases} 1 & \text{if } a \leq x \leq x_1, \\ e^{-k(x-x_1)} & \text{if } x_1 < x \leq b. \end{cases} \tag{3.115}$$

The sequence of functions $\{ h_k \}_{k\in\mathbb{N}}$ converges to $h(x)$ and satisfies $0 \leq h_k(x) \leq 1$. Due to the dominated convergence theorem (see, e.g., Ref. [170]), we have

$$\lim_{k\to\infty} \int_a^b f(x) h_k(x)\,\mathrm{d}x = \int_a^{x_1} f(x)\,\mathrm{d}x. \tag{3.116}$$

The same result holds for g.

Since both $f(x)/g(x)$ and $\tilde{h}(x)/h_k(x)$ are non-decreasing, we can apply Theorem 2 of Ref. [169] to find

$$\int_a^b f(x)\tilde{h}(x)\,\mathrm{d}x \int_a^b g(x) h_k(x)\,\mathrm{d}x \geq \int_a^b f(x) h_k(x)\,\mathrm{d}x \int_a^b g(x)\tilde{h}(x)\,\mathrm{d}x. \tag{3.117}$$

After taking the limit $k \to \infty$ on both sides, we obtain

$$\int_a^b f(x)\,\mathrm{d}x \int_a^{x_1} g(x)\,\mathrm{d}x \geq \int_a^{x_1} f(x)\,\mathrm{d}x \int_a^b g(x)\,\mathrm{d}x, \tag{3.118}$$

which finishes the proof. □

With this Lemma at hand, we are ready to present the promised results concerning the uniqueness of the van Dam and Hayden embezzling family.

Theorem 3.2.11: Necessary and sufficient conditions for regular universal embezzling families

Let f be a positive non-increasing function such that $f(x)/x^\alpha$ is asymptotically monotonic for all $\alpha \in \mathbb{R}$ and let $\{\chi_n\}_{n\in\mathbb{N}}$ be the regular family of states associated to it (see Definition 3.2.6). Then $\{\chi_n\}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})\ \forall \varepsilon > 0$ and $\sum_{x=1}^{\infty} f(x) = \infty$. Furthermore, if $f \notin \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for at least one $\varepsilon > 0$, then $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$, where $\left\{n_j\right\}_{j\in\mathbb{N}}$ is any sequence of natural numbers, is not a universal embezzling family.

Proof. Before we start, we notice that if $\lim_{j\to\infty} n_j = J < \infty$, then $\lim_{j\to\infty} F_{n_j} = \sum_{x=1}^{J} f(x) < \infty$, which implies that the family of states $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is not an embezzling family (see Corollary 3.2.7). In this proof we will thus assume, w.l.o.g., that $\lim_{j\to\infty} n_j = \infty$.

Necessary condition — Let $\{\chi_n\}_{n\in\mathbb{N}}$ be a universal embezzling family with the corresponding function f satisfying the assumptions above. We already proved in Corollary 3.2.7 that if $\{\chi_n\}_{n\in\mathbb{N}}$ is a universal embezzling family, then $\sum_{x=1}^{\infty} f(x) = +\infty$. To show the remainder, let

$$R = \left\{ \alpha \in \mathbb{R} \,\middle|\, \lim_{x\to\infty} \frac{f(x)}{x^\alpha} = 0 \right\}, \quad L = \left\{ \alpha \in \mathbb{R} \,\middle|\, \lim_{x\to\infty} \frac{f(x)}{x^\alpha} = +\infty \right\}. \tag{3.119}$$

Since $\lim_{x\to\infty} f(x)/x^\alpha = 0$ for $\alpha > 0$, we have that $R \neq \emptyset$. Now we prove, by contradiction, that $L \neq \emptyset$ as well. Let us assume that L is empty. This implies that $\lim_{x\to\infty} f(x)/x^\alpha = l_\alpha < \infty$ for all α. If there exists an α such that $l_\alpha \neq 0$, then $\lim_{x\to\infty} f(x)/x^{\alpha-1} = l_\alpha(\lim_{x\to\infty} x) = +\infty$. Thus, $\alpha - 1 \in L$, and $L \neq \emptyset$, leading to the desired contradiction. If instead $\lim_{x\to\infty} f(x)/x^\alpha = 0$ for all α, then $\lim_{x\to\infty} f(x)/x^{-2} = 0$. This implies that $f(x)/x^{-2}$ converges monotonically to zero for large x, i.e., there exists an N such that $f(x) < x^{-2}$ for $x > N$. Thus, $\sum_{x=N}^{\infty} f(x) < \sum_{x=N}^{\infty} x^{-2} < \infty$. According to Corollary 3.2.7, this contradicts the hypothesis that the family under consideration is universally embezzling. We have therefore shown that $L \neq \emptyset \neq R$.

The next step is to prove that $\inf R = \sup L$. From the definition of R and L, it follows that $\inf R \geq \sup L$. Let us assume that $\inf R > \sup L$, i.e., that there exists an $\alpha \in \mathbb{R}$ such that $\sup L < \alpha < \inf R$. Since $\alpha \notin L \cup R$, there exists a positive real number l such that $\lim_{x\to\infty} f(x)/x^\alpha = l$. Now pick α_1 such that $\sup L < \alpha_1 < \alpha < \inf R$. We observe that $\lim_{x\to\infty} f(x)/x^{\alpha_1} = l(\lim_{x\to\infty} x^\alpha/x^{\alpha_1}) = +\infty$, thus $\alpha_1 \in L$. This is in contradiction to the choice $\alpha_1 > \sup L$. We therefore showed that $\sup L = \inf R$.

So far, we have shown that if $\{\chi_n\}$ is a universal embezzling family satisfying our assumptions, then there exists a unique α such that $f(x) \in o(x^{\alpha+\varepsilon}) \cap \omega(x^{\alpha-\varepsilon})$ for every $\varepsilon > 0$. What is left to show is that if $\alpha \neq -1$, then the family of states corresponding to f cannot be universally embezzling. From the above discussion, we know that $\{ \alpha > 0 \} \subseteq R$, thus we can focus on $\alpha \leq 0$ and split our discussion into two scenarios, $\alpha < -1$ and $-1 < \alpha \leq 0$.

If $\alpha < -1$, then there exists an $\varepsilon > 0$ such that $\alpha + \varepsilon < -1$. Since $f \in o(x^{\alpha+\varepsilon})$ and $f(x)/x^{\alpha+\varepsilon}$ converges monotonically to zero for large x, there exists an N such that $f(x) < x^{\alpha+\varepsilon}$ for $x > N$. This implies that $\sum_{x=N}^{\infty} f(x) < \sum_{x=N}^{\infty} x^{\alpha+\varepsilon} < \infty$ and therefore $\sum_{x=1}^{\infty} f(x) < \infty$, which, according to Corollary 3.2.7, contradicts the hypothesis that $\{ \chi_n \}$ is a universal embezzling family. Let $\{ n_j \}_{j\in\mathbb{N}}$ be any sequence of natural number such that $\lim_{j\to\infty} n_j = \infty$. Then $\lim_{j\to\infty} \sum_{x=1}^{n_j} f(x) = \sum_{x=1}^{\infty} f(x) < \infty$. This implies, again due to Corollary 3.2.7, that $\{ \chi_{n_j} \}_{j\in\mathbb{N}}$ is not a universal embezzling family.

If $-1 < \alpha \leq 0$, then there exists an $\varepsilon > 0$ such that $\alpha - \varepsilon > -1$. We notice that $f(x)/x^{\alpha-\varepsilon}$ diverges to infinity for large x, because $f \in \omega(x^{\alpha-\varepsilon})$. This implies that there exists an N such that $f(x) \geq x^{\alpha-\varepsilon}$ for $x > N$ and therefore $\sum_{x=N}^{\infty} f(x) \geq \sum_{x=N}^{\infty} x^{\alpha-\varepsilon} = \infty$, which allows us to use the results of Section 3.2.1. Since $f(x)/x^{\alpha-\varepsilon}$ is non-decreasing for $x > N$, for all y such that $y/m > N$, we can use Lemma 3.2.10 to obtain

$$\int_N^y f(x)\,\mathrm{d}x \int_N^{y/m} x^{\alpha-\varepsilon}\,\mathrm{d}x \geq \int_N^{y/m} f(x)\,\mathrm{d}x \int_N^y x^{\alpha-\varepsilon}\,\mathrm{d}x. \tag{3.120}$$

This implies that

$$\frac{(y/m)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}}{(y)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}} \geq \frac{\int_N^{y/m} f(x)\,\mathrm{d}x}{\int_N^y f(x)\,\mathrm{d}x}. \tag{3.121}$$

Next we introduce $\delta := \alpha - \varepsilon + 1 > 0$ and take on both sides the limit inferior $y \longrightarrow \infty$, leading to

$$\begin{aligned} \frac{1}{m^\delta} &= \liminf_{y\to\infty} \frac{(y/m)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}}{(y)^{\alpha-\varepsilon+1} - N^{\alpha-\varepsilon+1}} \\ &\geq \liminf_{y\to\infty} \frac{\int_N^{y/m} f(x)\,\mathrm{d}x}{\int_N^y f(x)\,\mathrm{d}x} \\ &= \liminf_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x}, \end{aligned} \tag{3.122}$$

where in the last equality, we used that $\int_1^N f(x)\,\mathrm{d}x$ is finite, whilst $\int_N^\infty f(x)\,\mathrm{d}x$ diverges. As the last step, we observe that according to Proposition 3.2.8,

$$\begin{aligned} \liminf_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) &\geq \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \{ \int_a^{am} f(x)\,\mathrm{d}x \}}{\int_1^y f(x)\,\mathrm{d}x} \\ &\geq \liminf_{y\to\infty} \frac{\int_{y/m}^y f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \\ &\geq 1 - \frac{1}{m^\delta} \\ &> 0. \end{aligned} \tag{3.123}$$

Using the comment after Definition 3.2.6, we obtain that for $-1 < \alpha \leq 0$, there are no sequences $\{ n_j \}_{j\in\mathbb{N}}$ such that $\{ \chi_{n_j} \}_{j\in\mathbb{N}}$ is a universal embezzling family, contradicting the hypothesis. Once we combine the results for $\alpha < -1$ and $-1 < \alpha \leq 0$, we have that if $\{ \chi_n \}$ is a universal embezzling family satisfying our assumptions, $f(x) \in o(x^{-1+\varepsilon}) \cap \omega(x^{-1-\varepsilon})$ for all $\varepsilon > 0$. We have also shown that if $f \notin o(x^{-1+\varepsilon}) \cap \omega(x^{-1-\varepsilon})$ for at least one $\varepsilon > 0$, then there are no sequences $\{ n_j \}_{j \in \mathbb{N}}$ such that the family $\{ \chi_{n_j} \}_{j \in \mathbb{N}}$ is a universal embezzling family. This concludes the first part of the proof.

Sufficient condition — Let f be a function satisfying our assumptions. Since $f(x)/x^{-1}$ is asymptotically monotonic by assumption, $\lim_{x\to\infty} f(x)/x^{-1}$ exists in $[0, +\infty]$. A priori, it can be either 0, $0 < l \in \mathbb{R}$, or $+\infty$. If $\lim_{x\to\infty} f(x)/x^{-1} = l > 0$, then by definition of the limit, for any $l > \tilde{\varepsilon} > 0$ there exists an N such that $\left| \frac{f(x)}{x^{-1}} - l \right| < \tilde{\varepsilon}$ for all $x > N$. This is equivalent to $(l - \tilde{\varepsilon})x^{-1} < f(x) < (l + \tilde{\varepsilon})x^{-1}$ for all $x > N$. Since $\sum_{x=N}^{\infty} f(x) \geq (l - \tilde{\varepsilon}) \sum_{x=N}^{\infty} x^{-1} = +\infty$, we can compute the star conversion distance using Proposition 3.2.8, and find

$$\begin{aligned}
&\lim_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\
&\quad = \lim_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\quad \leq \lim_{y\to\infty} \frac{\max_{1\leq a< N} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\} + \max_{N\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\quad = \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\quad \leq \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_N^y f(x)\,\mathrm{d}x} \\
&\quad \leq \frac{l+\tilde{\varepsilon}}{l-\tilde{\varepsilon}} \lim_{y\to\infty} \frac{\max_{N\leq a\leq y/m} \left\{ \int_a^{am} x^{-1}\,\mathrm{d}x \right\}}{\int_N^y x^{-1}\,\mathrm{d}x} \\
&\quad = \frac{l+\tilde{\varepsilon}}{l-\tilde{\varepsilon}} \lim_{y\to\infty} \frac{\log m}{\log y - \log N} \\
&\quad = 0.
\end{aligned} \tag{3.124}$$

Thus if $\lim_{x\to\infty} f(x)/x^{-1} = l \neq 0$, then $\{ \chi_n \}$ is a universal embezzling family.

Suppose now $\lim_{x\to\infty} f(x)/x^{-1} = 0$. Since $f(x)/x^{-1}$ is non-increasing for large x, there exists an $\tilde{N}$ such that $f(x) < x^{-1}$ for all $x > \tilde{N}$. For any fixed $\varepsilon > 0$, by hypothesis, $\lim_{x\to\infty} f(x)/x^{-1-\varepsilon} = +\infty$. Furthermore, since by assumption $\int_1^\infty f(x)\,\mathrm{d}x = \infty$, while $\int_1^\infty x^{-1-\varepsilon}\,\mathrm{d}x < \infty$, there exists an $\tilde{N}_\varepsilon$, such that $\int_1^y f(x)\,\mathrm{d}x \geq \int_1^y x^{-1-\varepsilon}\,\mathrm{d}x$ for $y > \tilde{N}_\varepsilon$. Let $N_\varepsilon = \max\left\{ \tilde{N}, \tilde{N}_\varepsilon \right\}$. Using again Proposition 3.2.8 and performing the same steps as in Eq. (3.124), we obtain

$$\begin{aligned}
\limsup_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\leq \limsup_{y\to\infty} \frac{\max_{N_\varepsilon\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\leq \limsup_{y\to\infty} \frac{\max_{N_\varepsilon\leq a\leq y/m} \left\{ \int_a^{am} x^{-1}\,\mathrm{d}x \right\}}{\int_1^y x^{-1-\varepsilon}\,\mathrm{d}x} \\
&= \log m \lim_{y\to\infty} \frac{\varepsilon}{1 - y^{-\varepsilon}} \\
&= \varepsilon \log m.
\end{aligned} \tag{3.125}$$

Since this is true for all $\varepsilon > 0$,

$$\limsup_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) \leq \lim_{\varepsilon\to 0} \varepsilon \log m = 0. \tag{3.126}$$

This implies that $\lim_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) = 0$, i.e., $\{\chi_n\}$ is a universal embezzling family.

Let us now consider the case $\lim_{x\to\infty} f(x)/x^{-1} = \infty$. From this immediately follows that $\sum_{x=1}^{\infty} f(x) = +\infty$, so we can use the results of Section 3.2.1. Furthermore, there exists an N such that $f(x)/x^{-1}$ is non-decreasing for $x \geq N$. We write $f(x) = x^{-1}h(x)$ and thus $h(x)$ is non-decreasing for $x \geq N$. Computing the derivative of $\int_a^{am} f(x)\,\mathrm{d}x$ for $a \geq N$, we find

$$\frac{\mathrm{d}\int_a^{am} f(x)\,\mathrm{d}x}{\mathrm{d}a} = m f(am) - f(a) = \frac{h(am) - h(a)}{a} \geq 0. \tag{3.127}$$

Thus, $\max_{N\leq a\leq y/m} \int_a^{am} f(x)\,\mathrm{d}x = \int_{y/m}^{y} f(x)\,\mathrm{d}x$ and, by following the same steps as in Eq. (3.124), we obtain

$$\begin{aligned} \limsup_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) &\leq \limsup_{y\to\infty} \frac{\max_{N\leq a\leq y/m} \left\{ \int_a^{am} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\ &= \limsup_{y\to\infty} \frac{\int_{y/m}^{y} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x}. \end{aligned} \tag{3.128}$$

Let us fix $\varepsilon > 0$. Since $f(x)/x^{-1+\varepsilon}$ is non-increasing for $x > N_\varepsilon$ by hypothesis, $x^{-1+\varepsilon}/f(x)$ is non-decreasing for $x > N_\varepsilon$. By applying Lemma 3.2.10 (for y large enough), we obtain

$$\int_{N_\varepsilon}^{y} x^{-1+\varepsilon}\,\mathrm{d}x \int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x \geq \int_{N_\varepsilon}^{y/m} x^{-1+\varepsilon}\,\mathrm{d}x \int_{N_\varepsilon}^{y} f(x)\,\mathrm{d}x. \tag{3.129}$$

From this follows that

$$\frac{\int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x}{\int_{N_\varepsilon}^{y} f(x)\,\mathrm{d}x} \geq \frac{(y/m)^\varepsilon - N_\varepsilon^\varepsilon}{y^\varepsilon - N_\varepsilon^\varepsilon}. \tag{3.130}$$

Taking the lim sup on both sides, and adding the finite contributions $\int_1^{N_\varepsilon} f(x)\,\mathrm{d}x$ to the diverging integrals $\int_{N_\varepsilon}^{y/m} f(x)\,\mathrm{d}x$ and $\int_{N_\varepsilon}^{y} f(x)\,\mathrm{d}x$, we obtain

$$\limsup_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \geq \frac{1}{m^\varepsilon}. \tag{3.131}$$

Combining Eq. (3.128) and Eq. (3.131) we get

$$\limsup_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) \leq 1 - \limsup_{y\to\infty} \frac{\int_1^{y/m} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \leq 1 - \frac{1}{m^\varepsilon}. \tag{3.132}$$

This is true for all $\varepsilon > 0$, thus taking the limit $\varepsilon \longrightarrow 0$ we obtain the desired result

$$\limsup_{n\to\infty} T_\star(\chi_n \longrightarrow \chi_n \otimes \Phi_m) \leq \lim_{\varepsilon\to 0} 1 - \frac{1}{m^\varepsilon} = 0. \tag{3.133}$$

This concludes the proof of the sufficient condition. □

The above result on the uniqueness of the van Dam and Hayden family can be expressed as follows: Any regular family of states $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ satisfying the conditions of Theorem 3.2.11 is a universal embezzling family if and only if f, the function associated to $\{\, \chi_n \,\}_{n\in\mathbb{N}}$, is asymptotically close to x^{-1}, the function associated to the van Dam and Hayden family, where asymptotically close means $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for all $\varepsilon > 0$. For examples of such embezzling families, one can see Ref. [53]. All the families introduced in that work satisfy the conditions in Theorem 3.2.11. To the best of our knowledge, there are no embezzling families in literature that do not satisfy the conditions in Theorem 3.2.11.

The assumption that $f(x)/x^{\alpha}$ is asymptotically monotonic for every $\alpha \in \mathbb{R}$ is crucial for our proof and does not follow from the monotonicity of f. There are functions that are non-increasing but oscillate asymptotically when multiplied by powers of x. An example is the function

$$f(x) = \frac{1 + (1 + \sin \ln \ln x) \ln x}{x}, \tag{3.134}$$

which is non-increasing, but $x f(x)$ oscillates between 1 and ∞. Theorem 3.2.11 does not provide any information about families of states associated with such functions, and it cannot be used to determine whether such families are universally embezzling or not.

Next, we prove two related propositions concerning regular families of states associated with non-decreasing functions.

Proposition 3.2.12: Sufficient conditions for regular universal embezzling families - I

Let f be a positive non-decreasing function such that $f(x)/x^{\alpha}$ is asymptotically non-increasing for at least one $\alpha > 0$. Then the regular family of states $\{\, \chi_n \,\}_{n\in\mathbb{N}}$ associated to f (see Definition 3.2.6) is not a universal embezzling family. Furthermore, there are no sequences $\{\, n_j \,\}_{j\in\mathbb{N}}$ such that $\{\, \chi_{n_j} \,\}_{j\in\mathbb{N}}$ is a universal embezzling family.

Proof. Since $f(x)$ is non-decreasing and $\sum_{x=1}^{\infty} f(x)$ diverges, we can use Proposition 3.2.8 and Lemma 3.2.10, to obtain

$$\begin{aligned}
\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\geq \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{ \int_a^{am} g(x,y)\,\mathrm{d}x \right\}}{\int_1^y g(x,y)\,\mathrm{d}x} \\
&= \liminf_{y\to\infty} \frac{\max_{1\leq a\leq y/m} \left\{ \int_{y-am+1}^{y-a+1} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq \liminf_{y\to\infty} \frac{\left\{ \int_1^{y(1-1/m)+1} f(x)\,\mathrm{d}x \right\}}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq \liminf_{y\to\infty} \frac{\left\{ \int_1^{y(1-1/m)+1} x^{\alpha}\,\mathrm{d}x \right\}}{\int_1^y x^{\alpha}\,\mathrm{d}x} \\
&= \left(1 - \frac{1}{m}\right)^{\alpha+1}.
\end{aligned} \tag{3.135}$$

The last inequality is based on Lemma 3.2.10 and is derived as in the previous cases. Using the comment after Definition 3.2.6, we obtain that $\{ \chi_n \}_{n\in\mathbb{N}}$ is not a universal embezzling family. Furthermore, there are no sequences $\left\{ n_j \right\}_{j\in\mathbb{N}}$ such that $\left\{ \chi_{n_j} \right\}_{j\in\mathbb{N}}$ is a universal embezzling family. □

Proposition 3.2.13: Sufficient conditions for regular universal embezzling families - II

Let f be a positive non-decreasing function such that $f(x)/e^{kx}$ is asymptotically non-decreasing for at least one $k > 0$. Then, the regular family of states $\{ \chi_n \}_{n\in\mathbb{N}}$ associated with f is not a universal embezzling family. Furthermore, there are no sequences $\left\{ n_j \right\}_{j\in\mathbb{N}}$ such that $\left\{ \chi_{n_j} \right\}_{j\in\mathbb{N}}$ is a universal embezzling family.

Proof. From Eq. (3.46) follows that

$$\begin{aligned}
\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &\geq \left(1 - \frac{1}{m}\right) \liminf_{n\to\infty} p_1^{(n)} \\
&= \left(1 - \frac{1}{m}\right) \liminf_{n\to\infty} \frac{f(n)}{\sum_{x=1}^{n} f(x)} \\
&= \left(1 - \frac{1}{m}\right) \liminf_{y\to\infty} \frac{f(y)}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq \left(1 - \frac{1}{m}\right) \liminf_{y\to\infty} \frac{\int_{y-1}^{y} f(x)\,\mathrm{d}x}{\int_1^y f(x)\,\mathrm{d}x} \\
&\geq \left(1 - \frac{1}{m}\right) \liminf_{y\to\infty} \frac{\int_{y-1}^{y} e^{kx}\,\mathrm{d}x}{\int_1^y e^{kx}\,\mathrm{d}x} \\
&= \left(1 - \frac{1}{m}\right) (1 - e^{-k}).
\end{aligned} \tag{3.136}$$

Also, here, the last inequality is due to Lemma 3.2.10, and the family of states is not a universal embezzling family. Furthermore, there are no sequences $\left\{ n_j \right\}_{j\in\mathbb{N}}$ such that $\left\{ \chi_{n_j} \right\}_{j\in\mathbb{N}}$ is a universal embezzling family. □

3.2.4 Asymptotically Regular Families

In the definition of universal embezzling families, Definition 3.2.1, and in all the results about embezzlement, only the asymptotic behaviour of a family of states $\{ \chi_n \}_{n\in\mathbb{N}}$ is relevant. This motivates the following definition, which is a generalization of regular families.

Definition 3.2.14: Asymptotically regular family

A family of states $\{ \chi_n \}_{n\in\mathbb{N}}$ is called asymptotically regular if there exists an asymptotically monotonic function $f\colon \mathbb{N} \to (0,\infty)$ such that

$$|\chi_n\rangle = \frac{1}{\sqrt{F_n}} \sum_{x=1}^{n} \sqrt{f(x)}\, |xx\rangle \quad \text{for all} \quad n \in \mathbb{N}, \tag{3.137}$$

where $F_n = \sum_{x=1}^{n} f(x)$.

We next demonstrate that our results also hold for asymptotically regular families. To this end, we start with the following theorem.

Theorem 3.2.15: Regular versus asymptotically regular embezzling families

Let $\{ \chi_n \}_{n\in\mathbb{N}}$ be an asymptotically regular family and let f be the function associated to it (see Definition 3.2.14). Then one can construct a function $\tilde{f}$ that satisfies

1. $\tilde{f}$ corresponds to a regular family $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ (see Definition 3.2.6),
2. $\lim_{x\to\infty} \tilde{f}(x)/f(x) = 1$,
3. $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $\{ \chi_n \}_{n\in\mathbb{N}}$ is a universal embezzling family,
4. $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ contains a universal embezzling subfamily if and only if $\{ \chi_n \}_{n\in\mathbb{N}}$ contains a universal embezzling subfamily.

Proof. Since f is by assumption asymptotically monotonic, $\lim_{x\to\infty} f(x)$ exists in the extended domain $[0,\infty]$. We first study the case $\lim_{x\to\infty} f(x) = l$, with $0 < l \in \mathbb{R}$. In this case, we choose $\tilde{f}(x) = l$, which corresponds to a regular family. Moreover, from the definition of $\tilde{f}$ follows that $\lim_{x\to\infty} f(x)/\tilde{f}(x) = 1$. We also notice that for every $0 < \varepsilon < l$, there exists an N such that $l - \varepsilon < f(x) < l + \varepsilon$ for all $x \geq N$. This implies that $\sum_{x=1}^{\infty} f(x) \geq \sum_{x=N}^{\infty} f(x) \geq \sum_{x=N}^{\infty} (l - \varepsilon) = \infty$. Similarly, $\sum_{x=1}^{\infty} \tilde{f}(x) = \sum_{x=1}^{\infty} l = \infty$. Since $\lim_{n\to\infty} F_n^{-1} = \lim_{n\to\infty} \tilde{F}_n^{-1} = 0$, we obtain that $\lim_{n\to\infty} p_1^{(n)} = 0$ (where $\mathbf{p}^{(n)}$ are the Schmidt coefficients of χ_n), and therefore, according to Eq. (3.46),

$$\begin{aligned}
\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) &= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\|\mathbf{p}^{(n)}\right\|_{(k)} - \left\|\mathbf{p}^{(n)}\right\|_{(a_k)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in\{N,\ldots,n\}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} - \frac{b_k}{m} p^{(n)}_{a_k+1} \right\} \\
&= \liminf_{n\to\infty} \max_{k\in\{N,\ldots,n\}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\},
\end{aligned} \tag{3.138}$$

where $a_k = \lfloor k/m \rfloor$ and $b_k = k - m a_k$. The Schmidt coefficients are, by definition, non-increasing, i.e., obtained by reordering $\{ f(x)/F_n \}_{x\in\mathbb{N}}$. Since there are at most N natural numbers x that do not satisfy the condition $l - \varepsilon < f(x) < l + \varepsilon$, there are at most N Schmidt coefficients that do not satisfy $\frac{l-\varepsilon}{F_n} < p_x^{(n)} < \frac{l+\varepsilon}{F_n}$. We call the set of indices corresponding to these Schmidt coefficients A (thus $|A| \leq N$) and observe that for any $a, b \in \{ N, \dots, n \}$ such that $a \leq b$,

$$\begin{aligned}\sum_{x=a}^{b} p_x^{(n)} &= \sum_{x \in \{ a,\dots,b \} \setminus A} p_x^{(n)} + \sum_{x \in \{ a,\dots,b \} \cap A} p_x^{(n)} \\ &= \sum_{x \in \{ a,\dots,b \} \setminus A} p_x^{(n)} + \sum_{x \in \{ a,\dots,b \} \cap A} \frac{l}{F_n} + \sum_{x \in \{ a,\dots,b \} \cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right).\end{aligned} \tag{3.139}$$

This implies that

$$\begin{aligned}\sum_{x=a}^{b} \frac{(l-\varepsilon)}{F_n} + \sum_{x \in \{ a,\dots,b \} \cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right) &< \sum_{x=a}^{b} p_x^{(n)} \\ &< \sum_{x=a}^{b} \frac{(l+\varepsilon)}{F_n} + \sum_{x \in \{ a,\dots,b \} \cap A} \left(p_x^{(n)} - \frac{l}{F_n} \right),\end{aligned} \tag{3.140}$$

and therefore

$$\begin{aligned}\liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} \frac{l-\varepsilon}{F_n} \right\} &\leq \liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} \\ &\leq \liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} \frac{l+\varepsilon}{F_n} \right\},\end{aligned} \tag{3.141}$$

which implies

$$\begin{aligned}(l-\varepsilon) \liminf_{n\to\infty} \frac{n - \lfloor n/m \rfloor}{F_n} &\leq \liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} \\ &\leq (l+\varepsilon) \liminf_{n\to\infty} \frac{n - \lfloor n/m \rfloor}{F_n}.\end{aligned} \tag{3.142}$$

Observing that $\sum_{x=1}^{N} f(x) + (l-\varepsilon)(n-N) < F_n < \sum_{x=1}^{N} f(x) + (l+\varepsilon)(n-N)$, we obtain

$$\frac{l-\varepsilon}{l+\varepsilon} \left(1 - \frac{1}{m} \right) \leq \liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} \leq \frac{l+\varepsilon}{l-\varepsilon} \left(1 - \frac{1}{m} \right). \tag{3.143}$$

Since Eq. (3.143) holds for every $\varepsilon > 0$, we conclude that

$$\liminf_{n\to\infty} \max_{k \in \{ N,\dots,n \}} \left\{ \sum_{x=a_k+1}^{k} p_x^{(n)} \right\} = 1 - \frac{1}{m}. \tag{3.144}$$

Inserting this result into Eq. (3.138), we have

$$\liminf_{x\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) = 1 - \frac{1}{m}. \tag{3.145}$$

The function $\tilde{f}$ is a rescaling of $f(x) = x^0$, which we already studied in Section 3.2.2. This implies that the family $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ is equal to the family $\left\{ \chi_n^{(0)} \right\}_{n\in\mathbb{N}}$ and

$$\begin{aligned}\liminf_{x\to\infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m) &= \liminf_{x\to\infty} T_\star(\chi_n^{(0)} \to \chi_n^{(0)} \otimes \Phi_m) \\ &= 1 - \frac{1}{m} \\ &= \liminf_{x\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m).\end{aligned} \tag{3.146}$$

This proves that neither f nor $\tilde{f}$ corresponds to families with universally embezzling subfamilies (and are therefore also not universally embezzling themselves).

We now consider the case $\lim_{x\to\infty} f(x) = 0$; thus f is asymptotically non-increasing. Let N be such that f is non-increasing on (N, ∞) and let $a = \min_{x\in\{ 1,\dots,N \}} f(x)$. Since f is positive, $a > 0$. Let $M > N$ be such that $f(x) < a$ for every $x \geq M$ (such M exists because $\lim_{x\to\infty} f(x) = 0$). In this case, we define $\tilde{f}$ as

$$\tilde{f}(x) = \begin{cases} f(M) & \text{if } x \leq M, \\ f(x) & \text{if } x > M. \end{cases} \tag{3.147}$$

Clearly, the family $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ associated to it is regular and $\lim_{x\to\infty} f(x)/\tilde{f}(x) = 1$. Furthermore, the ordered Schmidt coefficients satisfy for all $x \geq M$

$$p_x^{(n)} = \frac{f(x)}{F_n} = \frac{\tilde{f}(x)}{F_n} = \tilde{p}_x^{(n)} \frac{\tilde{F}_n}{F_n}. \tag{3.148}$$

If F_n converges, $\tilde{F}_n$ converges too and by Corollary 3.2.5 neither $\{ \chi_n \}_{n\in\mathbb{N}}$ nor $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ are universal embezzling families (and do not contain universal embezzling subfamilies). If instead F_n diverges, so does $\tilde{F}_n$ and

$$\lim_{n\to\infty} \frac{\tilde{F}_n}{F_n} = \lim_{n\to\infty} \frac{\sum_{x=1}^n \tilde{f}(x)}{\sum_{x=1}^n f(x)} = \lim_{n\to\infty} \frac{\sum_{x=M}^n \tilde{f}(x)}{\sum_{x=M}^n f(x)} = 1. \tag{3.149}$$

Using Eq. (3.46) and Eq. (3.148), we obtain

$$\begin{aligned}&\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\ &\quad= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \\ &\quad= \liminf_{n\to\infty} \max_{k\in\{ mM,\dots,n \}} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \\ &\quad= \liminf_{n\to\infty} \frac{\tilde{F}_n}{F_n} \max_{k\in\{ mM,\dots,n \}} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}_{a_k+1}^{(n)} \right\} \\ &\quad= \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}_{a_k+1}^{(n)} \right\} \\ &\quad= \liminf_{n\to\infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m),\end{aligned} \tag{3.150}$$

where again $a_k = \lfloor k/m \rfloor$ and $b_k = k - ma_k$. In this case as well, we therefore proved that $\{ \chi_n \}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ is a universal embezzling family (and the same holds for subfamilies).

Lastly, we consider the case when $\lim_{x\to\infty} f(x) = \infty$, and f is asymptotically non-decreasing. Analogously to the previous case, let N be such that f is non-decreasing for $x \in (N, \infty)$. Let a be the maximum of $f(x)$ for $x \in \{ 1, \dots, N \}$, and let $M > N$ be such that $f(x) > a$ for $x \geq M$. Also here, we define $\tilde{f}$ via

$$\tilde{f} = \begin{cases} f(M) & \text{if } x \leq M, \\ f(x) & \text{if } x > M. \end{cases} \tag{3.151}$$

The family of states $\{ \tilde{\chi}_n \}_{n\in\mathbb{N}}$ is regular, $\lim_{x\to\infty} \tilde{f}(x)/f(x) = 1$, and $\lim_{x\to\infty} \tilde{F}_n/F_n = 1$. The Schmidt coefficients associated to χ_n and $\tilde{\chi}_n$ are related by

$$p_x^{(n)} = \frac{f(n+1-x)}{F_n} = \frac{\tilde{f}(n+1-x)}{F_n} = \tilde{p}_x^{(n)} \frac{\tilde{F}_n}{F_n} \quad \forall x \leq n - M. \tag{3.152}$$

Using Eq. (3.46) again and the relation between Schmidt coefficients derived in Eq. (3.152) we obtain

$$\begin{aligned} &\liminf_{n\to\infty} T_\star(\chi_n \to \chi_n \otimes \Phi_m) \\ &\quad = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \\ &\quad = \liminf_{n\to\infty} \max_{k\in[n-M]} \left\{ \left\| \mathbf{p}^{(n)} \right\|_{(k)} - \left\| \mathbf{p}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} p_{a_k+1}^{(n)} \right\} \\ &\quad = \liminf_{n\to\infty} \frac{\tilde{F}_n}{F_n} \max_{k\in[n-M]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}_{a_k+1}^{(n)} \right\} \\ &\quad = \liminf_{n\to\infty} \max_{k\in[n]} \left\{ \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(k)} - \left\| \tilde{\mathbf{p}}^{(n)} \right\|_{(a_k)} - \frac{b_k}{m} \tilde{p}_{a_k+1}^{(n)} \right\} \\ &\quad = \liminf_{n\to\infty} T_\star(\tilde{\chi}_n \to \tilde{\chi}_n \otimes \Phi_m). \end{aligned} \tag{3.153}$$

This proves the theorem. □

Thanks to Theorem 3.2.15, Theorem 3.2.11 also holds for asymptotically regular families.

Corollary 3.2.16: Necessary and sufficient conditions for asymptotically regular universal embezzling families

Let f be a positive asymptotically non-increasing function such that $f(x)/x^\alpha$ is asymptotically monotonic for all $\alpha \in \mathbb{R}$ and let $\{ \chi_n \}_{n\in\mathbb{N}}$ be the asymptotically regular family of states associated to f (see Definition 3.2.14). Then $\{ \chi_n \}_{n\in\mathbb{N}}$ is a universal embezzling family if and only if $f \in \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for all $\varepsilon > 0$ and $\sum_{x=1}^{\infty} f(x) = \infty$. Furthermore, if $f \notin \omega(x^{-1-\varepsilon}) \cap o(x^{-1+\varepsilon})$ for at least one $\varepsilon > 0$, then $\{ \chi_{n_j} \}_{j\in\mathbb{N}}$, where $\{ n_j \}_{j\in\mathbb{N}}$ is any sequence of natural numbers, is not a universal embezzling family.

Proof. Combine Theorem 3.2.15 and Theorem 3.2.11. □

For the same reasons, Proposition 3.2.12 and Proposition 3.2.13 also hold for asymptotically regular families.

Corollary 3.2.17: Sufficient conditions for asymptotically regular universal embezzling families - I

Let f be a positive asymptotically non-decreasing function such that $f(x)/x^{\alpha}$ is asymptotically non-increasing for at least one $\alpha > 0$. Then, the asymptotically regular family of states $\{\chi_n\}_{n\in\mathbb{N}}$ associated with f is not a universal embezzling family. Furthermore, there are no sequences $\left\{n_j\right\}_{j\in\mathbb{N}}$ such that $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is a universal embezzling family.

Corollary 3.2.18: Sufficient conditions for asymptotically regular universal embezzling families - II

Let f be a positive asymptotically non-decreasing function such that $f(x)/e^{kx}$ is asymptotically non-decreasing for at least one $k > 0$. Then, the asymptotically regular family of states $\{\chi_n\}_{n\in\mathbb{N}}$ associated with f is not a universal embezzling family. Furthermore, there are no sequences $\left\{n_j\right\}_{j\in\mathbb{N}}$ such that $\left\{\chi_{n_j}\right\}_{j\in\mathbb{N}}$ is a universal embezzling family.

Chapter 4

Cooling and Heating in the Resource Theory of Thermodynamics

In Section 2.3, we presented the resource theory of thermodynamics or athermality [33–35, 107, 171–188]. This is an operational approach to quantum thermodynamics based on the partition of channels and states into free and resourceful. As a reminder, in the case of thermodynamics, states in equilibrium with a given large bath (or environment) are free, while states out of thermal equilibrium are resources. Thermal operations consist of 1) tensoring with a thermal state, 2) performing an energy-preserving evolution, and 3) discarding subsystems. As explained in Section 2.3, free operations are the closure of thermal operations. The interest in quantum thermodynamics is more than a futile intellectual exercise. There is a growing interest in heat engines operating in the quantum regime motivated by the promise that quantum engines may grant an advantage over classical ones [189–196]. Moreover, quantum thermodynamics is also important from a foundational perspective since it has applications ranging from the small scales of biochemistry to the large scales of black hole physics [197–200].

Using the framework of quantum resource theories, in this Chapter, we derive important results about two of the most notable thermodynamic tasks: cooling and heating. The first question that we have to answer is what it means to cool and heat a quantum system A [201]. One approach is that we convert the system from a state at equilibrium with a bath at a temperature T (which we assume to be the temperature of the environment) to a state corresponding to the equilibrium with another bath at temperature $\tilde{T}$. Following this approach, we derive expressions for the highest and lowest temperature to which a system A can be heated and cooled using a given resource and only closed thermal operations. These expressions can be easily computed analytically when $|A| = 2$ and numerically in all other cases. Another common interpretation of cooling that is independent of a notion of temperature is to increase the overlap with A's ground state, which is considered, for example, in algorithmic heat bath cooling [192, 202–205]. Once again, we find the maximum overlap that can be achieved using a resource and closed thermal operations. Importantly, when cooling qubits, the two interpretations are equivalent and reduce to increasing the population of the ground state. Similarly, heating a qubit reduces to increasing the population of its higher energy level. The last result that we present in Section 4.1 is that the ability to cool and heat qubits fully characterizes the convertibility between quasi-classical states, i.e., states that do not exhibit coherence between different energy eigenstates.

Indeed, a state ρ^R can be converted into a quasi-classical state σ^S with closed thermal operations if and only if ρ^R heats and cools qubits to higher and lower temperatures than σ^S does.

In Section 4.2, we change perspective. Instead of asking to what temperatures a qubit can be heated and cooled, we determine, given a desired temperature $\tilde{T}$, which qubits, identified by the energy gap between the ground and the excited state, can be heated and cooled to such a temperature using a given resource and closed thermal operation. Furthermore, we show again that the ability to cool and heat qubits is critical for characterizing state conversion. Indeed, a state ρ^R can be converted into a quasi-classical state σ^S with closed thermal operations if and only if, for any target temperature $\tilde{T}$, ρ^R can cool or heat to the temperature $\tilde{T}$ at least as many different qubits as σ^S.

4.1 Cooling and Heating

Two central tasks in classical thermodynamics are cooling and heating. In this Section, we study cooling and heating in the context of quantum thermodynamics. First, there is the system that we want to heat and cool, which we denote with the letter A. The Hamiltonian of the system A, H^A, is not completely degenerate; otherwise, cooling and heating would be trivial. Indeed, if the Hamiltonian had only one energy level, no transformation would change the population of that single energy level. We assume that A is initially at equilibrium with the environment. As discussed in Section 2.3, this state is the Gibbs state

$$\gamma^A = \frac{e^{-\beta H^A}}{\mathrm{Tr}_A[e^{-\beta H^A}]}, \tag{4.1}$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature of the environment. Second, we need to characterize what it means to cool and heat system A. From the zeroth law of thermodynamics, we have a notion of temperature as equivalence classes of states at equilibrium. As such, we can say that the system has been cooled or heated to a temperature $\tilde{T}$ if the new system is at equilibrium with a bath at temperature $\tilde{T}$. But this new state is nothing else than the Gibbs state

$$\tilde{\gamma}^A = \frac{e^{-\tilde{\beta} H^A}}{\mathrm{Tr}_A[e^{-\tilde{\beta} H^A}]}, \tag{4.2}$$

where $\tilde{\beta} = \frac{1}{k_B \tilde{T}}$. Moreover, note that, in Section 2.3, we have shown that if we consider γ^A and $\tilde{\gamma}^A$ as statistical ensembles, then T and $\tilde{T}$ are the temperatures of the ensembles, respectively. Therefore, we can describe the cooling and heating of a system A as a process that converts the state γ^A into the state $\tilde{\gamma}^A$. Third, we use a system R, with Hamiltonian H^R, in a state ρ^R as the resources for the cooling and heating. Lastly, we assume that only closed thermal operations can be used in the heating and cooling protocols. Therefore, the question becomes: Can the state ρ^R be used to convert the state γ^A into the state $\tilde{\gamma}^A$ using only CTO? That is, is the conversion

$$(\rho^R \otimes \gamma^A, \gamma^R \otimes \gamma^A) \xrightarrow{CTO} (\gamma^R \otimes \tilde{\gamma}^A, \gamma^R \otimes \gamma^A) \tag{4.3}$$

possible? Observe that we have assumed that the final state of system R is γ^R. This assumption can be made without loss of generality because if it is possible to convert $\rho^R \otimes \gamma^A$ to $\sigma^R \otimes \tilde{\gamma}^A$, then it is also possible to convert $\rho^R \otimes \gamma^A$ into $\gamma^R \otimes \tilde{\gamma}^A$. Indeed, remember that sequential and parallel composition of free operations are free, and so are partial traces and preparing a system at equilibrium with the environment. Therefore:

$$(\rho^R \otimes \gamma^A, \gamma^R \otimes \gamma^A) \xrightarrow{CTO} (\sigma^R \otimes \tilde{\gamma}^A, \gamma^R \otimes \gamma^A) \xrightarrow{\mathrm{Tr}_R} (\tilde{\gamma}^A, \gamma^A) \xrightarrow{\gamma^R \otimes} (\gamma^R \otimes \tilde{\gamma}^A, \gamma^R \otimes \gamma^A). \tag{4.4}$$

With a similar reasoning, we can simplify Eq. (4.3):

$$\begin{tikzcd}
(\rho^R \otimes \gamma^A, \gamma^R \otimes \gamma^A) \arrow[r, "\mathrm{CTO}"] \arrow[d, shift right, "\mathrm{Tr}_A"'] & (\gamma^R \otimes \tilde{\gamma}^A, \gamma^R \otimes \gamma^A) \arrow[d, shift right, "\mathrm{Tr}_R"'] \\
(\rho^R, \gamma^R) \arrow[r, "\mathrm{CTO}"] \arrow[u, shift right, "\otimes\gamma^A"'] & (\tilde{\gamma}^A, \gamma^A) \arrow[u, shift right, "\gamma^R\otimes"']
\end{tikzcd} \tag{4.5}$$

Eq. (4.5) shows that the conversion problem in Eq. (4.3) is equivalent to:

$$(\rho^R, \gamma^R) \xrightarrow{CTO} (\tilde{\gamma}^A, \gamma^A). \tag{4.6}$$

As explained in Section 2.3, there is a nice characterization for the conversion problem in the resource theory of athermality only for quasi-classical states. In our case, $\tilde{\gamma}^A$ is quasi-classical, but no assumption has been made about ρ^R. However, for every ρ^R, the quasi-classical state $\mathcal{P}^{R\to R}(\rho^R)$ is equivalent to ρ^R in cooling and heating, where $\mathcal{P}^{R\to R}$ is the pinching or twirling channel

$$\mathcal{P}^{R\to R}(\cdot) = \sum_x \Pi_x^R (\cdot) \Pi_x^R, \tag{4.7}$$

and Π_x^R are the projectors onto the eigenspaces of H^R.

Lemma 4.1.1: Equivalence of quasi-classical and non quasi-classical states in CTO conversions with quasi-classical output

Assume (σ^A, γ^A) is quasi-classical. Then

$$(\rho^R, \gamma^R) \xrightarrow{\mathrm{CTO}} (\sigma^A, \gamma^A) \Leftrightarrow (\mathcal{P}^{R\to R}(\rho^R), \gamma^R) \xrightarrow{\mathrm{CTO}} (\sigma^A, \gamma^A). \tag{4.8}$$

Proof. Remember that the pinching channel is a free operation [98]. We first assume that $(\mathcal{P}^{R\to R}(\rho^R), \gamma^R) \xrightarrow{\mathrm{CTO}} (\sigma^A, \gamma^A)$ holds, i.e., there exists an $\mathcal{M}^{R\to A} \in \mathrm{CTO}(R \to A)$ such that

$$\mathcal{M}^{R\to A}(\mathcal{P}^{R\to R}(\rho^R)) = \sigma^A. \tag{4.9}$$

Due to Proposition 2.3.2, $\mathcal{N}^{R\to A} := \mathcal{M}^{R\to A} \circ \mathcal{P}^{R\to R} \in \mathrm{CTO}(R \to R)$. Moreover, $\mathcal{N}^{R\to A}(\rho^R) = \sigma^A$, i.e., $(\rho^B, \gamma^B) \xrightarrow{\mathrm{CTO}} (\sigma^A, \gamma^A)$.

For the reverse, assume that $(\rho^R, \gamma^R) \xrightarrow{\mathrm{CTO}} (\sigma^A, \gamma^A)$, i.e., that there exists a sequence of thermal operations $\mathcal{M}_n^{R\to A} \in \mathrm{TO}(R \to A)$ such that

$$\mathcal{M}^{R\to A}(\rho^R) := \lim_{n\to\infty} \mathcal{M}_n^{R\to A}(\rho^R) = \sigma^A. \tag{4.10}$$

Now, every thermal operation is covariant in the sense that [176]

$$\mathcal{P}^{A\to A} \circ \mathcal{M}_n^{R\to A} = \mathcal{M}_n^{R\to A} \circ \mathcal{P}^{R\to R}, \tag{4.11}$$

and since the set of covariant operations is closed, also $\mathcal{M}^{R\to A}$ is covariant. Therefore

$$\sigma^A = \mathcal{P}^{A\to A}(\sigma^A) = \mathcal{P}^{A\to A} \circ \mathcal{M}^{R\to A}(\rho^R) = \mathcal{M}^{R\to A} \circ \mathcal{P}^{R\to R}(\rho^R), \tag{4.12}$$

and $(\mathcal{P}^{R\to R}(\rho^R), \gamma^R) \xrightarrow{\text{CTO}} (\sigma^A, \gamma^A)$. □

This Lemma significantly simplifies our treatment of cooling and heating because we can use the quasi-classical state $\mathcal{P}^{R\to R}(\rho^R)$ and all the results about relative majorization and Lorenz curves presented in Section 2.3.

Now, we want to find the highest and lowest (inverse) temperature to which ρ^R can heat and cool γ^A:

$$\begin{aligned}\tilde{\beta}_{\min} &= \min\left\{ \tilde{\beta} \mid (\rho^R, \gamma^R) \xrightarrow{CTO} (\tilde{\gamma}^A(\tilde{\beta}), \gamma^A) \right\}\\ \tilde{\beta}_{\max} &= \max\left\{ \tilde{\beta} \mid (\rho^R, \gamma^R) \xrightarrow{CTO} (\tilde{\gamma}^A(\tilde{\beta}), \gamma^A) \right\}.\end{aligned} \tag{4.13}$$

Observe that we wrote $\tilde{\gamma}^A(\tilde{\beta})$ to make the dependency on $\tilde{\beta}$ explicit. Thanks to Lemma 4.1.1, we can rewrite the equations above as:

$$\begin{aligned}\tilde{\beta}_{\min} &= \min\left\{ \tilde{\beta} \mid (\mathcal{P}^{R\to R}(\rho^R), \gamma^R) \xrightarrow{CTO} (\tilde{\gamma}^A(\tilde{\beta}), \gamma^A) \right\}\\ \tilde{\beta}_{\max} &= \max\left\{ \tilde{\beta} \mid (\mathcal{P}^{R\to R}(\rho^R), \gamma^R) \xrightarrow{CTO} (\tilde{\gamma}^A(\tilde{\beta}), \gamma^A) \right\}.\end{aligned} \tag{4.14}$$

As detailed in Section 2.3, it is possible to associate a pair of probability vectors with each quasi-classical state. Let $(\mathbf{r}^R, \mathbf{g}^R)$ be the vectors associated with $(\mathcal{P}^{R\to R}(\rho^R), \gamma^R)$, and let $(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)$ be the probability vectors associated with $(\tilde{\gamma}^A(\tilde{\beta}), \gamma^A)$. Hence, thanks to Theorem 2.3.4, we can define $\tilde{\beta}_{\max}$ and $\tilde{\beta}_{\min}$, as

$$\begin{aligned}\tilde{\beta}_{\min} &= \min\left\{ \tilde{\beta} \mid (\mathbf{r}^R, \mathbf{g}^R) \succ (\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A) \right\}\\ \tilde{\beta}_{\max} &= \max\left\{ \tilde{\beta} \mid (\mathbf{r}^R, \mathbf{g}^R) \succ (\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A) \right\},\end{aligned} \tag{4.15}$$

or in terms of Lorenz curves, as

$$\begin{aligned}\tilde{\beta}_{\min} &= \min\left\{ \tilde{\beta} \mid \alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(y) \geq \alpha^{(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)}(y), \forall y \in [0,1] \right\}\\ \tilde{\beta}_{\max} &= \max\left\{ \tilde{\beta} \mid \alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(y) \geq \alpha^{(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)}(y), \forall y \in [0,1] \right\}.\end{aligned} \tag{4.16}$$

Now, the task is to determine how the Lorenz curve of $(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)$ changes with $\tilde{\beta}$. Let $\left\{ |i\rangle^A \right\}_i$ be an orthonormal basis for A such that $H^A = \sum_i h_i |i\rangle\langle i|^A$, and $h_{|A|} \geq \cdots \geq h_1$. In this basis,

$$\gamma^A = \sum_i g_i^A |i\rangle\langle i|^A, \text{ and } \quad \tilde{\gamma}_i^A = \sum_i \tilde{g}_i^A |i\rangle\langle i|^A, \tag{4.17}$$

where

$$g_i^A = \frac{e^{-\beta h_i}}{Z(\beta)}, \text{ and } \quad \tilde{g}_i^A = \frac{e^{-\tilde{\beta} h_i}}{Z(\tilde{\beta})}, \tag{4.18}$$

and $Z(\lambda) = \sum_i e^{-\lambda h_i}$. To construct the Lorenz curves, we first have to order $\left\{ \tilde{g}_i^A / g_i^A \right\}$ in non-increasing order. Observe that

$$\frac{\tilde{g}_i^A}{g_i^A} = \frac{Z(\tilde{\beta})}{Z(\beta)} e^{-(\tilde{\beta}-\beta)h_i}. \tag{4.19}$$

In the case of cooling $\tilde{\beta} \geq \beta$, and therefore, $\frac{\tilde{g}_i^A}{g_i^A}$ decreases with i, and in the case of heating $\tilde{\beta} \leq \beta$ and therefore $\frac{\tilde{g}_i^A}{g_i^A}$ increases with i. We treat the two cases separately. We start with cooling because the set $\left\{ \tilde{g}_i^A / g_i^A \right\}$ is already ordered properly. In this case, the x and y values of the k-th elbow of the Lorenz curve are obtained by summing the first k entries of $\tilde{\mathbf{g}}^A$ and $\mathbf{g}^A$, respectively:

$$(x_k^{(\tilde{\mathbf{g}}^A,\mathbf{g}^A)}, y_k^{(\tilde{\mathbf{g}}^A,\mathbf{g}^A)}) = \left(\left\|\tilde{\mathbf{g}}^A\right\|_{(k)}, \left\|\mathbf{g}^A\right\|_{(k)} \right). \tag{4.20}$$

The y value of these elbows does not vary with $\tilde{\beta}$; therefore, when $\tilde{\beta}$ changes, the elbows can only move horizontally. For the x value, observe that for $k < |A|$,

$$\left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} = \frac{\sum_{i=1}^{k} e^{-\tilde{\beta}h_i}}{\sum_{i=1}^{|A|} e^{-\tilde{\beta}h_i}}, \tag{4.21}$$

which increases with $\tilde{\beta}$. Therefore, the more A is cooled, the more the elbows move to the right. We do not need to consider $k = |A|$ because, in that case, the elbow is $(1, 1)$. With this geometric intuition, we can compute $\tilde{\beta}_{\max}$ by finding how much we can move the elbows of $\alpha^{(\tilde{\mathbf{g}}^A(\tilde{\beta}),\mathbf{g}^A)}$ while keeping the Lorenz curve on the left of $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}$. Thanks to the concavity of $\alpha^{(\cdot,\cdot)}$, it is enough to check if all the elbows of $\alpha^{(\tilde{\mathbf{g}}^A(\tilde{\beta}),\mathbf{g}^A)}$ are on the left of $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}$ [119]. That is

$$\tilde{\beta}_{\max} = \max \left\{ \tilde{\beta} \mid \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(y) \geq \alpha^{(\tilde{\mathbf{g}}^A(\tilde{\beta}),\mathbf{g}^A)}(y), \forall y \in [0,1] \right\} \tag{4.22}$$

is equivalent to

$$\tilde{\beta}_{\max} = \max \left\{ \tilde{\beta} \mid \left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} \leq \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right), \forall k \in [|A|-1] \right\}. \tag{4.23}$$

Since $\left\|\tilde{\mathbf{g}}(\tilde{\beta})\right\|_{(k)}$ increases with $\tilde{\beta}$, there exists at least one k for which

$$\left\|\tilde{\mathbf{g}}^A(\tilde{\beta}_{\max})\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right). \tag{4.24}$$

To compute $\tilde{\beta}_{\max}$, we can first solve $\left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right)$ for each k. Then, $\tilde{\beta}_{\max}$ must be one of the β_k, and specifically, it must be the smallest, because if there exists a $\tilde{\beta}_k < \tilde{\beta}_{\max}$, then

$$\left\|\tilde{\mathbf{g}}^A(\tilde{\beta}_{\max})\right\|_{(k)} > \left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right), \tag{4.25}$$

which contradicts, Eq. (4.23). Therefore,

$$\tilde{\beta}_{\max} = \min_{k \in [|A|-1]} \left\{ \tilde{\beta}_k \mid \left\|\tilde{\mathbf{g}}^A(\tilde{\beta}_k)\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right) \right\} \tag{4.26}$$

What is left to show is whether $\tilde{\beta}_k$ exists and is unique. First, observe that $\left\|\tilde{\mathbf{g}}^A\right\|_{(k)} = 1 - \sum_{i=k+1}^{n} e^{-\tilde{\beta}h_i}$ is strictly increasing with $\tilde{\beta}$. Therefore if a solution for $\left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(\left\|\mathbf{g}^A\right\|_{(k)} \right)$ exists, it is unique.

Then, notice that $\lim_{\tilde{\beta}\to\infty} \left\|\tilde{\mathbf{g}}^A\right\|_{(k)} = 1$. This means that in the case of cooling, where $\tilde{\beta} \geq \beta$, $\left\|\tilde{\mathbf{g}}^A\right\|_{(k)}$ can assume any value between $\left\|\mathbf{g}^A\right\|_{(k)}$ (included) and 1 (excluded). From the geometry of Lorenz curves, it immediately follows that $\left\|\mathbf{g}^A\right\|_{(k)} \leq \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right) \leq 1$. Adding this to what we know about $\left\|\tilde{\mathbf{g}}^A\right\|_{(k)}$, we deduce that

$$\left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right) \tag{4.27}$$

has a unique solution if $\left\|\mathbf{g}^A\right\|_{(k)} \leq \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right) < 1$, and no solution if $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right) = 1$. With a slight abuse, we say that $\tilde{\beta}_k = +\infty$ is a solution of Eq. (4.27) in this last case. Given that $\left\|\tilde{\mathbf{g}}^A(\tilde{\beta})\right\|_{(k)}$ is strictly monotonic, finding a numerical solution is rather easy, once the value of $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right)$ is known.

Since we know how to draw the lower Lorenz curve associated with a pair of vectors, we can easily compute the value of $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}\left(\left\|\mathbf{g}^A\right\|_{(k)}\right)$. Recall that $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}$ is obtained by connecting the elbows $\left\{ (x_k^{(\mathbf{r}^R,\mathbf{g}^R)}, y_k^{(\mathbf{r}^R,\mathbf{g}^R)}) \right\}_k$ with straight lines. For $y_k^{(\mathbf{r}^R,\mathbf{g}^R)} \leq y < y_{k+1}^{(\mathbf{r}^R,\mathbf{g}^R)}$, it thus holds that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(y) = x_k^{(\mathbf{r}^R,\mathbf{g}^R)} + \frac{x_{k+1}^{(\mathbf{r}^R,\mathbf{g}^R)} - x_k^{(\mathbf{r}^R,\mathbf{g}^R)}}{y_{k+1}^{(\mathbf{r}^R,\mathbf{g}^R)} - y_k^{(\mathbf{r}^R,\mathbf{g}^R)}}\left(y - y_k^{(\mathbf{r}^R,\mathbf{g}^R)}\right), \tag{4.28}$$

since this describes the straight line through the two points $(x_k^{(\mathbf{r}^R,\mathbf{g}^R)}, y_k^{(\mathbf{r}^R,\mathbf{g}^R)})$ and $(x_{k+1}^{(\mathbf{r}^R,\mathbf{g}^R)}, y_{k+1}^{(\mathbf{r}^R,\mathbf{g}^R)})$. By choosing l_k such that $y_{l_k}^{(\mathbf{r}^R,\mathbf{g}^R)} \leq \left\|\mathbf{g}^A\right\|_{(k)} < y_{l_k+1}^{(\mathbf{r}^R,\mathbf{g}^R)}$, it thus holds that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(\left\|\mathbf{g}^A\right\|_{(k)}) = x_{l_k}^{(\mathbf{r}^R,\mathbf{g}^R)} + \frac{x_{l_k+1}^{(\mathbf{r}^R,\mathbf{g}^R)} - x_{l_k}^{(\mathbf{r}^R,\mathbf{g}^R)}}{y_{l_k+1}^{(\mathbf{r}^R,\mathbf{g}^R)} - y_{l_k}^{(\mathbf{r}^R,\mathbf{g}^R)}}\left(\left\|\mathbf{g}^A\right\|_{(k)} - y_{l_k}^{(\mathbf{r}^R,\mathbf{g}^R)}\right), \tag{4.29}$$

This shows that the $\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(\left\|\mathbf{g}^A\right\|_{(k)})$ can be directly calculated once the initial resource and γ^A are specified.

We can summarize these findings with the following Theorem.

Theorem 4.1.2: Cooling in the resource theory of athermality

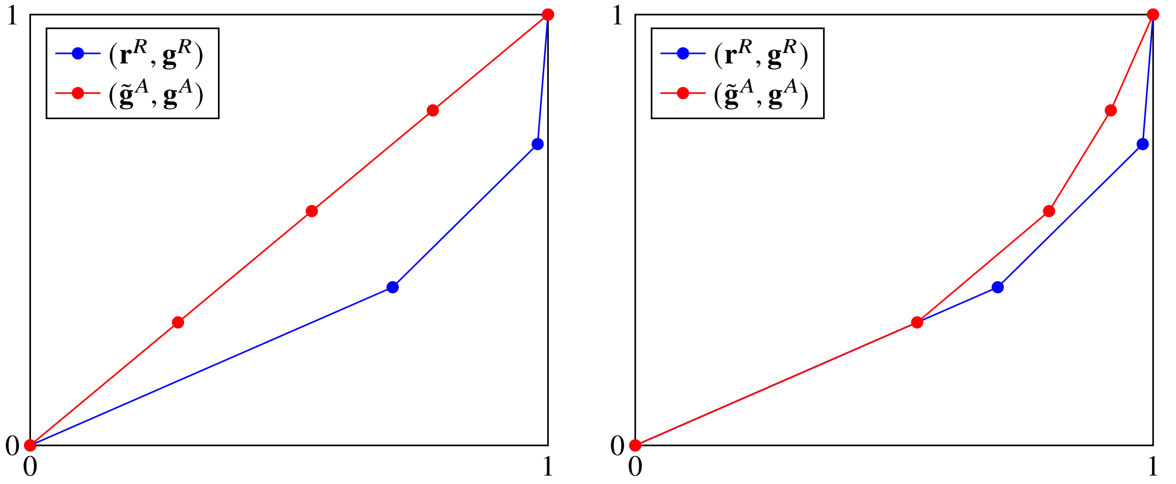
Let A and R be quantum systems with Hamiltonians H^A and H^R, respectively. Let $\rho^R \in \mathfrak{D}(R)$ and let β be the temperature of the environment. The maximal $\tilde{\beta}$ to which one can cool the system A, initially at equilibrium with the environment, using only CTO and the resource (ρ^R, γ^R) is

$$\tilde{\beta}_{\max} = \min_{k\in[|A|-1]} \left\{ \tilde{\beta}_k | \left\|\tilde{\mathbf{g}}^A(\tilde{\beta}_k)\right\|_{(k)} = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(\left\|\mathbf{g}^A\right\|_{(k)}) \right\}, \tag{4.30}$$

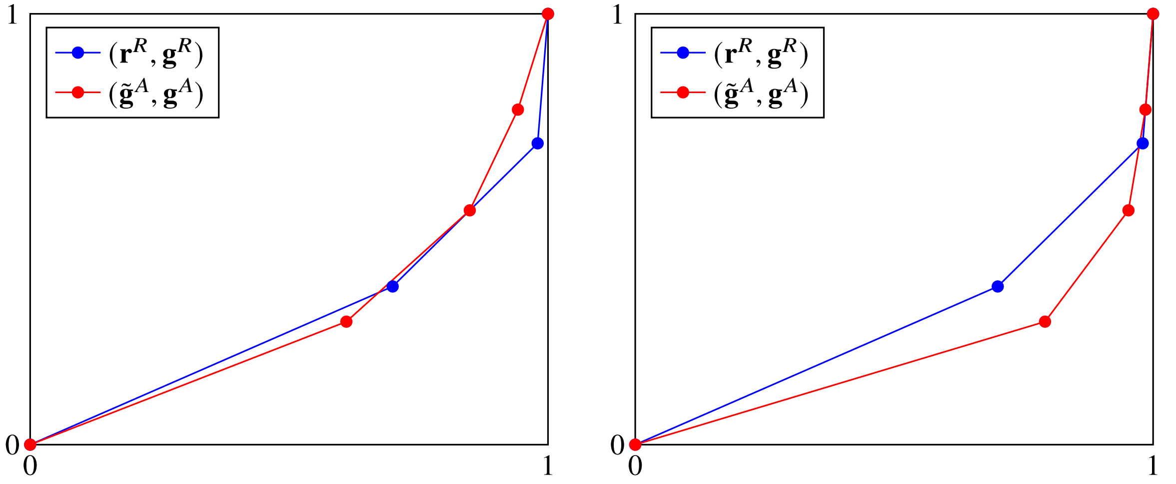
where $(\mathbf{r}^R, \mathbf{g}^R)$ are the vectors associated with $(\mathcal{P}^{R\to R}(\rho^R), \gamma^R)$, and $(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)$ the vectors associated with $(\tilde{\gamma}^A(\tilde{\beta}), \gamma^A)$.

For qubits, there is only one equation to solve, which can be solved analytically. Let $E > 0$ denote the energy gap between the ground and excited level of the qubit. Then,

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(\left\|\mathbf{g}^A\right\|_{(1)}) = \left\|\tilde{\mathbf{g}}^A(\tilde{\beta}_{\max})\right\|_{(1)} = \frac{1}{1 - e^{-\tilde{\beta}_{\max}E}}. \tag{4.31}$$



(a) $\tilde{\beta} = 1$ (b) $\tilde{\beta}_1 \sim 7.59$



(c) $\tilde{\beta}_2 \sim 9.41$ (d) $\tilde{\beta}_3 \sim 15.92$

Figure 4.1: Geometric visualization of the proof for the value of $\tilde{\beta}_{\max}$. To reproduce these graphs, fix $\beta = 1$. The eigenvalues of H^A are $\{ 0, 0.1, 0.2, 0.25 \}$, and the eigenvalues of H^R are $\{ 0, 0.1, 0.2 \}$. The x values of the elbows of $(\mathbf{r}^R, \mathbf{g}^R)$ are $\{ 0.7, 0.9, 1 \}$. The y values are obtained by H^R and β without any reordering. To find $\tilde{\beta}_{\max}$, the starting point is to find the Lorenz curves associated with $(\mathbf{r}^R, \mathbf{g}^R)$ and $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ when $\tilde{\beta} = \beta$ (Figure a). The elbows of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ move horizontally when $\tilde{\beta}$ increases. Starting with the first elbow, we want to find the x value in the Lorenz associated with $(\mathbf{r}^R, \mathbf{g}^R)$ corresponding to the y value of the first elbow of the Lorenz curve associated with $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$. From the x value, one can deduce the temperature $\tilde{\beta}_1$ and draw the Lorenz curve based on $(\tilde{\mathbf{g}}^A(\tilde{\beta}_1), \mathbf{g}^A)$. (Figure b). Similarly, one can draw the Lorenz curves associated with the temperatures at which the second and third elbow of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ intersect the Lorenz curve of $(\mathbf{r}^R, \mathbf{g}^R)$. CTO-conversion is possible if and only if the Lorenz curve of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ is completely on the left of the Lorenz curve of $(\mathbf{r}^R, \mathbf{g}^R)$. This happens only for the smallest of the $\tilde{\beta}_k$. Therefore, $\tilde{\beta}_{\max} \sim 7.59$.

From the equation above, it immediately follows that

$$\tilde{\beta}_{\max} = \frac{1}{E} \ln \left(\frac{\alpha^{\left(\mathbf{r}^R, \mathbf{g}^R\right)}(g_1^A)}{1 - \alpha^{\left(\mathbf{r}^R, \mathbf{g}^R\right)}(g_1^A)} \right), \tag{4.32}$$

where $g_1^A = \frac{1}{1-e^{-\beta E}}$ and $\alpha^{\left(\mathbf{r}^R, \mathbf{g}^R\right)}(g_1^A)$ can be computed as in Eq. (4.29).

We now turn our attention to the heating process, in which $\tilde{\beta} \leq \beta$. As the reader may recall, in this case, the entries of $\left\{ \frac{\tilde{g}_i^A}{g_i^A} \right\}$ are in non-decreasing order. Therefore, the x and y values of the k-th elbow of the Lorenz curve are obtained by summing the last k entries of $\tilde{\mathbf{g}}^A$ and $\mathbf{g}^A$, respectively:

$$(x_k^{(\tilde{\mathbf{g}}^A, \mathbf{g}^A)}, y_k^{(\tilde{\mathbf{g}}^A, \mathbf{g}^A)}) = \left(\sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A, \sum_{i=|A|-k+1}^{|A|} g_i^A \right) = \left(\sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A, 1 - \left\| \mathbf{g}^A \right\|_{(k)} \right). \tag{4.33}$$

where we understand that $\|\mathbf{g}\|_{(0)} = 0$. Note that we do not replace $\sum_{j=n-k+1}^{n} \tilde{g}_j$ with $1 - \left\| \tilde{\mathbf{g}}^A \right\|_{(n-k)}$ because they differ for negative $\tilde{\beta}$, which we allow. Negative $\tilde{\beta}$ are associated with population inversion. That is, when $\tilde{\beta} > 0$, the low energy levels have higher population than the high energy levels. When $\tilde{\beta}$ goes to zero, all the levels have the same population. When $\tilde{\beta} < 0$, the population of high energy levels is higher than the population of low energy levels. With positive $\tilde{\beta}$, the probability $\tilde{g}_i^A$ decreases with i; the higher the energy eigenvalue, the lower the probability. Instead, with negative $\tilde{\beta}$, the probability increases. In these scenarios, we talk of population inversion [206, 207]. With the same considerations done in the case of cooling leading to Eq. (4.23), $\tilde{\beta}_{\min}$ can be expressed as

$$\tilde{\beta}_{\min} = \min \left\{ \tilde{\beta} \mid \sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A \leq \alpha^{(\mathbf{r}^R, \mathbf{g}^R)} \left(1 - \left\| \mathbf{g}^A \right\|_{(k)} \right), \forall k \in [|A|-1] \right\}. \tag{4.34}$$

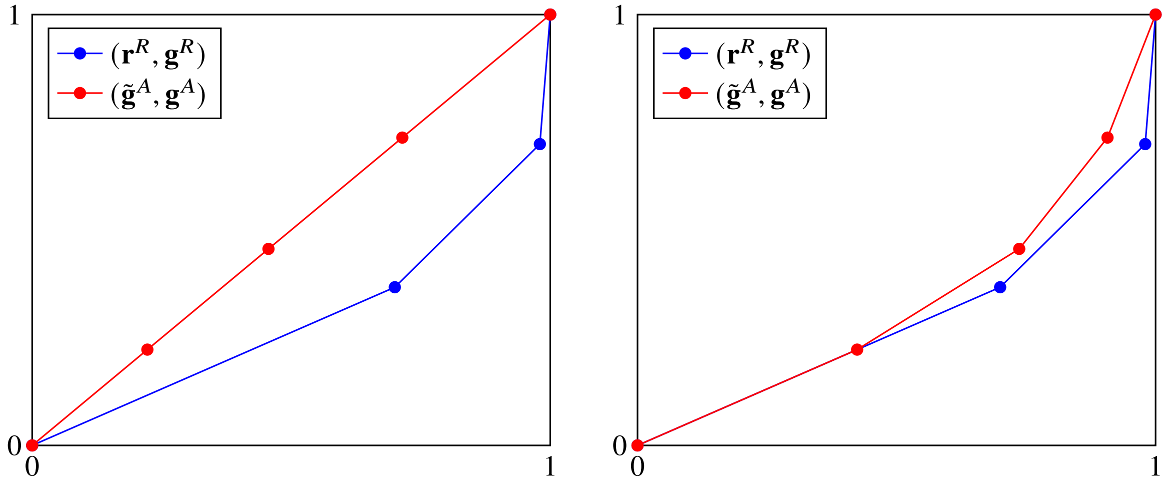
Once again, we observe that the y coordinate of the elbows is constant, while the x coordinate increases when $\tilde{\beta}$ decreases:

$$x_k^{(\tilde{\mathbf{g}}^A, \mathbf{g}^A)} = \sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A. \tag{4.35}$$

This means that the more the system is heated, the more the elbows move to the right (note that the elbows in the case of heating are different from the elbows in the case of cooling). Therefore, $\tilde{\beta}_{\min}$ has to satisfy $\sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A(\tilde{\beta}_{\min}) = \alpha^{(\mathbf{r}^R, \mathbf{g}^R)} \left(1 - \left\| \mathbf{g}^A \right\|_{(k)} \right)$ for at least one k. As in the case of cooling, we can solve all the k equations

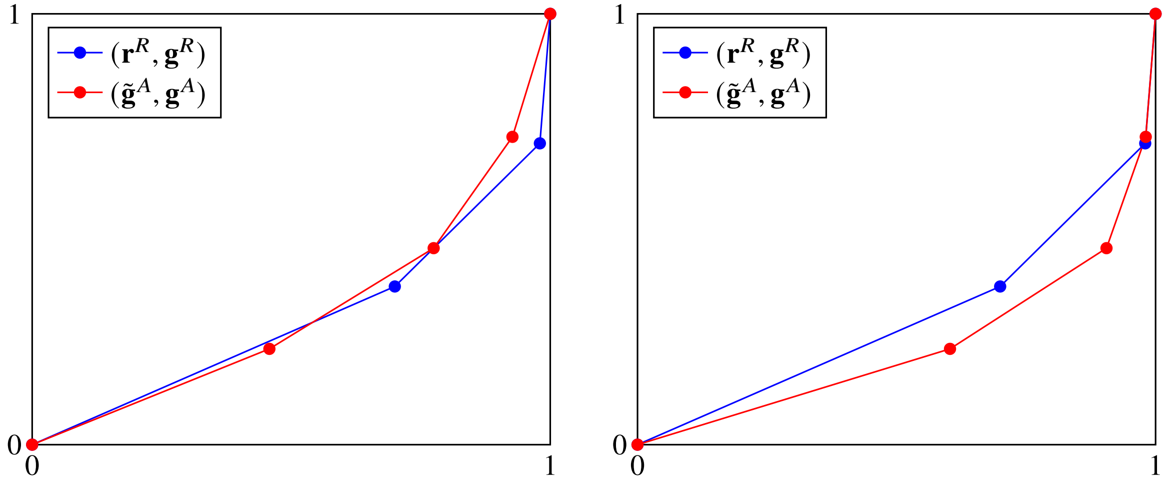
$$\sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A = \alpha^{(\mathbf{r}^R, \mathbf{g}^R)} \left(1 - \left\| \mathbf{g}^A \right\|_{(k)} \right). \tag{4.36}$$

If the solution exists, it is unique, and we denote it with $\tilde{\beta}_k$; if not, we say that $\tilde{\beta}_k = -\infty$. In this case, since $\sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A$ increases when $\tilde{\beta}$ decreases, $\tilde{\beta}_{\min}$ must be smallest β_k. We can summarize the results about heating with the following theorem.



(a) $\tilde{\beta} = 1$

(b) $\tilde{\beta}_1 \sim -6.08$



(c) $\tilde{\beta}_2 \sim -7.35$

(d) $\tilde{\beta}_3 \sim -13.83$

Figure 4.2: Geometric visualization of the proof for the value of $\tilde{\beta}_{\min}$. These graphs have been created using the same values as in Figure 4.1. To find $\tilde{\beta}_{\min}$, the starting point is to find the Lorenz curves associated with $(\mathbf{r}^R, \mathbf{g}^R)$ and $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ when $\tilde{\beta} = \beta$ (Figure a). Observe that the elbows of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ are the reflection of the elbows in Figure 4.1a through $(1/2, 1/2)$. The elbows of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ move horizontally when $\tilde{\beta}$ decreases. Starting with the first elbow, we want to find the x value in the Lorenz curve associated with $(\mathbf{r}^R, \mathbf{g}^R)$ corresponding to the y value of the first elbow of the Lorenz curve associated with $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$. From the x value, one can deduce the temperature $\tilde{\beta}_1$ and draw the Lorenz curve based on $(\tilde{\mathbf{g}}^A(\tilde{\beta}_1), \mathbf{g}^A)$. (Figure b). Similarly, one can draw the Lorenz curves associated with the temperatures at which the second and third elbow of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ intersect the Lorenz curve of $(\mathbf{r}^R, \mathbf{g}^R)$. CTO-conversion is possible if and only if the Lorenz curve of $(\tilde{\mathbf{g}}^A, \mathbf{g}^A)$ is completely on the left of the Lorenz curve of $(\mathbf{r}^R, \mathbf{g}^R)$. This happens only for the highest of the $\tilde{\beta}_k$. Therefore, $\tilde{\beta}_{\min} \sim -6.08$.

Theorem 4.1.3: Heating in the resource theory of athermality

Let A and R be quantum systems with Hamiltonians H^A and H^R, respectively. Let $\rho^R \in \mathfrak{D}(R)$ and let β be the temperature of the environment. The minimal $\tilde{\beta}$ to which one can heat the system A, initially at equilibrium with the environment, using only CTO and the resource (ρ^R, γ^R) is

$$\tilde{\beta}_{\min} = \max_{k\in[|A|-1]} \left\{ \tilde{\beta}_k \middle| \sum_{i=|A|-k+1}^{|A|} \tilde{g}_i^A(\tilde{\beta}_k) = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(1 - \left\|\mathbf{g}^A\right\|_{(k)}\right) \right\}, \tag{4.37}$$

where $(\mathbf{r}^R, \mathbf{g}^R)$ are the vectors associated with $(\mathcal{P}^{R\to R}(\rho^R), \gamma^R)$, and $(\tilde{\mathbf{g}}^A(\tilde{\beta}), \mathbf{g}^A)$ the vectors associated with $(\tilde{\gamma}^A(\tilde{\beta}), \gamma^A)$.

We can find an analytical solution for qubits, as we did in the case of cooling. The only non-trivial equation is

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(1 - g_1^A\right) = \tilde{g}_i^A = \frac{e^{-\tilde{\beta}_{\min}E}}{1 + e^{-\tilde{\beta}_{\min}E}}, \tag{4.38}$$

from which, one gets,

$$\tilde{\beta}_{\min} = \frac{1}{E} \ln \left(\frac{1 - \alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(1 - g_1^A\right)}{\alpha^{(\mathbf{r}^R,\mathbf{g}^R)} \left(1 - g_1^A\right)} \right). \tag{4.39}$$

4.1.1 A Different Notion of Cooling

Another common interpretation of cooling that is independent of a notion of temperature is to increase the overlap with A's ground state, which is considered, for example, in algorithmic heat bath cooling [192, 202–205]. The idea is to use a resource (ρ^R, γ^R) to take system A out of equilibrium to a state that has a greater overlap with the ground states, that is, if Π^A is the projector over the (possibly degenerate) ground state of the Hamiltonian, we want to convert γ^A to a state τ^A such that $\mathrm{Tr}_A[\Pi^A, \tau^A] \geq \mathrm{Tr}_A[\Pi^A \gamma^A]$. We denote with $O_{\max}$ the maximal overlap achieve given access (ρ^R, γ^R) and CTO:

$$O_{\max}(\rho^R, \gamma^R, \gamma^A) = \max\{\mathrm{Tr}_A \left[\Pi^A \tau^A\right] : (\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\tau^A, \gamma^A)\}. \tag{4.40}$$

With the following Proposition, we find a closed-form expression for $O_{\max}$.

Proposition 4.1.4: Maximal ground state overlap

With $\mathbf{r}^R, \mathbf{g}^R, \mathbf{g}^A$ denoting the probability vectors corresponding to $\mathcal{P}_{\gamma^R}(\rho^R), \gamma^R, \gamma^A$, respectively,

$$O_{\max}(\rho^R, \gamma^R, \gamma^A) = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(\mathrm{Tr}[\Pi^A \gamma^A]). \tag{4.41}$$

Proof. Let $\tau_\star^A$ be an optimizer of this optimization problem. Since $\mathrm{Tr}_A \left[\Pi^A \tau_\star^A\right] = \mathrm{Tr}\left[\Pi^A \mathcal{P}^{A\to A}(\tau_\star^A)\right]$, we can restrict ourselves to quasi-classical target states and therefore assume without loss of generality that (ρ^R, γ^R) is quasi-classical too (see Lemma 4.1.1). Assuming that the ground state has degeneracy d, this implies that there exists an orthonormal basis $\{|i\rangle\}$ of A such that

$$\begin{aligned} \tau_\star^A &= \sum_{i=1}^{A} t_i^A |i\rangle\langle i|, \\ H^A &= \sum_{i=1}^{A} h_i^A |i\rangle\langle i| \,:\; h_1 = \ldots = h_d < h_{d+1} \leq \ldots \leq h_{|A|} \end{aligned} \tag{4.42}$$

and thus $O_{\max}(\rho^R, \gamma^R, \gamma^A) = \sum_{i=1}^d t_i^A$. Next, we notice that we can further restrict ourselves to the case $t_1^A = t_2^A = \ldots = t_d^A$: Any unitary acting non-trivially only on the support of Π^A is a thermal operation. By applying such unitary transformations uniformly at random (which is in CTO, since CTO is convex, see Ref. [175, App. C], Ref. [208, Prop. 4], and Ref. [98, Thm. II.1]), for $i \in 1, \ldots, d$, we map t_i^A to $\sum_{i=1}^d t_i^A/d$. Obviously, this does not change $O_{\max}$.

For the moment, assume $O_{\max} < 1$ and let

$$\mathbf{t}^{\tilde{A}} = \frac{1}{1 - dt_1^A} \begin{pmatrix} t_{d+1}^A \\ t_{d+2}^A \\ \vdots \\ t_{|A|}^A \end{pmatrix}, \; \mathbf{g}^{\tilde{A}} = \frac{1}{1 - dg_1^A} \begin{pmatrix} g_{d+1}^A \\ g_{d+2}^A \\ \vdots \\ g_{|A|}^A \end{pmatrix}, \tag{4.43}$$

which by construction are valid probability vectors. Since $(\mathbf{t}^{\tilde{A}}, \mathbf{g}^{\tilde{A}}) \succ (\mathbf{g}^{\tilde{A}}, \mathbf{g}^{\tilde{A}})$, by definition, there exists a column stochastic matrix $\tilde{E}$ such that $\tilde{E}\mathbf{t}^{\tilde{A}} = \mathbf{g}^{\tilde{A}}$ and $\tilde{E}\mathbf{g}^{\tilde{A}} = \mathbf{g}^{\tilde{A}}$. Then

$$E = \begin{pmatrix} I_d & 0 \\ 0 & \tilde{E} \end{pmatrix}, \tag{4.44}$$

where I_d is the identity matrix of dimension d, is column stochastic too, and, with $\lambda = \frac{1-dt_1^A}{1-dg_1^A}$,

$$E\mathbf{g}^A = E \begin{pmatrix} g_1^A \\ \vdots \\ g_d^A \\ (1 - dg_1^A)\mathbf{g}^{\tilde{A}} \end{pmatrix} = \mathbf{g}^A, \tag{4.45}$$

$$E\mathbf{t}^A = E \begin{pmatrix} t_1^A \\ \vdots \\ t_d^A \\ (1 - dt_1^A)\mathbf{t}^{\tilde{A}} \end{pmatrix} = \begin{pmatrix} t_1^A \\ \vdots \\ t_d^A \\ (1 - dt_1^A)\mathbf{g}^{\tilde{A}} \end{pmatrix} = \begin{pmatrix} t_1^A \\ \vdots \\ t_d^A \\ \lambda g_{d+1}^A \\ \lambda g_{d+2}^A \\ \vdots \\ \lambda g_{|A|}^A \end{pmatrix} =: \tilde{\mathbf{t}}^A(t_1^A), \tag{4.46}$$

where we used in the notation that $t_1^A = t_2^A = \ldots = t_d^A$. This implies that $(\mathbf{t}^A, \mathbf{g}^A) \succ (\tilde{\mathbf{t}}^A(t_1^A), \mathbf{g}^A)$. Moreover, $O_{\max}(\rho^R, \gamma^R, \gamma^A) = d\tilde{t}_1^A$, and thus we can and will restrict our considerations to targets of the form $\tilde{\mathbf{t}}^A(t_1^A)$, which also includes the special case $O_{\max}(\rho^R, \gamma^R, \gamma^A) = 1$ again.

Since

$$(\mathbf{r}^R, \mathbf{g}^R) \succ (\mathbf{g}^A, \mathbf{g}^A), \tag{4.47}$$

we find that $O_{\max}(\rho^R, \gamma^R, \gamma^A) = d\tilde{t}_1^A \geq dg_1^A$. We thus have that $\tilde{t}_i^A / g_i^A$ is ordered non-increasingly for an ideal $\tilde{\mathbf{t}}^A$, which implies that

$$\left(x_k^{(\tilde{\mathbf{t}}^A, \mathbf{g}^A)}, y_k^{(\tilde{\mathbf{t}}^A, \mathbf{g}^A)}\right) = \left(\tilde{t}_1^A \min\{k, d\} + \lambda \sum_{j=d+1}^{k} g_j^A, \sum_{j=1}^{k} g_j^A\right). \tag{4.48}$$

Now, we remember that between quasi-classical states, we can use relative majorization to characterize CTO conversion. Thus $O_{\max}$ can be written as

$$O_{\max}(\rho^R, \gamma^R, \gamma^A) = d \max\left\{t_1^A : (\mathbf{r}^R, \mathbf{g}^R) \succ (\tilde{\mathbf{t}}^A(t_1^A)^A, \mathbf{g}^A)\right\}, \tag{4.49}$$

or equivalently as

$$O_{\max}(\rho^R, \gamma^R, \gamma^A) = d \max\left\{t_1^A : \alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(y) \geq \alpha^{(\tilde{\mathbf{t}}^A(t_1^A)^A, \mathbf{g}^A)}(y), \forall y \in [0, 1]\right\}. \tag{4.50}$$

The Lorenz curve $\alpha^{(\tilde{\mathbf{t}}^A(t_1^A)^A, \mathbf{g}^A)}$ is very easy to draw. Looking at the points in Eq. (4.48), one notices that $\alpha^{(\tilde{\mathbf{t}}^A(t_1^A)^A, \mathbf{g}^A)}$ is composed of two segments. The first connects $(0, 0)$ with $(d\tilde{t}_1^A, dg_1^A)$. The second connects the latter point with $(1, 1)$. The largest possible dt_1^A is the one associated with the Lorenz curve that has the elbow $(d\tilde{t}_1^A, dg_1^A)$ on the Lorenz curve $\alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(y)$. That is, $\alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(dg_1^A) = d\tilde{t}_1^A$, which gives

$$O_{\max}(\rho^R, \gamma^R, \gamma^A) = \alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(\mathrm{Tr}[\Pi^A \gamma^A]). \tag{4.51}$$

□

Note that this Proposition provides a closed-form expression for arbitrary dimensions of A and not only for qubits. Moreover, it is straightforward to see that for qubits, the two interpretations of cooling coincide.

4.1.2 Cooling Qubits with Qubits

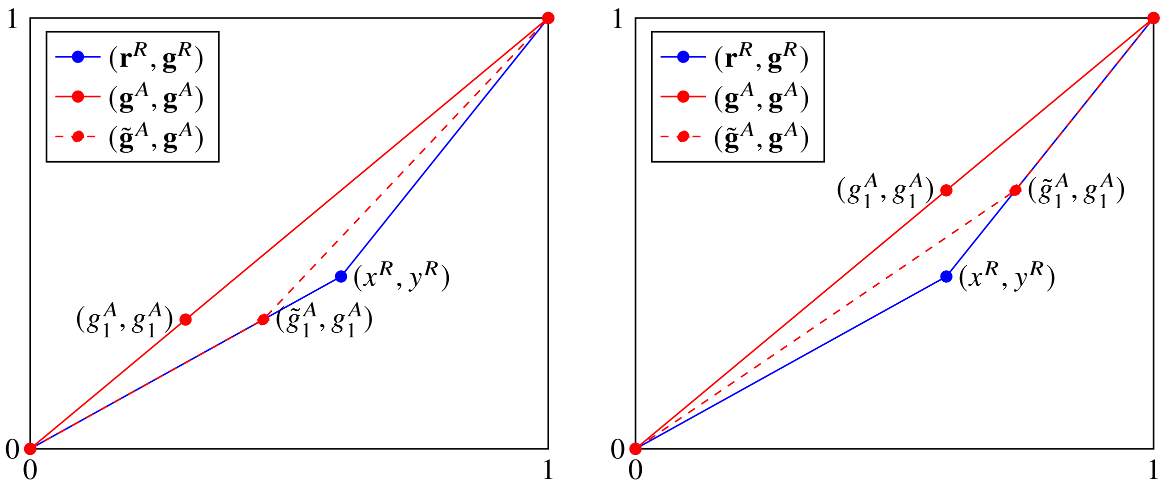
In Ref. [2], we derived an analytical expression for the maximal $\tilde{\beta}$ to which a qubit can be cooled using another resource. Orr [4], under the supervision of C. M. Scandolo and me, found the GPO that achieves the optimal qubit cooling using another qubit as a resource.

Let $(\mathbf{r}^R, \mathbf{g}^R)$ denote the probability vectors associated with the qubit resource, and let $\mathbf{g}^A$ denote the probability vector associated with the Gibbs state of the qubit that gets cooled. As shown in Figure 4.3, we can split the derivation of the GPO that achieves the optimal cooling in two scenarios: $0 \leq g_1^A < y^R$ and $y^R \leq g_1^A < 1$, where (x^R, y^R) is the only non-trivial elbow of the Lorenz curve associated with $(\mathbf{r}^R, \mathbf{g}^R)$.

Case $0 < g_1^A < y^R$.

In this case, the point $(\tilde{g}_1^A, g_1^A)$ lies on the line connecting $(0, 0)$ and (x^R, y^R), therefore

$$\tilde{g}_1^A = \frac{x^R}{y^R} g_1^A. \tag{4.52}$$



(a) $0 < g_1^A < y^R$

(b) $y^R \leq g_1^A < 1$

Figure 4.3: Cooling of $(\mathbf{g}^A, \mathbf{g}^A)$ with the resource $(\mathbf{r}^R, \mathbf{g}^R)$. The elbow of $(\mathbf{r}^R, \mathbf{g}^R)$ is denoted with (x^R, y^R). The elbow of $(\mathbf{g}^A, \mathbf{g}^A)$ is (g_1^A, g_1^A), as detailed below Eq. (4.19). The point $(\tilde{\mathbf{g}}_1^A, g_1^A)$ can either be on the first or the second segment of the Lorenz curve associated with $(\mathbf{r}^R, \mathbf{g}^R)$. It is on the first segment if $0 < g_1^A < y^R$ (see Figure a), and on the second segment if $y^R \leq g_1^A < 1$ (see Figure b).

Since $(\mathbf{r}^R, \mathbf{g}^R) \succ (\tilde{\mathbf{g}}^A, \mathbf{g}^A)$, there exists a stochastic matrix S such that $S\mathbf{r}^R = \tilde{\mathbf{g}}^A$ and $S\mathbf{g}^R = \mathbf{g}^A$. We denote the entries of S as

$$S = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}, \tag{4.53}$$

where $0 \leq a, b \leq 1$. The relative majorization conditions can be written as

$$\begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix} \begin{pmatrix} x^R \\ 1-x^R \end{pmatrix} = \begin{pmatrix} \tilde{g}_1^A \\ 1-\tilde{g}_1^A \end{pmatrix}, \qquad \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix} \begin{pmatrix} y^R \\ 1-y^R \end{pmatrix} = \begin{pmatrix} g_1^A \\ 1-g_1^A \end{pmatrix}. \tag{4.54}$$

The conditions originating from the second line of S depend linearly on the conditions from the first line. We can rewrite Eq. (4.54) as the system of equations

$$\begin{cases} ax^R + b(1-x^R) = \frac{x^R}{y^R} g_1^A, \\ ay^R + b(1-y^R) = g_1^A. \end{cases} \tag{4.55}$$

The solution of this system is $a = \frac{g_1^A}{y^R}$, $b = 0$. Thus,

$$S = \begin{pmatrix} \frac{g_1^A}{y^R} & 0 \\ 1 - \frac{g_1^A}{y^R} & 1 \end{pmatrix}. \tag{4.56}$$

Let $\mathcal{S}^{R\to A}$ denote a GPO associated with S. Since

$$S\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{g_1^A}{y^R} \\ 1 - \frac{g_1^A}{y^R} \end{pmatrix}, \text{ and } \quad S\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.57}$$

the action of $\mathcal{S}^{R\to A}$ on quasi-classical states is completely determined by

$$\begin{aligned} \mathcal{S}^{R\to A}\left(|0\rangle\langle 0|^R\right) &= \frac{g_1^A}{y^R}|0\rangle\langle 0|^A + \left(1-\frac{g_1^A}{y^R}\right)|1\rangle\langle 1|^A, \text{ and} \\ \mathcal{S}^{R\to A}\left(|1\rangle\langle 1|^R\right) &= |1\rangle\langle 1|^A. \end{aligned} \tag{4.58}$$

A channel that satisfies the conditions above is the channel with Kraus operators $K_1 = |1\rangle^A\langle 1|^R$, $K_2 = \sqrt{\frac{g_1^A}{y^R}}\,|0\rangle^A\langle 0|^R$, and $K_3 = \sqrt{1-\frac{g_1^A}{y^R}}\,|1\rangle^A\langle 0|^R$ (note that $\{\, K_i \,\}_i$ are indeed Kraus operators because they satisfy the condition $\sum_i K_i^\dagger K_i = \mathbb{1}^R$).

Case $y^R \le g_1^A < 1$.

Here we follow the same steps as the previous case. The main difference is that the point $(\tilde{\mathbf{g}}_1^A, g_1^A)$ lies on the line connecting (x^R, y^R) and $(1,1)$, therefore

$$\tilde{g}_1^A = \frac{(1-x^R)g_1^A - y^R + x^R}{1-y^R}. \tag{4.59}$$

The entries of the stochastic matrix S are the solutions of the following system of equations

$$\begin{cases} a x^R + b(1-x^R) = \frac{(1-x^R)g_1^A - y^R + x^R}{1-y^R}, \\ a y^R + b(1-y^R) = g_1^A. \end{cases} \tag{4.60}$$

That is, $a = 1$, $b = \frac{g_1^A - y^R}{1-y^R}$. Thus,

$$S = \begin{pmatrix} 1 & \frac{g_1^A - y^R}{1-y^R} \\ 0 & \frac{1-g_1^A}{1-y^R} \end{pmatrix}. \tag{4.61}$$

The action of the GPO $\mathcal{S}^{R\to A}$ associated with S on quasi-classical states is completely determined by

$$\begin{aligned} \mathcal{S}^{R\to A}\left(|0\rangle\langle 0|^R\right) &= |0\rangle\langle 0|^A, \text{ and} \\ \mathcal{S}^{R\to A}\left(|1\rangle\langle 1|^R\right) &= \frac{g_1^A - y^R}{1-y^R}|0\rangle\langle 0|^A + \frac{1-g_1^A}{1-y^R}|1\rangle\langle 1|^A. \end{aligned} \tag{4.62}$$

In this case, a channel that satisfies the conditions above is the channel with Kraus operators $K_1 = |0\rangle^A\langle 0|^R$, $K_2 = \sqrt{\frac{g_1^A - y^R}{1-y^R}}\,|0\rangle^A\langle 1|^R$, and $K_3 = \sqrt{\frac{1-g_1^A}{1-y^R}}\,|1\rangle^A\langle 1|^R$ (note that $\{\, K_i \,\}_i$ satisfy the condition $\sum_i K_i^\dagger K_i = \mathbb{1}^R$).

4.1.3 State Transformations From Cooling and Heating

Let us consider two resources, (ρ^R, γ^R) and (σ^S, γ^S), and let us denote with $\tilde{\beta}_{\max}(\rho, \gamma; \beta, E)$ the maximal inverse temperature to which the resource (ρ, γ) can cool a qubit with energy gap E, initially at equilibrium with a bath at temperature β. If $(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S)$ we can use (ρ^R, γ^R) to reach every $\tilde{\beta}$ that we can reach with (σ^S, γ^S) by first converting (ρ^R, γ^R) to (σ^S, γ^S) and then using (σ^S, γ^S) to reach $\tilde{\beta}$ (since CTO is closed under concatenation, see Proposition 2.3.2). Therefore, $\tilde{\beta}_{\max}(\rho^R, \gamma^R; \beta, E) \geq \tilde{\beta}_{\max}(\sigma^S, \gamma^S; \beta, E)$. Therefore, the function

$$C_\beta^E(\rho, \gamma) := \tilde{\beta}_{\max}(\rho, \gamma; \beta, E) - \beta \tag{4.63}$$

satisfies the condition of Definition 2.1.7 and is a resource measure. Similarly,

$$H_\beta^E(\rho, \gamma) := \beta - \tilde{\beta}_{\min}(\rho, \gamma; \beta, E) \tag{4.64}$$

is a resource measure as well. Interestingly, they are a complete family of resource measures for quasi-classical states under CTO [26, 64, 67–69, 109, 139, 174, 209–212], in the sense that, if they hold for any fixed β and every $E > 0$, then $(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S)$.

Theorem 4.1.5: A complete set of resource measures

Let (σ^S, γ^S) be quasi-classical. The following statements are equivalent:

1. $(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S)$.
2. For any fixed $\beta > 0$ and for all $E \in (0, \infty)$, it holds that

$$\begin{aligned} C_\beta^E(\rho^R, \gamma^R) &\geq C_\beta^E(\sigma^S, \gamma^S), \\ H_\beta^E(\rho^R, \gamma^R) &\geq H_\beta^E(\sigma^S, \gamma^S). \end{aligned} \tag{4.65}$$

Proof. First, we remember that according to Lem. 4.1.1, we can, without loss of generality, assume that (ρ^R, γ^R) is quasi-classical too, since both (σ^S, γ^S) and the Gibbs qubit states that we are considering are quasi-classical. As before, let thus $\mathbf{r}^R, \mathbf{g}^R, \mathbf{s}^S, \mathbf{g}^S$ denote the probability vectors corresponding to $\mathcal{P}^{R\to R}(\rho^R), \gamma^R, \sigma^S, \gamma^S$, respectively, and we can reduce our analysis to relative majorization.

We first note that *2* follows from *1* due to the monotonicity of C_β^E and H_β^E, as explained above the Theorem.

Let us next assume that *2* holds and note that this is equivalent to

$$\tilde{\beta}_{\max}(\rho^R, \gamma^R; \beta, E) \geq \tilde{\beta}_{\max}(\sigma^S, \gamma^S; \beta, E) \tag{4.66}$$

$$\tilde{\beta}_{\min}(\rho^R, \gamma^R; \beta, E) \leq \tilde{\beta}_{\min}(\sigma^S, \gamma^S; \beta, E) \tag{4.67}$$

for the (by assumption) fixed β and all $E \in (0, \infty)$. According to Eq. (4.20), when cooling a qubit, the non-trivial elbow of the lower boundary of its testing region is given by

$$(\tilde{g}_1^A, g_1^A). \tag{4.68}$$

We first consider a fixed energy gap E, i.e., a fixed $g_1^A = g_1^A(\beta, E)$, and remember that how far we can cool the qubit using (ρ^R, γ^R) is determined by the largest $\tilde{g}_1^A$ such that the elbow in Eq. (4.68) is still inside the testing region corresponding to $(\mathbf{r}^R, \mathbf{g}^R)$: Using the argument that led to Eq. (4.32), we have that

$$\alpha^{(\mathbf{r}^R, \mathbf{g}^R)}(g_1^A(\beta, E)) = \tilde{g}_1^A(\tilde{\beta}_{\max}(\rho^R, \gamma^R; \beta, E), E). \tag{4.69}$$

If for a given E, with (ρ^R, γ^R) we can cool the qubit more than with (σ^S, γ^S), we thus find

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(g_1^A) = \tilde{g}_1^A(\tilde{\beta}_{\max}(\rho^R, \gamma^R; \beta, E), E) \geq \tilde{g}_1^A(\tilde{\beta}_{\max}(\sigma^S, \gamma^S; \beta, E), E) = \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(g_1^A), \tag{4.70}$$

since $\tilde{g}_1$ increases if $\tilde{\beta}$ increases (we suppressed the dependence of g_1^A on β and E for readability). We now note that by varying the energy gap E between zero and infinity, we vary $g_1^A(\beta, E)$ between 1 and $1/2$.

Next, we consider heating: According to Eq. (4.33), the non-trivial elbow is given by

$$(\tilde{g}_2^A, g_2^A). \tag{4.71}$$

With fixed E and analogous arguments, we then find that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(g_2^A) = \tilde{g}_2^A(\tilde{\beta}_{\min}(\rho^R, \gamma^R; \beta, E), E) \geq \tilde{g}_2^A(\tilde{\beta}_{\min}(\sigma^S, \gamma^S; \beta, E), E) = \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(g_2^A), \tag{4.72}$$

where we remember that we allowed for population inversions. This time, by varying E between zero and infinity, we vary $g_2^A(\beta, E)$ between 0 and $1/2$.

Combining the heating and cooling cases, we have shown that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(y) \geq \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(y) \tag{4.73}$$

for y between 0 and 1. This implies relative majorization and, therefore, convertibility under CTO, i.e., we have shown that *1* follows from *2*.

□

The above Theorem shows that the performance in two of the most elementary thermodynamic tasks, namely cooling and heating of the simplest systems, i.e., qubits, fully determines the convertibility between and thus the relative value of quasi-classical athermality states. For a fixed (ρ^R, γ^R), (σ^S, γ^S), one only needs $|S| - 1$ monotones, corresponding to the $|S| - 1$ non-trivial elbows of $\alpha^{(\mathbf{s}^S,\mathbf{g}^S)}$. To show this, we start with the assumption that there is no $k \in [|S| - 1]$ such that $y_k^{(\mathbf{s}^S,\mathbf{g}^S)} = \frac{1}{2}$: Let (σ^S, γ^S) be quasi-classical, $\beta > 0$ fixed, we define the set of indices associated with cooling and heating, respectively, as

$$\begin{aligned} \mathcal{S}_c &:= \left\{ k \in [|S| - 1] : y_k^{(\mathbf{s}^S,\mathbf{g}^S)} > \frac{1}{2} \right\}, \\ \mathcal{S}_h &:= \left\{ k \in [|S| - 1] : y_k^{(\mathbf{s}^S,\mathbf{g}^S)} < \frac{1}{2} \right\} \end{aligned} \tag{4.74}$$

and define the $|S| - 1$ energies

$$E_k = \begin{cases} \frac{1}{\beta} \ln \frac{y_k^{(\mathbf{s}^S,\mathbf{g}^S)}}{1 - y_k^{(\mathbf{s}^S,\mathbf{g}^S)}} & \text{if } k \in \mathcal{S}_c, \\ \frac{1}{\beta} \ln \frac{1 - y_k^{(\mathbf{s}^S,\mathbf{g}^S)}}{y_k^{(\mathbf{s}^S,\mathbf{g}^S)}} & \text{if } k \in \mathcal{S}_h. \end{cases} \tag{4.75}$$

By definition, $E_k > 0$. We will now show that $(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S)$ iff

$$\begin{aligned} C_\beta^{E_k}(\rho^R, \gamma^R) &\geq C_\beta^{E_k}(\sigma^S, \gamma^S) \text{ for } k \in \mathcal{S}_c, \\ H_\beta^{E_k}(\rho^R, \gamma^R) &\geq H_\beta^{E_k}(\sigma^S, \gamma^S) \text{ for } k \in \mathcal{S}_h. \end{aligned} \tag{4.76}$$

The forward direction follows directly from Theorem 4.1.5. For the reverse, assume that Eqs. (4.76) hold. According to Eq. (4.70), this implies that for $k \in \mathcal{S}_c$,

$$\alpha(\mathbf{r}^R, \mathbf{g}^R)(y_k^{(\mathbf{s}^S,\mathbf{g}^S)}) = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(g_1^A(\beta, E_k)) \geq \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(g_1^A(\beta, E_k)) = \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(y_k^{(\mathbf{s}^S,\mathbf{g}^S)}) = x_k^{(\mathbf{s}^S,\mathbf{g}^S)}. \quad (4.77)$$

Moreover, according to Eq. (4.72), it implies that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(y_k^{(\mathbf{s}^S,\mathbf{g}^S)}) = \alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(g_2^A(\beta, E_k)) \geq \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(g_2^A(\beta, E_k)) = \alpha^{(\mathbf{s}^S,\mathbf{g}^S)}(y_k^{(\mathbf{s}^S,\mathbf{g}^S)}) = x_k^{(\mathbf{s}^S,\mathbf{g}^S)} \quad (4.78)$$

for $k \in \mathcal{S}_h$. In summary, we have thus shown that

$$\alpha^{(\mathbf{r}^R,\mathbf{g}^R)}(y_k^{(\mathbf{s}^S,\mathbf{g}^S)}) \geq x_k^{(\mathbf{s}^S,\mathbf{g}^S)} \ \forall k \in |S| - 1. \quad (4.79)$$

This implies relative majorization and, therefore, convertibility under CTO, which finishes our proof. If there exists a $k \in [|S| - 1]$ such that $y_k^\star(\mathbf{s}^S, \mathbf{g}^S) = \frac{1}{2}$, we can apply the above criteria to check convertibility to states arbitrarily close to the target state. Since CTO allows for an arbitrarily small error anyway, this is sufficient.

To check convertibility to a single quasi-classical target state, using $|S| - 1$ monotones is thus sufficient. However, which monotones we must choose depends on the specific target state (via the energies E_k given in Eq. (4.75)). For different target states, we thus need different monotones, and it is easy to see that a complete set of monotones that allows to check convertibility to an arbitrary quasi-classical target state necessarily includes all $E \in (0, \infty)$. This is not surprising since it has been shown recently that at least in the limit of infinite temperature, there cannot exist a finite complete set of monotones [139].

4.2 Alternative Characterization of State Transformations

So far, we answered the question of how much we can cool or heat specific qubits. However, in many applications, we want to reach a specific target inverse temperature $\tilde{\beta}$. A dual question of interest is, therefore, which qubits (characterized by their energy gap E) we can cool or heat to this $\tilde{\beta}$ given a resource (ρ^R, γ^R) and background inverse temperature $\beta > 0$ [107, 213]. In other words, we want to find the set $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta})$ of all energy gaps E for which we can achieve the desired transformation. Since we still investigate the heating and cooling of qubits, i.e., are concerned with quasi-classical target states, due to Lemma 4.1.1, we can again assume without loss of generality that (ρ^R, γ^R), i.e., the initial resource, is quasi-classical too and reduce our analysis to relative majorization. To this end, let $\mathbf{r}^R, \mathbf{g}^R$ be the probability vectors corresponding to $\mathcal{P}(\rho^R), \gamma^R$ respectively.

Now let A be a qubit with energy gap E. Its Gibbs state corresponding to the inverse background temperature β is denoted by γ^A and corresponds to the probability vector

$$\mathbf{g}^A = \begin{pmatrix} g_1^A \\ 1 - g_1^A \end{pmatrix}, \ g_1^A = \frac{1}{1 + e^{-\beta E}}. \quad (4.80)$$

Cooling or heating the qubit A means that we want to create the athermality state $(\tilde{\gamma}^A, \gamma^A)$, where $\tilde{\gamma}$ is the Gibbs state corresponding to the target inverse temperature $\tilde{\beta}$, i.e.,

$$\tilde{\mathbf{g}}^A = \begin{pmatrix} \tilde{g}_1^A \\ 1 - \tilde{g}_1^A \end{pmatrix}, \ \tilde{g}_1^A = \frac{1}{1 + e^{-\tilde{\beta} E}}. \quad (4.81)$$

According to Eq. (4.20), when cooling the qubit ($\tilde{\beta} > \beta$), the non-trivial elbow of the lower boundary of its testing region is given by

$$(\tilde{g}_1^A, g_1^A) \tag{4.82}$$

and when heating (see Eq. (4.33), $\tilde{\beta} < \beta$) by

$$(\tilde{g}_2^A, g_2^A). \tag{4.83}$$

Remembering that the background temperature is assumed to be finite and non-negative, i.e., $\beta > 0$, which is physically well motivated, we introduce $w := w(E) := e^{-\beta E}$, where $0 < w \leq 1$, since without loss of generality, we only consider non-negative energy gaps. With $a = \tilde{\beta}/\beta$, this allows us to rewrite the elbows as

$$\begin{aligned}(\tilde{g}_1^A, g_1^A) &= \left(\frac{1}{1+e^{-\tilde{\beta}E\frac{\beta}{\beta}}}, \frac{1}{1+e^{-\beta E}}\right) = \left(\frac{1}{1+w^a}, \frac{1}{1+w}\right), \\ (\tilde{g}_2^A, g_2^A) &= \left(\frac{w^a}{1+w^a}, \frac{w}{1+w}\right). \end{aligned} \tag{4.84}$$

In summary, the elbows of $(\tilde{\gamma}^A, \gamma^A)$ lie on the curve

$$F_a(w) = \begin{cases} \left(\frac{1}{1+w^a}, \frac{1}{1+w}\right) & \text{for } a > 1, \\ \left(\frac{w^a}{1+w^a}, \frac{w}{1+w}\right) & \text{for } a < 1. \end{cases} \tag{4.85}$$

Note that the case $a = 1$ is trivial since it corresponds to leaving the qubit in its initial equilibrium state, which is, of course, always possible, independent of E.

If $E = 0$, then $w = 1$, and $F_a(1) = \left(\frac{1}{2}, \frac{1}{2}\right)$. If instead $E \to \infty$, then $w \to 0^+$, and

$$\lim_{w \to 0^+} F_a(w) = \begin{cases} (1,1) & \text{for } a > 1, \\ (0,0) & \text{for } 0 < a < 1, \\ \left(\frac{1}{2}, 0\right) & \text{for } a = 0, \\ (1,0) & \text{for } a < 0. \end{cases} \tag{4.86}$$

For all fixed $a \neq 1$, the continuous curves describing the position of the elbows in terms of E thus start at $\left(\frac{1}{2}, \frac{1}{2}\right)$ (corresponding to $E = 0$) and go towards the boundary of the square with corners $(0,0)$, $(0,1)$, $(1,1)$, $(1,0)$ for $E \to \infty$ (see Figure 4.4).

From Figure 4.4, it is plausible that for $a > 0$, $a \neq 1$, there exist (quasi-classical) athermality states (ρ^R, γ^R) with lower boundaries of their associated testing regions that cross the curve $F_a(w)$ in multiple points. Since $F_a(w)$ determines the position of the qubit's elbow in terms of E and $\tilde{\beta}$, this will imply the following proposition.

Proposition 4.2.1: Energy gaps in cooling and heating

For every $\beta, \tilde{\beta} > 0, \beta \neq \tilde{\beta}$ there exist initial resources (ρ^R, γ^R) such that $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta})$ is *not an interval*, i.e., there exist $E_1 < E_2 < E_3$ such that $E_1, E_3 \in \mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta})$ and $E_2 \notin \mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta})$.

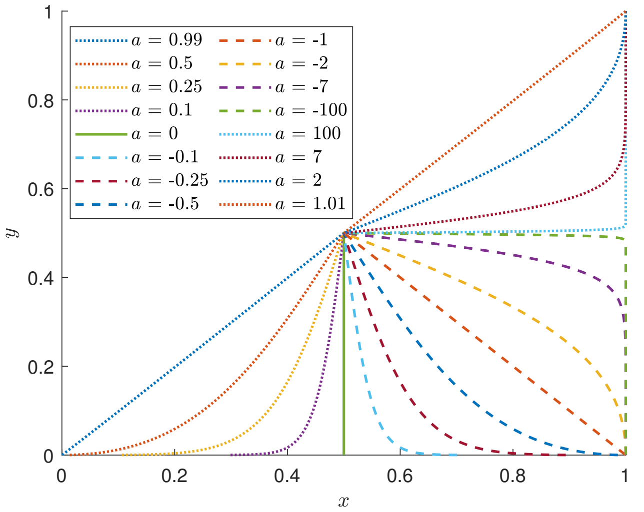


Figure 4.4: The curves $F_a(w)$ for various choices of $a \neq 1$. Independent of a, $F_a(1) = (\frac{1}{2}, \frac{1}{2})$. Curves corresponding to $a > 0$ and, therefore, finite positive target temperatures are dotted. The curves on the top right correspond to cooling, and for $a \to 1^+$, (vanishing cooling), we approach a straight line from $(\frac{1}{2}, \frac{1}{2})$ to $(0, 1)$. The curves in the lower left correspond to heating, and for $a \to 1^-$ (vanishing heating), we approach a straight line from $(\frac{1}{2}, \frac{1}{2})$ to $(0, 0)$. The special case $a = 0$, drawn as a solid straight line, corresponds to heating to infinite temperature. Negative a corresponding to population inversions in the target state are dashed. For $a \to \pm\infty$, we approach vertical lines at $x = 1$.

Proof. First, we consider the case $0 < a < 1$. Let $f_a(x)$ be the affine function that passes through $(1, 1)$ and is tangent to $F_a(w)$. As seen in Figure 4.4, this function always exists. Denoting by $(x_0, f_a(x_0))$ the tangent point, clearly $0 < x_0, f_a(x_0) < \frac{1}{2}$ and $f_a(0) < 0$. Let now $\tilde{f}_a(x)$ be the affine function passing through $(1, 1)$ and $(0, f_a(0)/2)$. By construction, this function crosses $F_a(w)$ in two points, and we denote the corresponding x-values by $x_2 < x_3$. Furthermore, $\tilde{f}_a(x)$ crosses the x-axis at $x_4 < x_2$. With $x_1 = (x_2 - x_4)/2$, the qubit athermality state with corresponding elbow $(x_1, \tilde{f}_a(x_1))$ satisfies our claims. Moreover, it can be shown that for this elbow, there is a third crossing for an $x < x_1$ (see Figure 4.5). Even if it is not obvious from Figure 4.4, this third intersection always exists (even for a arbitrarily close to 1) because the function $F_a(w)$ is flat close to the origin. Indeed, by setting $x = \frac{w^a}{1+w^a}$, we can give the y values of the curve $F_a(w)$ as a function of the x values,

$$\tilde{F}_a(x) = \frac{x^{1/a}}{(1-x)^{1/a} + x^{1/a}}. \tag{4.87}$$

Then

$$\tilde{F}'_a(x) = \frac{1}{a} \frac{1}{(1-x)^{1/a} + x^{1/a}} \left(x^{1/a-1} - x^{1/a} \frac{x^{1/a-1} - (1-x)^{1/a-1}}{(1-x)^{1/a} + x^{1/a}} \right) \tag{4.88}$$

and (remember $0 < a < 1$)

$$\lim_{x \to 0^+} \tilde{F}'_a(x) = 0, \tag{4.89}$$

from which follows the claim. Notice that this also implies that for $E \to \infty$, heating to any non-trivial target temperature becomes impossible.

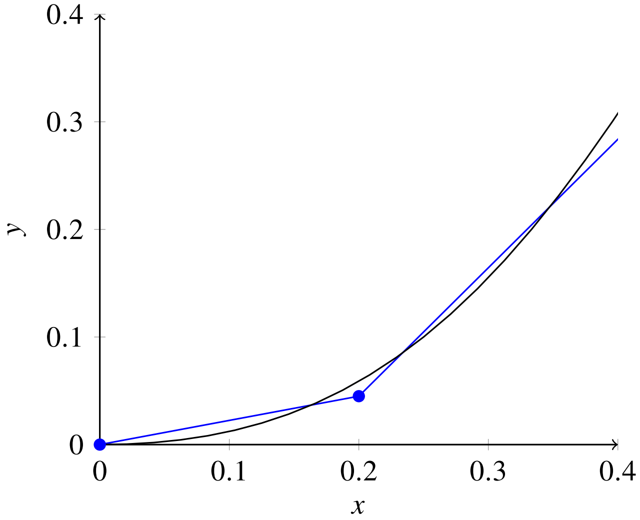


Figure 4.5: The lower boundary of the curve associated with a qubit athermality state (ρ^R, γ^R) (blue) which intersects $F_{1/2}(w)$ (black, solid) three times.

With this construction, we created a resource with only one elbow. This can either be interpreted as a qubit or as a resource with all the elbows lying on two segments. However, for a generic system R with dimension $|R| > 2$, it is also possible to construct additional intersections: Starting as in the qubit case, one determines $(x_1, \tilde{f}_a(x_1))$ and then considers the affine function that goes through that point and is tangent to $F_a(w)$ at a point $x < x_1$. Analogously, one can then construct another elbow, leading to additional intersections. From the maximal number of elbows being $|R| - 1$ follows the maximal possible number of intersections.

We now turn to the case of $a > 1$. An analogous construction is possible, now starting with the affine function that passes through $(0, 0)$ and is tangent to $F_a(w)$. $\square$

Proposition 4.2.1 may seem surprising since, naively, one might assume that if one can, e.g., heat a qubit with energy gap E_3 to $\tilde{\beta}$, one should be able to heat a qubit with energy gap $E_2 < E_3$ to $\tilde{\beta}$ too. However, the change of the expectation value of the energy of a qubit with energy gap E when cooling/heating it from β to $\tilde{\beta}$ is given by

$$\Delta \langle H(E) \rangle = \left(\frac{e^{-\tilde{\beta}E}}{1 + e^{-\tilde{\beta}E}} - \frac{e^{-\beta E}}{1 + e^{-\beta E}} \right) E, \tag{4.90}$$

which, as shown in Figure 4.6, is not a monotonic function of E. This provides an intuition for Proposition 4.2.1: If the initial resource cannot provide/absorb enough energy, heating/cooling to a given $\tilde{\beta}$ is impossible. Even if in Figure 4.6, there are no energy restrictions if $E \to \infty$, heating becomes impossible for $E \to \infty$, as discussed in the proof of Proposition 4.2.1. Indeed, it is well known that in the quantum regime, many second laws exist [174], i.e., (free) energy considerations alone are not sufficient to determine if a transformation is possible or not.

We conclude this Section by showing that the ability to cool and heat qubits to fixed temperatures again fully characterizes the conversions between quasi-classical states.

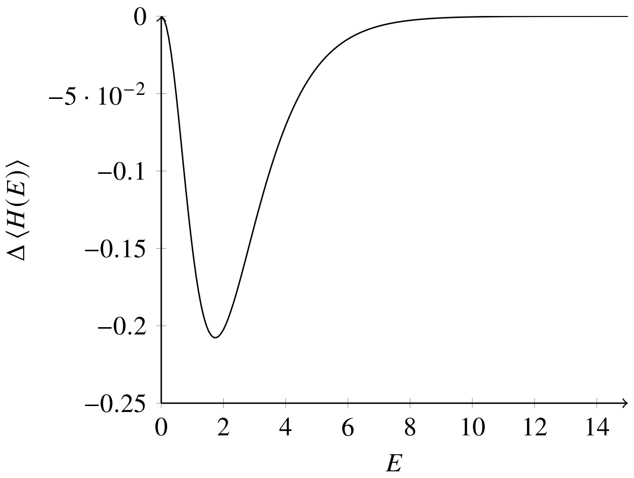


Figure 4.6: The change of the expectation value of the energy $\Delta \langle H(E) \rangle$ of a qubit when cooling it from $\beta = 1$ to $\tilde{\beta} = 2$ against its energy gap E.

Theorem 4.2.2: A generalized resoure measure for quasi-classical states under CTO

Let (σ^S, γ^S) be quasi-classical. Then $(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S)$ if and only if

$$\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}) \supseteq \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta}) \tag{4.91}$$

for any fixed $\beta > 0$ and all $\tilde{\beta} \in (-\infty, \infty)$.

Proof. One direction is again trivial. If

$$(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S) \tag{4.92}$$

and

$$(\sigma^S, \gamma^S) \xrightarrow{\text{CTO}} (\tilde{\gamma}^A, \gamma^A) \tag{4.93}$$

then also

$$(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\tilde{\gamma}^A, \gamma^A), \tag{4.94}$$

since CTO is closed under concatenation. Thus $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}) \supseteq \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta})$. Moreover, due to Lemma 4.1.1, we can again assume without loss of generality that (ρ^R, γ^R) is quasi-classical too, and identify $\mathbf{r}^R, \mathbf{g}^R, \mathbf{s}^S, \mathbf{g}^S$ with $\mathcal{P}^{R\to R}(\rho^R), \gamma^R, \sigma^S, \gamma^S$, respectively.

Assume now that $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}) \supseteq \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta})$ for a fixed $\beta > 0$ and all $\tilde{\beta} \in (-\infty, \infty)$ and remember that

$$(\rho^R, \gamma^R) \xrightarrow{\text{CTO}} (\sigma^S, \gamma^S) \tag{4.95}$$

if and only if $(\mathbf{r}^R, \mathbf{g}^R) \succ (\mathbf{s}^S, \mathbf{g}^S)$, i.e., if the testing region associated with $(\mathbf{r}^R, \mathbf{g}^R)$ contains the testing region associated with $(\mathbf{s}^S, \mathbf{g}^S)$. Now suppose that the testing region associated with $(\mathbf{s}^S, \mathbf{g}^S)$ is not contained in the testing region associated with $(\mathbf{r}^R, \mathbf{g}^R)$, i.e., that there exists at least one point (x_0, y_0) outside the testing region corresponding to $(\mathbf{r}^R, \mathbf{g}^R)$ and inside the testing region corresponding to $(\mathbf{s}^S, \mathbf{g}^S)$.

Using the notation introduced previously, we then choose a_0, w_0 such that $(x_0, y_0) = F_{a_0}(w_0)$. For $y_0 \notin \{0, 1/2\}$ and $x_0 \neq 1$, such a pair always exists, see Figure 4.4. If $y_0 \in \{0, 1/2\}$ or $x_0 = 1$, by continuity of the boundaries of the testing regions, there always exists a point (x_0', y_0') close to (x_0, y_0) that also satisfies that it is outside of the testing region corresponding to $(\mathbf{r}^R, \mathbf{g}^R)$ and inside the testing region corresponding to $(\mathbf{s}^S, \mathbf{g}^S)$ such that $y_0' \notin \{0, 1/2\}$ and $x_0' \neq 1$ and we use that point for our argument instead.

With E_0 such that $w_0 = e^{-\beta E_0}$ and $\tilde{\beta}_0$ such that $a_0 = \tilde{\beta}_0/\beta$, we then find that $E_0 \notin \mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}_0)$ and $E_0 \in \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta}_0)$. This is clearly a contradiction to the assumption that $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}) \supseteq \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta})$ for all $\tilde{\beta} \in (-\infty, \infty)$. □

Whilst the $\mathfrak{E}_\beta$ represent sets and are therefore no resource measures in the strict sense, the equation $\mathfrak{E}_\beta(\rho^R, \gamma^R; \tilde{\beta}) \supseteq \mathfrak{E}_\beta(\sigma^S, \gamma^S; \tilde{\beta})$ can be seen as a generalized resource measure [214–216] that implies an operational order, namely that (ρ^R, γ^R) can be used to cool/heat more qubits to the desired target inverse temperature $\tilde{\beta}$ than (σ^S, γ^S).

Chapter 5

Choi-Defined Resource Theories

In Chapter 3 and Chapter 4, we focused on two specific resource theories and derived results that are valid only within those theories. However, one of the advantages of having a common framework is that it allows one to expand the focus and find properties that are common to many resource theories. This is our goal for this Chapter. We start with the observation that the resource theories of separable and non-positive partial transpose entanglement [54–64], magic states [36–40], imaginarity [41–43], and non-negativity of quantum amplitudes [217] share two interesting properties:

1. A channel is free if and only if its Choi matrix (properly renormalized) is a free state,
2. Free operations coincide with the set of completely resource-non-generating operations.

We call resource theories that exhibit the first property *Choi-defined* resource theories (CDRTs), and we will show that property *2* is a necessary condition for property *1*. Being a CDRT is a valuable property: Every problem associated with free quantum channels can be converted into a problem involving free states. This leads to computational convenience, especially in convex resource theories, where many problems can be stated as conic optimization problems over the cone generated by the free states.

Given the advantages of CDRTs, one wonders when it is possible to construct such resource theories from a given set of free states. In this Chapter, we answer this question by providing necessary and sufficient conditions for the set of free states. Moreover, since the free operations in a CDRT are all and only the completely resource-non-generating operation, this result provides a constructive definition of CRNG operations and generalizes what has already been observed in the resource theories of magic states, imaginarity, NPT, SEP, and non-negativity of quantum amplitudes. Next, we introduce resource measures, a complete family of monotones, and conversion distances in CDRTs, which can all be computed with conic linear programs (CLP) [83, 218] whenever the set of free states is convex and closed, or even with semidefinite programs (SDP) [219].

5.1 The Choi Isomorphism

In this Section, we present the mathematical tool that is at the base of this Chapter: The Choi isomorphism [220]. This is an isomorphism between operators in $\mathfrak{L}(A \to B)$ and bipartite matrices in $\mathfrak{L}(BA')$, where A' is a copy of A. The bipartite matrix $M^{B:A'}$ associated with the map $\mathcal{M}^{A\to B}$ is called Choi matrix of $\mathcal{M}^{A\to B}$ and it is defined as

$$M^{B:A'} = \mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'}(\Phi^{AA'}), \tag{5.1}$$

where $\mathcal{I}^{A}$ is the identity channel, $\Phi^{AA'} = \sum_{x,y} |x\rangle\langle y|^{A} \otimes |x\rangle\langle y|^{A'}$ is the *unnormalized* Choi state on AA', and $\left\{ |x\rangle^{A} \right\}$ is an orthonormal basis for A. Here, we use the notation ':' to keep track of input and output. Moreover, $M^{B:A'}$ is positive semidefinite if and only if $\mathcal{M}^{A\to B}$ is completely positive-preserving and $\mathrm{Tr}_B\, M^{B:A'} = \mathbb{1}^{A'}$ if and only $\mathcal{M}^{A\to B}$ is trace preserving [66, 221]. From $M^{B:A'}$, one reconstructs the action of the quantum channel $\mathcal{M}^{A\to B}$ on a state ρ^{A} with the inverse Choi isomorphism:

$$\begin{aligned}\mathcal{M}^{A\to B}(\rho^{A}) &= \mathrm{Tr}_{A'A}[(M^{B:A'} \otimes \rho^{A})(\mathbb{1}^{B} \otimes \Phi^{A'A})]\\ &= \mathrm{Tr}_{A}[M^{B:A}(\mathbb{1}^{B} \otimes (\rho^{A})^{T})].\end{aligned} \tag{5.2}$$

Note that here, one needs to know which system is the input of the original channel and which is the output. If such information is missing, a matrix M^{BA} could be associated with different linear maps, as shown in Subsection 5.1.2.

Observe that, if we divide the Choi matrix $M^{B:A'}$ of the channel $\mathcal{M}^{A\to B}$ by the dimension d_A of A, we obtain a quantum state $\mu^{B:A'}$ such that

$$\mathrm{Tr}_B\, \mu^{B:A'} = \frac{1}{d_A}\mathbb{1}^{A'}. \tag{5.3}$$

Consequently, a bipartite state is the *renormalized Choi matrix* of a quantum channel if it satisfies the condition in Eq. (5.3).

We point out that, Eq. (5.1), the choice of tensoring $\mathcal{M}^{A\to B}$ with the identity channel on the right is arbitrary. One can tensor on the left and define the Choi matrix of a channel $\mathcal{M}^{A\to B}$ as:

$$M^{A':B} = (\mathcal{I}^{A'} \otimes \mathcal{M}^{A\to B})(\Phi^{A'A}). \tag{5.4}$$

In this Thesis, we choose Eq. (5.1) as the definition of the Choi matrix of the channel $\mathcal{M}^{A\to B}$, but the treatment can be adapted to the other definition with minimal effort.

5.1.1 The Link Product and Swap Tensor Product

A very useful operation between Choi matrices is the link product, defined as

$$N^{C:B'} * M^{B:A'} = \mathrm{Tr}_{B'B}[(N^{CB'} \otimes M^{BA'})(\mathbb{1}^{C} \otimes \Phi^{B'B} \otimes \mathbb{1}^{A})]. \tag{5.5}$$

Indeed, if $M^{B:A'}$ and $N^{C:B'}$ are the Choi matrices of the quantum channels $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{B\to C}$, respectively, then the matrix $T_{C:A'} = N^{C:B'} * M^{B:A'}$ is the Choi matrix of the quantum channel $\mathcal{T}_{A\to C} = \mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$ [66]. It is worth noticing that if $\mathcal{M}^{A\to B}$ is a preparation channel, i.e., $A = \mathbb{C}$ and $\mathcal{M}^{\mathbb{C}\to B}(1) = \rho^{B}$, then $M^{B:\mathbb{C}'} = \rho^{B}$ and $N^{C:B'} * M^{B:\mathbb{C}'} = \mathcal{N}^{B\to C}(\rho^{B})$.

The link product is the operation between Choi matrices associated with the sequential composition of channels. Now, we focus on parallel composition. Let $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{C\to D}$ be quantum channels, and let $M^{B:A'}$ and $N^{D:C'}$ be the Choi matrices associated with them. Let $\mathcal{T}_{AC\to BD} = \mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$. The Choi matrix of the tensor product of channels is often either overlooked or claimed to be equal to $M^{B:A'} \otimes N^{D:C'}$.

The latter approach has its merits if the order of the systems is not relevant. However, when it is crucial to keep track of input and output systems, like in this work, we expect the Choi matrix of $\mathcal{T}_{AC\to BD}$ to be a matrix $T_{BD:A'C'}$ which is not equal to $M^{B:A'} \otimes N^{D:C'}$ because the order of the systems is not the same. We define here an operation $\boxtimes$, which we call *swap tensor product*, that takes care of the order of systems:

$$M^{B:A'} \boxtimes N^{D:C'} := (\mathcal{I}^B \otimes \mathcal{S}_{A'D\to DA'} \otimes \mathcal{I}^{C'})(M^{B:A'} \otimes N^{D:C'}). \tag{5.6}$$

We show now that $M^{B:A'} \boxtimes N^{D:C'}$ is the Choi matrix $T_{BD:A'C'}$ of $\mathcal{T}_{AC\to BD} = \mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$.

$$\begin{aligned} T_{BD:A'C'} &= (\mathcal{T}_{AC\to BD} \otimes \mathcal{I}^{A'C'})(\Phi^{ACA'C'}) \\ &= [(\mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D} \otimes \mathcal{I}^{A'C'}) \circ (\mathcal{I}^A \otimes \mathcal{S}_{A'C\to CA'} \otimes \mathcal{I}^{C'})](\Phi^{AA'} \otimes \Phi^{CC'}) \\ &= [(\mathcal{I}^B \otimes \mathcal{S}_{A'D\to DA'} \otimes \mathcal{I}^{C'}) \circ (\mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'} \otimes \mathcal{N}^{C\to D} \otimes \mathcal{I}^{C'})](\Phi^{AA'} \otimes \Phi^{CC'}) \\ &= (\mathcal{I}^B \otimes \mathcal{S}_{A'D\to DA'} \otimes \mathcal{I}^{C'})(M^{B:A'} \otimes N^{D:C'}) \\ &= M^{B:A'} \boxtimes N^{D:C'}. \end{aligned} \tag{5.7}$$

5.1.2 Ambiguity of the Choi Matrix

In this Section, we demonstrate that one must be cautious when dealing with Choi matrices, as multiple linear maps can have the same Choi matrix. This is the key motivation for using rigorous formalism and notation for Choi matrices. Consider the matrix

$$M = \frac{1}{5}\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & -2 & 1 \end{pmatrix}. \tag{5.8}$$

It is Hermitian and positive semidefinite. Therefore, it could be the Choi matrix of a quantum map (CP linear operator) [66]. However, if we have no information about the bipartition of such a matrix, we can find different quantum maps with M as their Choi matrix. M could be the Choi matrix of the quantum map that prepares the supernormalized state M. Another easy choice is to consider M as the Choi matrix of the effect M^T, which acts on a 4×4 matrix N as $\mathrm{Tr}(M^T N)$.

The last choice is to consider M as the Choi matrix of a quantum map from a two-dimensional system to a two-dimensional system. In this case, there is some ambiguity as well. Let A and B be two-dimensional complex Hilbert spaces. We observe that we can write M as a linear combination of elementary tensors in $A \otimes B$:

$$M^{AB} = \frac{1}{5}|0\rangle\langle 0|^A \otimes \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}_B + \frac{1}{5}|1\rangle\langle 1|^A \otimes \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}_B. \tag{5.9}$$

As detailed in Eq. (5.1) and Eq. (5.4), there are two conventions for the definition of the Choi matrix, i.e., $M^{A:B} = (\mathcal{M}^{B'\to A} \otimes \mathcal{I}^B)(\Phi^{B'B})$ and $M^{A:B} = (\mathcal{I}^A \otimes \mathcal{M}^{A'\to B})(\Phi^{AA'})$ (we relabelled the systems to match with M^{AB}). Unsurprisingly, using different conventions, one finds different quantum maps associated with M^{AB}. If we consider the former, which is the one that we use in this Thesis, we obtain the map that acts on $|0\rangle\langle 0|^{B'}$ as

$$\mathcal{M}^{B'\to A}(|0\rangle\langle 0|^{B'}) = \frac{1}{5}\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}^A. \tag{5.10}$$

While with the latter, $|0\rangle\langle 0|^{A'}$ is mapped into

$$\mathcal{M}^{A'\to B}(|0\rangle\langle 0|^{A'}) = \frac{1}{5}\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}^{B}. \tag{5.11}$$

Sometimes, if we know that a Choi matrix is the Choi matrix of a quantum channel, we can resolve this ambiguity. Indeed, the trace of the Choi matrix of a quantum channel is equal to the dimension of the input system. Therefore, since $\operatorname{Tr} M = 2$, we can rule out the first two quantum maps associated with M, so we must have the qubit-to-qubit map because it satisfies this condition. Another thing to consider is that if M^{AB} is the Choi matrix of a quantum channel, then the marginal on the input system is the identity matrix. Consequently, we can check if either $\operatorname{Tr}_A M^{AB} = \mathbb{1}^B$ or $\operatorname{Tr}_B M^{AB} = \mathbb{1}^A$ to solve the ambiguity. We obtain

$$\begin{aligned}\operatorname{Tr}_A M &= \frac{1}{5}\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}^{B} + \frac{1}{5}\begin{pmatrix}4 & -2\\ -2 & 1\end{pmatrix}^{B} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}^{B},\\ \operatorname{Tr}_B M &= |0\rangle\langle 0|^A + |1\rangle\langle 1|^A = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}^{A}.\end{aligned} \tag{5.12}$$

This test is inconclusive for M since both marginals are the identity matrix. Therefore, even if we know that we have the Choi matrix of a quantum channel, we still have the ambiguity about the convention used to compute it.

This indicates that a significant amount of information is required to reconstruct a quantum channel from its Choi matrix. It is not enough to know just the entries of a Choi matrix; one needs to know the dimensions of the input and output systems and their order in the Choi matrix. While the dimensions of the systems can always be deduced when dealing with quantum channels, the same cannot be said for the order of the systems. If one is fortunate, one can deduce it by computing partial traces. However, this is not always the case, as demonstrated for M.

We point out that if we consider a *renormalized* Choi matrix instead of working with a Choi matrix, we lose even the information about the dimension of the input system because all renormalized Choi matrices have trace one. For example, let μ^{ABC} be a normalized quantum state such that $\operatorname{Tr}_A \mu^{ABC} = \frac{1}{d_B d_C}\mathbb{1}^{BC}$, then also $\operatorname{Tr}_{AB} \mu^{ABC} = \frac{1}{d_C}\mathbb{1}^{C}$, which implies that both

$$\begin{aligned}&\mathcal{M}^{BC\to A}(\rho^{BC}) =\\ &\quad \operatorname{Tr}_{BCB'C'}[(d_B d_C \mu^{ABC} \otimes \rho^{B'C'})(\mathbb{1}^A \otimes \Phi^{BCB'C'})],\end{aligned} \tag{5.13}$$

and

$$\mathcal{M}^{C\to AB}(\rho^{C}) = \operatorname{Tr}_{CC'}[(d_C \mu^{ABC} \otimes \rho^{C'})(\mathbb{1}^{AB} \otimes \Phi^{CC'})] \tag{5.14}$$

are quantum channels associated with μ^{ABC}.

5.2 Definition and Properties of Choi-Defined Resource Theories

Remember that a resource theory provides a partition of all quantum channels into free and resourceful ones with the following properties (see Section 2.1):

1. the identity channel is free,
2. the swap channel is free,
3. discarding a system is free,
4. sequential composition of free channels is free,
5. parallel composition of free channels is free.

Moreover, a state is free if it can be prepared with free operations. If one applies a free channel to a free state, one obtains a free state, meaning that obtaining a resource from free objects is impossible. This is sometimes called the golden rule of resource theories [84]. Consequently, every set of free states satisfies the following properties: It is closed under tensor product, partial tracing, and system swapping.

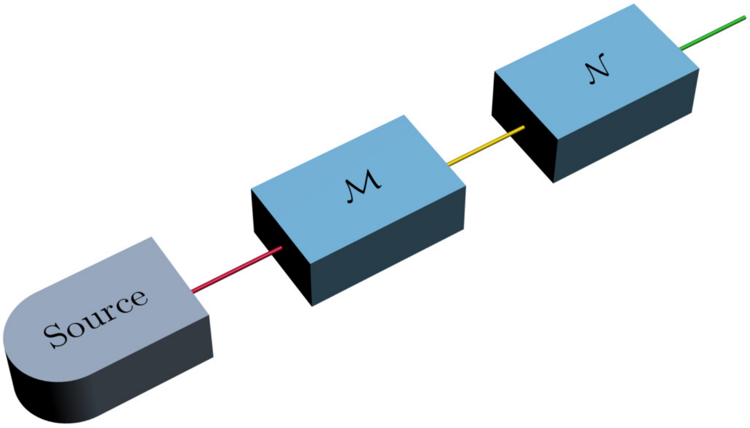


Figure 5.1: Sequential composition of operations. A source emits a quantum system, e.g., a photon (coloured in red). A first operation $\mathcal{M}$ is performed on this quantum system and produces a new quantum system as output. A second operation $\mathcal{N}$ is performed on the new quantum system. The overall operation done on the first (red) quantum system is the sequential composition of $\mathcal{M}$ and $\mathcal{N}$, denoted with $\mathcal{N} \circ \mathcal{M}$. The sequential composition is translated into the link product when dealing with Choi matrices.

Remark 5.2.1

One of the constraints of the resource theory of bipartite entanglement is that the two physically separated agents cannot exchange quantum systems. This may appear to conflict with the condition that the swap channel is a free operation. However, in this scenario, the systems in question are bipartite systems [85], i.e., $S_1 := A_1B_1$ and $S_2 := A_2B_2$. The physical separation is, as usual, between systems labelled with A and systems labelled with B. The swap channel only swaps the order of the two bipartite systems S_1 and S_2, i.e., $A_1B_1A_2B_2 \leftrightarrow A_2B_2A_1B_1$. If we organize the systems according to the physical separation, the action of the swap channel is $(A_1A_2)(B_1B_2) \to (A_2A_1)(B_2B_1)$. Therefore, the two agents are *locally* exchanging the order of their own quantum systems, and there is no exchange of quantum systems between them.

As noted at the beginning of this Chapter, the resource theories of magic states, imaginarity, SEP, and NPT share an interesting property: A channel is free if and only if its renormalized Choi matrix is a free state. Now, we generalize such resource theories by introducing the construction of Choi-defined operations and Choi-defined resource theories.

Definition 5.2.2: Choi-defined operations

The *Choi-defined* (CD) operations associated with a set of free states are all and only the quantum channels such that their renormalized Choi matrix is a free state.

With this definition, we are ready to define the Choi-defined resource theories.

Definition 5.2.3: Choi-defined resource theory

A quantum resource theory is a *Choi-defined resource theory* (CDRT) if its free operations coincide with the Choi-defined operations associated with its set of free states.

Not all quantum resource theories are CDRTs. For example, in the resource theory of athermality [33, 34], there is only one free state in $\mathfrak{F}(BA')$ but many free operations in $\mathfrak{F}(A \to B)$. Therefore, there is no one-to-one correspondence between free states and free channels. A more interesting example is the resource theory of entanglement. In this case, the free states are the separable states, but the construction of CD operations produces *all* separable operations, which are a larger set than LOCC [222, 223]. As a result, the resource theory of LOCC entanglement is not a CDRT.

It is a natural question to characterize when, given a set of free states, the construction of a CDRT is allowed. In every well-defined resource theory, the identity is a free operation. As a consequence, its renormalized Choi state $\frac{1}{d_A}\Phi^{AA'}$ is free in every CDRT.

Remark 5.2.4

The renormalized Choi state is *not* the maximally entangled state in the resource theories of entanglement, SEP entanglement, or NPT entanglement. Indeed, in these resource theories, each system is a pair of physical systems [85], i.e., $S := AB$, and $S' := A'B'$, and the spatial separation is between systems AA' and BB'. Let $\{\,|a\rangle^A\,\}$ and $\{\,|b\rangle^B\,\}$ be orthogonal bases for A and B, respectively, then

$$\frac{1}{d_S}\Phi^{SS'} = \frac{1}{d_A d_B}\sum_{a,b,\tilde{a},\tilde{b}} |ab\rangle\!\langle\tilde{a}\tilde{b}|^{AB} \otimes |ab\rangle\!\langle\tilde{a}\tilde{b}|^{A'B'}. \tag{5.15}$$

If the systems are reorganized according to the spatial separation, the renormalized Choi state becomes

$$\begin{aligned} &\frac{1}{d_A d_B}\sum_{a,\tilde{a}} |aa\rangle\!\langle\tilde{a}\tilde{a}|^{AA'} \otimes \sum_{b,\tilde{b}} |bb\rangle\!\langle\tilde{b}\tilde{b}|^{BB'} \\ &= \frac{1}{d_A}\Phi^{AA'} \otimes \frac{1}{d_B}\Phi^{BB'}. \end{aligned} \tag{5.16}$$

This shows that the state $\frac{1}{d_S}\Phi^{SS'}$ is separable with respect to the spatial separation between AA' and BB'. Therefore, it is free in the resource theories mentioned above, so it is *not* the maximally entangled state on $(AA')(BB')$.

When we require that resource theories be closed under the sequential composition of free channels (Figure 5.1), we find a less trivial condition for the set of free states. In Ref. [66, 221], the authors introduced an operation between bipartite matrices called link product, defined as

$$\begin{aligned} &N^{C:B'} * M^{B:A'} \\ &= \mathrm{Tr}_{BB'}[(N^{C:B'} \otimes M^{B:A'})(\mathbb{1}^C \otimes \Phi^{B'B} \otimes \mathbb{1}^{A'})] \\ &= \mathrm{Tr}_B[(N^{C:B} \otimes \mathbb{1}^{A'})(\mathbb{1}^C \otimes (M^{B:A})^{T_B})], \end{aligned} \tag{5.17}$$

where $\cdot^{T_B}$ is the partial transpose on system B. This operation translates the sequential composition of channels into an operation between Choi matrices. Thanks to the link product, we can restate the golden rule of resource theories as a condition on the set of free states. That is, the condition that $\mathcal{M}^{A\to B}(\rho^A)$ be free whenever $\mathcal{M}^{A\to B}$ and ρ^A are free becomes that $d_A\mu^{B:A'} * \rho^A$ be a free state if ρ^A and $\mu^{B:A'}$ are free states and $\mu^{B:A'}$ is the renormalized Choi matrix of a quantum channel.

We now state our first main result.

Theorem 5.2.5: Necessary and sufficient conditions for CDRTs construction

It is possible to construct a CDRT associated with a set of free states if and only if for all systems A and B

1. the state $\frac{1}{d_A}\Phi^{AA'}$ is free,
2. if ρ^A and $\mu^{B:A'}$ are free states, and $\mu^{B:A'}$ is the renormalized Choi matrix of a quantum channel, then $d_A\mu^{B:A'} * \rho^A$ is a free state.

Proof. First, we prove the necessary conditions. Assume that the free states and the free operations form a well-defined CDRT.

1. The identity channel is a free operation in every resource theory. Its Choi matrix is $(\mathcal{I}^A \otimes \mathcal{I}^{A'})(\Phi^{AA'}) = \Phi^{AA'}$, therefore the renormalized Choi matrix $\frac{1}{d_A}\Phi^{AA'}$ is a free state.
2. Let $\mathcal{M}^{A\to B}$ be the CD operation associated with $\mu^{B:A'}$. Then, by Definition 5.2.3, $\mathcal{M}^{A\to B}$ is free. Therefore, $d_A \mu^{B:A'} * \rho^A = \mathcal{M}^{A\to B}(\rho^A)$ is free, because both $\mathcal{M}^{A\to B}$ and ρ^A are free.

We now prove that conditions 1 and 2 are sufficient for a well-defined resource theory when the free operations are the Choi-defined operations. Recall that we assume that the set of free states under consideration is compatible with a minimal resource theory, i.e., closed under tensor product, partial tracing, and system swapping. Now, we demonstrate that the five conditions for a resource theory listed in Section 2.1 are satisfied.

1. The identity channel is free. Indeed, since $\frac{1}{d_A}\Phi^{AA'}$ is free for all A, then the CD operation associated with it, the identity channel on A, is free.
2. The swap channel is free. Indeed, from condition 1, one has that $\frac{1}{d_A d_B}\Phi^{ABA'B'}$ is a free state. Since the set of free states is closed under system swapping, one obtains that $\frac{1}{d_A d_B}\Phi^{BA:A'B'} := \frac{1}{d_A d_B}(\mathcal{S}^{AB\to BA} \otimes \mathcal{I}^{A'B'})(\Phi^{ABA'B'})$ is free too. This state is, by definition, the renormalized Choi matrix of the swap channel, which is, therefore, free.
3. Discarding a system is free. Indeed, since $\frac{1}{d_A}\Phi^{AA'}$ is free for all A and the set of free states is closed under partial tracing, then $(\mathrm{Tr}_A \otimes \mathcal{I}^{A'})(\frac{1}{d_A}\Phi^{AA'}) = \frac{1}{d_A}\mathbb{1}^A$ is free. Once again, this is, by definition, the Choi matrix of the discarding channel on A, which is therefore free.
4. Parallel composition of free channels is free. Indeed, let $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{C\to D}$ be free channels, and let $\mu^{BA'}$ and $\nu^{DC'}$ be their renormalized Choi matrices, respectively. The state $\tau^{BDA'C'} = (\mathcal{I}^B \otimes \mathcal{S}_{A'D\to DA'} \otimes \mathcal{I}^{C'})(\mu^{BA'} \otimes \nu^{DC'})$, where $\mathcal{S}_{A'D\to DA'}$ is the swap channel, is free because the set of free states is closed under tensor product and system swapping. As shown in Appendix 5.1.1, the CD operation associated with $\tau^{BDA'C'}$ is $\mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$, which is therefore free. This proves that the parallel composition of channels is free.
5. Sequential composition of free channels is free. Let $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{B\to C}$ be free channels, let $\mu^{B:A'}$ and $\nu^{C:B'}$ be their renormalized Choi matrices. These Choi matrices are free states. We point out that condition 2 does not immediately imply that $d_B \nu^{C:B'} * \mu^{B:A'}$ is free. Indeed, condition 2 only applies when the system of the second state in the link product matches the second system in the bipartition of the first state. In the case of the link product between $\nu^{C:B'}$ and $\mu^{B:A'}$, BA' is not a copy of B'. Here, we want to construct a free state $\tilde{\nu}_{CA':B'A}$ such that $d_B d_A \tilde{\nu}_{CA':B'A} * \mu^{BA'}$ is the renormalized Choi matrix of $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$. If such construction is possible, then we can apply condition 2 of the Theorem and the definition of Choi defined resource theories to deduce that $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$ is free. Let $\tilde{\nu}_{CA':AB'}$ be the renormalized Choi matrix of $\mathcal{N}^{B\to C} \otimes \mathcal{I}^{A'}$, that is $\tilde{\nu}^{CA':B'A} = (\mathcal{I}^C \otimes \mathcal{S}_{B'A'\to A'B'} \otimes \mathcal{I}^A)(\nu^{C:B'} \otimes \frac{1}{d_A}\Phi^{A':A})$. Such a state satisfies the conditions above:

- It is free. Indeed, both $\nu^{CB'}$ and $\frac{1}{d_A}\Phi^{A'A}$ are free, and the set of free states is closed under tensor product and system swapping.
- $d_B d_A \tilde{\nu}_{CA':B'A} * \mu^{BA'}$ is the renormalized Choi matrix of $\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$:

$$\begin{aligned} d_B d_A \tilde{\nu}_{CA':B'A} * \mu^{BA'} &= (\mathcal{N}^{B\to C} \otimes \mathcal{I}^{A'})(\mu^{BA'}) \\ &= \frac{1}{d_A}(\mathcal{N}^{B\to C} \otimes \mathcal{I}^{A'}) \circ (\mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'})(\Phi^{AA'}) \\ &= \frac{1}{d_A}\left[\left(\mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}\right) \otimes \mathcal{I}^{A'}\right](\Phi^{AA'}). \end{aligned} \tag{5.18}$$

□

We observe that point 3 of the proof of the sufficient conditions implies the following Corollary.

Corollary 5.2.6: Maximally mixed state

In every CDRT, the maximally mixed state is free.

An interesting common property of the examples of CDRTs presented in the introduction is that their free operations coincide with the CRNG operations [39, 40, 43, 55, 56, 64]. As we show in the following Theorem, this is a general feature of every CDRT.

Theorem 5.2.7: CRNG operations

In every CDRT, free operations coincide with CRNG operations.

Proof. In every resource theory, free operations are CRNG (see, e.g., Ref. [84]). Suppose $\mathcal{M}^{A\to B}$ is a CRNG operation. Since $\mathcal{M}^{A\to B}$ is CRNG and $\frac{1}{d_A}\Phi^{AA'}$ is free as a consequence of Theorem 5.2.5, then $(\mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'})(\frac{1}{d_A}\Phi^{AA'})$ is a free state. This state is the renormalized Choi matrix of $\mathcal{M}^{A\to B}$. Therefore, it follows from Definition 5.2.3 that $\mathcal{M}^{A\to B}$ is a CD operation. □

Notably, this Theorem provides an easy construction for CRNG operations when the set of free states satisfies the conditions of Theorem 5.2.5. Two examples of resource theories that are not CDRTs, and for which Theorem 5.2.7 does not hold, are the resource theory of athermality [224] and the resource theory of k-unextendibility [225]. In athermality, the CRNG operations are the Gibbs-preserving operations, but their renormalized Choi matrices are not free states. In k-unextendibility, an attempt to construct a CDRT from k-extendible states was made in Ref. [226], but the resulting operations were not resource-non-generating. Indeed, k-extendible states do not satisfy condition 1 of Theorem 5.2.5.

In the next Subsection, we present resource theories that satisfy the conditions of Theorem 5.2.5. We start with the resource theory of asymmetry, where, to the best of our knowledge, there is no explicit result yet in this direction [44, 45].

5.2.1 Examples of CDRTs

Resource Theories of Asymmetry

In the resource theory of asymmetry [44], two parties have access to the same quantum systems but do not share a reference frame. The relation between reference frames is described by an element g of a compact group G determined by the symmetries of the problem. Free states are those that the two parties are certain to describe in the same way, that is, G-invariant states. In other words, as state ρ^A is free if $\mathcal{U}_g^A(\rho^A) = U_g^A \rho^A (U_g^A)^\dagger = \rho^A$ for all $g \in G$, where U_g^A is the unitary representation of $g \in G$ on the Hilbert space $\mathcal{H}_A$. Note that the unitary representation on the product space $\mathcal{H}_A \otimes \mathcal{H}_B$ is $U_g^{AB} = U_g^A \otimes U_g^B$. Similarly, a channel $\mathcal{M}^{A\to B}$ is free if it is G-covariant, that is, $\mathcal{U}_g^B \circ \mathcal{M}^{A\to B} \circ (\mathcal{U}_g^A)^{-1} = \mathcal{M}^{A\to B}$ for all $g \in G$.

Theorem 5.2.8: CDRTs of asymmetry

A resource theory of asymmetry is a CDRT if and only if the renormalized Choi state is free.

Proof. Necessity is trivial because of Theorem 5.2.5. For sufficiency, we have only to show that the set of free states is closed under the link product whenever the Choi state is free. To this end, let $\mu^{BA'}$ and ρ^A be free states such that $d_{A'} \operatorname{Tr}_B \mu^{BA'} = \mathbb{1}^{A'}$. Then,

$$\begin{aligned}
\mathcal{U}_g^B(d_A \mu^{B:A'} * \rho^A) &= \mathcal{U}_g^B(d_A \operatorname{Tr}_{A'A}[(\mu^{BA'} \otimes \rho^A)(\mathbb{1}^B \otimes \Phi^{A'A})]) \\
&= \mathcal{U}_g^B(d_A \operatorname{Tr}_{A'A}\{[(\mathcal{U}_g^{BA'A})^{-1}(\mu^{BA'} \otimes \rho^A)](\mathbb{1}^B \otimes \Phi^{A'A})\}) \\
&= \mathcal{U}_g^B(d_A \operatorname{Tr}_{A'A}\{[((\mathcal{U}_g^B)^{-1} \otimes (\mathcal{U}_g^{A'A})^{-1})(\mu^{BA'} \otimes \rho^A)] \\
&\quad (\mathbb{1}^B \otimes \Phi^{A'A})\}) \\
&= d_A \operatorname{Tr}_{A'A}\{[\mathcal{I}^B \otimes (\mathcal{U}_g^{A'A})^{-1}(\mu^{BA'} \otimes \rho^A)](\mathbb{1}^B \otimes \Phi^{A'A})\} \\
&= d_A \operatorname{Tr}_{A'A}\{(\mu^{BA'} \otimes \rho^A)[\mathbb{1}^B \otimes \mathcal{U}_g^{A'A}(\Phi^{A'A})]\} \\
&= d_A \operatorname{Tr}_{A'A}[(\mu^{BA'} \otimes \rho^A)(\mathbb{1}^B \otimes \Phi^{A'A})] \\
&= d_A \mu^{B:A'} * \rho^A.
\end{aligned} \tag{5.19}$$

□

In addition, the Choi state is G-invariant if and only if all the unitary representations are real in the Choi basis. To prove this, it is easier to work with the Choi vector $|\phi\rangle^{AA'} = \sum_j |jj\rangle^{AA'}$ rather than with the Choi matrix $\Phi^{AA'} = |\phi\rangle\langle\phi|^{AA'}$. Indeed, for all matrices M_A, we have

$$\begin{aligned}(M_A \otimes \mathbb{1}^{A'})\,|\phi\rangle^{AA'} &= (M_A \otimes \mathbb{1}^{A'}) \sum_j |jj\rangle^{AA'} \\ &= \sum_{j,k,l} M^{k,l} (|k\rangle\langle l|^A \otimes \mathbb{1}^{A'})\,|jj\rangle^{AA'} \\ &= \sum_{j,k} M^{k,j}\,|kj\rangle^{AA'} \\ &= \sum_{j,k,l} M^{l,j} (\mathbb{1}^A \otimes |j\rangle\langle l|^{A'})\,|kk\rangle^{AA'} \\ &= (\mathbb{1}^A \otimes M^T_{A'}) \sum_k |kk\rangle^{AA'} \\ &= (\mathbb{1}^A \otimes M^T_{A'})\,|\phi\rangle^{AA'}\,.\end{aligned} \tag{5.20}$$

It is straightforward to see that if U_g^A is real for every $g \in G$, then $(U_g \otimes U_g)\,|\phi\rangle^{AA'} = (\mathbb{1}^A \otimes U_g^A (U_g^A)^T)\,|\phi\rangle^{AA'} = |\phi\rangle^{AA'}$ and therefore $\Phi^{AA'}$ is G-invariant. If instead we assume that $|\phi\rangle^{AA'}$ is G-invariant, then for all $g \in G$

$$\begin{aligned}|\phi\rangle^{AA'} &= (U_g^A \otimes U_g^{A'})\,|\phi\rangle^{AA'} \\ &= (\mathbb{1}^A \otimes U_g^{A'} (U_g^{A'})^T)\,|\phi\rangle^{AA'}\,.\end{aligned} \tag{5.21}$$

As a consequence,

$$\begin{aligned}|j\rangle^{A'} &= (\langle j|^A \otimes \mathbb{1}^{A'})\,|\phi\rangle^{AA'} \\ &= (\langle j|^A \otimes U_g^{A'} (U_g^{A'})^T)\,|\phi\rangle^{AA'} \\ &= U_g^{A'} (U_g^{A'})^T\,|j\rangle^{A'}\end{aligned} \tag{5.22}$$

for all elements of the Choi basis. This proves that $U_g^{A'} (U_g^{A'})^T = \mathbb{1}^{A'}$, and therefore the representation is real in the Choi basis.

These results are similar to those presented in the Supplementary Information of Ref. [227] about completely symmetry-preserving operations. This similarity is not surprising because CD and CRNG operations coincide in CDRTs as a consequence of Theorem 5.2.7. In the context of asymmetry, an example of a CDRT is the resource theory of parity generated by the group $\mathbb{Z}_2$ [44].

Resource Theory of Magic States

The resource theory of magic states [36–40, 228] or non-stabilizer quantum computation is based on the Gottesman-Knill theorem [21], which states that quantum circuits composed of Clifford group gates, measurements of Pauli group operators and classically conditioned Clifford group operations acting on a finite number, n, of qubits initially in their ground states can be efficiently simulated classically. As a reminder, the Pauli group on n qubits, $\mathcal{P}_n$, is the group generated by the tensor products of the single-qubit Pauli matrices $\{ \mathbb{1}_2, \sigma_x, \sigma_y, \sigma_z \}$, up to overall phases $\{ \pm 1, \pm i \}$. For example, $\mathcal{P}_1 = \{ \pm\mathbb{1}_2, \pm i\mathbb{1}_2, \pm\sigma_x, \pm i\sigma_x, \pm\sigma_y, \pm i\sigma_y, \pm\sigma_z, \pm i\sigma_z \}$. The Clifford group on n qubits is the set of unitary transformations that normalize $\mathcal{P}_n$, that is, the set containing all the unitary transformations U such that $U\mathcal{P}_n U^\dagger = \mathcal{P}_n$. The operations listed in this paragraph are called *stabilizer* operations. The free states in the resource theory of magic are the states that can be prepared with a circuit composed solely of stabilizer operations.

We prove now that the set of magic states satisfies the conditions of Theorem 5.2.5. For the first condition, we denote the number of qubits in A with n and the elements of the computational basis with $|x_1 \dots x_n\rangle$, where $x_j = 0, 1$. The normalized Choi state $|\phi\rangle = \frac{1}{2^{n/2}} \sum_{x_1 \dots x_n} |x_1 \dots x_n\rangle |x_{1'} \dots x_{n'}\rangle$ is obtained by preparing the n entangled states $\frac{1}{\sqrt{2}} \sum_{x_j} \left|x_j x_{j'}\right\rangle$ and by swapping the order of the systems. The j-th 2-qubit entangled state can be obtained with a Hadamard and a CNOT gate acting on $\left|0^j 0^{j'}\right\rangle$, and the swap channels can be decomposed into three CNOT gates (see, e.g., Ref. [11]). All these operations are stabilizer operations; therefore, the renormalized Choi state $\frac{1}{d_A}\Phi^{AA'}$ is a stabilizer state. This proves that condition 1 of Theorem 5.2.5 is satisfied.

We now consider the closure under the link product of the set of stabilizer states. Let $\mu^{BA'}$ and ρ^A be free states such that $d_{A'} \operatorname{Tr}_B \mu^{BA'} = \mathbb{1}^{A'}$, where A is a n-qubit system and B is a m-qubit system. This implies that there exists a protocol to create $\mu^{BA'} \otimes \rho^A$ for free. By Eq. (5.2.5), the state $d_A \mu^{B:A'} * \rho^A$ is obtained by a protocol with postselection:

1. Prepare $\mu^{BA'} \otimes \rho^A$.
2. Perform a measurement in the Bell basis for each of the pairs of qubits in $A'A$
3. Postselect on the measurement outcome associated with $\frac{1}{d_A}\Phi^{A'A}$. After the postselection, the system B is in the state $d_A \mu^{B:A'} * \rho^A$, as requested.

Each of the steps of this protocol is a free operation. Indeed, the measurement in the Bell basis can be converted into a measurement in the computational basis by adding Hadamard and CNOT gates, which are stabilizer operations. Therefore the state $d_A \mu^{B:A'} * \rho^A$ is free, as requested by Theorem 5.2.5. The set of stabilizer states satisfies all the conditions of Theorem 5.2.5, and the CDRT constructed from it is the resource theory of magic states [40].

Resource Theory of Imaginarity

We now focus on the resource theory of imaginarity [41–43, 228]. This resource theory aims to evaluate the role of imaginarity in quantum mechanics. A state is free if its density matrix, expressed with respect to a fixed basis, is real. States with density matrices containing non-real numbers are resources. The Choi state $\frac{1}{d_A}\Phi^{AA'}$ relative to the fixed basis has only real coefficients and, therefore, is free. This proves condition 1 of Theorem 5.2.5. Condition 2 follows trivially because it can be expressed using only multiplications and traces of real matrices, from which it is possible to obtain only real matrices. This proves that it is possible to construct a CDRT from the set of real matrices, and it coincides with the resource theory of imaginarity, where the free operations are all the CRNG operations [43].

Resource Theory of Separable Entanglement

In the theory of entanglement [23, 27, 122, 124, 125, 229], systems are bipartite systems, and free states are separable states. We have already shown in Remark 5.2.4 that the Choi state is separable and, therefore, free. We now prove that the separable states are closed under the link product. To align with the convention used in entanglement-related resource theories, we denote the first pair of systems with $S_1 := A_1 B_1$ and the second with $S_2 := A_2 B_2$. Let $\mu^{S_2 S_1'}$ be a separable state, i.e., there exists density matrices $\mu_j^{A_2 A_1'}$ and $\tilde{\mu}_j^{B_2 B_1'}$

such that $\mu^{A_2A_1'B_2B_1'} = \sum_j p_j\mu_j^{A_2A_1'} \otimes \tilde{\mu}^j_{B_2B_1'}$, where $p_j \geq 0$ and $\sum_j p_j = 1$. Moreover, let $\mu^{S_2S_1'}$ be such that $d_{S_1'} \operatorname{Tr}_{S_2} \mu^{S_2S_1'} = \mathbb{1}^{S'}$. Let $\rho^{S_1} = \sum_k q_k\rho_k^{A_1} \otimes \tilde{\rho}_k^{B_1}$, where $q_k \geq 0$ and $\sum_k q_k = 1$, be another separable state. Then

$$\begin{aligned} \mu^{S_2:S_1'} * \rho^{S_1} &= \operatorname{Tr}_{(A_1B_1)}[\mu^{(A_2B_2)(A_1B_1)}(\mathbb{1}^{(A_2B_2)} \otimes (\rho^{(A_1B_1)})^T)] \\ &= \sum_{j,k} p_jq_k\sigma_{j,k}^{A_2} \otimes \tilde{\sigma}_{j,k}^{B_2}. \end{aligned} \tag{5.23}$$

where $\sigma_{j,k}^{A_2} := \operatorname{Tr}_{A_1}\{\mu_j^{A_2A_1}[\mathbb{1}^{A_2} \otimes (\rho_k^{A_1})^T]\}$, and $\tilde{\sigma}_{j,k}^{B_2} := \operatorname{Tr}_{B_1}\{\tilde{\mu}_j^{B_2B_1}[\mathbb{1}^{B_2} \otimes (\tilde{\rho}_k^{B_1})^T]\}$. This proves that $d_{S_1}\mu^{S_2S_1'} * \rho^{S_1}$ is a separable state whenever $\mu^{S_2S_1'}$ and ρ^{S_1} are separable, therefore condition 2 of Theorem 5.2.5 is satisfied. This implies that it is possible to construct a CDRT from the set of separable states, which coincides with SEP [55, 56].

Resource Theory of NPT Entanglement

Looking again at the resource theory of entanglement, one notices that all separable states have positive partial transpose [57, 58]. One can thus consider the bipartite states with positive partial transpose to be free states. This set is strictly larger than the set of separable states due to bound entanglement [59], which is of great relevance in quantum information [57–63]. Since the set of PPT states contains separable states, it automatically satisfies condition 1 of Theorem 5.2.5. Let $\mu^{(A_2B_2)(A_1B_1)'}$ and $\rho^{(A_1B_1)}$ be free states such that $d_{(A_1B_1)'} \operatorname{Tr}_{(A_2,B_2)} \mu^{(A_2B_2)(A_1B_1)'} = \mathbb{1}^{(A_1B_1)'}$. We write $(\mu^{(A_2B_2)(A_1B_1)'} * \rho^{(A_1B_1)})^{T_{B_2}}$ as

$$\begin{aligned} &(\mu^{(A_2B_2):(A_1B_1)'} * \rho^{(A_1B_1)})^{T_{B_2}} \\ &\quad= \operatorname{Tr}_{A_1B_1}((\mu^{(A_2B_2)(A_1B_1)})^{T_{B_2}}(\mathbb{1}^{(A_2B_2)} \otimes (\rho^{(A_1B_1)})^T)) \\ &\quad= \operatorname{Tr}_{A_1B_1}[(\mu^{(A_2B_2)(A_1B_1)})^{T_{B_2B_1}}(\mathbb{1}^{(A_2B_2)} \otimes (\rho^{(A_1B_1)})^{T_{A_1}})]. \end{aligned} \tag{5.24}$$

Since $\rho^{(A_1B_1)}$ has positive partial transpose, there exists and eigenbasis $\left\{ |\varphi_j\rangle^{(A_1B_1)} \right\}$ and positive eigenvalues λ_j such that $(\rho^{(A_1B_1)})^{T_{A_1}} = \sum_j \lambda_j|\varphi_j\rangle\langle\varphi_j|^{(A_1B_1)}$. This implies that, for all $|\psi\rangle^{(A_2B_2)}$,

$$\begin{aligned} &\left\langle\psi^{(A_2B_2)}\right|(\mu^{(A_2B_2):(A_1B_1)'} * \rho^{(A_1B_1)})^{T_{B_2}}\left|\psi^{(A_2B_2)}\right\rangle = \\ &\sum_j \lambda_j \left\langle\psi^{(A_2B_2)} \otimes \varphi_j^{(A_1B_1)}\right|(\mu^{(A_2B_2)(A_1B_1)})^{T_{B_2B_1}}\left|\psi^{(A_2B_2)} \otimes \varphi_j^{(A_1B_1)}\right\rangle \\ &\quad\geq 0, \end{aligned} \tag{5.25}$$

where the last inequality follows from the fact that $(\mu^{(A_2B_2)(A_1B_1)})^{T_{B_2B_1}} \geq 0$, as it is free, and $\lambda_i > 0$. This proves that $d_A(\mu^{(A_2B_2):(A_1B_1)'} * \rho^{(A_1B_1)})^{T_{B_2}} \geq 0$, and that the set of states with positive partial transpose is closed under link product. In this case, as well, we have shown that the set of free states satisfies the conditions of Theorem 5.2.5. The CDRT constructed coincides with the theory of NPT [60–63].

Resource Theory of Non-Negativity of Quantum Amplitudes

In the resource theory of non-negativity of quantum amplitudes [217], there is a fixed orthonormal basis $\{\,|j\rangle^A\,\}$ for every system A. For consistency, the orthonormal basis on composite systems AB is $\{\,|jk\rangle^{AB}\,\}$, where $\{\,|j\rangle^A\,\}$ and $\{\,|k\rangle^B\,\}$ are the fixed orthonormal basis for systems A and B, respectively. In the resource theory of non-negativity of quantum amplitudes, a normalized pure state $|\psi\rangle^A$ is free if and only if, up to a global phase, it is a positive linear combination of elements of the fixed basis, i.e., if and only if there exists θ such that $|\psi\rangle^A = e^{i\theta}\sum_j a_j\,|j\rangle^A$, $a_j \geq 0$ for all j, and $\sum_j a_j^2 = 1$. If $|\psi\rangle^A$ satisfies this condition we write $|\psi\rangle^A \geq 0$. From this definition, it immediately follows that $\frac{1}{\sqrt{d_A}}\sum_j |jj\rangle^{AA'}$ is a free pure state. In other words, the renormalized Choi state $\frac{1}{d_A}\Phi^{AA'}$ is free.

A mixed state is free, if it is a convex combination of free pure states, i.e., ρ^A is free if and only if $\rho^A = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|^A$, where $|\psi_\alpha\rangle^A \geq 0$, $p_\alpha \geq 0$ for all j, and $\sum_\alpha p_\alpha = 1$. Note that a free mixed state expressed in the orthonormal basis has the form

$$\sum_{\alpha,k,l} p_\alpha a_k^\alpha a_l^\alpha |k\rangle\langle l|^A, \tag{5.26}$$

where $p_\alpha \geq 0$, $a_j^\alpha \geq 0$, and $\sum_\alpha p_\alpha = \sum_j (a_j^\alpha)^2 = 1$.

To show that the set of free states satisfies the conditions of Theorem 5.2.5, we need to show that $d_A\mu^{B:A'} * \rho^A$ is free whenever $\mu^{B:A'}$ and ρ^A are free and $d_A \operatorname{Tr}_B \mu^{BA} = \mathbb{1}^A$. Let $\mu^{BA} = \sum_{\alpha,j,k,l,m} p_\alpha a_{j,k}^\alpha a_{l,m}^\alpha |jk\rangle\langle lm|^{BA}$, with $p_\alpha \geq 0$, $a_{j,k}^\alpha \geq 0$, $\sum_\alpha p_\alpha = \sum_{j,k}(a_{j,k}^\alpha)^2 = 1$, and let $\rho^A = \sum_{\beta,x,y} q_\beta b_x^\beta b_y^\beta |x\rangle\langle y|^A$, with $q_\beta \geq 0$, $b_x^\beta \geq 0$ and $\sum_\beta q_\beta = \sum_x (b_x^\beta)^2 = 1$. Then,

$$\begin{aligned}
d_A\mu^{B:A'} * \rho^A &= d_A \operatorname{Tr}_A[\mu^{BA}(\mathbb{1}^B \otimes (\rho^A)^T)] \\
&= d_A \sum_{\alpha,\beta,j,k,l,m,x,y} \operatorname{Tr}_A[p_\alpha a_{j,k}^\alpha a_{l,m}^\alpha |jk\rangle\langle lm|^{BA} \\
&\qquad (\mathbb{1}^B \otimes q_\beta b_x^\beta b_y^\beta |x\rangle\langle y|^A)] \\
&= d_A \sum_{\alpha,\beta,j,k,l,m} p_\alpha a_{j,k}^\alpha a_{l,m}^\alpha q_\beta b_k^\beta b_m^\beta |j\rangle\langle l|^B \\
&= d_A \sum_{\alpha,\beta,j,l} p_\alpha q_\beta \Big(\sum_k a_{j,k}^\alpha b_k^\beta\Big)\Big(\sum_m a_{l,m}^\alpha b_m^\beta\Big)|j\rangle\langle l|^B.
\end{aligned} \tag{5.27}$$

This expression suggests the definition of the coefficients

$$\begin{aligned}
r_{\alpha,\beta} &:= d_A p_\alpha q_\beta, \\
c_j^{\alpha,\beta} &:= \sum_k a_{j,k}^\alpha b_k^\beta.
\end{aligned} \tag{5.28}$$

With such coefficients, the expression above becomes

$$d_A\mu^{B:A'} * \rho^A = \sum_{\alpha,\beta,j,l} r_{\alpha,\beta} c_j^{\alpha,\beta} c_l^{\alpha,\beta} |j\rangle\langle l|^B. \tag{5.29}$$

However, these coefficients do not satisfy the desired conditions $\sum_{\alpha,\beta} r_{\alpha,\beta} = \sum_j (c_j^{\alpha,\beta})^2 = 1$. If one defines the renormalized coefficients

$$\begin{aligned}
s_{\alpha,\beta} &:= r_{\alpha,\beta}\sum_k (c_k^{\alpha,\beta})^2, \\
d_j^{\alpha,\beta} &:= \frac{c_j^{\alpha,\beta}}{\sqrt{\sum_k (c_k^{\alpha,\beta})^2}},
\end{aligned} \tag{5.30}$$

it immediately follows that

$$\begin{aligned} d_A\mu^{B:A'} * \rho^A &= \sum_{\alpha,\beta,j,l} r_{\alpha,\beta} c_j^{\alpha,\beta} c_l^{\alpha,\beta} |j\rangle\!\langle l|^B \\ &= \sum_{\alpha,\beta,j,l} s_{\alpha,\beta} d_j^{\alpha,\beta} d_l^{\alpha,\beta} |j\rangle\!\langle l|^B, \end{aligned} \tag{5.31}$$

and

$$\sum_j (d_j^{\alpha,\beta})^2 = 1. \tag{5.32}$$

What is left to show is that $\sum_{\alpha,\beta} s_{\alpha,\beta} = 1$. To this end, notice that $d_A \operatorname{Tr}_B \mu^{BA} = \mathbb{1}^A$ implies

$$\begin{aligned} \delta_{k,l} &= \langle k \,|\, d_A \operatorname{Tr}_B \mu^{BA} \,|\, l \rangle_A \\ &= \sum_j d_A \langle jk \,|\, \mu^{BA} \,|\, jl \rangle_{BA} \\ &= \sum_j d_A \langle jk| \Big(\sum_{\alpha,x,y,w,z} p_\alpha a_{x,y}^\alpha a_{w,z}^\alpha |xy\rangle\!\langle wz|^{BA} \Big) |jl\rangle^{BA} \\ &= \sum_{\alpha,j} d_A p_\alpha a_{j,k}^\alpha a_{j,l}^\alpha. \end{aligned} \tag{5.33}$$

Therefore

$$\begin{aligned} \sum_{\alpha,\beta} s_{\alpha,\beta} &= \sum_{\alpha,\beta} r_{\alpha,\beta} \sum_j c_j^{\alpha,\beta} c_j^{\alpha,\beta} \\ &= \sum_{\alpha,\beta} d_A p_\alpha q_\beta \sum_{j,k,l} a_{j,k}^\alpha b_k^\beta a_{j,l}^\alpha b_l^\beta \\ &= \sum_{\beta,k,l} q_\beta b_k^\beta b_l^\beta \sum_{\alpha,j} d_A p_\alpha a_{j,k}^\alpha a_{j,l}^\alpha \\ &= \sum_{\beta,k,l} q_\beta b_k^\beta b_l^\beta \delta_{k,l} \\ &= \sum_\beta q_\beta \sum_k b_k^\beta b_k^\beta \\ &= \sum_\beta q_\beta \\ &= 1. \end{aligned} \tag{5.34}$$

The state $d_A\mu^{B:A'} * \rho^A$ can be written as a convex sum of positive pure vectors and is therefore free. The CDRT generated from this set of pure states is the resource theory of non-negativity of quantum amplitudes.

5.3 Optimization Problems in CDRTs

In CDRTs, every optimization problem over the set of free channels can be converted into an optimization problem over the set of free states. That is,

$$\max_{\mathcal{M}^{A\to B} \in \mathfrak{F}(A\to B)} f(\mathcal{M}^{A\to B}) \tag{5.35}$$

becomes

$$\max_{\substack{\mu^{BA'}\in\mathfrak{F}(BA'),\\ d_{A'}\operatorname{Tr}_B\mu^{BA'}=\mathbb{1}^{A'}}} \tilde{f}(\mu^{BA'}). \tag{5.36}$$

Here, f is a function from the set of quantum channels to the $\mathbb{R}$, $\mathfrak{F}(A \to B)$ and $\mathfrak{F}(BA')$ are the sets of free channels from A to B and free states on BA' and $\tilde{f}$ is the composition of f with the inverse Choi isomorphism defined in Eq. (5.2). It satisfies $\tilde{f}(\mu^{BA'}) = f(\mathcal{M}^{A\to B})$ whenever $\mu^{BA'}$ is the renormalized Choi matrix of $\mathcal{M}^{A\to B}$. Moreover, if the set of free states is convex and $\tilde{f}$ is linear, the optimization problem can be expressed as a CLP (see Subsection 5.3.3 for details) or even as an SDP.

If one has a set of free states that does not satisfy the conditions of Theorem 5.2.5, one can still consider the CRNG operations. However, in this case, there is no universal way of expressing the constraints in optimization problems as a single constraint in terms of free states.

In the following Subsections, we will introduce several quantities that can be computed with optimization problems like the one in Eq. (5.36) and the conic linear programs to solve them.

5.3.1 Quantifying Resources in CDRTs

CDRTs are constructed from a given set of free states that satisfies the conditions of Theorem 5.2.5. Consequently, they inherit all the resource measures that depend only on the set of free states. Examples of such resource measures are robustness, generalized robustness, and all the measures based on quantum divergences (see Ref. [84] for a review). Among these, the max-relative entropy [230] is particularly interesting in the context of CDRTs. It is defined as

$$D_{\max}(\rho^A\|\sigma^A) = \log\min\left\{\lambda \mid \rho^A \le \lambda\sigma^A\right\}, \tag{5.37}$$

when $\operatorname{supp}\rho^A \subseteq \operatorname{supp}\sigma^A$ and $+\infty$ otherwise. The resource measure based on the max-relative entropy is

$$D^{\mathfrak{F}}_{\max}(\rho^A) = \inf_{\omega^A\in\mathfrak{F}(A)} D_{\max}(\rho^A\|\omega^A). \tag{5.38}$$

The extensions of these quantities to quantum channels are [231, 232]

$$\begin{aligned} D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) &= \\ &\sup_{\rho^{RA}} D_{\max}(\mathcal{I}^R\otimes\mathcal{M}^{A\to B}(\rho^{RA})\|\mathcal{I}^R\otimes\mathcal{N}^{A\to B}(\rho^{RA})), \\ D^{\mathfrak{F}}_{\max}(\mathcal{M}^{A\to B}) &= \inf_{\mathcal{N}^{A\to B}\in\mathfrak{F}(A\to B)} D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}). \end{aligned} \tag{5.39}$$

Notably, the max-relative entropy for quantum channels satisfies [233]

$$D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) = D_{\max}(\mu^{BA'}\|\nu^{BA'}), \tag{5.40}$$

where $\mu^{BA'}$ and $\nu^{BA'}$ are the renormalized Choi matrices of $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{A\to B}$, respectively. In a CDRT, this property is particularly relevant because it allows us to write the optimization problem for $D^{\mathfrak{F}}_{\max}(\mathcal{M}^{A\to B})$ as

$$D^{\mathfrak{F}}_{\max}(\mathcal{M}^{A\to B}) = \inf_{\substack{\omega^{BA'}\in\mathfrak{F}(BA'),\\ d_{A'}\operatorname{Tr}_B\omega^{BA'}=\mathbb{1}^{A'}}} D_{\max}(\mu^{BA'}\|\omega^{BA'}). \tag{5.41}$$

Notice that, in every CDRT, $D_{\max}^{\mathfrak{F}}$ is finite for both states and channels because the set of free states always contains a full rank state, see Corollary 5.2.6. The same is not true for all resource theories.

As a direct consequence of the results in Ref. [233], the resource measures based on the max-relative entropy satisfy the following inequality.

Proposition 5.3.1: Inequality for the max-relative entropy of states and channels

For all states ρ^A and channels $\mathcal{M}^{A\to B}$

$$D_{\max}^{\mathfrak{F}}(\mathcal{M}^{A\to B}(\rho^A)) \leq D_{\max}^{\mathfrak{F}}(\rho^A) + D_{\max}^{\mathfrak{F}}(\mathcal{M}^{A\to B}). \tag{5.42}$$

Proof. We copy Eq. (52) and Eq. (85) of Ref. [233], adapted to our notation, which will be useful for the next step. Eq. (52) is the definition of $D_{\max}^{A}$:

$$\begin{aligned} &D_{\max}^{A}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) := \\ &\sup_{\rho^{RA},\sigma^{RA}} [D_{\max}(\mathcal{I}^R\otimes\mathcal{M}^{A\to B}(\rho^{RA})\|\mathcal{I}^R\otimes\mathcal{N}^{A\to B}(\sigma^{RA})) \\ &\qquad - D_{\max}(\rho^{RA}\|\sigma^{RA})]. \end{aligned} \tag{5.43}$$

Eq. (85) shows the relationship between $D_{\max}^{A}$ and $D_{\max}$:

$$D_{\max}^{A}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) = D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}). \tag{5.44}$$

These equations imply that:

$$\begin{aligned} D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) &= \sup_{\rho^{RA},\sigma^{RA}} [D_{\max}(\mathcal{I}^R\otimes\mathcal{M}^{A\to B}(\rho^{RA})\|\mathcal{I}^R\otimes\mathcal{N}^{A\to B}(\sigma^{RA})) - D_{\max}(\rho^{RA}\|\sigma^{RA})] \\ &\geq D_{\max}(\mathcal{I}^R\otimes\mathcal{M}^{A\to B}(\tau_R\otimes\rho^A)\|\mathcal{I}^R\otimes\mathcal{N}^{A\to B}(\tau_R\otimes\omega^A)) - D_{\max}(\tau_R\otimes\rho^A\|\tau_R\otimes\omega^A) \\ &= D_{\max}(\tau_R\otimes\mathcal{M}^{A\to B}(\rho^A)\|\tau_R\otimes\mathcal{N}^{A\to B}(\omega^A)) - D_{\max}(\tau_R\otimes\rho^A\|\tau_R\otimes\omega^A) \\ &= D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\mathcal{N}^{A\to B}(\omega^A)) - D_{\max}(\rho^A\|\omega^A), \end{aligned} \tag{5.45}$$

where the last equality follows from the additivity of the max-relative entropy, and ρ^A, ω^A, and τ_R are arbitrary quantum states. As a consequence, for every quantum state ρ^A and ω^A and every channel $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{A\to B}$, the following holds

$$\begin{aligned} &D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) + D_{\max}(\rho^A\|\omega^A) \\ &\qquad \geq D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\mathcal{N}^{A\to B}(\omega^A)). \end{aligned} \tag{5.46}$$

By taking the infimum on both sides over all free channels $\mathcal{N}^{A\to B}$ and free states ω^A, one obtains:

$$\begin{aligned} &\inf_{\mathcal{N}^{A\to B}\in\mathfrak{F}(A\to B)} D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B}) \\ &\qquad + \inf_{\omega^A\in\mathfrak{F}(A)} D_{\max}(\rho^A\|\omega^A) \geq \\ &\inf_{\substack{\mathcal{N}^{A\to B}\in\mathfrak{F}(A\to B),\\ \omega^A\in\mathfrak{F}(A)}} D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\mathcal{N}^{A\to B}(\omega^A)), \end{aligned} \tag{5.47}$$

where the infimum on the left-hand side splits because $D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B})$ does not depend on ω^A and $D_{\max}(\rho^A\|\omega^A)$ does not depend on $\mathcal{N}^{A\to B}$. Moreover, since both $\mathcal{N}^{A\to B}$ and ω^A are free, then $\mathcal{N}^{A\to B}(\omega^A)$ is free as well, and therefore

$$\begin{aligned}&\inf_{\substack{\mathcal{N}^{A\to B}\in\mathfrak{F}(A\to B),\\ \omega^A\in\mathfrak{F}(A)}} D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\mathcal{N}^{A\to B}(\omega^A))\\ &\qquad\geq \inf_{\omega^B\in\mathfrak{F}(B)} D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\omega^B).\end{aligned}\tag{5.48}$$

Eq. (5.47) and Eq. (5.48) imply

$$\begin{aligned}&\inf_{\mathcal{N}^{A\to B}\in\mathfrak{F}(A\to B)} D_{\max}(\mathcal{M}^{A\to B}\|\mathcal{N}^{A\to B})+\\ &\inf_{\omega^A\in\mathfrak{F}(A)} D_{\max}(\rho^A\|\omega^A)\geq \inf_{\omega^B\in\mathfrak{F}(B)} D_{\max}(\mathcal{M}^{A\to B}(\rho^A)\|\omega^B).\end{aligned}\tag{5.49}$$

This is equivalent to

$$D_{\max}^{\mathfrak{F}}(\mathcal{M}^{A\to B})+D_{\max}^{\mathfrak{F}}(\rho^A)\geq D_{\max}^{\mathfrak{F}}(\mathcal{M}^{A\to B}(\rho^A)).\tag{5.50}$$

□

Since in a CDRT $D_{\max}^{\mathfrak{F}}$ is always finite and computable with a CLP, this inequality, valid for all resource theories, is particularly meaningful in the context of CDRTs. When we interpret $D^{\mathfrak{F}}$ as the value of a resource, it implies that the value of the output of a quantum operation applied to a quantum state is always less than the sum of the values of the operation and the input state. Equivalently, for every ρ^A, σ^B, and $\mathcal{M}^{A\to B}$ such that $\sigma^B=\mathcal{M}^{A\to B}(\rho)$, the cost of preparing σ^B is always less or equal than the cost of preparing ρ^A and then converting it to σ^B with the channel $\mathcal{M}^{A\to B}$. An inequality similar to the one in Eq. (5.42) based on the quantum relative entropy and its regularization was derived in Ref. [234].

5.3.2 Resource Conversion in CDRTs

The formalism of CDRTs provides significant insights into all quantities connected with resource conversion or manipulation via free operations. Indeed, at the core of every resource theory, there is a preorder of resources defined as $\rho^A\succ^{\mathfrak{F}}\sigma^B$ if there exists a free channel $\mathcal{M}^{A\to B}$ such that $\sigma^B=\mathcal{M}^{A\to B}(\rho^A)$. In a CDRT, this condition is equivalent to requiring the existence of free states μ^{BA}, such that $\mathrm{Tr}_B\,\mu^{BA}=\frac{1}{d_A}\mathbb{1}^A$ and $\sigma^B=d_A\mu^{B:A'}*\rho^A$.

Consequently, many quantities or results involving free channels can be redefined in terms of the link product of free states. A first example is the complete family of monotones defined in Ref. [68] for convex and closed resource theories, which in the case of a convex and closed CDRT can be expressed as

$$f_{\tau_B}(\rho^A)=\max_{\substack{\omega^{BA}\in\mathfrak{F}(BA),\\ d_A\,\mathrm{Tr}_B\,\omega^{BA}=\mathbb{1}^A}} \mathrm{Tr}_{BA}[\omega^{BA}(\tau_B\otimes(\rho^A)^T)].\tag{5.51}$$

Being a complete family of monotones, it satisfies $f_{\tau_B}(\rho^A)\geq f_{\tau_B}(\sigma^B)$ for all density matrices τ_B if and only if ρ^A can be converted to σ^B with a free operation.

As a second example, we present conversion distances in CDRTs (see Section 2.1). These distances estimate how close to a target state an initial state can be converted using only free operations. In a CDRT, they satisfy

$$\begin{aligned} D(\rho^A \to \sigma^B) &= \inf_{\mathcal{M}^{A\to B}\in\mathfrak{F}(A\to B)} D(\sigma^B, \mathcal{M}^{A\to B}(\rho^A)) \\ &= \inf_{\substack{\mu^{BA'}\in\mathfrak{F}(BA'),\\ d_{A'}\operatorname{Tr}_B \mu^{BA'}=\mathbb{1}^{A'}}} D(\sigma^B, d_A\mu^{B:A'} * \rho^A), \end{aligned} \tag{5.52}$$

where $D(\cdot,\cdot)$ is a function on quantum states that satisfies the data processing inequality [84].

5.3.3 Conic Linear Programs and Semidefinite Programs

At the beginning of this Section, we have shown that in a CDRT, an optimization problem over free channels can be expressed as an optimization problem over free states. That is,

$$\max_{\mathcal{M}^{A\to B}\in\mathfrak{F}(A\to B)} f(\mathcal{M}^{A\to B}) \tag{5.53}$$

becomes

$$\max_{\substack{\mu^{BA'}\in\mathfrak{F}(BA'),\\ d_{A'}\operatorname{Tr}_B \mu^{BA'}=\mathbb{1}^{A'}}} \tilde{f}(\mu^{BA'}). \tag{5.54}$$

As a reminder, f is a function from the set of quantum channels to the $\mathbb{R}$, $\mathfrak{F}(A\to B)$ and $\mathfrak{F}(BA')$ are the sets of free channels from A to B and free states on BA' and $\tilde{f}$ is the composition of f with the inverse Choi isomorphism defined in Eq. (5.2). It satisfies $\tilde{f}(\mu^{BA'}) = f(\mathcal{M}^{A\to B})$ whenever $\mu^{B:A'}$ is the renormalized Choi matrix of $\mathcal{M}^{A\to B}$. As in the case of the max-relative entropy (Section 5.3.1), it is often possible to find simpler expressions for $\tilde{f}$ that do not rely on the inverse Choi isomorphism.

If the set of free states is convex and $\tilde{f}$ is linear, the optimization problem can be expressed as a conic linear program (CLP) or even as a semidefinite program (SDP). Sometimes, these optimization problems are hard to solve when the dimension of the systems involved is large, for example, in the case of separable entanglement [235, 236]. However, they are still useful for systems of low dimensions.

Here, we present the techniques used to write CLPs and provide some examples. In this Thesis, we use the notation presented in Ref. [83], which is particularly useful in quantum information. In our setting, a *primal* conic linear program is an optimization problem that can be expressed as

$$\begin{aligned} \text{Find} \quad & \alpha := \inf \operatorname{Tr}[XH_1] \\ \text{subject to} \quad & \mathcal{N}(X) - H_2 \in \mathfrak{K}_2 \\ & X \in \mathfrak{K}_1 \end{aligned} \tag{5.55}$$

where $\mathfrak{K}_1$ is a closed convex cone in V_1, a subspace of the Hermitian matrix on a Hilbert space $\mathcal{H}_1$. Similarly, $\mathfrak{K}_2 \subseteq V_2 \subseteq \operatorname{Herm}(\mathcal{H}_2)$. Moreover, $\mathcal{N} : V_1 \to V_2$ is a linear map, and $H_1 \in V_1$, $H_2 \in V_2$. We say that $X \in \mathfrak{K}_1$ is a *feasible* solution if $\mathcal{N}(X) - H_2 \in \mathfrak{K}_2$. If no feasible solution exists, then α is set to $+\infty$. A feasible solution is *optimal* if $\operatorname{Tr}[XH_1] = \alpha$.

The *dual* of the conic linear program in Eq. (5.55) is

$$\begin{aligned} \text{Find} \quad & \beta := \sup \operatorname{Tr}[Y H_2] \\ \text{subject to} \quad & H_1 - \mathcal{N}^\dagger(Y) \in \mathfrak{K}_1^* \\ & Y \in \mathfrak{K}_2^*, \end{aligned} \tag{5.56}$$

where $\mathfrak{K}^*$ is the dual cone of $\mathfrak{K}$, and $\mathcal{N}^\dagger$ is the adjoint of $\mathcal{N}$. As in the case above, $Y \in \mathfrak{K}_2^*$ is a dual feasible solution if $H_1 - \mathcal{N}^\dagger(Y) \in \mathfrak{K}_1^*$. If no dual feasible solution exists, then $\beta = -\infty$. A dual feasible solution is optimal if $\operatorname{Tr}[Y H_2] = \beta$. The dual CLP is often relevant because it provides a lower bound for the primal CLP, that is, $\alpha \geq \beta$. This relation is called weak duality. Strong duality happens if $\alpha = \beta$.

In this Subsection, $V_1 \subseteq \operatorname{Herm}(\mathcal{H}_1)$ will always be the vector space spanned by block-diagonal Hermitian matrices, where the dimensions of the blocks are fixed. We will denote such matrices as $X = (X_1, \ldots, X_n)$. Each of the matrices X_i is a Hermitian matrix on an orthogonal subspace A_i of $\mathcal{H}_1$. As a consequence, $V_1 \cong \oplus_{i=1}^n \operatorname{Herm}(A_i)$. An analogous statement holds for V_2.

CLPs for a Complete Family of Monotones in CDRTs

We start with the complete family of monotones $f_{\tau_B}(\rho^A) = \max_{\substack{\omega^{BA} \in \mathfrak{F}(BA), \\ d_A \operatorname{Tr}_B \omega^{BA} = \mathbb{1}^A}} \operatorname{Tr}_{BA}[\omega^{BA}(\tau_B \otimes (\rho^A)^T)]$, where τ_B and ρ^A are density matrices. To compute one of such monotones with a conic linear program, the first step is to define the cone generated by the free states.

$$\mathfrak{K}_1 = \left\{ (\lambda, \lambda\omega^{BA}) \mid \lambda \geq 0, \omega^{BA} \in \mathfrak{F}(BA) \right\} \subseteq V_1 \cong \mathbb{R} \oplus \operatorname{Herm}(BA). \tag{5.57}$$

The conditions $\omega^{BA} \in \mathfrak{F}(BA)$ and $d_A \operatorname{Tr}_B \omega^{BA} = \mathbb{1}^A$ are equivalent to $(\lambda, \lambda\omega^{BA}) \in \mathfrak{K}_1$ and $d_A \operatorname{Tr}_B \lambda\omega^{BA} - \lambda\mathbb{1}^A = 0$. The last condition is implemented in a CLP by defining $\mathcal{N}[(x, X^{BA})] = d_A \operatorname{Tr}_B X^{BA} - x\mathbb{1}^A$, $H_2 = 0$, and $\mathfrak{K}_2 = \left\{ 0^A \right\} \subseteq V_2 = \operatorname{Herm}(A)$. The conic linear program is

$$\begin{aligned} \text{Find} \quad & -f_{\tau_B}(\rho^A) := \min \operatorname{Tr}[(x, X^{BA})(0, -\tau_B \otimes \rho_A^T)] \\ \text{subject to} \quad & \mathcal{N}[(x, X^{BA})] = d_A \operatorname{Tr}_B X^{BA} - x\mathbb{1}^A = 0^A \\ & (x, X^{BA}) \in \mathfrak{K}_1. \end{aligned} \tag{5.58}$$

Now, we turn our attention to the dual of the CLP above. Since $H_2 = 0^A$, then β (the result of the dual CLP) is 0 if a dual feasible solution exists or $-\infty$ if no dual feasible solution exists. In the former case, $-f_{\tau_B}(\rho^A) = \alpha \geq \beta = 0$, and this would imply that $f_{\tau_B}(\rho^A) = 0$ because $f_{\tau_B}(\rho^A)$ is positive by definition. Interestingly, in every CDRT, we can exclude this case because $\frac{1}{d_A d_B}\mathbb{1}^{BA}$ is free, and therefore

$$f_{\tau_B}(\rho^A) = \max_{\substack{\omega^{BA} \in \mathfrak{F}(BA), \\ d_A \operatorname{Tr}_B \omega^{BA} = \mathbb{1}^A}} \operatorname{Tr}_{BA}[\omega^{BA}(\tau_B \otimes (\rho^A)^T)] \geq \frac{1}{d_A d_B} \operatorname{Tr}_{BA}[\tau_B \otimes (\rho^A)^T] = \frac{1}{d_B d_A} \tag{5.59}$$

This implies that in every CDRT $\beta = -\infty$, and weak duality does not provide a meaningful bound for $f_{\tau_B}(\rho^A)$. However, writing the dual of the CLP in Eq. (5.58) is still useful. The first step is to characterize the dual cones of $\mathfrak{K}_1$ and $\mathfrak{K}_2$:

$$\begin{aligned} \mathfrak{K}_1^* &= \left\{ (x, X^{BA}) \in V_1 \mid \left\langle (x, X^{BA}) \,\middle|\, (\lambda, \lambda\omega^{BA}) \right\rangle_{HS} \geq 0,\ \forall \lambda \geq 0,\ \forall \omega^{BA} \in \mathfrak{F}_{BA} \right\} \\ &= \left\{ (x, X^{BA}) \in V_1 \mid \left\langle (x, X^{BA}) \,\middle|\, (1, \omega^{BA}) \right\rangle_{HS} \geq 0,\ \forall \omega^{BA} \in \mathfrak{F}_{BA} \right\}, \\ \mathfrak{K}_2^* &= \left\{ X^A \in V_2 \mid \left\langle X^A \,\middle|\, 0^A \right\rangle_{HS} \geq 0 \right\} = V_2, \end{aligned} \tag{5.60}$$

where $\langle \cdot \mid \cdot \rangle_{HS}$ is the Hilbert-Schmidt inner product defined as $\langle X \mid Y \rangle_{HS} = \operatorname{Tr}[X^\dagger Y]$, i.e.,

$$\left\langle (x, X^{BA}) \,\middle|\, (y, Y^{BA}) \right\rangle_{HS} = xy + \operatorname{Tr}_{BA}[X^{BA} Y^{BA}]. \tag{5.61}$$

$\mathcal{N}^\dagger : V_2 \longrightarrow V_1$ is the unique map that satisfies

$$\left\langle (x, X^{BA}) \,\middle|\, \mathcal{N}^\dagger(X^A) \right\rangle_{HS} = \left\langle \mathcal{N}[(x, X^{BA})] \,\middle|\, X^A \right\rangle_{HS} \tag{5.62}$$

for all $(x, X^{BA}) \in V_1 \cong \mathbb{R} \oplus \operatorname{Herm}(BA)$ and $X^A \in V_2 = \operatorname{Herm}(A)$. We denote the components of $\mathcal{N}^\dagger$ with $\mathcal{N}_1^\dagger$ and $\mathcal{N}_2^\dagger$, that is, $\mathcal{N}^\dagger(X^A) = (\mathcal{N}_1^\dagger(X^A), \mathcal{N}_2^\dagger(X^A))$. To find the expression of $\mathcal{N}^\dagger$, we evaluate Eq. (5.62) for each member of the basis $\left\{ (1, 0^{BA}) \right\} \cup \left\{ (0, \eta_B^k \otimes \eta_A^j) \right\}$ of V_1, where $\left\{ \eta_A^j \right\}$ and $\left\{ \eta_B^k \right\}$ are orthonormal bases for $\operatorname{Herm}(A)$ and $\operatorname{Herm}(B)$, respectively. Moreover, we denote with $[\mathcal{N}_2^\dagger(X^A)]^{kj}$, the components of $\mathcal{N}_2^\dagger(X^A)$ in the $\left\{ \eta_B^k \otimes \eta_A^j \right\}$ basis, that is $\mathcal{N}_2^\dagger(X^A) = \sum_{j,k} [\mathcal{N}_2^\dagger(X^A)]^{kj} \eta_B^k \otimes \eta_A^j$. Similarly, we denote with X_j^A the components of X^A in the $\left\{ \eta_A^j \right\}$ basis.

$$\begin{aligned} \mathcal{N}_1^\dagger(X^A) \quad &= \left\langle (1, 0^{BA}) \,\middle|\, (\mathcal{N}_1^\dagger(X^A), \mathcal{N}_2^\dagger(X^A)) \right\rangle_{HS} \\ &= \left\langle (1, 0^{BA}) \,\middle|\, \mathcal{N}^\dagger(X^A) \right\rangle_{HS} \\ &= \left\langle \mathcal{N}[(1, 0^{BA})] \,\middle|\, X^A \right\rangle_{HS} \\ &= -\left\langle \mathbb{1}^A \,\middle|\, X^A \right\rangle_{HS} \\ &= -\operatorname{Tr}_A X^A, \\ [\mathcal{N}_2^\dagger(X^A)]^{kj} \quad &= \left\langle (0, \eta_B^k \otimes \eta_A^j) \,\middle|\, (\mathcal{N}_1^\dagger(X^A), \mathcal{N}_2^\dagger(X^A)) \right\rangle_{HS} \\ &= \left\langle (0, \eta_B^k \otimes \eta_A^j) \,\middle|\, \mathcal{N}^\dagger(X^A) \right\rangle_{HS} \\ &= \left\langle \mathcal{N}[(0, \eta_B^k \otimes \eta_A^j)] \,\middle|\, X^A \right\rangle_{HS} \\ &= \left\langle d_A \eta_A^j \operatorname{Tr}_B \eta_B^k \,\middle|\, X^A \right\rangle_{HS} \\ &= d_A X_j^A \operatorname{Tr}_B \eta_B^k. \end{aligned} \tag{5.63}$$

Therefore,

$$\mathcal{N}_2^\dagger(X^A) = \sum_{j,k} [\mathcal{N}_2^\dagger(X^A)]^{kj} \eta_B^k \otimes \eta_A^j = d_A \sum_i (\operatorname{Tr}_B \eta_B^k) \eta_B^k \otimes \sum_j X_j^A \eta_A^j = d_A \mathbb{1}^B \otimes X^A, \tag{5.64}$$

where $\sum_i (\operatorname{Tr}_B \eta_B^k) \eta_B^k = \mathbb{1}^B$ because $\operatorname{Tr}_B(\eta_B^k \mathbb{1}^B)$ are the components of $\mathbb{1}^B$ in the $\left\{ \eta_B^k \right\}$ basis.

Now, we can write the dual CLP as

$$\begin{aligned} \text{Find} \quad & \beta := \max 0 \\ \text{subject to} \quad & H_1 - \mathcal{N}^\dagger(X^A) = (0, -\tau_B \otimes (\rho^A)^T) - (-\operatorname{Tr}_A X^A, d_A \mathbb{1}^B \otimes X^A) \in \mathfrak{K}_1^* \\ & X^A \in \operatorname{Herm}(A), \end{aligned} \tag{5.65}$$

where $\mathfrak{K}_1^* = \left\{ (x, X^{BA}) \in V_1 \mid \left\langle (x, X^{BA}) \middle| (1, \omega^{BA}) \right\rangle_{HS} \geq 0,\ \forall \omega^{BA} \in \mathfrak{F}_{BA} \right\}$ (see Eq. (5.60)). The condition $H_1 - \mathcal{N}^\dagger(X^A) \in \mathfrak{K}_1^*$ is equivalent to:

$$\operatorname{Tr}_A X^A - \operatorname{Tr}_{BA}[(\tau_B \otimes (\rho^A)^T)\omega^{BA}] + d_A \operatorname{Tr}_{BA}[(\mathbb{1}^B \otimes X^A)\omega^{BA}] \geq 0, \quad \forall \omega^{BA} \in \mathfrak{F}(BA). \tag{5.66}$$

It is worth checking that no dual feasible solution exists for any CDRT. Indeed, in a CDRT, $\omega^{BA} = \frac{1}{d_A d_B}\mathbb{1}^{BA} \in \mathfrak{F}(A)$, and for all X^A we have that

$$\begin{aligned}
&\operatorname{Tr}_A X^A - \operatorname{Tr}_{BA}[(\tau_B \otimes (\rho^A)^T)\omega^{BA}] + d_A \operatorname{Tr}_{BA}[(\mathbb{1}^B \otimes X^A)\omega^{BA}] \\
&\quad = \operatorname{Tr}_A X^A - \frac{1}{d_A d_B} \operatorname{Tr}_{BA}[\tau_B \otimes (\rho^A)^T] - \frac{1}{d_B} \operatorname{Tr}_{BA}[\mathbb{1}^B \otimes X^A] \\
&\quad = -\frac{1}{d_B d_A}.
\end{aligned} \tag{5.67}$$

Therefore, $H_1 - \mathcal{N}^\dagger(X^A) \notin \mathfrak{K}_1^*$ for all $X^A \in \operatorname{Herm}(A)$, and no dual feasible solution exists.

CLPs for the Max-Relative Entropy of a State

With the same techniques used in the previous case, we write the conic linear program to compute $D_{\max}^{\mathfrak{F}}(\rho^A)$. First, we notice that

$$2^{D_{\max}^{\mathfrak{F}}(\rho^A)} = \min_{\substack{\omega^A \in \mathfrak{F}(A) \\ \rho^A \leq \lambda\omega^A}} \lambda = \min_{\substack{\omega^A \in \mathfrak{F}(A) \\ \lambda\omega^A - \rho^A \geq 0}} \lambda. \tag{5.68}$$

Second, we define the cone

$$\mathfrak{K}_1 = \left\{ (\lambda, \lambda\omega^A) \mid \lambda \geq 0, \omega^A \in \mathfrak{F}(A) \right\} \subseteq V_1 \cong \mathbb{R} \oplus \operatorname{Herm}(A), \tag{5.69}$$

and the cone $\mathfrak{K}_2$ as the cone of positive semidefinite matrices on A, i.e., $\mathfrak{K}_2 = \operatorname{Pos}(A) \subseteq V_2 = \operatorname{Herm}(A)$. Lastly, $\mathcal{N}[(x, X^A)] = X^A$ and $H_2 = \rho^A$. The conic linear program for the max relative entropy is

$$\begin{aligned}
\text{Find} \quad & 2^{D_{\max}^{\mathfrak{F}}(\rho^A)} := \inf \operatorname{Tr}[(1, 0^A)(x, X^A)] = \inf x \\
\text{subject to} \quad & \mathcal{N}[(x, X^A)] - H_2 = X^A - \rho^A \geq 0, \\
& (x, X^A) \in \mathfrak{K}_1.
\end{aligned} \tag{5.70}$$

As before, to write the dual CLP, we start with the dual cones:

$$\begin{aligned}
&\mathfrak{K}_1^* = \left\{ (x, X^A) \in V_1 \mid \left\langle (x, X^A) \middle| (1, \omega^A) \right\rangle_{HS} \geq 0,\ \forall \omega^A \in \mathfrak{F}(A) \right\}, \\
&\mathfrak{K}_2^* = (\operatorname{Pos}(A))^* = \operatorname{Pos}(A).
\end{aligned} \tag{5.71}$$

The components of the adjoint map are

$$\begin{aligned}
\mathcal{N}_1^\dagger(X^A) \quad &= \left\langle (1, 0^A) \,\middle|\, (\mathcal{N}_1^\dagger(X^A), \mathcal{N}_2^\dagger(X^A)) \right\rangle_{HS} \\
&= \left\langle (1, 0^A) \,\middle|\, \mathcal{N}^\dagger(X^A) \right\rangle_{HS} \\
&= \left\langle \mathcal{N}[(1, 0^A)] \,\middle|\, X^A \right\rangle_{HS} \\
&= -\left\langle 0^A \,\middle|\, X^A \right\rangle_{HS} \\
&= 0, \\
[\mathcal{N}_2^\dagger(X^A)]^j \quad &= \left\langle (0, \eta_A^j) \,\middle|\, (\mathcal{N}_1^\dagger(X^A), \mathcal{N}_2^\dagger(X^A)) \right\rangle_{HS} \\
&= \left\langle (0, \eta_A^j) \,\middle|\, \mathcal{N}^\dagger(X^A) \right\rangle_{HS} \\
&= \left\langle \mathcal{N}[(0, \eta_A^j)] \,\middle|\, X^A \right\rangle_{HS} \\
&= \left\langle \eta_A^j \,\middle|\, X^A \right\rangle_{HS} \\
&= X_j^A,
\end{aligned} \tag{5.72}$$

where, as before, $\left\{ \eta_A^j \right\}$ is an orthonormal basis for $\mathrm{Herm}(A)$. Thus, $\mathcal{N}^\dagger(X^A) = (0, X^A)$. The dual CLP is

$$\begin{aligned}
\text{Find} \quad & \beta := \sup \mathrm{Tr}[X^A \rho^A] \\
\text{subject to} \quad & H_1 - \mathcal{N}^\dagger(X^A) = (1, -X^A) \in \mathfrak{K}_1^* \\
& X^A \geq 0,
\end{aligned} \tag{5.73}$$

where $\mathfrak{K}_1^* = \left\{ (x, X^A) \in V_1 \mid \left\langle (x, X^A) \,\middle|\, (1, \omega^A) \right\rangle_{HS} \geq 0, \ \forall \omega^A \in \mathfrak{F}(A) \right\}$ (see Eq. (5.71)). The condition $(1, -X^A) \in \mathfrak{K}_1^*$ is equivalent to $\mathrm{Tr}_A[X^A \omega^A] \leq 1$ for all $\omega^A \in \mathfrak{F}(A)$. In this case, a feasible solution exists: The zero matrix 0^A is a feasible solution.

CLPs for the Max-Relative Entropy of a Channel in CDRTs

Here, we compute the CLP for

$$2^{D_{\max}^{\mathfrak{F}}(\mathcal{M}^{A \to B})} = \min_{\substack{\omega^{BA'} \in \mathfrak{F}(BA'), \\ d_{A'} \mathrm{Tr}_B\, \omega^{BA'} = \mathbb{1}^{A'}, \\ \mu^{BA'} \leq \lambda \omega^{BA'}}} \lambda, \tag{5.74}$$

where $\mu^{B:A'}$ is the renormalized Choi matrix of $\mathcal{M}^{A \to B}$. Once again, we define

$$\mathfrak{K}_1 = \left\{ (\lambda, \lambda \omega^{BA'}) \mid \lambda \geq 0, \omega^{BA'} \in \mathfrak{F}(BA') \right\} \subseteq V_1 \cong \mathbb{R} \oplus \mathrm{Herm}(BA'). \tag{5.75}$$

The linear transformation $\mathcal{N} : \mathbb{R} \oplus \mathrm{Herm}(BA') \to \mathrm{Herm}(BA') \oplus \mathrm{Herm}(A')$ is defined as the map that sends $(x, X^{BA'})$ into $(X^{BA'}, d_{A'} \mathrm{Tr}_B\, X^{BA'} - x \mathbb{1}^{A'})$. Moreover, we define $H_2 = (\mu^{BA'}, 0^{A'})$, and

$$\mathfrak{K}_2 = \left\{ (X^{BA'}, 0^{A'}) \mid X^{BA'} \geq 0 \right\} \subseteq V_2 \cong \mathrm{Herm}(BA') \oplus \mathrm{Herm}(A'). \tag{5.76}$$

The conic linear program is

$$
\begin{aligned}
\text{Find}\quad & 2^{D^{\mathfrak{F}}_{\max}(\mathcal{M}^{A\to B})} := \inf \operatorname{Tr}[(1, 0^{BA'})(x, X^{BA'})] = \inf x \\
\text{subject to}\quad & \mathcal{N}[(x, X^{BA'})] - H_2 = (X^{BA'} - \mu^{BA'}, d_{A'} \operatorname{Tr}_B X^{BA'} - x\mathbb{1}^{A'}) \in \mathfrak{K}_2, \\
& (x, X^{BA'}) \in \mathfrak{K}_1.
\end{aligned}
\tag{5.77}
$$

To write the dual CLP, we first need the dual cones:

$$
\begin{aligned}
\mathfrak{K}_1^* &= \left\{ (x, X^{BA'}) \in V_1 \mid \left\langle (x \middle| X^{BA'}) \middle| (1, \omega^{BA'}) \right\rangle_{HS} \geq 0,\ \forall \omega^{BA'} \in \mathfrak{F}(BA') \right\}, \\
\mathfrak{K}_2^* &= \left\{ (X^{BA'}, Y^{A'}) \in V_2 \mid \left\langle (X^{BA'}, Y^{A'}) \middle| (Z_{BA'}, 0^{A'}) \right\rangle_{HS} \geq 0,\ \forall Z_{BA'} \geq 0 \right\} \\
&= \left\{ (X^{BA'}, Y^{A'}) \in V_2 \mid X^{BA'} \geq 0 \right\}.
\end{aligned}
\tag{5.78}
$$

,

With the techniques introduced in the previous subsections, we compute the components of $\mathcal{N}^\dagger$.

$$
\begin{aligned}
\mathcal{N}_1^\dagger[(X^{BA'}, Y^{A'})] \quad &= \left\langle (1, 0^{BA'}) \middle| (\mathcal{N}_1^\dagger[(X^{BA'}, Y^{A'})], \mathcal{N}_2^\dagger[(X^{BA'}, Y^{A'})]) \right\rangle_{HS} \\
&= \left\langle (1, 0^{BA'}) \middle| \mathcal{N}^\dagger[(X^{BA'}, Y^{A'})] \right\rangle_{HS} \\
&= \left\langle \mathcal{N}[(1, 0^{BA'})] \middle| (X^{BA'}, Y^{A'}) \right\rangle_{HS} \\
&= -\left\langle (0^{BA'}, \mathbb{1}^{A'}) \middle| (X^{BA'}, Y^{A'}) \right\rangle_{HS} \\
&= -\operatorname{Tr}_{A'} Y^{A'}, \\
\{\mathcal{N}_2^\dagger[(X^{BA'}, Y^{A'})]\}^{kj} \quad &= \left\langle (0, \eta_B^k \otimes \eta_{A'}^j) \middle| (\mathcal{N}_1^\dagger[(X^{BA'}, Y^{A'})], \mathcal{N}_2^\dagger[(X^{BA'}, Y^{A'})]) \right\rangle_{HS} \\
&= \left\langle (0, \eta_B^k \otimes \eta_{A'}^j) \middle| \mathcal{N}^\dagger[(X^{BA'}, Y^{A'})] \right\rangle_{HS} \\
&= \left\langle \mathcal{N}[(0, \eta_B^k \otimes \eta_{A'}^j)] \middle| (X^{BA'}, Y^{A'}) \right\rangle_{HS} \\
&= \left\langle (\eta_B^k \otimes \eta_A^j, d_{A'} (\operatorname{Tr}_B \eta_B^k) \eta_{A'}^j) \middle| (X^{BA'}, Y^{A'}) \right\rangle_{HS} \\
&= X_{kj}^{BA'} + d_{A'} Y_j^{A'} (\operatorname{Tr}_B \eta_B^k).
\end{aligned}
\tag{5.79}
$$

Thus, $N^\dagger[(X^{BA'}, Y^{A'})] = (-\operatorname{Tr}_{A'} Y^{A'}, X^{BA'} + d_{A'} \mathbb{1}^B \otimes Y^{A'})$. The dual linear program is

$$
\begin{aligned}
\text{Find}\quad & \beta := \sup \operatorname{Tr}[(X^{BA'}, Y^{A'})(\mu^{BA'}, 0^{A'})] = \operatorname{Tr}_{BA'}[X^{BA'} \mu^{BA'}] \\
\text{subject to}\quad & H_1 - \mathcal{N}^\dagger[(X^{BA'}, Y^{A'})] = (1, 0^{BA'}) - (-\operatorname{Tr}_{A'} Y^{A'}, X^{BA'} + d_{A'} \mathbb{1}^B \otimes Y^{A'}) \in \mathfrak{K}_1^* \\
& X^{BA'} \geq 0,\ Y^{A'} \in \operatorname{Herm}(A'),
\end{aligned}
\tag{5.80}
$$

where $\mathfrak{K}_1^* = \left\{ (x, X^{BA'}) \in V_1 \mid \left\langle (x \middle| X^{BA'}) \middle| (1, \omega^{BA'}) \right\rangle_{HS} \geq 0,\ \forall \omega^{BA'} \in \mathfrak{F}(BA') \right\}$ (see Eq. (5.78)). The condition $(1 + \operatorname{Tr}_{A'} Y^{A'}, -X^{BA'} - d_{A'} \mathbb{1}^B \otimes Y^{A'}) \in \mathfrak{K}_1^*$ is equivalent to $\operatorname{Tr}_{BA'}[X^{BA'} \omega^{BA'}] + d_{A'} \operatorname{Tr}_{BA'}[(\mathbb{1}^B \otimes Y^{A'}) \omega^{BA'}] \leq 1 + \operatorname{Tr}_{A'} Y^{A'}$ for all $\omega^{BA'} \in \mathfrak{F}(BA')$. It is straightforward to see that $(0^{BA'}, 0^{A'})$ is a dual feasible solution.

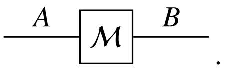
5.4 A Diagrammatic Proof of Theorem 5.2.5

String diagrams have been frequently used in quantum information [85, 237–242], and they provide a useful representation of the Choi isomorphism, from which many of its properties follow naturally. Indeed, we first proved all our results diagrammatically, and later, we converted the diagrammatic proofs into algebraic ones. In this Section, we introduce the string diagram formalism and provide a diagrammatic proof of Theorem 5.2.5.

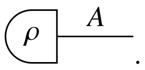
5.4.1 A Diagrammatic Approach to Quantum Information

Quantum information processes can be described using string diagrams [85, 237–242]. These diagrams are not only a graphical representation of quantum processes but can also be used for rigorous proofs. Indeed, one can associate an algebraic expression with every diagram and a diagram with every algebraic expression. Moreover, diagrams can be manipulated as algebraic expressions.

The fundamental component of string diagrams is the diagrammatic symbol for quantum channels. A channel $\mathcal{M}^{A\to B}$ is represented in a diagram as a box:

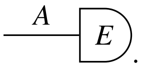


This symbol represents the channel acting on an input quantum system A and producing an output quantum system B. When a channel has n input systems and m output systems it is represented with n input wires and m output wires. There are two special types of channels. The first is the preparation channel, which has no quantum input, represented with no wire, and outputs a state ρ^A. The symbol for this channel is

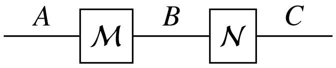


Since there is a one-to-one correspondence between preparation channels and quantum states, we say that this is the symbol for the state ρ^A. With a slight abuse of notation, we will use this symbol to represent positive semidefinite matrices, too. However, these two cases can easily be distinguished because we denote states with lowercase Greek letters, e.g., ρ, and matrices with capital Latin letters, e.g., M.

The second special type of channel is associated with positive operator-valued measurements, or POVMs (see, e.g., Ref. [11] for details). There is no quantum output in this case, so the channel is represented without the output wire. When E^A is an element of a POVM on A, also known as effect, i.e., $0 \leq E \leq \mathbb{1}^A$, we represent it with the symbol

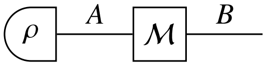


From these building blocks, one can construct more complex diagrams by connecting a symbol's output wire with another's input wire. For example, by connecting two channels in sequence, one obtains



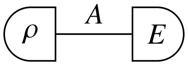
$$= \mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}\,. \tag{5.81}$$

The special case when the first channel is a state is represented as



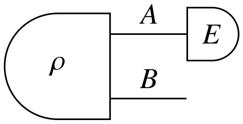
$$= \mathcal{M}^{A\to B}(\rho^A)\,. \tag{5.82}$$

On the other hand, an element of a POVM E^A maps a quantum state ρ^A into $\mathrm{Tr}_A(E^A\rho^A)$, i.e., the probability of E^A occurring on ρ^A, and this is the algebraic expression associated with the diagram below



$$= \mathrm{Tr}_A(E^A\rho^A). \tag{5.83}$$

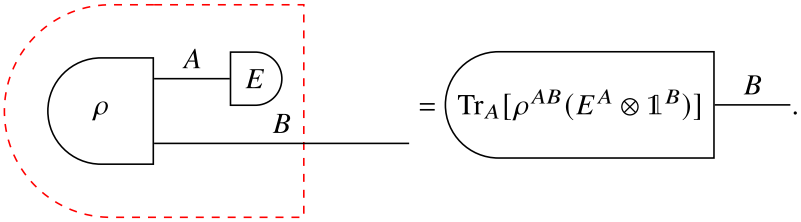
More generally, an effect E^A could act on only a part of a state ρ^{AB}: In this case, the diagram and the associated algebraic expression are



$$= \mathrm{Tr}_A[\rho^{AB}(E^A\otimes \mathbb{1}^B)]. \tag{5.84}$$

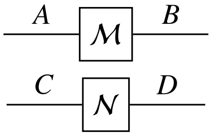
Therefore, every time an effect is composed with another diagram, the algebraic expression associated with it contains the trace over the common systems.

The expression above allows us to introduce one of the benefits of string diagrams. The diagram in Eq. (5.84) shows only one output wire. Therefore, it can be simplified and written as the diagram of a (possibly unnormalized) state:



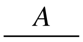
$$\tag{5.85}$$

The tensor product of states, channels, and effects, also known as parallel composition, is represented by writing one symbol below the other, e.g.,

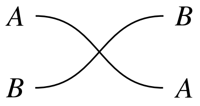


$$= \mathcal{M}^{A\to B}\otimes \mathcal{N}^{C\to D}. \tag{5.86}$$

It is useful to denote two common channels with special diagrams. The identity channel $\mathcal{I}^A$ on A, and the swap channel $\mathcal{S}^{AB\to BA}$ from AB to BA are represented as



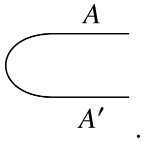
$$= \mathcal{I}^A\,,$$



$$= \mathcal{S}^{AB\to BA}. \tag{5.87}$$

The choice of these symbols is intuitive in that they convey that nothing is done on system A in the former case and that the systems A and B are swapped in the latter case.

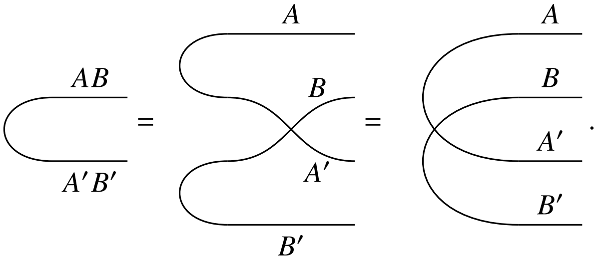
Of particular interest is the symbol for the Choi state $\Phi^{AA'} = \sum_{x,y} |xx\rangle\langle yy|^{AA'}$, where $\{\, |x\rangle \,\}_A$ is a fixed orthonormal basis for A and A' is a copy of A:



It is particularly useful in proofs to differentiate A and A': This ensures that all the systems are properly accounted for. However, A' being a copy of A, we will use either $\Phi^{AA'}$ or $\Phi^{A'A}$ depending on the need. In case of a bipartite system AB, the Choi state is $\sum_{a,b,a',b'} |abab\rangle\!\langle a'b'a'b'|^{ABA'B'}$, where $\{\,|a\rangle^A\,\}$ and $\{\,|b\rangle^B\,\}$ are orthonormal basis for A and B, respectively. This state can be rewritten as

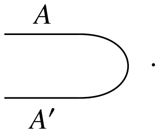
$$\begin{aligned}\Phi^{ABA'B'} &= \sum_{a,b,a'b'} |abab\rangle\!\langle a'b'a'b'|^{ABA'B'} \\ &= \sum_{a,b,a',b'} |a\rangle\!\langle a'|^{A} \otimes |b\rangle\!\langle b'|^{B} \otimes |a\rangle\!\langle a'|^{A'} \otimes |b\rangle\!\langle b'|^{B'} \\ &= \sum_{a,b,a',b'} (\mathcal{I}^{A} \otimes \mathcal{S}^{A'B\to BA'} \otimes \mathcal{I}^{B'})(|a\rangle\!\langle a'|^{A} \otimes |a\rangle\!\langle a'|^{A'} \otimes |b\rangle\!\langle b'|^{B} \otimes |b\rangle\!\langle b'|^{B'}) \\ &= (\mathcal{I}^{A} \otimes \mathcal{S}^{A'B\to BA'} \otimes \mathcal{I}^{B'})(\Phi^{AA'} \otimes \Phi^{BB'}).\end{aligned} \tag{5.88}$$

This is easily represented diagrammatically as

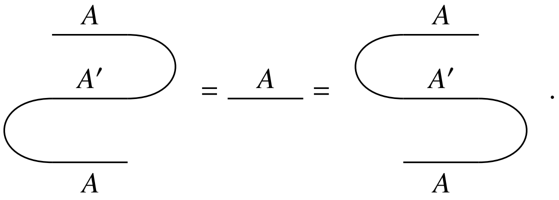


(5.89)

To understand the choice of the symbol associated with the Choi state, we need first to introduce the Choi effect $\Phi^{AA'}$, which has the same matrix as the Choi state and acts on a state ρ^A as $\mathrm{Tr}_A[\Phi^{AA'}(\rho^A \otimes \mathbb{1}^{A'})]$. The action on a state $\rho^{A'}$ is analogous. The symbol for the Choi effect is



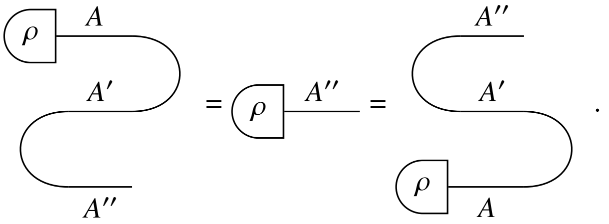
Two useful identities involving the Choi state and effect are (see, e.g., Refs [85, 237–242])



(5.90)

These identities are called snake identities [243] and motivate the choice of the symbol for the Choi state and effect. Indeed, by connecting a Choi state to a Choi effect, one obtains a plain wire, i.e., the identity channel.

To prove the snake identities, we use a test state ρ^A and, to track all systems properly, we change the system in the output wire to A''. That is, we prove



(5.91)

We start with the diagram on the left

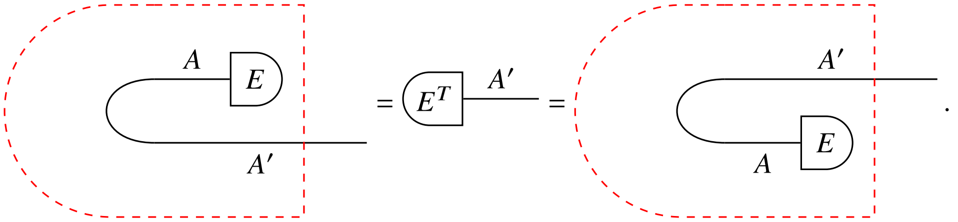
$$
\begin{aligned}
\mathrm{Tr}_{AA'}[(\Phi^{AA'}\otimes \mathbb{1}^{A''})(\rho^A\otimes\Phi^{A'A''})] &= \sum_{x,y,a,b}\mathrm{Tr}_{AA'}[(|xx\rangle\langle yy|^{AA'}\otimes\mathbb{1}^{A''})(\rho^A\otimes|aa\rangle\langle bb|^{A'A''})]\\
&= \sum_{x,y,a,b}\langle yy\,|\,(\rho\otimes|\,a\rangle\langle b\,|)\,|\,xx\rangle^{AA'}\,|a\rangle\langle b|^{A''}\\
&= \sum_{x,y,a,b}\delta_{b,x}\delta_{a,y}\,\langle y\,|\,\rho\,|\,x\rangle_A\,|a\rangle\langle b|^{A''}\\
&= \sum_{a,b}\langle a\,|\,\rho\,|\,b\rangle_A\,|a\rangle\langle b|^{A''}\\
&= \rho^{A''}
\end{aligned}
\tag{5.92}
$$

The proof of the right identity is analogous. This is an example of the simplicity of the diagrammatic approach in comparison to the algebraic expression. In many of the upcoming proofs, instead of writing multiple lines of algebraic simplifications, we will simply 'yank' a wire [244, 245]. It is important to notice again that this simplification does not reduce the rigour of a proof; it is equivalent to an algebraic proof.

Other useful diagrammatic identities involving the Choi state are the following. If we apply an element $E = \sum_{a,b} E_{a,b}|a\rangle\langle b|^A$ of a POVM to one end of the Choi state, we obtain

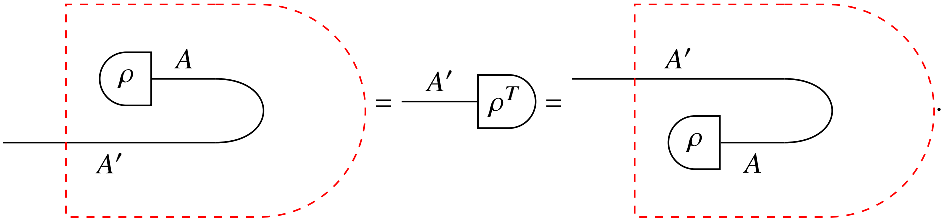
$$
\begin{aligned}
\mathrm{Tr}_A[\Phi^{AA'}(E^A\otimes\mathbb{1}^{A'})] &= \sum_{x,y,a,b}E_{a,b}\,\mathrm{Tr}_A[|xx\rangle\langle yy|^{AA'}(|a\rangle\langle b|^A\otimes\mathbb{1}^{A'})]\\
&= \sum_{x,y,a,b}E_{a,b}\,\langle y\,|\,(\,|\,a\rangle\langle b\,|\,)\,|\,x\rangle_A\,|x\rangle\langle y|^{A'}\\
&= \sum_{a,b}E_{a,b}|b\rangle\langle a|^{A'} = E^T_{A'}.
\end{aligned}
\tag{5.93}
$$

Therefore, the Choi state maps an element E of a POVM into the unnormalized state E^T. This is represented diagrammatically as



(5.94)

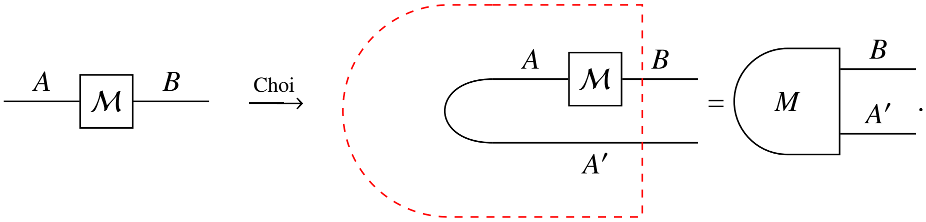
With a similar proof, one obtains



(5.95)

5.4.2 The Choi Isomorphism

The Choi isomorphism maps quantum channels (or, more generally, linear operators that act on matrices) to bipartite matrices. Diagrammatically, there is an intuitive way to map a channel (one input and one output wire) into a bipartite matrix (two output wires): To transform an input wire into an output wire is enough to bend it in a way that its free end is directed towards the right, as depicted below.



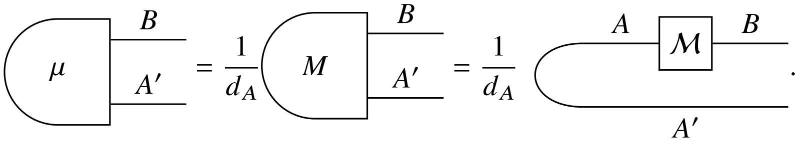
(5.96)

These diagrams coincide with the definition of Choi isomorphism that we gave in Section 5.1, i.e., the Choi matrix $M^{B:A'}$ associated with a channel $\mathcal{M}^{A\to B}$ is defined as

$$M^{B:A'} = (\mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'})(\Phi^{AA'})\,. \tag{5.97}$$

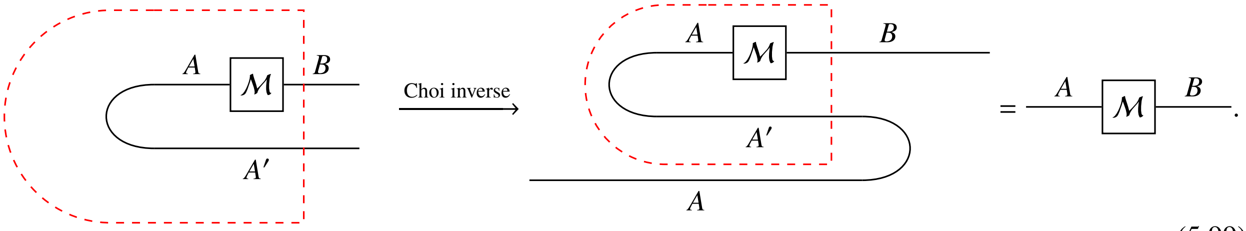
Note that we decided to bend the input wire below the channel. Similarly, one could have chosen to bend the input wire above the channel, producing a different but still well-defined Choi matrix $M^{A:B} = (\mathcal{I}^{A} \otimes \mathcal{M}^{A'\to B})(\Phi^{AA'})$. Throughout this Thesis, we always use the convention that a Choi matrix is defined as in Eq. (5.97), i.e., by bending the wire below.

The Choi matrix of a quantum channel is Hermitian, positive semidefinite, and the marginal on the input system is the identity matrix, i.e., $\mathrm{Tr}_B\, M^{B:A'} = \mathbb{1}^{A'}$ [66]. Therefore, a Choi matrix is not a quantum state, as its trace is not 1. However, one can easily define a quantum state as the renormalized Choi matrix $\mu^{B:A'} = \frac{1}{d_A} M^{B:A'}$. Diagrammatically, one obtains



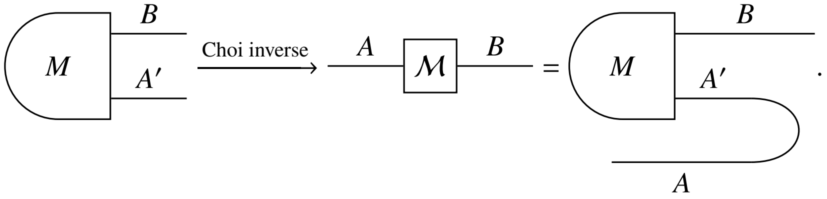
(5.98)

One of the advantages of using string diagrams to approach the Choi isomorphism is that the definition of the inverse Choi isomorphism becomes straightforward. Indeed, one can undo the Choi isomorphism by bending the second wire back and using the snake identities in Eq. (5.90)



(5.99)

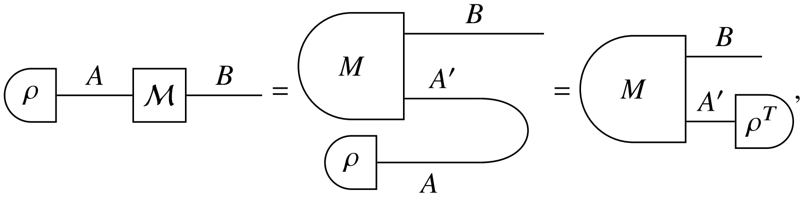
Therefore, the inverse Choi isomorphism transforms a Choi matrix $M^{B:A'}$ in the channel $\mathcal{M}^{A\to B}$ defined as



(5.100)

With this definition, one quickly obtains the formula for the action of the channel $\mathcal{M}^{A\to B}$ on a state ρ^A

$\mathcal{M}^{A\to B}(\rho^A) =$



(5.101)

These diagrams correspond to the algebraic expressions for the inverse Choi map commonly used in literature:

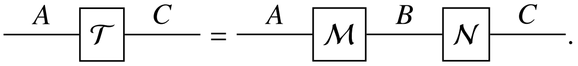
$$\mathcal{M}^{A\to B}(\rho^A) = \mathrm{Tr}_{A'A}[(M^{B:A'} \otimes \rho^A)(\mathbb{1}^B \otimes \Phi^{A'A})] = \mathrm{Tr}_{A'}[M^{B:A'}(\mathbb{1}^B \otimes \rho^T_{A'})]. \tag{5.102}$$

Thanks to the identities in Eq. (5.90), it is straightforward to see that the Choi map is indeed an isomorphism. In addition, diagrammatically, one immediately notices that the Choi matrix of a state ρ^A is the state itself since there is no input wire to bend, and the Choi matrix of an effect is the transpose of the matrix of the effect, see Eq. (5.94).

5.4.3 The Link Product

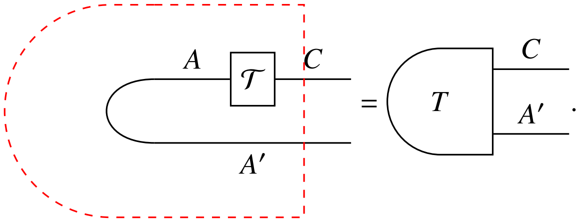
Once the Choi isomorphism is well-defined, the natural step is translating operations between channels into operations between Choi matrices. Once again, we approach this problem with string diagrams. Then, we show that the operation that naturally arises with this approach is the algebraic operation between Choi matrices, known as link product.

Let us consider two channels $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{B\to C}$. We call the channel resulting from their sequential composition $\mathcal{T}_{A\to C} = \mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$. Diagrammatically,



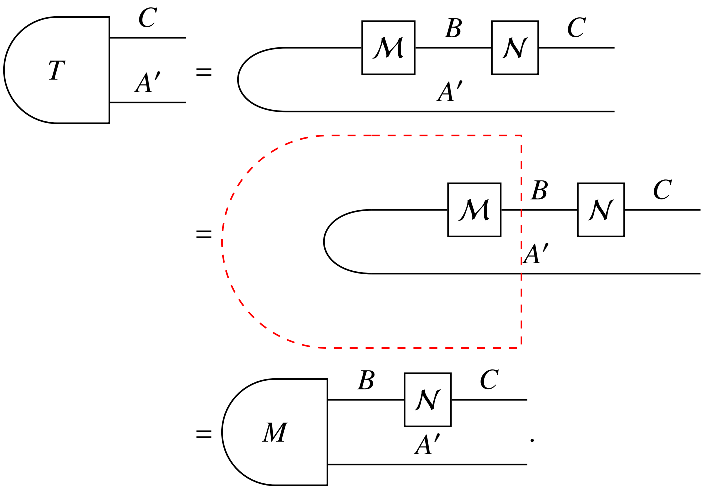
(5.103)

The Choi matrix $T_{C:A'}$ of $\mathcal{T}_{A\to C}$ is defined as



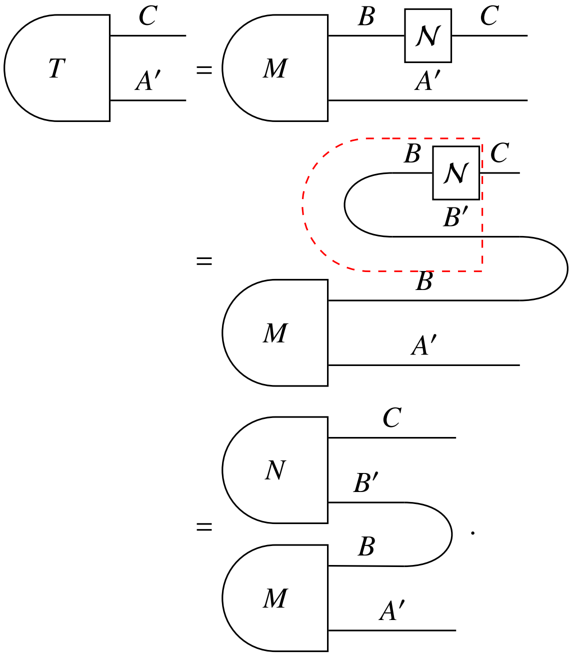
(5.104)

From Eq. (5.103) we have that



(5.105)

So far, we have been able to express the Choi matrix of $\mathcal{T}_{A\to C}$ in terms of the Choi matrix of $\mathcal{M}^{A\to B}$. To replace $\mathcal{N}^{B\to C}$ with its Choi matrix, we insert one of the identities in Eq. (5.90) between $M^{B:A'}$ and $\mathcal{N}^{B\to C}$:



(5.106)

The algebraic expression associated with this diagram is

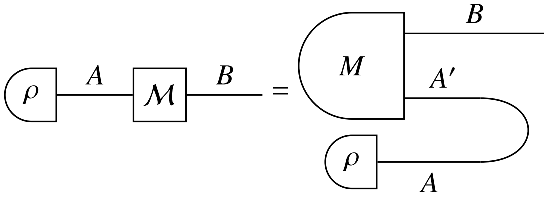
$$T_{C:A'} = \mathrm{Tr}_{BB'}[(N^{C:B'} \otimes M^{B:A'})(\mathbb{1}^{C} \otimes \Phi^{B'B} \otimes \mathbb{1}^{A'})]. \tag{5.107}$$

This operation is known as link product, and it is denoted with $N^{C:B'} * M^{B:A'}$ [66]. The name ‘link’ product is well-motivated from a diagrammatic point of view because this operation links the Choi matrices $N^{C:B'}$ and $M^{B:A'}$. Indeed, diagrammatically, it is represented as a ‘bridge’ that links two bipartite matrices without any wire crossing, i.e., the second system of the first matrix is linked to the first system of the second matrix.

Eq. (5.106) shows that the link product of the Choi matrices of two channels is the Choi matrix of the sequential composition of these channels. Moreover, the link product of Choi matrices of channels always produces the Choi matrix of a channel, that is, a positive semidefinite matrix with the identity as marginal on the input state [66].

As a final note, in Eq. (5.101) and Eq. (5.102), we expressed the action of a channel $\mathcal{M}^{A\to B}$ on a state ρ in terms of its Choi matrix $M^{B:A'}$. That expression is nothing but the link product $M^{B:A'} * \rho^A$, where ρ^A is seen as the bipartite Choi matrix on A and the trivial system, corresponding to the Hilbert space $\mathbb{C}$, associated with the channel that prepares ρ^A.

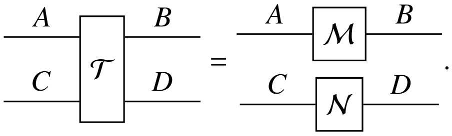
$$\mathcal{M}^{A\to B}(\rho^A) =$$



$$= M^{B:A'} * \rho^A. \tag{5.108}$$

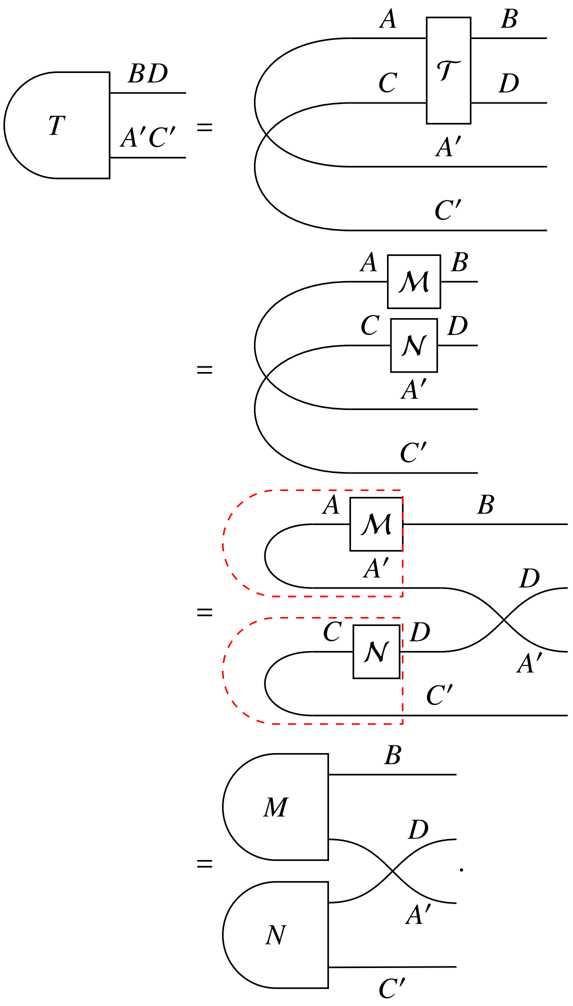
5.4.4 The Swap Tensor Product

As seen in the previous subsection, the sequential composition of channels is translated into the link product of Choi matrices. Parallel composition of channels is often overlooked or translated into the tensor product of Choi matrices [66, 221]. However, one needs to be careful with the order of the systems, as we show here. We start with two channels, $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{C\to D}$, and analogously to what we did before, we define $\mathcal{T}_{AC\to BD} = \mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$. Diagrammatically, we have



(5.109)

As before, we compute the Choi matrix $T_{BD:A'C'}$ of $\mathcal{T}_{AC\to BD}$ with string diagrams:



(5.110)

The diagrammatic approach indicates that the tensor product of channels is *not* associated with the tensor product of Choi matrices, but with the swap tensor product that we introduced in Subsection 5.1.1:

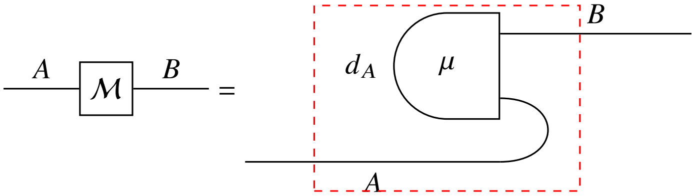
$$T_{BD:A'C'} = M^{B:A'} \boxtimes N^{D:C'} = (\mathcal{I}^{B} \otimes \mathcal{S}^{A'D\to DA'} \otimes \mathcal{I}^{C'})(M^{B:A'} \otimes N^{D:C'}). \tag{5.111}$$

As in the case of the link product, the Choi matrix of the tensor product of channels is the swap tensor product of the Choi matrices by construction, and every time we take the swap tensor product of two Choi matrices, of quantum channels, we obtain the Choi matrix of a quantum channel, that is, a positive semidefinite matrix such that its marginal on the input systems is the identity matrix. This follows directly from the equalities in Eq. (5.110).

5.4.5 Choi-Defined Resource Theories

As a reminder, the CD operations associated with a set of free states are all and only the quantum channels such that their renormalized Choi matrix is a free state. CD operations are easily constructed from free states with the inverse Choi map, i.e., if $\mu^{B:A'}$ is a free state and a renormalized Choi matrix, which means

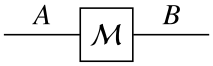
$\mathrm{Tr}_B\,\mu^{B:A'} = \frac{1}{d_A}\mathbb{1}^{A'}$, then the channel $\mathcal{M}^{A\to B}$ defined as



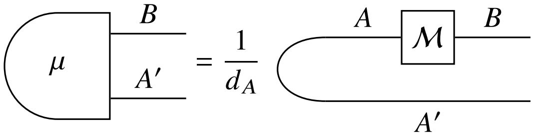
(5.112)

is a CD operation.

A quantum resource theory is a CDRT if its free operations coincide with the Choi-defined operations associated with its set of free states. This diagrammatically means that,



is free if and only if

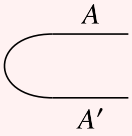


is free. (5.113)

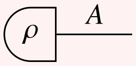
We are now ready to prove the main result of this work diagrammatically.

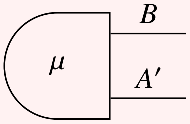
Theorem 5.4.1: Necessary and sufficient conditions for CDRTs construction

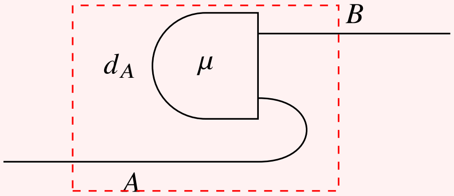
It is possible to construct a CDRT associated with a set of free states if and only if for all systems A, B

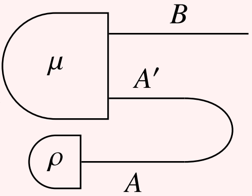
1. $\frac{1}{d_A}$ 

is free,

2. if 

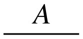
and 

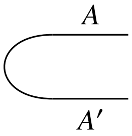
are free states, and 

is a quantum channel, then d_A 

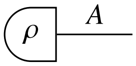
is a free state.

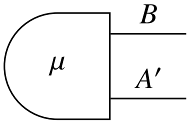
Proof. We prove first the necessary conditions. We assume that the free states and the free operations form a well-defined Choi-defined resource theory.

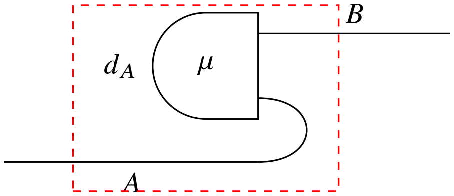
1. 

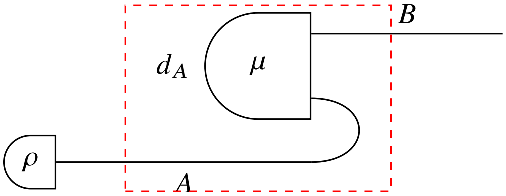
is a free operation in every resource theory. Therefore, in a CDRT, $\frac{1}{d_A}$ 

is free as well, see Eq. (5.113).

2. If 

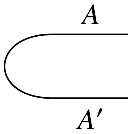
 and 

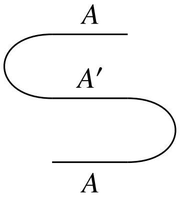
 are free states, then 

 is a free quantum channel (from Eq. (5.113)), and 

 is a free state (it is the output of a free channel applied to a free state).

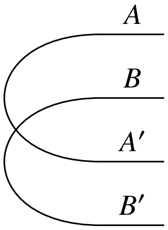
We prove now that conditions 1 and 2 are sufficient to have a well-defined resource theory when the free operations are the Choi-defined operations. Recall that we assume that the set of free states under consideration is compatible with a minimal resource theory, i.e., closed under tensor product, partial tracing, and system swapping. Now, we demonstrate that the five conditions for a resource theory listed in Section 2.1 are satisfied.

1. The identity channel is free. Indeed, since $\frac{1}{d_A}$ 

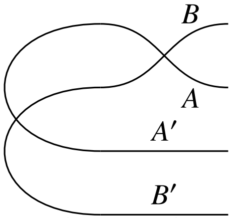
 is free for all A, then 

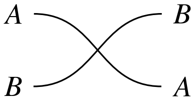
 $=$ 

 is a free operation (Eq. (5.112) and Eq. (5.90)).

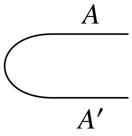
2. The swap channel is free. Indeed, from condition 1, one has that $\frac{1}{d_A d_B}$ 

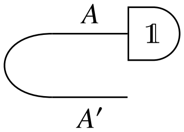
 is a free state.

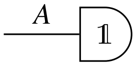
   Since the set of free states is closed under system swapping, one obtains that $\frac{1}{d_A d_B}$ 

 is free too. From Eq. (5.112) and Eq. (5.90), it immediately follows that the CD operation associated with this state is 

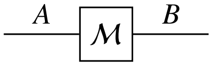
.

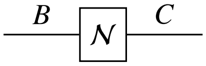
3. Discarding a system is free. Once again, $\frac{1}{d_A}$ 

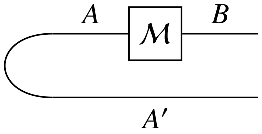
 is free for all A, and since the set of free operations is closed under partial tracing, then $\frac{1}{d_A}$ 

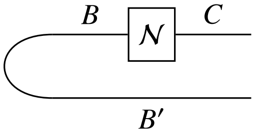
is free as well. As above, from Eq. (5.112) and Eq. (5.90), one obtains that the CD operation associated with this state is 

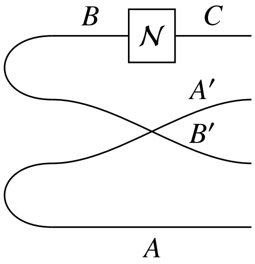
, that is, the discarding channel.

4. Sequential composition of free channels is free. Let 

and 

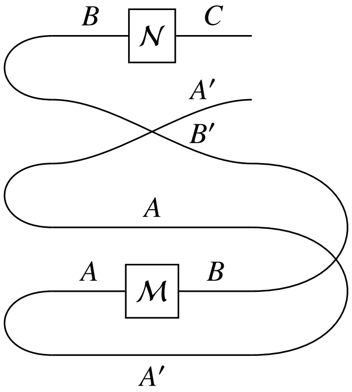
be free channels. Since the free operations are the CD operations, the states $\frac{1}{d_A}$ 

and $\frac{1}{d_B}$ 

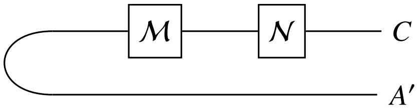
are free. From condition 1 and the closure of the set of free states under tensor product and system swapping, it follows that the state $\frac{1}{d_A d_B}$ 

is free as well. Condition 2 implies that the state

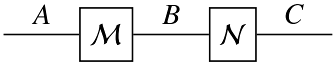
$$d_A d_B \frac{1}{d_A d_B} \frac{1}{d_A}$$



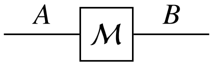
$$= \frac{1}{d_A}$$

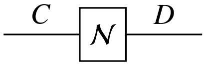


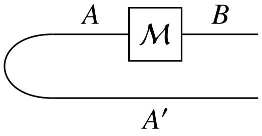
(5.114)

is free. The equality above follows from multiple uses of Eq. (5.90). The CD operation associated with the state in Eq. (5.114) is 

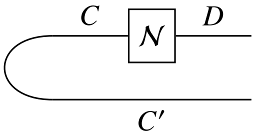
.

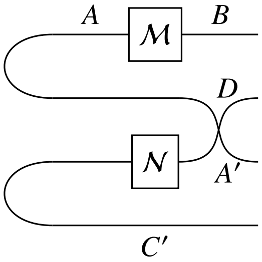
5. Parallel composition of free channels is free. Let 

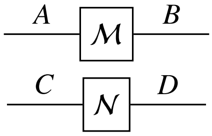
and 

be free channels. Since the free operations are the CD operations, the states $\frac{1}{d_A}$ 

and

$\frac{1}{d_C}$ 

are free. Since the set of free states is closed under tensor product and system swapping, the state $\frac{1}{d_A d_C}$ 

is free as well. Eq. (5.110) implies that the CD operation associated with this state is 

.

□

Chapter 6

Higher-Order Quantum Maps

The focus of the works presented so far in this Thesis has been on resource theories of states. At the core of these resource theories is the conversion of states using free channels. Considering that states and channels provide a complete description of quantum systems and their evolution, one may think that there is no need to push the framework of quantum resource theories further. On the other hand, there are tasks that, although they have a resource-theoretical flavour, do not fit within the framework presented so far. For example, the quantum teleportation protocol [16] can be interpreted as a simulation of the identity channel using the maximally entangled state. In other words, using only LOCC, a static resource, the maximally entangled state is converted into a dynamical resource, the identity channel. This kind of problem sounds very similar to what we are used to for static resource theories. Indeed, Gour and Scandolo [64, 67] propose the resource theory of dynamical entanglement, where the key problem is the conversion of channels (of which states are a special case) using free superchannels [65]. This idea of shifting from static resources to dynamic ones [68] was quickly applied to other resource theories, such as coherence [69], and magic [70], and ideas from static resource theories, such as distillation [71] and reversibility [72] were transposed to resource theories of channels.

In these dynamical resource theories, the key ingredients are superchannels, first formalized by Chiribella *et al.* [65, 66]. As a quantum channel from system A to system B can be axiomatically defined as a linear operator that maps every possible quantum state of A to a valid quantum state of B, a quantum superchannel from operators $A \to B$ (from system A to system B) to operators $C \to D$ is a linear operator that maps every possible quantum channel $A \to B$ to a valid channel $C \to D$. This definition points towards a hierarchy of maps: there are quantum states, quantum channels (i.e., maps that preserve quantum states), and quantum superchannels (i.e., maps that preserve quantum channels). However, one can push this concept even further and have higher-order quantum maps, such as quantum networks [66, 221, 246]. Once again, one may ask what the benefit of going beyond quantum superchannels is. It turns out that new and exciting properties emerge when we consider higher orders. Indeed, superchannels can be decomposed into circuits of channels with memory, adding nothing significantly new. However, the same is not true for maps of higher order, which are, therefore, genuinely different from quantum channels. One example of such a map is the quantum switch [73], which maps two channels $\mathcal{M}$ and $\mathcal{N}$ into a channel that acts as the superposition of $\mathcal{M} \circ \mathcal{N}$ and $\mathcal{N} \circ \mathcal{M}$. The quantum switch was the first example of a map with indefinite causal order [247]. As such, it opened the path to numerous results about causality in quantum mechanics [248–254], some of which were investigated with the framework of quantum resource theories [255, 256]. Moreover, maps with indefinite causal order provide an advantage in communication [74], computation [75–78], and channel discrimination [79]. Other research areas, such as the theory of non-Markovian quantum processes [257–260], were stimulated as well by this attempt to go beyond channels and superchannels to a theory of higher-order quantum maps.

With the flourishing of applications for higher-order quantum maps, Bisio and Perinotti [80] proposed a theoretical framework to describe all higher-order maps. At the core of their work, there is what they call *type system*, a way of labelling maps of any order. For example, a state of the system A is labelled with A, a channel from the system A to the system B is labelled with $(A \to B)$, etc. Assisted by this type system, they use the Choi isomorphism to reduce any quantum map to a matrix, and they characterize some of the properties of these Choi matrices of quantum maps. However, the collapse of any quantum map to matrices makes it challenging to work with this framework and almost impossible to provide a clear definition of the parallel composition of maps. The work presented in this Chapter starts with the definition of types given by Bisio and Perinotti [80] but diverges immediately after with a double goal. First, we aim to describe higher-order quantum maps without reducing them to matrices. Second, we seek a framework that naturally incorporates the evaluation of a map at an input, as well as sequential and parallel composition. We divide this Chapter into three sections. We start with a description of our type system, inspired by that of Bisio and Perinotti [80] but enriched with new operations and concepts. In the second Section, we characterize the spaces of linear maps associated with a given type, for example, the spaces that generalize the spaces of Hermitian-preserving and completely positive-preserving channels. Moreover, we present the operations between maps mentioned above, with a focus on parallel composition, which is the less trivial one. In the third Section, we introduce the Choi isomorphism as an operation that reduces the order of a map in the hierarchy of higher-order quantum maps rather than as a way of transforming a higher-order map into a matrix. Among the most interesting results of this Section is the fact that the Choi isomorphism preserves the generalizations of Hermitian-preserving operations and completely positive-preserving operations, and therefore, it can be used to reduce the order of a map while keeping some of its properties.

6.1 Types

The overall goal of this Chapter is to give a unified description of quantum maps of every order, and in this Section, we start with the basics. If one wants to define a function, the first step is to state the domain and co-domain. Usually, one writes $f : X \to Y$ to say that f has X as domain and Y as co-domain. In this Section, we will introduce an analogous way to label functions based on type theory. A quantum state that acts on a Hilbert space A is usually denoted as ρ^A. Therefore, one could label it with A. A quantum channel that maps operators on A to operators on B is denoted with $\mathcal{M}^{A\to B}$. Therefore, we could associate the label $A \to B$ with it. A superchannel that maps channels to channels would be labelled with $(A \to B) \to (C \to D)$, and so on.

We build all these possible labels, which we call types, with a bottom-up approach. At the lowest level of this theory of types, there is the alphabet of symbols. The reader should consider these symbols as labels for all the possible physical systems of the theory that are not the composition of two or more other systems.

Definition 6.1.1: Alphabet

An alphabet $\mathcal{A}$ is a non-empty set of symbols.

We usually denote the symbols in $\mathcal{A}$ with capital letters, e.g., A, B, C. From the symbols in an alphabet, one can construct strings. A string is a finite sequence of elements of $\mathcal{A}$ (repetitions are allowed). We denote with $\mathcal{A}^n$ the set of all strings of length $n \geq 0$. For example, if $\mathcal{A} = \{ A, B, C \}$, then $\mathcal{A}^2 = \{ AA, AB, AC, BA, BB, BC, CA, CB, CC \}$. In this example, if A and B are labels for two distinct physical systems, then AB is the label for the composite system. The set $\mathcal{A}^0$, which includes all the sequences of 0 symbols, is of special interest. Following the convention used in formal languages (e.g., Ref. [261]), we say that $\mathcal{A}^0$ contains only one element, the empty string, which we denote with I (we assume that $I \notin \mathcal{A}$, otherwise, one can choose as symbol for the empty string any symbol not in the alphabet). Continuing with the parallelism to quantum mechanics, one should think of the empty string as the trivial quantum system. We are ready to define the set of all strings of finite length formed by symbols in $\mathcal{A}$.

Definition 6.1.2: Elementary types

The set of elementary types of an alphabet $\mathcal{A}$, denoted with $\mathrm{EleTypes}_{\mathcal{A}}$, is the set of all finite sequences of elements of $\mathcal{A}$. That is,

$$\mathrm{EleTypes}_{\mathcal{A}} = \bigcup_{n \geq 0} \mathcal{A}^n. \tag{6.1}$$

As usually done for strings, we introduce an operation '|': $\mathrm{EleTypes}_{\mathcal{A}} \times \mathrm{EleTypes}_{\mathcal{A}} \to \mathrm{EleTypes}_{\mathcal{A}}$ defined as follows. Let $E, G \in \mathrm{EleTypes}_{\mathcal{A}}$, which means $E = e_1 \dots e_m$, $G = g_1 \dots g_n$, for some $e_i, g_j \in \mathcal{A}$ and $m, n \geq 0$. Then $E|G = e_1 \dots e_m g_1 \dots g_m$. This operation is usually called concatenation. Later, we will see an extension of this operation, which we will name 'type product'. For elementary types, we will use both concatenation and product. Moreover, we will drop the symbol '|' whenever there is no ambiguity, e.g., we will write EG instead of $E|G$.

We introduced the product of elementary types here to give further insights into the role of the empty string. Indeed, concatenating any type E with the empty string gives E: $EI = E = IE$. This is what we expect from the trivial quantum system. Considering the composition of a quantum system E and the trivial system, the resulting system is E. Because of this analogy, we will often refer to I as the trivial type.

So far, we have only introduced elementary types equivalent to quantum systems. We want to define the equivalent of quantum channels, superchannels, etc. What we want is a set that contains expressions that can be used to label maps, such as $(A \to B)$ or $((A \to B) \to (C \to D))$, but does not contain expressions like $)($ or $A \to)$. This is achieved with the following definition.

Definition 6.1.3: Types

The set of $\mathrm{Types}_{\mathcal{A}}$ of an alphabet $\mathcal{A}$ is the set of strings of symbols in $\mathcal{A} \cup \{ (,), \rightarrow, I \}$ recursively defined as follows:

- Each element $E \in \mathrm{EleTypes}_{\mathcal{A}}$ is an element of $\mathrm{Types}_{\mathcal{A}}$.
- If $x, y \in \mathrm{Types}_{\mathcal{A}}$, then $(x \rightarrow y) \in \mathrm{Types}_{\mathcal{A}}$, where $(x \rightarrow y)$ is the concatenation of '(' with the symbols of x, followed by the concatenation with '$\rightarrow$', the symbols in y, and ')'.

An element $z \in \mathrm{Types}_{\mathcal{A}}$ is either an elementary type or the concatenation of an open round bracket, another type x, an arrow, another type y, and a closed round bracket. Both x and y are, in turn, either elementary or a concatenation of symbols, as described above. Since a string is a finite sequence of symbols, this recursion ends, and what is left is a concatenation of elementary types with elements of $\{ (,), \rightarrow \}$. For example, consider the string $z = ((A \rightarrow B) \rightarrow C)$. We want to check if it is a type. Clearly $z = (x \rightarrow y)$, with $x = (A \rightarrow B)$ and $y = C$. Here, y is an elementary type, while x is not. From this, we deduce that y is a type, while we still do not know if x is a type. However, we observe that $x = (A \rightarrow B)$, where both A and B are elementary types (and thus types). By Definition 6.1.2, x is a type. Since x and y are types and $z = (x \rightarrow y)$, then z is also a type. On the other hand, let $z =)(A \rightarrow)$. In this case, z is not elementary, and the first symbol is not (, therefore $z \notin \mathrm{Types}_{\mathcal{A}}$.

Definition 6.1.3 allows for the recursive construction of types. We assume that any type can be obtained only with a finite number of recursive steps. We do not study the case of types obtained with infinite recursion.

Our goal was to construct a set of strings that can be used to label quantum maps. If $z \in \mathrm{EleTypes}_{\mathcal{A}}$, then z is the label for a quantum system, if $z \in \mathrm{Types}_{\mathcal{A}} \setminus \mathrm{EleTypes}_{\mathcal{A}}$, then $z = (x \rightarrow y)$ is the label from a map that has 'the maps of type x' as domain and 'the maps of type y' as co-domain. We will formalize this intuition in the next Section. For example, the type associated with a channel is $(A \rightarrow B)$ because it sends operators on A to operators on B. The type for a superchannel is $((A \rightarrow B) \rightarrow (C \rightarrow D))$ because it maps channels of type $(A \rightarrow B)$ to channels of type $(C \rightarrow D)$.

6.1.1 Trees

So far, types are very abstract objects. In this Section, we introduce a graphical representation of types that will be beneficial for understanding and deriving many properties of types. Each type in $\mathrm{Types}_{\mathcal{A}}$ can be graphically represented as a full binary tree with leaves labelled by elements of $\mathrm{EleTypes}_{\mathcal{A}}$. Here, we adapt the definition from Ref. [262] to our scenario.

Definition 6.1.4: Full binary tree (cf. [262], Sec. 5.3, Def. 5)

The set of full binary trees with leaves labelled by elements of $\mathrm{EleTypes}_{\mathcal{A}}$, denoted with $\mathrm{Trees}_{\mathcal{A}}$, is defined recursively by these steps:

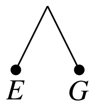
- There is a full binary tree for each $E \in \mathrm{EleTypes}_{\mathcal{A}}$ consisting only of a single vertex labelled with E.
- If T_1 and T_2 are full binary trees, then the tree $T_1 \bullet T_2$, consisting of a root, an edge connecting the root with the root of T_1 on the left, and another edge connecting the root with the root of T_2 on the right, is a full binary tree.

A full binary tree labelled by elements in $\mathrm{EleTypes}_{\mathcal{A}}$ is a tree such that every node has exactly zero or two children. If a node has zero children, we say that it is a leaf and it is labelled by an element of $\mathrm{EleTypes}_{\mathcal{A}}$. If a node has no parent, it is the tree's root. Each full binary tree has a single root.

Graphically, we represent the full binary tree associated with an elementary type E as

.

If T_1 and T_2 are the trees associated with the elementary types E and G, then the tree $T_1 \bullet T_2$ is



.

With these elements, one can recursively construct all the full binary trees labelled by elements of $\mathrm{EleTypes}_{\mathcal{A}}$.

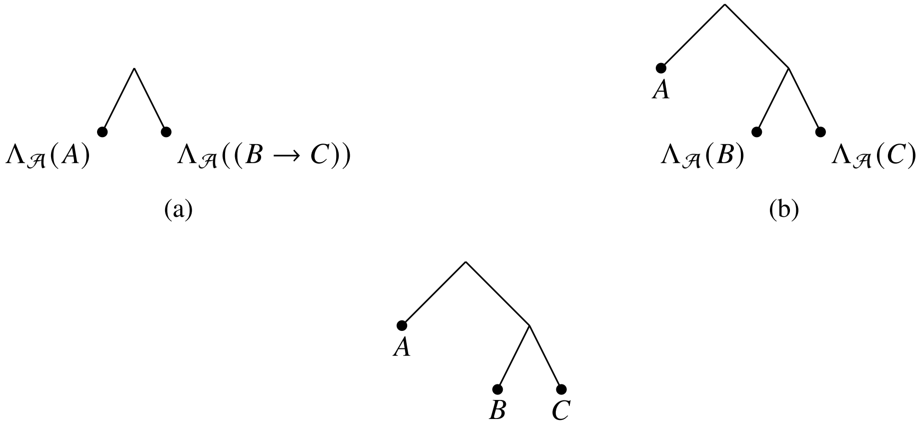
The reader may have noticed that the definitions of $\mathrm{Types}_{\mathcal{A}}$ and $\mathrm{Trees}_{\mathcal{A}}$ are very similar. Indeed, the two sets are isomorphic, as we will show now. To do so, we show that there exists an invertible map from $\mathrm{Types}_{\mathcal{A}}$ to $\mathrm{Trees}_{\mathcal{A}}$. We define $\Lambda_{\mathcal{A}} : \mathrm{Types}_{\mathcal{A}} \to \mathrm{Trees}_{\mathcal{A}}$ recursively as follows:

- If $E \in \mathrm{EleTypes}_{\mathcal{A}}$, then $\Lambda_{\mathcal{A}}(E)$ is the full binary tree associated with the elementary type E.
- If $z \in \mathrm{Types}_{\mathcal{A}} \setminus \mathrm{EleTypes}_{\mathcal{A}}$, then $z = (x \to y)$ for some $x \in \mathrm{Types}_{\mathcal{A}}$ and $y \in \mathrm{Types}_{\mathcal{A}}$ and $\Lambda_{\mathcal{A}}(z) = \Lambda_{\mathcal{A}}(x) \bullet \Lambda_{\mathcal{A}}(y)$.

We denote our candidate for the inverse of $\Lambda_{\mathcal{A}}$ with $\Lambda_{\mathcal{A}}^{-1}$ (a hopeful abuse of notation), and we define it recursively as follows:

- If $T \in \mathrm{Trees}_{\mathcal{A}}$ is the tree associated with an elementary type E, then $\Lambda_{\mathcal{A}}^{-1}(T) = E$.
- If $T = T_1 \bullet T_2$, then $\Lambda_{\mathcal{A}}^{-1}(T_1 \bullet T_2) = (\Lambda_{\mathcal{A}}^{-1}(T_1) \to \Lambda_{\mathcal{A}}^{-1}(T_2))$.

To prove that $\Lambda_{\mathcal{A}}^{-1}$ is both the left and the right inverse of $\Lambda_{\mathcal{A}}$, we use a technique called structural induction (see Ref. [262]). In the base step, one has to show that the result holds for elements in the base step of the definition of a set (in our case, for elementary types). In the recursive step, one proves that if the



(a) (b) (c)

Figure 6.1: Construction of the tree associated with $(A \to (B \to C))$.

The tree associated with $(A \to (B \to C))$ is the tree that has $\Lambda_{\mathcal{A}}(A)$ as left subtree and $\Lambda_{\mathcal{A}}((B \to C))$ as right subtree, represented in Figure a. Then we observe that $\Lambda_{\mathcal{A}}(A)$ is the node labelled with A, while the tree $\Lambda_{\mathcal{A}}((B \to C))$ has $\Lambda_{\mathcal{A}}(B)$ as left subtree, and $\Lambda_{\mathcal{A}}(C)$ as right subtree. This step is represented in Figure b. Lastly, we observe that $\Lambda_{\mathcal{A}}(B)$ and $\Lambda_{\mathcal{A}}(C)$ are the trees with only one node labelled with B and C, respectively. This concludes the construction of the tree associated with $(A \to (B \to C))$, shown in Figure c.

statement is true for each element used to construct a new element in the recursive step of the definition, then the result holds for the new element as well.

Theorem 6.1.5: Types and trees isomorphism

$\text{Types}_{\mathcal{A}}$ and $\text{Trees}_{\mathcal{A}}$ are isomorphic and $\Lambda_{\mathcal{A}}$ is an isomorphism.

Proof. We start by showing that $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(z) = z$ for each $z \in \text{Types}_{\mathcal{A}}$. *Basic step:* Let $E \in \text{EleTypes}_{\mathcal{A}}$, then $\Lambda_{\mathcal{A}}(E)$ is the tree associated with the elementary type E. Therefore $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(E) = E$ by definition of $\Lambda_{\mathcal{A}}^{-1}$. *Recursive step:* Let $z = (x \to y)$ for some $x, y \in \text{Types}_{\mathcal{A}}$. We observe that $\Lambda_{\mathcal{A}}(z) = \Lambda_{\mathcal{A}}(x) \bullet \Lambda_{\mathcal{A}}(y)$. By definition of $\Lambda_{\mathcal{A}}^{-1}$, we obtain that $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(z) = (\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(x) \to \Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(y))$. However, by induction hypothesis, $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(x) = x$ and $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(y) = y$. Therefore, $\Lambda_{\mathcal{A}}^{-1} \circ \Lambda_{\mathcal{A}}(z) = (x \to y) = z$.

The proof of $\Lambda_{\mathcal{A}} \circ \Lambda_{\mathcal{A}}^{-1}(z) = z$ for each $z \in \text{Types}_{\mathcal{A}}$ is analogous. □

With this theorem, we have validated the intuition that we can use trees to represent types graphically. We point out that there are many isomorphisms between types and trees. For example, we could have defined a map $\tilde{\Lambda}_{\mathcal{A}}$ with a recursive step such that $\tilde{\Lambda}_{\mathcal{A}}((x \to y)) = \tilde{\Lambda}_{\mathcal{A}}(y) \bullet \tilde{\Lambda}_{\mathcal{A}}(x)$. The difference between $\Lambda_{\mathcal{A}}$ and $\tilde{\Lambda}_{\mathcal{A}}$ is that the leaves of $\Lambda_{\mathcal{A}}(z)$ are filled with the elementary types that appear in z from left to right, while the leaves of $\tilde{\Lambda}_{\mathcal{A}}(z)$ are filled from right to left. Throughout this Thesis, we will always use $\Lambda_{\mathcal{A}}$ as the isomorphism from types to trees. That is, $\Lambda_{\mathcal{A}}(x)$ is the tree associated with type x, and $\Lambda_{\mathcal{A}}^{-1}(T)$ is *the* type associated with the tree T.

Thanks to Theorem 6.1.5, we can define the following objects associated with types.

Definition 6.1.6: Order and structure of a type

The *order* of a type $x \in \mathrm{Types}_{\mathcal{A}}$, denoted with $\mathrm{ord}(x)$, is the maximum number of distinct edges (traversed only once) connecting the root of the tree associated with x with a leaf. The *structure* of a type $x \in \mathrm{Types}_{\mathcal{A}}$, denoted with $\mathrm{struct}(x)$, is the full unlabelled binary tree obtained by removing all labels from the tree associated with x.

The order of a type corresponds to the *depth* of the associated tree; we will use the word depth when referring to trees and order when referring to types. With an abuse of language, we often write the structure of a tree $T \in \mathrm{Trees}_{\mathcal{A}}$ when we mean the structure of the type associated with the tree T.

The order of a type represents the order in the map hierarchy. For example, a state ρ^A is associated with the type A, a type or order zero. A channel, $\mathcal{M}^{A\to B}$, is associated with the type $(A \to B)$, which is a type of order one. A superchannel $\Theta^{(A\to B)\to(C\to D)}$ is associated with a type of order two, and so on. Therefore, the order is increased by one at every step in the hierarchy. We formalize this idea with the following Lemma.

Lemma 6.1.7: Properties of the order of a type

If $E \in \mathrm{EleTypes}_{\mathcal{A}}$, then $\mathrm{ord}(E) = 0$. If $z = (x \to y) \in \mathrm{Types}_{\mathcal{A}}$, then $\mathrm{ord}(z) = \max\{\mathrm{ord}(x), \mathrm{ord}(y)\} + 1$.

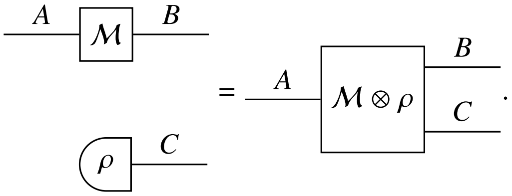
Proof. Let $E \in \mathrm{EleTypes}_{\mathcal{A}}$, then the tree associated with E consists of a single node, both the root and the only leaf of the tree. Therefore, no edges connect the root to the leaf, and $\mathrm{ord}(E) = 0$. Let $z = (x \to y)$ for some $x, y \in \mathrm{Types}_{\mathcal{A}}$. By definition, the tree associated with z has a root connected on the left with an edge to the tree associated with x and connected on the right with an edge to the tree associated with y. Therefore, the longest path from root to leaf on the left subtree is $1 + \mathrm{ord}(x)$, while the longest path from root to a leaf on the right subtree is $1 + \mathrm{ord}(y)$. Overall, the longest path from root to leaf is $\max\{\mathrm{ord}(x), \mathrm{ord}(y)\} + 1$. □

The order of a type will be a powerful tool for inductive proofs. In the base step, we will show that a statement is true for types of order zero, that is, for elementary types. In the inductive step, we will assume that a statement is true for types of order n (or up to order n) and prove that it is true for types of order $n + 1$.

6.1.2 Composition of Types

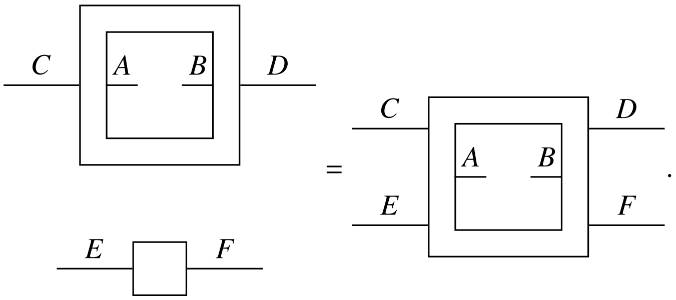
The goal of our type system is to label maps. If we have two maps, there are various ways to compose them. Let us consider the case of quantum channels as a reference. If $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{B\to C}$ are two quantum channels, then one can compose them sequentially to obtain the channel $\mathcal{T}^{A\to C} = \mathcal{N}^{B\to C} \circ \mathcal{M}^{A\to B}$. In our type system, if one sequentially composes a map of type $(x \to y)$ with a map of type $(y \to z)$, one obtains a map of type $(x \to z)$. Moreover, one can apply the channel $\mathcal{M}^{A\to B}$ to the state ρ^A and obtain the state $\sigma^B = \mathcal{M}^{A\to B}(\rho^A)$. Once again, this is easily translated into the type system. If one applies a map of type $(x \to y)$ to a map of type x, one obtains a map of type y. Lastly, if one composes in parallel the channel $\mathcal{M}^{A\to B}$ with the channel $\mathcal{N}^{C\to D}$, one obtains the channel $\mathcal{T}^{AC\to BD} = \mathcal{M}^{A\to B} \otimes \mathcal{N}^{C\to D}$. If we try to translate this into the type language, we would say that if one composes in parallel a map of type $(a \to b)$ with a map of type $(c \to d)$, one obtains a map of type $(ac \to bd)$. However, there are some problems with this translation.

First, observe that ac and bd are not well-defined types. Indeed, so far, we have not given any meaning to writing two types next to each other (unless they are elementary). Second, this is not the operation that we always want. For example, when one composes in parallel a channel $\mathcal{M}^{A\to B}$ with a state ρ^C, what they usually mean is that they compose $\mathcal{M}^{A\to B}$ with the channel that prepares ρ^C. Therefore, the resulting map is of type $(A \to BC)$. Diagrammatically [85, 237–242]:



(6.2)

Similarly if we composed in parallel a superchannel of type $((A \to B) \to (C \to D))$ with a channel of type $(E \to F)$ we would expect as outcome a superchannel of type $((A \to B) \to (CE \to DF))$ [263]:



(6.3)

In all these examples, the type of lower order ends up on the right of the type associated with the parallel composition of two maps. This motivates the following definition.

Definition 6.1.8: Product of types

For $x, y \in \mathrm{Types}_{\mathcal{A}}$, the product of x and y, denoted with $x|y$, is defined inductively as follows.

- If $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\} = 0$, then $x = E$, $y = G$ for some $E, G \in \mathrm{EleTypes}_{\mathcal{A}}$ and $x|y = E|G$, where $E|G$ is the concatenation of elementary type introduced below Definition 6.1.2.
- If $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\} > 0$ and
  - $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $x = (a \to b)$, $y = (c \to d)$ and $x|y = (a|c \to b|d)$, where $a, b, c, d \in \mathrm{Types}_{\mathcal{A}}$.
  - $\mathrm{ord}(x) > \mathrm{ord}(y)$, then $x = (a \to b)$ and $x|y = (a \to b|y)$, where $a, b \in \mathrm{Types}_{\mathcal{A}}$.
  - $\mathrm{ord}(x) < \mathrm{ord}(y)$, then $y = (c \to d)$ and $x|y = (c \to x|d)$, where $c, d \in \mathrm{Types}_{\mathcal{A}}$.

This definition covers both the case of the parallel composition of a channel with a channel, where we expect the inputs of both channels to be on the left and the outputs on the right of the type associated with parallel composition, and the composition of a channel with a state, or a channel with a superchannel, where we expect the type of the map of lower order to be on the right.

Note that we have the following relation between the order of the product of two types and the order of the two types.

Lemma 6.1.9: Order of product of types

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, the $\mathrm{ord}(x|y) = \max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\}$.

Proof. We prove this statement by induction on $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\}$. If $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\} = 0$, then the two types are elementary and their product is an elementary type. Now, assume that $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\} = n + 1 > 0$, and assume that the statement holds for all types up to order n. In this case, we have two split the proof in the three cases of Definition 6.1.8, and we use the same notation, that is, if $\mathrm{ord}(x) > 0$, we write $x = (a \to b)$ and if $\mathrm{ord}(y) > 0$, we write $y = (c \to d)$.

- If $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $\mathrm{ord}(a), \mathrm{ord}(b) \leq n$ and at least one of them is equal to n (see Lemma 6.1.7). Similarly, $\mathrm{ord}(c), \mathrm{ord}(d) \leq n$. The induction hypothesis implies that $\mathrm{ord}(a|c), \mathrm{ord}(b|d) \leq n$ and at least one is of order n. Therefore, the order of $x|y = (a|c \to b|d)$ is of order $n + 1$.

- If $n+1 = \mathrm{ord}(x) > \mathrm{ord}(y)$, then $\mathrm{ord}(a), \mathrm{ord}(b) \leq n$ and at least one of them is of order n. By induction hypothesis, $\mathrm{ord}(a), \mathrm{ord}(b|y) \leq n$ and at least one of them is of order n. Therefore, $x|y = (a \to b|y)$ is of order $n + 1$.

- If $n + 1 = \mathrm{ord}(y) > \mathrm{ord}(x)$, then $\mathrm{ord}(c), \mathrm{ord}(d) \leq n$ and at least one of them is of order n. As before, this implies that $x|y = (c \to x|d)$ is of order $n + 1$.

□

The isomorphism between types and trees induces an operation on trees that is compatible with the product of types and is defined as $T_1|T_2 = \Lambda_{\mathcal{A}}(\Lambda_{\mathcal{A}}^{-1}(T_1)|\Lambda_{\mathcal{A}}^{-1}(T_2))$. By definition $\Lambda_{\mathcal{A}}(x|y) = \Lambda_{\mathcal{A}}(x)|\Lambda_{\mathcal{A}}(y)$ and $\Lambda^{-1}(T_1|T_2) = \Lambda_{\mathcal{A}}^{-1}(T_1)|\Lambda_{\mathcal{A}}^{-1}(T_2)$. Moreover, we can explicitly write the action of '|' as follows.

- If T_1 and T_2 are trees of depth zero, then they are associated with $E, G \in \mathrm{EleTypes}_{\mathcal{A}}$, and $T_1|T_2 = \Lambda_{\mathcal{A}}(E|G)$ is the three associated with the elementary type $E|G$.

- If T_1 and T_2 are the trees of equal depth strictly greater than zero, then $T_1 = T_1^L \bullet T_1^R$ and $T_2 = T_2^L \bullet T_2^R$, where the L and R superscripts denote the left and right subtree. By definition of $\Lambda_{\mathcal{A}}$, $\Lambda_{\mathcal{A}}^{-1}$, and '|', one obtains
$$\begin{aligned} T_1|T_2 &= \Lambda_{\mathcal{A}}(\Lambda_{\mathcal{A}}^{-1}(T_1)|\Lambda_{\mathcal{A}}^{-1}(T_2)) \\ &= \Lambda_{\mathcal{A}}(\Lambda_{\mathcal{A}}^{-1}(T_1^L \bullet T_1^R)|\Lambda_{\mathcal{A}}^{-1}(T_2^L \bullet T_2^R)) \\ &= \Lambda_{\mathcal{A}}((\Lambda_{\mathcal{A}}^{-1}(T_1^L) \to \Lambda_{\mathcal{A}}^{-1}(T_1^R))|(\Lambda_{\mathcal{A}}^{-1}(T_2^L) \to \Lambda_{\mathcal{A}}^{-1}(T_2^R))) \\ &= \Lambda_{\mathcal{A}}((\Lambda_{\mathcal{A}}^{-1}(T_1^L)|\Lambda_{\mathcal{A}}^{-1}(T_2^L)) \to (\Lambda_{\mathcal{A}}^{-1}(T_1^R)|\Lambda_{\mathcal{A}}^{-1}(T_2^R))) \\ &= \Lambda_{\mathcal{A}}(\Lambda_{\mathcal{A}}^{-1}(T_1^L|T_2^L) \to \Lambda_{\mathcal{A}}^{-1}(T_1^R|T_2^R)) \\ &= (\Lambda_{\mathcal{A}} \circ \Lambda_{\mathcal{A}}^{-1}(T_1^L|T_2^L)) \bullet (\Lambda_{\mathcal{A}} \circ \Lambda_{\mathcal{A}}^{-1}(T_1^R|T_2^R)) \\ &= (T_1^L|T_2^L) \bullet (T_1^R|T_2^R) \end{aligned} \tag{6.4}$$

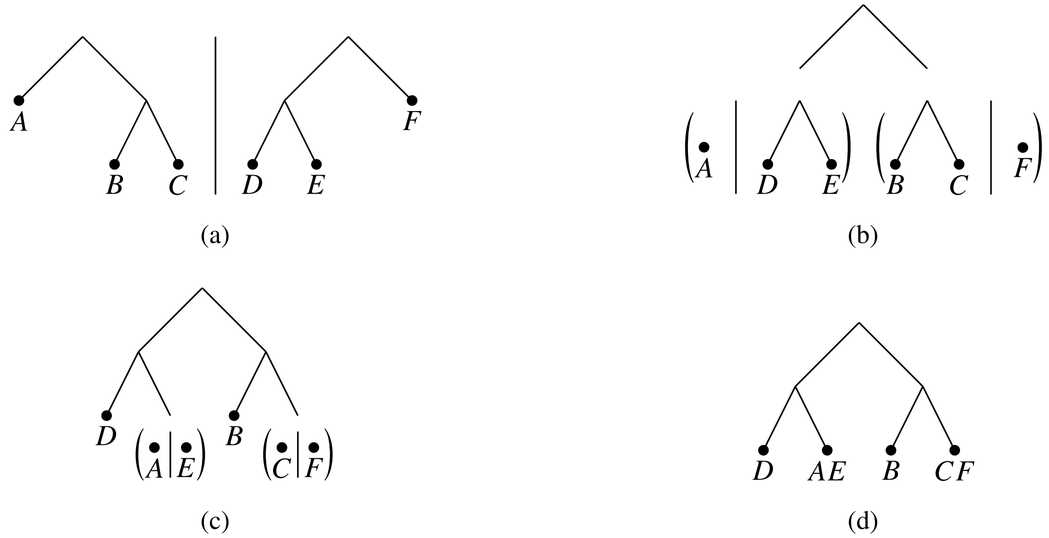


Figure 6.2: Product of the trees associated with $(A \rightarrow (B \rightarrow C))$ and $((D \rightarrow E) \rightarrow F)$.

Both trees have depth two. Therefore, the resulting tree has as the left subtree the product of the left subtrees of the two initial trees, and as the right subtree the product of the right subtrees (Figure a to Figure b). Both products are between a tree of depth zero and a tree of depth one. In this case, the tree with depth zero falls on the right of the tree of depth one (Figure b to Figure c). Lastly, the product of trees of order zero is simply a tree of order zero in which the labels are concatenated (Figure c to Figure d).

That is, $T_1|T_2$ is the tree that has as the left subtree the product of the left subtree of T_1 with the left subtree of T_2 and has as the right subtree the product of the right subtree of T_1 with the right subtree of T_2.

- If T_1 and T_2 are trees such that the depth of T_1 is strictly greater than the depth of T_2, then $T_1|T_2 = T_1^L \bullet (T_1^R|T_2)$. That is, $T_1|T_2$ is the tree that has as left subtree the left subtree of T_1 and as right subtree the product of the right subtree of T_1 with T_2 (same proof as above).

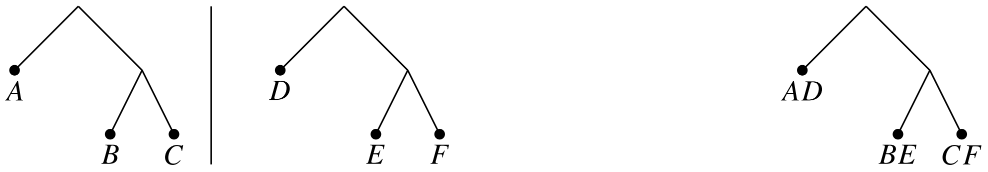
- If T_1 and T_2 are trees such that the depth of T_1 is strictly smaller than the depth of T_2, then $T_1|T_2 = T_2^L \bullet (T_1|T_2^R)$. That is, $T_1|T_2$ is the tree that has as left subtree the left subtree of T_2, and as right subtree the product of T_1 with the right subtree of T_2 (same proof as above).

In the following, we will drop the symbol '|' whenever there is no ambiguity. We will write xy to denote $x|y$ and T_1T_2 to denote $T_1|T_2$.

As we have discussed, the product of two types associated with channels gives a type associated with a channel. This is a special case of a more general property regarding the product of types with the same structure.

Lemma 6.1.10: Product of types with the same structure

Let $x, y \in \text{Types}_{\mathcal{A}}$ such that $\text{struct}(x) = \text{struct}(y)$. Then, $\text{struct}(xy) = \text{struct}(x)$ and each leaf of $\Lambda_{\mathcal{A}}(xy)$ is labelled with the concatenation of the elementary types that label the corresponding leaves of $\Lambda_{\mathcal{A}}(x)$ and $\Lambda_{\mathcal{A}}(y)$.



(a) (b)

Figure 6.3: Product of types with the same structure.

The product of two types with the same structure can be easily computed with trees. It is enough to copy the structure of the types and fill the leaves with concatenation of the labels in the corresponding positions in the two factors (Figure a to Figure b).

Proof. We prove it by induction on the order of x, which is equal to the order of y. If $\operatorname{ord}(x) = \operatorname{ord}(y) = 0$, then xy is just the concatenation of elementary types and $\operatorname{struct}(xy) = \operatorname{struct}(x)$. If $\operatorname{ord}(x) = \operatorname{ord}(y) = n + 1 > 0$, then $\Lambda_{\mathcal{A}}(xy)$ is by construction $T^L_x T^L_y \bullet T^R_x T^R_y$, where T^L_x is the left subtree of $\Lambda_{\mathcal{A}}(x)$, and T^L_y, T^R_x, and T^R_y are defined analogously. Since x and y have the same structure, their left subtrees have the same structure. Therefore, $T^L_x T^L_y$ has the same structure as T^L_x by induction hypothesis (we assume that the statement is true for all types up to order n). Analogously, $T^R_x T^R_y$ has the same structure as T^R_x. Since x and xy have the same left and right subtree structures, they have the same structure. Moreover, by induction hypothesis, the labels of the leaves in the left subtree are the concatenation of the corresponding leaves in the left subtrees of x and y. The same argument holds for the right subtree. Therefore, the labels in the leaves of xy are the concatenation of the labels in the corresponding positions in the trees associated with x and y. $\square$

6.1.3 Properties of the Product of Types

In this Section, we will list properties of the product of types. We start by proving that the product of types has a unit, the empty string I.

Lemma 6.1.11: Unit of product of types

The empty string I is the unit of the product of types, i.e., $Ix = x = xI$ for all $x \in \mathrm{Types}_{\mathcal{A}}$.

Proof. We prove this result by induction on the order of x. Let $\operatorname{ord}(x) = 0$, i.e., $x = E$, where $E \in \mathrm{EleTypes}_{\mathcal{A}}$. The concatenation of E with the empty string I gives E. Assume now that $\operatorname{ord}(x) > 0$, therefore $x = (a \longrightarrow b)$, where $\operatorname{ord}(b) < \operatorname{ord}(x)$. Since $\operatorname{ord}(I) = 0$, $xI = (a \longrightarrow bI)$ and by induction hypothesis $bI = b = Ib$. Therefore $xI = (a \longrightarrow b) = x$. Similarly $Ix = (a \longrightarrow Ib) = (a \longrightarrow b) = x$. $\square$

Once again, this matches the intuition from quantum information that composing a map of type x in parallel with the trivial system I gives a map of type x.

Now, we want to prove that the product of types is associative. To achieve this result, we first prove the following three lemmas. The goal of these lemmas is to provide abstract expressions for the product of two and three types that allow for an inductive proof. That is, we want to write the products xy, $(xy)z$, and $x(yz)$ in terms of products of types of lower order.

Lemma 6.1.12: Abstract expression for the product of two types

Let $x, y \in \mathrm{Types}_{\mathfrak{A}}$ such that $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y)\,\} > 0$. If $\mathrm{ord}(x) > 0$, let a, b be types such that $x = (a \to b)$. Similarly, if $\mathrm{ord}(y) > 0$, let c, d be types such that $y = (c \to d)$. Then $xy = (\alpha\gamma \to \beta\delta)$, where

$$\begin{aligned}
\alpha &= \begin{cases} I & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y), \\ a & \text{if } \mathrm{ord}(x) \geq \mathrm{ord}(y). \end{cases} &
\beta &= \begin{cases} x & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y), \\ b & \text{if } \mathrm{ord}(x) \geq \mathrm{ord}(y). \end{cases} \\
\gamma &= \begin{cases} c & \text{if } \mathrm{ord}(x) \leq \mathrm{ord}(y), \\ I & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y). \end{cases} &
\delta &= \begin{cases} d & \text{if } \mathrm{ord}(x) \leq \mathrm{ord}(y), \\ y & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y). \end{cases}
\end{aligned} \tag{6.5}$$

Proof. The proof follows from the definition of the product of types and of the empty string:

$$\begin{aligned}
xy &= \begin{cases} (ac \to bd) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ (a \to by) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ (c \to xd) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \\
&= \begin{cases} (ac \to bd) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ (aI \to by) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ (Ic \to xd) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases}
\end{aligned} \tag{6.6}$$

The expressions for α, β, γ, and δ follow from matching them with the types in the corresponding positions in the last equation. □

Lemma 6.1.13: Abstract expression for the product $(xy)z$

Let $x, y, z \in \mathrm{Types}_{\mathfrak{A}}$ be such that $\max\{\,\mathrm{ord}(x), \mathrm{ord}(y), \mathrm{ord}(z)\,\} = N > 0$. Let a, b, c, d, e, f be types such that $x = (a \to b)$ when $\mathrm{ord}(x) > 0$, $y = (c \to d)$ when $\mathrm{ord}(y) > 0$, and $z = (e \to f)$ when $\mathrm{ord}(z) > 0$. Then, $(xy)z = ((\alpha\gamma)\epsilon \to (\beta\delta)\phi)$, where

$$\begin{aligned}
\alpha &= \begin{cases} I & \text{if } \mathrm{ord}(x) < N, \\ a & \text{if } \mathrm{ord}(x) = N. \end{cases} &
\beta &= \begin{cases} x & \text{if } \mathrm{ord}(x) < N, \\ b & \text{if } \mathrm{ord}(x) = N. \end{cases} \\
\gamma &= \begin{cases} I & \text{if } \mathrm{ord}(y) < N, \\ c & \text{if } \mathrm{ord}(y) = N. \end{cases} &
\delta &= \begin{cases} y & \text{if } \mathrm{ord}(y) < N, \\ d & \text{if } \mathrm{ord}(y) = N. \end{cases} \\
\epsilon &= \begin{cases} I & \text{if } \mathrm{ord}(z) < N, \\ e & \text{if } \mathrm{ord}(z) = N. \end{cases} &
\phi &= \begin{cases} z & \text{if } \mathrm{ord}(z) < N, \\ f & \text{if } \mathrm{ord}(z) = N. \end{cases}
\end{aligned} \tag{6.7}$$

Proof. To prove this lemma we start with the expression for xy derived in Lemma 6.1.12, and then we compute $(xy)z$ splitting it into the cases $\mathrm{ord}(xy) = \mathrm{ord}(z)$, $\mathrm{ord}(xy) > \mathrm{ord}(z)$ and $\mathrm{ord}(xy) < \mathrm{ord}(z)$ as in

Definition 6.1.8.

$$(xy)z = \begin{cases} ((ac)e \to (bd)f) & \text{if } \mathrm{ord}(z) = \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) = \mathrm{ord}(y), \\ ((aI)e \to (by)f) & \text{if } \mathrm{ord}(z) = \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) > \mathrm{ord}(y), \\ ((Ic)e \to (xd)f) & \text{if } \mathrm{ord}(z) = \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) < \mathrm{ord}(y), \\ ((II)e \to (xy)f) & \text{if } \mathrm{ord}(z) > \mathrm{ord}(xy), \\ ((ac)I \to (bd)z) & \text{if } \mathrm{ord}(z) < \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) = \mathrm{ord}(y), \\ ((aI)I \to (by)z) & \text{if } \mathrm{ord}(z) < \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) > \mathrm{ord}(y), \\ ((Ic)I \to (xd)z) & \text{if } \mathrm{ord}(z) < \mathrm{ord}(xy) \text{ and } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \tag{6.8}$$

From this equation, one obtains the expressions for $\alpha, \beta, \ldots$. For example, $\mathrm{ord}(x) < N$ only in the third, fourth, and last expressions. In these cases, the first symbol is I. In all the other cases, the first symbol is a. This proves

$$\alpha = \begin{cases} I & \text{if } \mathrm{ord}(x) < N, \\ a & \text{if } \mathrm{ord}(x) = N. \end{cases} \tag{6.9}$$

□

Lemma 6.1.14: Abstract expression for the product $x(yz)$

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$ be such that $\max\{\mathrm{ord}(x), \mathrm{ord}(y), \mathrm{ord}(z)\} = N > 0$. Let a, b, c, d, e, f be types such that $x = (a \to b)$ when $\mathrm{ord}(x) > 0$, $y = (c \to d)$ when $\mathrm{ord}(y) > 0$, and $z = (e \to f)$ when $\mathrm{ord}(z) > 0$. Then, $x(yz) = (\alpha(\gamma\epsilon) \to \beta(\delta\phi))$, where

$$\begin{gathered} \alpha = \begin{cases} I & \text{if } \mathrm{ord}(x) < N, \\ a & \text{if } \mathrm{ord}(x) = N. \end{cases} \quad \beta = \begin{cases} x & \text{if } \mathrm{ord}(x) < N, \\ b & \text{if } \mathrm{ord}(x) = N. \end{cases} \\ \gamma = \begin{cases} I & \text{if } \mathrm{ord}(y) < N, \\ c & \text{if } \mathrm{ord}(y) = N. \end{cases} \quad \delta = \begin{cases} y & \text{if } \mathrm{ord}(y) < N, \\ d & \text{if } \mathrm{ord}(y) = N. \end{cases} \\ \epsilon = \begin{cases} I & \text{if } \mathrm{ord}(z) < N, \\ e & \text{if } \mathrm{ord}(z) = N. \end{cases} \quad \phi = \begin{cases} z & \text{if } \mathrm{ord}(z) < N, \\ f & \text{if } \mathrm{ord}(z) = N. \end{cases} \end{gathered} \tag{6.10}$$

Proof. Here, we first derive the expression for yz when $\mathrm{ord}(yz) > 0$:

$$yz = \begin{cases} (ce \to df) & \text{if } \mathrm{ord}(y) = \mathrm{ord}(z), \\ (cI \to dz) & \text{if } \mathrm{ord}(y) > \mathrm{ord}(z), \\ (Ie \to yf) & \text{if } \mathrm{ord}(y) < \mathrm{ord}(z). \end{cases} \tag{6.11}$$

Then we compute $x(yz)$ splitting it into the cases $\mathrm{ord}(x) = \mathrm{ord}(yz)$, $\mathrm{ord}(x) > \mathrm{ord}(yz)$ and $\mathrm{ord}(x) < \mathrm{ord}(yz)$

as in Definition 6.1.8.

$$(xy)z = \begin{cases} (a(ce) \to b(df)) & \text{if } \operatorname{ord}(x) = \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) = \operatorname{ord}(z), \\ (a(cI) \to b(dz)) & \text{if } \operatorname{ord}(x) = \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) > \operatorname{ord}(z), \\ (a(Ie) \to b(yf)) & \text{if } \operatorname{ord}(x) = \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) < \operatorname{ord}(z), \\ (a(II) \to b(yz)) & \text{if } \operatorname{ord}(x) > \operatorname{ord}(yz), \\ (I(ce) \to x(df)) & \text{if } \operatorname{ord}(x) < \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) = \operatorname{ord}(z), \\ (I(cI) \to x(dz)) & \text{if } \operatorname{ord}(x) < \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) > \operatorname{ord}(z), \\ (I(Ie) \to x(yf)) & \text{if } \operatorname{ord}(x) < \operatorname{ord}(yz) \text{ and } \operatorname{ord}(y) < \operatorname{ord}(z). \end{cases} \tag{6.12}$$

Here, we observe that $\operatorname{ord}(x) < N$ only in the last three expressions. In these cases, the first symbol is I. In all the other cases, the first symbol is a. This proves

$$\alpha = \begin{cases} I & \text{if } \operatorname{ord}(x) < N, \\ a & \text{if } \operatorname{ord}(x) = N. \end{cases} \tag{6.13}$$

Similarly, one finds the expressions for $\beta, \gamma, \ldots$ □

The expressions for $\alpha, \beta, \ldots,$ in Lemma 6.1.13 and in Lemma 6.1.14 are the same. This is what we need to prove associativity.

Theorem 6.1.15: Associativity of the product of types

The product of types is associative. That is, for all $x, y, z \in \mathrm{Types}_{\mathcal{A}}$,

$$(xy)z = x(yz). \tag{6.14}$$

Proof. We prove this theorem by induction on $\max \{ \operatorname{ord}(x), \operatorname{ord}(y), \operatorname{ord}(z) \}$. If $\max \{ \operatorname{ord}(x), \operatorname{ord}(y), \operatorname{ord}(z) \} = 0$ then all types are elementary. $(EG)F$ and $E(GF)$ are the sequence of symbols in E, followed by the symbols in G and the symbols in F. Therefore $(EG)F = E(GF)$. Let us assume now that $n > 0$. Then, $(xy)z = ((\alpha\gamma)\epsilon \to (\beta\delta)\phi)$ by Lemma 6.1.13. Since, $\operatorname{ord}(\alpha), \operatorname{ord}(\beta), \cdots < \max \{ \operatorname{ord}(x), \operatorname{ord}(y), \operatorname{ord}(z) \}$ we can use the induction hypothesis and we obtain

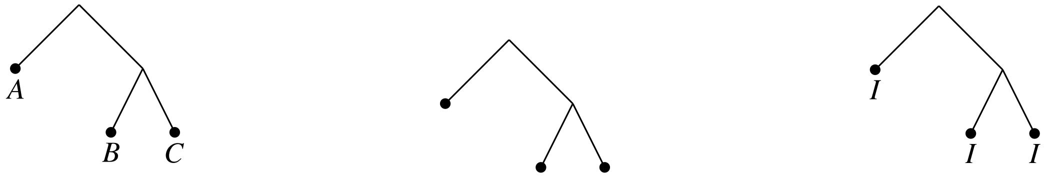
$$(xy)z = ((\alpha\gamma)\epsilon \to (\beta\delta)\phi) = (\alpha(\gamma\epsilon) \to \beta(\delta\phi)). \tag{6.15}$$

We conclude by noticing that the last expression is $x(yz)$, as proven in Lemma 6.1.14. □

Note that the set of types equipped with the product '|' is a monoid. Indeed, '|' is an associative binary operation on $\mathrm{Types}_{\mathcal{A}}$ and $I \in \mathrm{Types}_{\mathcal{A}}$ is the identity element of '|'.

6.1.4 Trivialization of Types

In the product of two types, the structure of the types plays a crucial role: It completely determines the structure of the product type. However, as we have discussed, the structure of a type is not a type; it is just an



(a) Tree associated with $x = (A \to (B \to C))$

(b) $\mathrm{struct}(x)$

(c) Tree associated with $\mathrm{triv}(x)$

Figure 6.4: Structure and trivialization of a type.

In Figure a, the tree associated with the type $x = (A \to (B \to C))$ is represented. In Figure b, the structure of x is represented, which is obtained by removing all the labels from the tree associated with x. In Figure c, the tree associated with $\mathrm{triv}(x)$ is represented. This tree is obtained by labelling the leaves of $\mathrm{struct}(x)$ with the trivial type I.

unlabelled tree. Here, we introduce for every type x a type that has the same structure as x but carries no information about the elementary types that appear in x.

Definition 6.1.16: Trivialization

The *trivialization* of a type $x \in \mathrm{Types}_{\mathcal{A}}$, denoted with $\mathrm{triv}(x)$, is the type associated with the tree obtained from $\Lambda_{\mathcal{A}}(x)$ by replacing all the leaves' labels with the empty string I.

With an abuse of language, we will sometimes write the trivialization of a tree T when we mean the trivialization of the type associated with T. In addition, we will refer to types obtained with this process as trivialized types.

One immediately deduces the following properties from the definition of $\mathrm{triv}(x)$.

Lemma 6.1.17: Properties of $\mathrm{triv}(x)$

The following expressions are true for every $x \in \mathrm{Types}_{\mathcal{A}}$.

- $\mathrm{struct}(\mathrm{triv}(x)) = \mathrm{struct}(x)$.
- $\mathrm{ord}(\mathrm{triv}(x)) = \mathrm{ord}(x)$.
- If $x = (a \to b)$, then $\mathrm{triv}(x) = (\mathrm{triv}(a) \to \mathrm{triv}(b))$.
- $\mathrm{triv}(\mathrm{triv}(x)) = \mathrm{triv}(x)$ (*Idempotence*).

In addition, the trivialization of a type has the following properties when in the product with another type.

Lemma 6.1.18: Product of trivialized types

The following expressions are true for all $x, y \in \mathrm{Types}_{\mathcal{A}}$.

- $\mathrm{triv}(x)x = x$ (*Absorption*).
- $\mathrm{triv}(y)x = x\,\mathrm{triv}(y)$ (*Centrality*).
- $\mathrm{triv}(xy) = \mathrm{triv}(x)\,\mathrm{triv}(y)$ (*Distributivity*).

Proof. Absorption follows from Lemma 6.1.10 and the definition of $\mathrm{triv}(x)$. Indeed, $\mathrm{triv}(x)$ and x have the same structure. Therefore, their product is a type with the same structure in which all the elementary types are the concatenation of the corresponding elementary types in $\mathrm{triv}(x)$ and x. Since all the elementary types in $\mathrm{triv}(x)$ are empty strings, the elementary types appearing in $\mathrm{triv}(x)x$ are the elementary types of x. Therefore, $\mathrm{triv}(x)x = x$.

We prove centrality by induction on $\mathrm{ord}(xy)$. If $\mathrm{ord}(xy) = 0$, then both x and y are elementary, $\mathrm{triv}(y) = I$, and we have shown in Lemma 6.1.11 that I is the unit of the product, therefore $\mathrm{triv}(y)x = x = x\,\mathrm{triv}(y)$. Now, we assume that $\mathrm{ord}(xy) > 0$ and that the centrality holds for all types up to order $\mathrm{ord}(xy) - 1$. Let $a, b, c, d \in \mathrm{Types}_{\mathcal{A}}$ be such that $x = (a \to b)$ when $\mathrm{ord}(x) > 0$ and $y = (c \to d)$ when $\mathrm{ord}(y) > 0$. From the definition of the product of types and the centrality of trivialized types of order lower than $\mathrm{ord}(xy)$, we have that

$$\begin{aligned}
\mathrm{triv}(y)x &= \begin{cases} (\mathrm{triv}(c)a \to \mathrm{triv}(d)b) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ (a \to \mathrm{triv}(y)b) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ (\mathrm{triv}(c) \to \mathrm{triv}(d)x) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \\
&= \begin{cases} (a\,\mathrm{triv}(c) \to b\,\mathrm{triv}(d)) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ (a \to b\,\mathrm{triv}(y)) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ (\mathrm{triv}(c) \to x\,\mathrm{triv}(d)) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \\
&= x\,\mathrm{triv}(y).
\end{aligned} \tag{6.16}$$

We also prove distributivity by induction on $\mathrm{ord}(xy)$. If $\mathrm{ord}(xy) = 0$, then x, y, and xy are elementary types. Therefore, $\mathrm{triv}(x)\,\mathrm{triv}(y) = II = I = \mathrm{triv}(xy)$. Now, we assume that $\mathrm{ord}(xy) > 0$ and that distributivity holds for the product of types up to $\mathrm{ord}(xy) - 1$. As before, let $a, b, c, d \in \mathrm{Types}_{\mathcal{A}}$ be such that $x = (a \to b)$ when $\mathrm{ord}(x) > 0$ and $y = (c \to d)$ when $\mathrm{ord}(y) > 0$. Then,

$$\mathrm{triv}(xy) = \begin{cases} \mathrm{triv}((ac \to bd)) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ \mathrm{triv}((a \to by)) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ \mathrm{triv}((c \to xd)) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \tag{6.17}$$

Using Lemma 6.1.17 we can write $\mathrm{triv}((\alpha \to \beta))$ as $(\mathrm{triv}(\alpha) \to \mathrm{triv}(\beta))$ and we obtain

$$\begin{cases} (\mathrm{triv}(ac) \to \mathrm{triv}(bd)) & \text{if } \mathrm{ord}(x) = \mathrm{ord}(y), \\ (\mathrm{triv}(a) \to \mathrm{triv}(by)) & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y), \\ (\mathrm{triv}(c) \to \mathrm{triv}(xd)) & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y). \end{cases} \tag{6.18}$$

Now, we can use distributivity for the types of order lower than $\mathrm{ord}(xy)$ together with $\mathrm{ord}(x) = \mathrm{ord}(\mathrm{triv}(x))$ to get

$$\begin{cases} (\mathrm{triv}(a)\,\mathrm{triv}(c) \to \mathrm{triv}(b)\,\mathrm{triv}(d)) & \text{if } \mathrm{ord}(\mathrm{triv}(x)) = \mathrm{ord}(\mathrm{triv}(y)), \\ (\mathrm{triv}(a) \to \mathrm{triv}(b)\,\mathrm{triv}(y)) & \text{if } \mathrm{ord}(\mathrm{triv}(x)) > \mathrm{ord}(\mathrm{triv}(y)), \\ (\mathrm{triv}(c) \to \mathrm{triv}(x)\,\mathrm{triv}(d)) & \text{if } \mathrm{ord}(\mathrm{triv}(x)) < \mathrm{ord}(\mathrm{triv}(y)). \end{cases} \tag{6.19}$$

This expression is the definition of the product of $\mathrm{triv}(x) = (\mathrm{triv}(a) \to \mathrm{triv}(b))$ with $\mathrm{triv}(y) = (\mathrm{triv}(c) \to \mathrm{triv}(d))$. This concludes the proof. □

Same useful identities based on the results of Lemma 6.1.17 and Lemma 6.1.18 are

$$\mathrm{triv}(x)\,\mathrm{triv}(x) = \mathrm{triv}(x) \tag{6.20}$$

$$\mathrm{triv}(xy) = \mathrm{triv}(yx) \tag{6.21}$$

$$\mathrm{triv}(x\,\mathrm{triv}(y)) = \mathrm{triv}(xy) \tag{6.22}$$

6.1.5 Promotion of Types

In this subsection, we present how to add to a type x the structure of another type y. More than being interesting per se, the content of this subsection will be beneficial when we define the 'tensor product' of maps of different types.

Definition 6.1.19: Type Promotion

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. The promotion of x with respect to y, denoted with x_y is defined as

$$x_y = x\,\mathrm{triv}(y). \tag{6.23}$$

Moreover, x is the *principal type*, and y the *subordinate type* of x_y.

For example, if one promotes the elementary type A with the type $B \to C$, one obtains $A|\,\mathrm{triv}(B \to C) = A|(I \to I) = I \to A$. That is, if one promotes a type associated with a state ρ^A with a type associated with a channel, one obtains the type of the channel that prepares ρ^A.

The following Lemma lists useful identities involving the promotion of a type.

Lemma 6.1.20: Properties of the promotion of a type

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$. The following statements are true.

1. $x_x = x_I = x$.
2. $x_{(y_z)} = (x_y)_z = x_{yz}$.
3. $x_z y = x(y_z) = (xy)_z = x_z y_z$.
4. $x_{yz} = x_{zy}$.
5. $x_{yy} = x_y$.

Proof. For the first property, note that $x_x = x\,\mathrm{triv}(x) = x$ by idempotence. Moreover, $x_I = x\,\mathrm{triv}(I) = xI = x$ because the empty string is the unit of the product.

For the second property, observe that

$$x_{y_z} = x_{y\,\mathrm{triv}(z)} = x\,\mathrm{triv}(y\,\mathrm{triv}(z)) = x\,\mathrm{triv}(y)\,\mathrm{triv}(z). \tag{6.24}$$

From this, if we associate on the left, we obtain the second expression of the equality. If instead we associate on the right and use distributivity of triv, we get the last expression in the equality.

For the third property, observe that

$$x_z y = x\,\mathrm{triv}(z)y. \tag{6.25}$$

Using the centrality of $\mathrm{triv}(z)$ one obtains $xy\,\mathrm{triv}(z)$. If we associate on the right, we get xy_z; on the left, we get $(xy)_z$. Lastly, observe that $\mathrm{triv}(z) = \mathrm{triv}(z)\,\mathrm{triv}(z)$. Therefore,

$$x\,\mathrm{triv}(z)y = x\,\mathrm{triv}(z)y\,\mathrm{triv}(z) = x_z y_z. \tag{6.26}$$

For the fourth property, we use Eq. (6.21): $x_{yz} = x\,\mathrm{triv}(yz) = x\,\mathrm{triv}(zy) = x_{zy}$.

Lastly, from distributivity and Eq. (6.20), we get $x_{yy} = x\,\mathrm{triv}(yy) = x\,\mathrm{triv}(y)\,\mathrm{triv}(y) = x\,\mathrm{triv}(y) = x_y$. □

This Lemma allows us to drop brackets when dealing with products of type extensions. Moreover, the order of the subordinate types does not affect the promotion of the principal type.

Thanks to the promotion of types, we can write the product xy in terms of a promotion of x and a promotion of y with the same structure as xy.

Corollary 6.1.21: Product of promotion of types

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, the following statements are true.

1. $xy = x_y y_x$.
2. $\mathrm{triv}(x_y) = \mathrm{triv}(y_x) = \mathrm{triv}(xy) = \mathrm{triv}(yx)$.
3. $\mathrm{struct}(x_y) = \mathrm{struct}(y_x) = \mathrm{struct}(xy)$.

Proof. From the previous Lemma, we have that

$$x_y y_x = (x y_x)_y = x y_{xy} = xy. \tag{6.27}$$

This proves the first property.

Using distributivity, idempotence, centrality, and absorption, we get

$$\begin{aligned}\operatorname{triv}(x_y) &= \operatorname{triv}(x \operatorname{triv}(y)) = \operatorname{triv}(x) \operatorname{triv}(y) = \operatorname{triv}(xy)\\ \operatorname{triv}(y_x) &= \operatorname{triv}(y \operatorname{triv}(x)) = \operatorname{triv}(y) \operatorname{triv}(x) = \operatorname{triv}(x) \operatorname{triv}(y) = \operatorname{triv}(xy).\end{aligned} \tag{6.28}$$

Together with Eq. (6.21), this proves the second property. The last property simply follows from observing that $\operatorname{struct}(\operatorname{triv}(z)) = \operatorname{struct}(z)$ for all $z \in \mathrm{Types}_{\mathcal{A}}$ (Lemma 6.1.17). □

Sometimes, promoting a type with another type leaves the first type unchanged. In a sense, the principal type already contains the structure of the subordinate type.

Definition 6.1.22: Inclusion of types

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, we say that y includes the structure of x (or the structure of x is included in y), and we write $x \prec y$ if $y_x = y$.

This relation is a preorder of types.

Lemma 6.1.23: Preorder of types

The inclusion of types is a preorder. That is,

- $x \prec x$,
- $x \prec y$ and $y \prec z$ imply $x \prec z$.

Proof. The first condition follows from $x_x = x$ (see Lemma 6.1.20).

To prove the second condition, observe that $z_x = (z_y)_x$ because $y \prec z$. By Lemma 6.1.20, this is equal to z_{y_x}. Since $x \prec y$, we have that $z_{y_x} = z_y$. Lastly, we use again $y \prec z$ to get $z_y = z$. This proves $z = z_x$. □

As one would expect, if $x \prec y$, and x is a trivialized type, i.e., $\operatorname{triv}(x) = x$, then the product xy is equal to y. Indeed, y already contains the structure of x, and all the elementary types that appear in x are empty strings; therefore, they do not contribute to the elementary types of xy. Interestingly, the converse holds as well. If $xy = y$, x must be a trivialized type and $x \prec y$.

Proposition 6.1.24: Trivial types and preorder of types

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. The following are equivalent:

1. $xy = y$.
2. $x \prec y$ and $\operatorname{triv}(x) = x$.

Proof. We start with *2⇒1*:

$$xy = \operatorname{triv}(x)y = y\operatorname{triv}(x) = y_x = y. \tag{6.29}$$

The first inequality follows from $x = \operatorname{triv}(x)$, the second from the centrality of $\operatorname{triv}(x)$, the third is the promotion definition, and the last follows from $x \prec y$.

Now, we prove *1⇒2*. First, observe that

$$y = xy = \operatorname{triv}(x)xy = \operatorname{triv}(x)y = y\operatorname{triv}(x) = y_x. \tag{6.30}$$

The first equality comes from *1*, the second from absorption, the third from *1* again, the fourth from centrality, and the last is the definition of the promotion of types. Therefore, $x \prec y$.

Second, we want to prove by contradiction that $\operatorname{triv}(x) = x$. Let us assume that a leaf in the tree associated with x is labelled by the elementary type $E \neq I$. Thus, there is a leaf in x_y labelled by E. The type x_y has the same structure as $y_x = y$. Let us denote with G the elementary type associated with the leaf in y corresponding to the leaf labelled by E in x_y. Now, $xy = x_y y_x$ and xy have the same structure as x_y and y_x (by Corollary 6.1.21). Therefore, the leaf of xy corresponding to the leaf labelled with E in x_y and G in $y_x = y$ is labelled by EG (Lemma 6.1.10). However, $EG \neq G$, because E is a non-empty string; thus, the number of symbols in EG is strictly greater than that in G. This contradicts $xy = y$. □

We conclude this Section by noticing that the promotion of I with the type x is equal to $\operatorname{triv}(x)$. Indeed, by definition $I_x = I\operatorname{triv}(x) = \operatorname{triv}(x)$. We will often write I_x instead of $\operatorname{triv}(x)$ in the following.

6.2 Higher-Order Linear Maps

In the previous Section, we introduced types with the promise of using them as labels for maps. In this Section, we formalize what we mean by a map of type $x \in \mathrm{Types}_{\mathcal{A}}$. As we have done multiple times so far and will do again throughout this Chapter, we will follow a bottom-up approach. We begin with symbols in the alphabet, then move on to elementary types, and finally discuss non-elementary types. We associate with every symbol $A \in \mathcal{A}$ a complex, finite-dimensional Hilbert space $\mathcal{H}_A$ with dimension strictly greater than 1. As in the previous chapters, we denote the dimension of $\mathcal{H}_A$ with $|\mathcal{H}_A|$. Each of these Hilbert spaces comes equipped with an inner product, denoted with $(\cdot,\cdot)_{\mathcal{H}_A}$, which is linear in the first component. We associate with every elementary type $E = E_1 \dots E_n$, where $E_i \in \mathcal{A}$, the finite dimensional complex Hilbert space $\mathcal{H}_{E_1} \otimes \cdots \otimes \mathcal{H}_{E_n}$. We denote with $(\cdot,\cdot)_{\mathcal{H}_E}$ the inner product defined on elementary tensors as $(u^1 \otimes \cdots \otimes u^n, v^1 \otimes \cdots \otimes v^n)_{\mathcal{H}_E} = (u^1, v^1)_{\mathcal{H}_{E_1}} \dots (u^n, v^n)_{\mathcal{H}_{E_n}}$, and extended by linearity in the first component and conjugate linearity in the second to all pairs of vectors in $\mathcal{H}_E$. The elementary type $E = I$ is not included in the definition above, and we associate with it the Hilbert space $\mathbb{C}$. Note that, with this definition, we have associated non-trivial quantum systems that cannot be decomposed into smaller subsystems with symbols of $\mathcal{A}$, all quantum systems, composite or not, with elementary types, and we have reserved the empty string I for the trivial quantum system. Following a convention common in quantum information, we will identify $\mathcal{H}_E \otimes \mathbb{C}$ and $\mathbb{C} \otimes \mathcal{H}_E$ with $\mathcal{H}_E$. We are now ready to define a map associated with a type $x \in \mathrm{Types}_{\mathcal{A}}$.

Definition 6.2.1: Vector space of maps of type x

The set of linear maps associated with type x, denoted with $L(x)$ is recursively defined as follows:

- If $x = E \in \text{EleTypes}_{\mathcal{A}}$, then $L(x) = \mathfrak{L}(E)$.
- If $x = (a \to b)$, where $a, b \in \text{Types}_{\mathcal{A}}$, then $L(x) = \mathfrak{L}(L(a), L(b))$.

As a reminder, we denote with $\mathfrak{L}(E)$ the set of linear endomorphisms on $\mathcal{H}_E$ and with $\mathfrak{L}(X, Y)$ the set of linear maps from X to Y. Once again, we follow a standard convention in quantum information, and we identify $L(I)$, which is formally the set of endomorphisms on $\mathbb{C}$, with $\mathbb{C}$, to which it is only isomorphic.

It is straightforward to see that each $L(x)$ is a complex vector space. It is true for elementary types, and the set of all linear maps from one complex vector space to another is a complex vector space.

Lemma 6.2.2: Dimension of $L(x)$

Let $x \in \text{Types}_{\mathcal{A}}$, then $|L(x)| = \prod_i \left|\mathcal{H}_{E_i}\right|^2$, where $\{ E_i \}$ are the elementary types labelling the leaves of the tree associated with x.

Proof. We prove it by induction on the order of x. If $\text{ord}(x) = 0$, then $x = E \in \text{EleTypes}_{\mathcal{A}}$ and $L(x)$ is the space of endomorphisms on $\mathcal{H}_E$, which has dimension $|\mathcal{H}_E|^2$. If $\text{ord}(x) > 0$, then $x = (a \to b)$, where $a, b \in \text{Types}_{\mathcal{A}}$, and $L(x)$ has dimension $|L(a)|\,|L(b)|$. By induction hypothesis, $|L(a)| = \prod_i \left|\mathcal{H}_{E_i}\right|^2$ and $|L(b)| = \prod_j \left|\mathcal{H}_{G_j}\right|^2$, where $\{ E_i \}$ are the labels of the leaves of the tree associated with a, and $\left\{ G_j \right\}$ are the labels of the leaves of the tree associated with b. Therefore,

$$|L(x)| = \prod_i \left|\mathcal{H}_{E_i}\right|^2 \prod_j \left|\mathcal{H}_{G_j}\right|^2 . \tag{6.31}$$

Since the tree associated with x has leaves labelled by $\{ E_i \} \cup \left\{ G_j \right\}$, the statement is true for types of order greater than zero, and this concludes the proof. □

We use the subscript $\cdot_x$ to denote that a linear map is in $L(x)$; that is, we write Γ^x to denote a linear map in $L(x)$. We will consistently use capital Greek letters for maps to avoid confusion with types, denoted mostly with lowercase and uppercase Latin letters, and abstract expressions of types, denoted with lowercase Greek letters (as in Lemma 6.1.12). The most immediate benefit of labelling a map with a type is that it makes it obvious if two maps can be composed and how.

- Sequential composition: If $\Gamma^{a\to b} \in L(a \to b)$ and $\Delta^{b\to c} \in L(b \to c)$, then $\Delta^{b\to c} \circ \Gamma^{a\to b}$ is the map in $L(a \to c)$ obtained as sequential composition of $\Delta^{b\to c}$ and $\Gamma^{a\to b}$.
- Evaluation: If $\Gamma^{a\to b} \in L(a \to b)$ and $u^a \in L(a)$, then $\Gamma^{a\to b}(u^a)$ is the map in $L(b)$ obtained by evaluating $\Gamma^{a\to b}$ at u^a. This notation quickly produces a large number of brackets, making most expressions unreadable. We decided to introduce a different notation to improve readability. We will write $\Gamma^{a\to b} @ u^a$, which one can read as ‘evaluated at’ or simply as ‘at’, as an alternative to $\Gamma^{a\to b}(u^a)$. We will use both notations, depending on the situation. Note that here, we used the lowercase Latin

letter u to denote a map and not an uppercase Greek letter. We will do so every time a map is used solely as an input for another map. We will use letters usually associated with vectors, like u, v, w, to give the idea that the linear operator $\Gamma^{a\to b}$ is evaluated at the vector u^a.

Now that we have well-defined vector spaces and notations for the composition of maps, we recursively define the inner product between two maps. That makes $L(x)$ a Hilbert space for all $x \in \mathrm{Types}_{\mathcal{A}}$. As usual, we give a recursive definition, which is a translation of the Hilbert-Schmidt inner product to our notation.

Definition 6.2.3: Inner product

For every $x \in \mathrm{Types}_{\mathcal{A}}$, the inner product $(\cdot,\cdot)_x$ is recursively defined as follows.

- If $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$: $(\Gamma^E, \Delta^E)_E = \sum_i (\Gamma^E u^i, \Delta^E u^i)_{\mathcal{H}_E}$, where $\{ u^i \}_{i=1}^{|\mathcal{H}_E|}$ is an orthonormal basis of $\mathcal{H}_E$.
- If $x = (a \to b)$, where $a, b \in \mathrm{Types}_{\mathcal{A}}$: $(\Gamma^x, \Delta^x)_x = \sum_i (\Gamma^{a\to b}(u_i^a), \Delta^{a\to b}(u_i^a))_b$, where $\{ u_i^a \}_{i=1}^{|L(a)|}$ is an orthonormal basis of $L(a)$.

Note that the Gram-Schmidt construction guarantees the existence of an orthonormal basis for $L(a)$ in the recursive definition of the inner product. Indeed, $L(a)$ is a vector space, $\mathrm{ord}(a) < \mathrm{ord}(x)$, and the recursive definition guarantees the existence of the inner product $(\cdot,\cdot)_a$. With that, one can construct an orthonormal basis for $L(a)$ starting from any basis of $L(a)$ using the Gram-Schmidt algorithm. Moreover, $(\Gamma^x, \Delta^x)_x$ satisfies all the conditions of an inner product because it is the Hilbert-Schmidt inner product of the operators $\Gamma^x, \Delta^x \in \mathfrak{L}(L(a), L(b))$, and $L(a)$ and $L(b)$ are Hilbert spaces by recursive construction.

When Hilbert spaces and quantum mechanics are involved, the reader expects to see bras and kets. We define them as follows.

Definition 6.2.4: Bra and ket

Let $x \in \mathrm{Types}_{\mathcal{A}}$, and $\Gamma^x \in L(x)$

- $|\Gamma^x\rangle^{I\to x} \in L(I \to x)$ is the linear map defined as $|\Gamma^x\rangle^{I\to x} @ 1 = \Gamma^x$.
- $\langle\Gamma^x|^{x\to I} \in L(x \to I)$ is the dual of Γ^x, that is the unique map that satisfies $\langle\Gamma^x|^{x\to I} @ \Delta^x = (\Delta^x, \Gamma^x)_x$ for all $\Delta^x \in L(x)$.

The Riesz representation theorem guarantees the existence and uniqueness of the dual. Following the physicist convention, we drop the symbol '$\circ$' when dealing with kets and bras:

- $\langle\Gamma^x \mid \Delta^x\rangle^{I\to I} := \langle\Gamma^x|^{x\to I} \circ |\Delta^x\rangle^{I\to x} \in L(I \to I)$,
- $|\Delta^y\rangle\langle\Gamma^x|^{x\to y} := |\Delta^y\rangle^{I\to y} \circ \langle\Gamma^x|^{x\to I} \in L(x \to y)$.

This implies that the relation between the bra-ket of two functions and the inner product is not the one physicists are used to, but $(\Delta^x, \Gamma^x)_x = \langle\Gamma^x \mid \Delta^x\rangle^{I\to I} @ 1$.

After the inner product, the bra, and the ket, we introduce the notation for the adjoint.

Definition 6.2.5: Adjoint

Let $x \in \text{Types}_{\mathcal{A}}$ and let $\Gamma^x \in L(x)$. The adjoint of Γ^x is defined as follows.

- If $x = E \in \text{EleTypes}_{\mathcal{A}}$, $(\Gamma^\dagger)^E \in L(E)$ is the unique linear operator such that $(u, (\Gamma^\dagger)^E(v))_{\mathcal{H}_E} = (\Gamma^E(u), v)_{\mathcal{H}_E}$ for all $u, v \in \mathcal{H}_E$.
- If $x = (a \to b)$, $(\Gamma^\dagger)^{b\to a} \in L(b \to a)$ is the unique linear operator such that $(u^a, (\Gamma^\dagger)^{b\to a}(v^b))_a = (\Gamma^{a\to b}(u^a), v^b)_b$ for all $u^a \in L(a)$, $v^b \in L(b)$.

We conclude this first part about linear maps associated with a type x and the corresponding notations with a proposition that shows how functions can move in and out of kets and bras.

Lemma 6.2.6: Properties of ket and bra

Let $x, y \in \text{Types}_{\mathcal{A}}$, $\Gamma^x \in L(x)$, and $\Delta^{x\to y} \in L(x \to y)$. Then,

1. $|\Delta^{x\to y} @ \Gamma^x\rangle^{I\to y} = \Delta^{x\to y} \circ |\Gamma^x\rangle^{I\to x}$,
2. $\langle\Delta^{x\to y} @ \Gamma^x|^{y\to I} = \langle\Gamma^x|^{x\to I} \circ (\Delta^\dagger)^{y\to x}$.

Proof. We start with the first statement. If we evaluate the left-hand side at 1, we obtain

$$|\Delta^{x\to y} @ \Gamma^x\rangle^{I\to y} @ 1 = \Delta^{x\to y} @ \Gamma^x. \tag{6.32}$$

For the right-hand side, we get

$$\Delta^{x\to y} \circ |\Gamma^x\rangle^{I\to x} @ 1 = \Delta^{x\to y} @ [|\Gamma^x\rangle^{I\to x} @ 1] = \Delta^{x\to y} @ \Gamma^x. \tag{6.33}$$

Since the two functions coincide when evaluated at 1, they coincide on every element of $L(1)$ by linearity, and therefore, they are equal.

We now prove the second statement. Let $\Theta^y \in L(y)$.

$$\langle\Delta^{x\to y} @ \Gamma^x|^{y\to I} @ \Theta^y = (\Theta^y, \Delta^{x\to y} @ \Gamma^x)_y, \tag{6.34}$$

and,

$$\begin{aligned} \langle\Gamma^x|^{x\to I} \circ (\Delta^\dagger)^{y\to x} @ \Theta^y &= \langle\Gamma^x|^{x\to I} @ [(\Delta^\dagger)^{y\to x} @ \Theta^y] \\ ((\Delta^\dagger)^{y\to x} @ \Theta^y, \Gamma^x)_x \\ &= (\Theta^y, \Delta^{x\to y} @ \Gamma^x)_y. \end{aligned} \tag{6.35}$$

□

6.2.1 Product of Maps associated with different types

So far, we have seen two ways of composing maps associated with different types: sequential composition and evaluation. However, there is a third kind of composition that we have not yet investigated: parallel composition. As we discussed when we introduced the product of two types, parallel composition does not always match the tensor product. If $f \in \mathfrak{L}(X, Y)$ and $g \in \mathfrak{L}(W, Z)$, the tensor product $f \otimes g$ is the linear map in $\mathfrak{L}(X \otimes W, Y \otimes Z)$ defined on elementary tensors as $f \otimes g(u \otimes v) = f(u) \otimes g(v)$ for all $u \in X$, $v \in W$. However, we have seen that $\mathcal{M}^{A \to B} \otimes \rho^{C}$ is commonly used to denote the tensor product of $\mathcal{M}^{A \to B}$ with the preparation channel $\mathcal{R}^{\mathbb{C} \to C}(1) = \rho^{C}$, and not the map in $\mathfrak{L}(\mathfrak{L}(A) \otimes \mathcal{H}_C, \mathfrak{L}(B) \otimes \mathcal{H}_C)$.

We introduced the product of types to correctly compute the type associated with the parallel composition of two maps. That is, if $\Gamma^{x} \in L(x)$ and $\Delta^{y} \in L(y)$, we expect the parallel composition of these maps to be in $L(xy)$. Following what was done for the product of types, we begin to define the parallel composition of two maps from the case in which the types associated with the maps have the same structure. Remember that, if $\mathrm{struct}(x) = \mathrm{struct}(y)$ and $\mathrm{ord}(x) = 0$, then both types are elementary, $x = E$ and $y = F$. In this case, the parallel composition is $\Gamma^{E} \otimes \Delta^{F}$, which satisfies the condition of being in $L(EF)$. If $\mathrm{ord}(x) > 0$, then $x = (a \to b)$, $y = (c \to d)$ and $xy = (ac \to bd)$. The tensor product $\Gamma^{a \to b} \otimes \Delta^{c \to d}$ is by definition a map in $\mathfrak{L}(L(a) \otimes L(c), L(b) \otimes L(b))$. In the following lemma, we show that $L(a) \otimes L(c) = L(ac)$ and $L(b) \otimes L(d) = L(bd)$. With that result, $\Gamma^{x} \otimes \Delta^{y} \in \mathfrak{L}(L(ac), L(bd)) = L(xy)$, as wanted.

Lemma 6.2.7: Product of maps associated with types with the same structure

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{struct}(x) = \mathrm{struct}(y)$. Then, $L(xy) = L(x) \otimes L(y)$, where

$$L(x) \otimes L(y) = \mathrm{Span}\,\{\, \Gamma^{x} \otimes \Delta^{y} \mid \Gamma^{x} \in L(x), \Delta^{y} \in L(y) \,\}\,. \tag{6.36}$$

Proof. We prove it by induction of $\mathrm{ord}(x) = \mathrm{ord}(y)$. *Base step*: $x = E$, and $y = G$, and $E, G \in \mathrm{EleTypes}_{\mathcal{A}}$. Then,

$$L(EG) = \mathfrak{L}(\mathcal{H}_{EG}) = \mathfrak{L}(\mathcal{H}_E \otimes \mathcal{H}_G) = \mathfrak{L}(\mathcal{H}_E) \otimes \mathfrak{L}(\mathcal{H}_G). \tag{6.37}$$

The first inequality follows from the definition of $L(EG)$, the second from the definition of $\mathcal{H}_{EG}$, and the last from the canonical identification

$$\mathfrak{L}(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{H}_3 \otimes \mathcal{H}_4) = \mathfrak{L}(\mathcal{H}_1, \mathcal{H}_3) \otimes \mathfrak{L}(\mathcal{H}_2, \mathcal{H}_4), \tag{6.38}$$

where $\mathcal{H}_i$ is a finite dimensional Hilbert spaces for $1 \leq i \leq 4$. Since $\mathfrak{L}(\mathcal{H}_E) = L(E)$ and $\mathfrak{L}(\mathcal{H}_G) = L(G)$, the statement holds for elementary types.

Inductive step: Let $x = (a \to b)$, $y = (c \to d)$. Since $\mathrm{struct}(x) = \mathrm{struct}(y)$, then $\mathrm{struct}(a) = \mathrm{struct}(c)$ and $\mathrm{struct}(b) = \mathrm{struct}(d)$. Then, using the induction hypothesis and Eq. (6.38), we obtain

$$\begin{aligned} L(xy) &= L(ac \to bd) \\ &= \mathfrak{L}(L(ac), L(bd)) \\ &= \mathfrak{L}(L(a) \otimes L(c), L(b) \otimes L(d)) \\ &= \mathfrak{L}(L(a), L(b)) \otimes \mathfrak{L}(L(c), L(d)) \\ &= L(x) \otimes L(y). \end{aligned} \tag{6.39}$$

This concludes the proof. □

Therefore, if two types have the same structure, the parallel composition of maps associated with them is computed with the tensor product. Now, we turn our attention to types with different structures. We notice that even if x and y are types with different structures, we can construct types x_y and y_x, which have the same structure by construction and satisfy $x_y y_x = xy$. Thus, if we can 'promote' a map Γ^x to a map Γ^{x_y} and a map Δ^y to a map Δ^{y_x}, then we can define the parallel composition of Γ^x and Δ^{y_x} as $\Gamma^{x_y} \otimes \Delta^{y_x}$ and we have a way of computing it. This is usually done when the parallel composition of $\mathcal{M}^{A\to B}$ and ρ^C is computed. Indeed, first one promotes ρ^C to the channel that prepares ρ^C, which with our notation is $\left|\rho^C\right\rangle^{I\to C}$. This map has type $I \to C$, which has the same structure as $A \to B$, the type associated with $\mathcal{M}^{A\to B}$. Second, one defines the parallel composition as $\mathcal{M}^{A\to B} \otimes \left|\rho^C\right\rangle^{I\to C}$, which is a map of type $A \to BC$, as expected.

Before giving the formal definition of 'promotion', we provide an intuitive explanation for it. The goal is to promote a map $\Gamma^x \in L(x)$ to its natural counterpart $\Gamma^{x_y} \in L(x_y)$, in the same way that a state ρ^C is promoted to the channel that prepares ρ^C. We start with the simplest case, as usual. Suppose that $x, y \in \text{EleTypes}_{\mathcal{A}}$. Then, the structure of x is equal to the structure of y and, therefore, $x = x_y$. This implies that the promoted map Γ^{x_y} is equal to the initial map Γ^x. If we call $\eta^{x\to x_y}$ the promotion map, then, in the case $x, y \in \text{EleTypes}_{\mathcal{A}}$, we have $\eta^{x\to x_y} = \mathrm{id}^{x\to x}$, the identity map in $L(x \to x)$, i.e., $\mathrm{id}^{x\to x}(\Gamma^x) = \Gamma^x$ for all $\Gamma^x \in L(x)$. In addition, we can introduce a 'demotion' map $\eta^{x_y\to x}$ that reverses the action of the promotion map. In this case, $\eta^{x_y\to x} = \mathrm{id}^{x\to x}$.

The next step in the recursive definition is to assume that $\mathrm{ord}(xy) > 0$, and we split the definition into three cases: $\mathrm{ord}(x) = \mathrm{ord}(y)$, $\mathrm{ord}(x) > \mathrm{ord}(y)$, and $\mathrm{ord}(x) < \mathrm{ord}(y)$, matching the three cases in the definition of the product of types. Whenever $\mathrm{ord}(x) > 0$, we have that $x = (a \to b)$, for some $a, b \in \text{Types}_{\mathcal{A}}$, and similarly, whenever $\mathrm{ord}(y) > 0$, we have that $y = (c \to d)$ for some $c, d \in \text{Types}_{\mathcal{A}}$. We start with the case $\mathrm{ord}(x) = \mathrm{ord}(y) > 0$. We want to promote a map Γ^x of type $x = (a \to b)$, to a map of type $x_y = (a_c \to b_d)$. Thanks to the recursive definition, we can assume that all the promotion and demotion maps of lower orders are well-defined. We observe that if we pre-compose $\Gamma^{a\to b}$ with the demotion map $\eta^{a_c\to a}$ we obtain $\Gamma^{a\to b} \circ \eta^{a_c\to a} \in L(a_c \to b)$. Similarly, if we post-compose the resulting map with the promotion map $\eta^{b\to b_d}$, we obtain $\eta^{b\to b_d} \circ \Gamma^{a\to b} \circ \eta^{a_c\to a} \in L(a_c \to b_d)$, as desired. With the same logic, one obtains the demotion of Δ^{x_y} is $\eta^{b_d\to b} \circ \Delta^{a_c\to b_d} \circ \eta^{a\to a_c}$. That is, if $\mathrm{ord}(x) = \mathrm{ord}(y)$, the promotion $\eta^{x\to x_y}$ and demotion $\eta^{x_y\to x}$ are completely defined in terms of the promotions and demotions of the left and right subtrees of x and x_y.

This is very similar to what happens in the case $\mathrm{ord}(x) > \mathrm{ord}(y)$. Indeed $x_y = (a \to b_y)$, and therefore Γ^x is promoted to $\eta^{b\to b_y} \circ \Gamma^{a\to b}$. With the same logic, $\Delta^{x_y} \in L(a \to b_y)$ is reduced to $\eta^{b_y\to b} \circ \Delta^{a\to b_y} \in L(a \to b)$.

So far, the promotion and demotion operators are either the identity or are recursively defined in terms of lower-order promotion and demotion operators. The case $\mathrm{ord}(x) < \mathrm{ord}(y)$ is the crucial case. We have already seen an example of this case when we promoted the state ρ^C to the preparation channel $1 \to \rho^C$. As already discussed, the preparation channel is $\left|\rho^C\right\rangle^{I\to C} \in L(I \to C)$. This is the key idea behind the definition of the promotion operator in the case $\mathrm{ord}(x) < \mathrm{ord}(y)$. Let $y = (c \to d)$, then $x_y = (I_c \to x_d)$. We start with a function $\Gamma^x \in L(x)$. If we evaluate $\eta^{x\to x_d}$ at Γ^x we obtain $\eta^{x\to x_d}(\Gamma^x) \in L(x_d)$. Now, we want to promote it to a map of higher order using the ket operator: $|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \in L(I \to x_d)$. We have almost obtained the desired map in $L(I_c \to x_d)$. What is left to do is pre-compose with $\eta^{I_c\to I}$ and we obtain $|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \in L(I_c \to x_d)$. Lastly, if one starts with Δ^{x_y}, the demotion to a map in $L(x)$ is obtained as follows. First, pre-compose with $\eta^{I\to I_c}$ to obtain $\Delta^{I_c\to x_d} \circ \eta^{I\to I_c} \in L(I \to x_d)$. If we evaluate this map at 1, the result is a map of type x_d: $\Delta^{I_c\to x_d}\circ\eta^{I\to I_c}@1\in L(x_d)$. Therefore, the only thing left to do is to reduce the resulting map to x: $\eta^{x_d\to x}[\Delta^{I_c\to x_d}\circ\eta^{I\to I_c}@1]\in L(x)$.

In conclusion, the promotion and demotion maps are recursive pre- and post-compositions of identities, ket promotions, or evaluations at 1. Below, we formalize the definition of promotion and demotion maps.

Definition 6.2.8: Promotion and demotion maps

Let $x, y\in \mathrm{Types}_{\mathcal{A}}$. Let $x=(a\to b)$ if $\mathrm{ord}(x)>0$ and let $y=(c\to d)$ if $\mathrm{ord}(y>0)$. Let $\Gamma^x\in L(x)$, $\Delta^{x_y}\in L(x_y)$. The promotion map $\eta^{x\to x_y}$ and demotion map $\eta^{x_y\to x}$ are recursively defined as follows.

- If $\mathrm{ord}(xy)=0$, then $x_y=x$ and $\eta^{x\to x_y}=\eta^{x_y\to x}=\mathrm{id}^{x\to x}$.
- If $\mathrm{ord}(xy)>0$ and
  - $\mathrm{ord}(x)=\mathrm{ord}(y)$: $x_y=(a_c\to b_d)$ and
$$\begin{aligned}\eta^{x\to x_y}(\Gamma^x)&=\eta^{b\to b_d}\circ\Gamma^{a\to b}\circ\eta^{a_c\to a}\in L(x_y),\\ \eta^{x_y\to x}(\Delta^{x_y})&=\eta^{b_d\to b}\circ\Delta^{a_c\to b_d}\circ\eta^{a\to a_c}\in L(x).\end{aligned}\tag{6.40}$$
  - $\mathrm{ord}(x)<\mathrm{ord}(y)$: $x_y=(I_c\to x_d)$, and
$$\begin{aligned}\eta^{x\to x_y}(\Gamma^x)&=|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d}\circ\eta^{I_c\to I}\in L(x_y),\\ \eta^{x_y\to x}(\Delta^{x_y})&=\eta^{x_d\to x}[\Delta^{I_c\to x_d}\circ\eta^{I\to I_c}(1)]\in L(x).\end{aligned}\tag{6.41}$$
  - $\mathrm{ord}(x)>\mathrm{ord}(y)$: $x_y=(a\to b_y)$, and
$$\begin{aligned}\eta^{x\to x_y}(\Gamma^x)&=\eta^{b\to b_y}\circ\Gamma^{a\to b}\in L(x_y),\\ \eta^{x_y\to x}(\Delta^{x_y})&=\eta^{b_y\to b}\circ\Delta^{a\to b_y}\in L(x).\end{aligned}\tag{6.42}$$

We have introduced the demotion map as a map that reverses the action of a promotion. Unsurprisingly, the demotion map is the inverse of the corresponding promotion map and vice versa, as we prove in the following Proposition.

Proposition 6.2.9: $\eta^{x\to x_y}$ is invertible

For all $x, y\in\mathrm{Types}_{\mathcal{A}}$, $(\eta^{x\to x_y})^{-1}=\eta^{x_y\to x}$.

Proof. We prove it by induction on $\mathrm{ord}(xy)$. The base case is trivially true, and so are the cases $\mathrm{ord}(x)=\mathrm{ord}(y)$

and $\mathrm{ord}(x) > \mathrm{ord}(y)$ in the induction step. What is left to prove is the case $\mathrm{ord}(x) < \mathrm{ord}(y)$.

$$\begin{aligned}
\eta^{x_y\to x} \circ \eta^{x\to x_y}(\Gamma^x) &= \eta^{x_y\to x}[|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I}] \\
&= \eta^{x_d\to d}[|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \circ \eta^{I\to I_c} @ 1] \\
&= \eta^{x_d\to d}[|\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} @ 1] \\
&= \eta^{x_d\to d}[\eta^{x\to x_d}(\Gamma^x)] \\
&= \Gamma^x,
\end{aligned} \tag{6.43}$$

were we have used the induction hypotheses $(\eta^{x\to x_d})^{-1} = \eta^{x_d\to x}$ and $(\eta^{I\to I_c})^{-1} = \eta^{I_c\to I}$. Similarly,

$$\begin{aligned}
\eta^{x\to x_y} \circ \eta^{x_y\to x}(\Delta^{x_y}) &= \eta^{x\to x_y}\{\eta^{x_d\to x}[\Delta^{I_c\to x_d} \circ \eta^{I\to I_c}(1)]\} \\
&= \left|\eta^{x\to x_d}\{\eta^{x_d\to x}[\Delta^{I_c\to x_d} \circ \eta^{I\to I_c} @ 1]\}\right\rangle^{I\to x_d} \circ \eta^{I_c\to I} \\
&= \left|\Delta^{I_c\to x_d} \circ \eta^{I\to I_c} @ 1\right\rangle^{I\to x_d} \circ \eta^{I_c\to I} \\
&= \Delta^{I_c\to x_d} \circ \eta^{I\to I_c} \circ \eta^{I_c\to I} \\
&= \Delta^{I_c\to x_d}.
\end{aligned} \tag{6.44}$$

In this proof, we have used the fact that for all $\Theta^{I\to z} \in L(I \to z)$, $\left|\Theta^{I\to z} @ 1\right\rangle^{I\to z} = \Theta^{I\to z}$. Indeed, $\left|\Theta^{I\to z} @ 1\right\rangle^{I\to z}$ is the linear map that maps $1 \in L(I)$ to $\Theta^{I\to z}(1) \in L(z)$, and this coincide with the action of $\Theta^{I\to z}$. □

The last remark of this proof will be helpful in other points. Therefore, we highlight it in the following Corollary.

Corollary 6.2.10: Ket and evaluation at 1

For all $x \in \mathrm{Types}_{\mathcal{A}}$ and $\Gamma^{I\to x} \in L(I \to x)$,

$$\left|\Gamma^{I\to x} @ 1\right\rangle^{I\to x} = \Gamma^{I\to x}. \tag{6.45}$$

Promotion maps are the 'natural' way of promoting a map $\Gamma^x \in L(x)$ to a map $\Gamma^{x_y} \in L(x_y)$. If $x_y = x$, we expect that the 'natural' promotion is the identity. That is, Γ^x is mapped to itself. Indeed, this is not only what happens but also the only case in which the promotion map is the identity.

Lemma 6.2.11: Promotion and demotion maps and preorder of types

The following are equivalent:

1. $y \prec x$.
2. $\eta^{x\to x_y} = \mathrm{id}^{x\to x}$.

Proof. 1⇒2: If $y \prec x$, the $x_y = x$. By induction on $\mathrm{ord}(x)$, we get that if $\mathrm{ord}(x) = 0$, then $\eta^{x\to x_y} = \mathrm{id}^{x\to x}$ by definition. For the case $\mathrm{ord}(x) > 0$, we notice that $\mathrm{ord}(y) \leq \mathrm{ord}(x)$. Indeed, $\mathrm{ord}(x) = \mathrm{ord}(x_y) =$

$\max\{\operatorname{ord}(x), \operatorname{ord}(y)\} \geq \operatorname{ord}(y)$. Therefore, $\eta^{x \to x_y}$ consists of pre- and post-composition with promotion and demotion maps of lower orders, which by induction hypothesis are all identity maps. This proves *1*⇒*2*.

2⇒*1*: If $\eta^{x \to x_y} = \mathrm{id}^{x \to x} \in L(x \to x)$, then $x_y = x$, and therefore $y \prec x$. ☐

The definition of promotion and reduction functions gives the idea that $\eta^{x \to x_y}$ depends on the type y with which x is promoted. However, this is not exactly true. More than depending on the type y itself, the promotion function depends on the target type x_y, that is, if $x_y = x_z$, then $\eta^{x \to x_y} = \eta^{x \to x_z}$. To prove that, we first show that promoting a function Γ^x with a type xy is the same as promoting Γ^x with just y.

Lemma 6.2.12: Promotion of a type with itself

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, then $\eta^{x \to x_{xy}} = \eta^{x \to x_y}$.

Proof. We prove it by induction on $\operatorname{ord}(xy)$. The base step is trivial. For the induction step, assume that $\operatorname{ord}(xy) > 0$, and $\operatorname{ord}(x) = \operatorname{ord}(xy)$. Let $x = (a \to b)$ and $xy = \alpha\gamma \to \beta\delta$ (Lemma 6.1.12). By definition,

$$\eta^{x \to x_{xy}}(fx) = \eta^{b_{\beta\delta} \to b} \circ \Gamma^{a \to b} \circ \eta^{a \to a_{\alpha\gamma}}. \tag{6.46}$$

Now, we separate the case $\operatorname{ord}(x) = \operatorname{ord}(xy)$ into $\operatorname{ord}(x) = \operatorname{ord}(y)$ and $\operatorname{ord}(x) > \operatorname{ord}(y)$. In the first, $\alpha = a$, $\beta = b$, $\gamma = c$ and $\delta = d$, where $c, d \in \mathrm{Types}_{\mathcal{A}}$ are such that $y = (c \to d)$. Therefore,

$$\begin{aligned} \eta^{x \to x_{xy}}(fx) &= \eta^{b \to b_{\beta\delta}} \circ \Gamma^{a \to b} \circ \eta^{a_{\alpha\gamma} \to a} \\ &= \eta^{b \to b_{bd}} \circ \Gamma^{a \to b} \circ \eta^{a_{ac} \to a} \\ &= \eta^{b \to b_d} \circ \Gamma^{a \to b} \circ \eta^{a_c \to a} \\ &= \eta^{x \to x_y}(\Gamma^x), \end{aligned} \tag{6.47}$$

where we have used the induction hypothesis and the definition of $\eta^{x \to x_y}$. In the case $\operatorname{ord}(x) > \operatorname{ord}(y)$, then $\alpha = a$, $\beta = b$, $\gamma = I$, $\delta = y$, and therefore

$$\begin{aligned} \eta^{x \to x_{xy}}(\Gamma^x) &= \eta^{b \to b_{\beta\delta}} \circ \Gamma^{a \to b} \circ \eta^{a_{\alpha\gamma} \to a} \\ &= \eta^{b \to b} \circ \Gamma^{a \to b_{by}} \circ \eta^{a_{aI} \to a} \\ &= \eta^{b \to b_y} \circ \Gamma^{a \to b} \circ \eta^{a_I \to a} \\ &= \eta^{b \to b_y} \circ \Gamma^{a \to b} \\ &= \eta^{x \to x_y}(\Gamma^x). \end{aligned} \tag{6.48}$$

Here, we have used Lemma 6.2.11 in addition to the induction hypothesis and the definition of $\eta^{x \to x_y}$.

The case $\operatorname{ord}(xy) > \operatorname{ord}(x)$ is logically equivalent to the case $\operatorname{ord}(y) > \operatorname{ord}(x)$ and we get

$$\begin{aligned} \eta^{x \to x_{xy}}(\Gamma^x) &= |\eta^{x \to x_{\beta\delta}}(\Gamma^x)\rangle^{I \to x_\delta} \circ \eta^{I_{\alpha\gamma} \to I} \\ &= |\eta^{x \to x_{xd}}(\Gamma^x)\rangle^{I \to x_d} \circ \eta^{I_{Ic} \to I} \\ &= |\eta^{x \to x_d}(\Gamma^x)\rangle^{I \to x_d} \circ \eta^{I_c \to I} \\ &= \eta^{x \to x_y}(\Gamma^x). \end{aligned} \tag{6.49}$$

☐

With this Lemma, the desired result $\eta^{x\to x_z} = \eta^{x\to x_y}$ when $x_z = x_y$ follows immediately.

Proposition 6.2.13: Dependencies of promotion maps

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$ be such that $x_y = x_z$. Then, $\eta^{x\to x_y} = \eta^{x\to x_z}$.

Proof. By the previous Lemma, $\eta^{x\to x_y} = \eta^{x\to x_{xy}}$, and $\eta^{x\to x_z} = \eta^{x\to x_{xz}}$. However, $x_z = x_y$ implies $\mathrm{triv}(xz) = \mathrm{triv}(xy)$, and therefore, $\eta^{x\to x_{xz}} = \eta^{x\to x_{xz}}$, from which $\eta^{x\to x_y} = \eta^{x\to x_z}$ follows. □

We have spent some pages discussing the fundamental properties of promotion functions. However, let us not forget our goal. We aim to develop a method for computing the parallel composition of operators associated with different types. Since we know that if two types have the same structure, then parallel composition is computed with the tensor product, we introduced promotion functions as a way to give maps associated with different types the same structure. Let $\Gamma^x \in L(x)$ and $\Delta^y \in L(y)$. Then, the types associated with $\eta^{x\to x_y}(\Gamma^x)$ and $\eta^{y\to y_x}(\Delta^y)$ have the same structure and the parallel composition of Γ^x and Δ^y can be computed as $\eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y)$. We refer to this operation as the square tensor product.

Definition 6.2.14: Square product

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. The map associated with the parallel composition of $\Gamma^x \in L(x)$ with $\Delta^y \in L(y)$ is $\Gamma^x \boxtimes \Delta^y := \eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y) \in L(xy)$.

The fact that $\Gamma^x \boxtimes \Delta^y \in L(xy)$ follows from Lemma 6.2.7 and the fact that x_y and y_x have the same structure together with $x_y y_x = xy$. Moreover, this definition includes the case we have already presented regarding types with the same structure. Indeed, if $\mathrm{struct}(x) = \mathrm{struct}(y)$, then $x_y = x$ and $y_x = y$ and promotion functions are identities (Lemma 6.2.11).

We have already discussed that elementary tensors $\Gamma^x \boxtimes \Delta^y$ are in $L(xy)$. We show now that every element of $L(xy)$ can be written as a linear combination of elements of elementary tensors.

Proposition 6.2.15: Product space

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. Then, $L(xy) = L(x) \boxtimes L(y)$, where

$$L(x) \boxtimes L(y) := \mathrm{Span}\,\{\, \Gamma^x \boxtimes \Delta^y \mid \Gamma^x \in L(x), \Delta^y \in L(y) \,\}\,. \tag{6.50}$$

Proof. The proof is quite straightforward and based on Lemma 6.2.7 and the fact that $\eta^{x\to x_y}$ are bijections.

$$\begin{aligned}
L(xy) &= L(x_y y_x) \\
&= L(x_y) \otimes L(y_x) \\
&= \mathrm{Span}\left\{\, \Gamma^{x_y} \otimes \Delta^{y_x} \mid \Gamma^{x_y} \in L(x_y), \Delta^{y_x} \in L(y_x) \,\right\} \\
&= \mathrm{Span}\,\{\, \eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y) \mid \Gamma^x \in L(x), \Delta^y \in L(y) \,\} \\
&= \mathrm{Span}\,\{\, \Gamma^x \boxtimes \Delta^y \mid \Gamma^x \in L(x), \Delta^y \in L(y) \,\} \\
&= L(x) \boxtimes L(y).
\end{aligned} \tag{6.51}$$

□

6.2.2 Properties of Promotion and Demotion Maps

In this subsection, we will present many properties of the promotion operators. Most of these properties, rather than being relevant on their own, will be helpful at a later stage to prove results about the parallel composition of functions and the Choi isomorphism. The reader who has little time may decide to rush through the results without reading the proofs or even skip this subsection. To such a reader, we say that promotion functions behave as one would expect them to. If the reader thinks that it is reasonable that a promotion function has a property, then it is very likely that it does. Indeed, many of the properties presented in this Section were found in this way. We asked ourselves: What are the properties that the most 'natural' promotion operator should have? Does our promotion operator have such a property? With such an introduction, we do not want to suggest that there is no value in reading this subsection. First, even if most properties are expected, it is far from trivial to show that our promotion functions satisfy them. Second, reading the proofs will make the reader more familiar with the formalism for types, which will simplify the understanding of many proofs later on.

We start by asking ourselves what happens if one promotes a product type, e.g., $\eta^{xy \to xy_z}$. It is reasonable to expect that one could split it into a promotion on x and a promotion on y. A first guess would be $\eta^{xy \to xy_z} = \eta^{x \to x_z} \boxtimes \eta^{y \to y_z}$. However, this map may not be in $L(xy \to xy_z)$ as we need. Indeed, if $\mathrm{ord}(x) > \mathrm{ord}(yz)$, then $(x \to x_z)|(y \to y_z) = (x \to x_z|(y \to y_z))$. To address this, we will employ a strategy that will quickly become the standard for achieving the results in this Chapter: we utilize the promoted types x_y and y_x. In this case, $\eta^{x_y \to x_{yz}} \boxtimes \eta^{y_x \to y_{xz}} \in L(xy \to xy_z)$. In the following lemma, we prove that $\eta^{xy \to xy_z}$ splits into a promotion on x_y and a promotion on y_x.

Proposition 6.2.16: Promotion of product types

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$. Then, $\eta^{xy \to xy_z} = \eta^{x_y \to x_{yz}} \boxtimes \eta^{y_x \to y_{xz}}$, and $\eta^{xy_z \to xy} = \eta^{x_{yz} \to x_y} \boxtimes \eta^{y_{xz} \to y_x}$.

Proof. Throughout this proof, whenever $\mathrm{ord}(x) > 0$, we denote with a and b the types such that $x = (a \to b)$. Similarly, we have $y = (c \to d)$, when $\mathrm{ord}(y) > 0$ and $z = (e \to f)$ when $\mathrm{ord}(z) > 0$. Moreover, we prove each statement by induction on $\mathrm{ord}(xyz)$.

Observe that $\mathrm{struct}(x_y \to x_{yz}) = \mathrm{struct}(y_x \to y_{xz})$, therefore $\eta^{x_y \to x_{yz}} \boxtimes \eta^{y_x \to y_{xz}} = \eta^{x_y \to x_{yz}} \otimes \eta^{y_x \to y_{xz}}$ and we can use all the properties of the tensor product.

Base step: All types are elementary; therefore, all the promotion functions are identities, and $\mathrm{id}^{xy \to xy} = \mathrm{id}^{x \to x} \otimes \mathrm{id}^{y \to y}$. *Induction step:* The promotion $\eta^{xy \to xy_z}$ acts on elements of $L(xy) = L(x_y y_z) = L(x_y) \otimes L(y_x)$. Therefore, we can investigate the action on elementary tensors $\Gamma^{x_y} \otimes \Delta^{y_x}$. We split the proof into the following cases:

- $\mathrm{ord}(x) = \mathrm{ord}(y) = \mathrm{ord}(z)$: The action on elementary tensors is

$$
\begin{aligned}
\eta^{xy\to xy_z}(\Gamma^{x_y} \otimes \Delta^{y_x}) &= \eta^{bd\to bd_f} \circ (\Gamma^{a_c\to b_d} \otimes \Delta^{c_a\to d_b}) \circ \eta^{ac_e\to ac} \\
\text{I.H.} &= (\eta^{b_d\to b_{df}} \otimes \eta^{d_b\to d_{bf}}) \circ (\Gamma^{a_c\to b_d} \otimes \Delta^{c_a\to d_b}) \\
&\quad \circ (\eta^{a_{ce}\to a_c} \otimes \eta^{c_{ae}\to c_a}) \\
&= (\eta^{b_d\to b_{df}} \circ \Gamma^{a_c\to b_d} \circ \eta^{a_{ce}\to a_c}) \otimes \\
&\quad (\eta^{d_b\to d_{bf}} \circ \Delta^{c_a\to d_b} \circ \eta^{c_{ae}\to c_a}) \\
&= \eta^{x_y\to x_{yz}}(\Gamma^{x_y}) \otimes \eta^{y_x\to x_{yz}}(\Delta^{y_z}) \\
&= (\eta^{x_y\to x_{yz}} \otimes \eta^{y_z\to y_{xz}})(\Gamma^{x_y} \otimes \Delta^{y_z}).
\end{aligned}
\tag{6.52}
$$

- $\mathrm{ord}(x) < \mathrm{ord}(y) = \mathrm{ord}(z)$:

$$
\begin{aligned}
\eta^{xy\to xy_z}(\Gamma^{x_y} \otimes \Delta^{y_x}) &= \eta^{xd\to xd_f} \circ (\Gamma^{I_c\to x_d} \otimes \Delta^{c_I\to d_x}) \circ \eta^{cI_e\to cI} \\
\text{I.H.} &= (\eta^{x_d\to x_{df}} \otimes \eta^{d_x\to d_{xf}}) \circ (\Gamma^{I_c\to x_d} \otimes \Delta^{c_I\to d_x}) \\
&\quad \circ (\eta^{I_{ce}\to I_c} \otimes \eta^{c_{Ie}\to c_I}) \\
&= (\eta^{x_d\to x_{df}} \circ \Gamma^{I_c\to x_d} \circ \eta^{I_{ce}\to I_c}) \otimes \\
&\quad (\eta^{d_x\to d_{xf}} \circ \Delta^{c_I\to d_x} \circ \eta^{c_{Ie}\to c_I}) \\
&= \eta^{x_y\to x_{yz}}(\Gamma^{x_y}) \otimes \eta^{y_x\to x_{yz}}(\Delta^{y_z}) \\
&= (\eta^{x_y\to x_{yz}} \otimes \eta^{y_z\to y_{xz}})(\Gamma^{x_y} \otimes \Delta^{y_z}).
\end{aligned}
\tag{6.53}
$$

- $\mathrm{ord}(y) < \mathrm{ord}(x) = \mathrm{ord}(z)$: Analogous to the previous case.

- $\mathrm{ord}(xy) < \mathrm{ord}(z)$:

$$
\begin{aligned}
\eta^{xy\to xy_z}(\Gamma^{x_y} \otimes \Delta^{y_x}) &= |\eta^{xy\to xy_f}(\Gamma^{x_y} \otimes \Delta^{y_x})\rangle^{I\to xy_f} \circ \eta^{I_e\to I} \\
\text{I.H.} &= |\eta^{x_y\to x_{yf}}(\Gamma^{x_y}) \otimes \eta^{y_x\to y_{xf}}(\Delta^{x_y})\rangle^{I\to xy_f} \circ \eta^{II_e\to II} \\
\text{I.H.} &= (|\eta^{x_y\to x_{yf}}(\Gamma^{x_y})\rangle^{I\to x_{yf}} \otimes |\eta^{y_x\to y_{xf}}(\Delta^{x_y})\rangle^{I\to y_{xf}}) \\
&\quad \circ (\eta^{I_e\to I} \otimes \eta^{I_e\to I}) \\
&= (|\eta^{x_y\to x_{yf}}(\Gamma^{x_y})\rangle^{I\to x_{yf}} \circ \eta^{I_e\to I}) \otimes \\
&\quad (|\eta^{y_x\to y_{xf}}(\Delta^{x_y})\rangle^{I\to y_{xf}} \circ \eta^{I_e\to I}) \\
&= \eta^{x_y\to x_{yz}}(\Gamma^{x_y}) \otimes \eta^{y_x\to x_{yz}}(\Delta^{y_z}) \\
&= (\eta^{x_y\to x_{yz}} \otimes \eta^{y_z\to y_{xz}})(\Gamma^{x_y} \otimes \Delta^{y_z}).
\end{aligned}
\tag{6.54}
$$

Here we have used the fact that for any $\alpha, \beta \in \mathrm{Types}_{\mathcal{A}}$ and functions $\Theta^{\alpha\beta}, \Xi^{\beta\alpha}$ we have $\left|\Theta^{\alpha\beta} \otimes \Xi^{\beta_\alpha}\right\rangle^{I\to\alpha\beta} = |\Theta^{\alpha_\beta}\rangle^{I\to\alpha_\beta} \otimes \left|\Xi^{\beta_\alpha}\right\rangle^{I\to\alpha_\beta}$. Indeed,

$$
\left|\Theta^{\alpha_\beta} \otimes \Xi^{\beta_\alpha}\right\rangle^{I\to\alpha_\beta}(1) = \Theta^{\alpha_\beta} \otimes \Xi^{\beta_\alpha},
\tag{6.55}
$$

and

$$
\begin{aligned}
|\Theta^{\alpha_\beta}\rangle^{I\to\alpha_\beta} \otimes \left|\Xi^{\beta_\alpha}\right\rangle^{I\to\alpha_\beta}(1) &= |\Theta^{\alpha_\beta}\rangle^{I\to\alpha_\beta} \otimes \left|\Xi^{I\to\beta_\alpha}\right\rangle^{\alpha_\beta}(1 \otimes 1) \\
&= |\Theta^{\alpha_\beta}\rangle^{I\to\alpha_\beta}(1) \otimes \left|\Xi^{\beta_\alpha}\right\rangle^{I\to\alpha_\beta}(1) \\
&= \Theta^{\alpha_\beta} \otimes \Xi^{\beta_\alpha}.
\end{aligned}
\tag{6.56}
$$

Therefore, the two functions coincide.

- $\mathrm{ord}(xy) > \mathrm{ord}(z)$: From Lemma 6.1.12 we have that there exists $\alpha, \beta, \gamma, \delta \in \mathrm{Types}_{\mathcal{A}}$ such that $x_y y_x = (\alpha\gamma \to \beta\delta)$. From the definition of α, β, γ, and δ immediately follows that $x_y = (\alpha_\gamma \to \beta_\delta)$ and $y_x = (\gamma_\alpha \to \delta_\beta)$. Therefore,

$$\begin{aligned}\eta^{xy \to xy_z}(\Gamma^{x_y} \otimes \Delta^{y_x}) &= \eta^{\beta\delta \to \beta\delta_z} \circ (\Gamma^{\alpha_\gamma \to \beta_\delta} \otimes \Delta^{\gamma_\alpha \to \delta_\beta}) \\ \text{I.H.} &= (\eta^{\beta_\delta \to \beta_{\delta z}} \otimes \eta^{\delta_\beta \to \delta_{\beta z}}) \circ (\Gamma^{\alpha_\gamma \to \beta_\delta} \otimes \Delta^{\gamma_\alpha \to \delta_\beta}) \\ &= (\eta^{\beta_\delta \to \beta_{\delta z}} \circ \Gamma^{\alpha_\gamma \to \beta_\delta}) \otimes (\eta^{\delta_\beta \to \delta_{\beta z}} \circ \Delta^{\gamma_\alpha \to \delta_\beta}) \\ &= \eta^{x_y \to x_{yz}}(\Gamma^{x_y}) \otimes \eta^{y_x \to x_{yz}}(\Delta^{y_z}) \\ &= (\eta^{x_y \to x_{yz}} \otimes \eta^{y_z \to y_{xz}})(\Gamma^{x_y} \otimes \Delta^{y_z}).\end{aligned} \tag{6.57}$$

Since $\eta^{xy \to xy_z} = \eta^{x_y \to x_{yz}} \otimes \eta^{y_x \to y_{xz}}$ for elementary tensors, the same is true for any element of $L(xy)$ by linearity.

Lastly, notice that $\eta^{x_{yz} \to x_y}$ is the inverse of $\eta^{x_y \to x_{yz}}$ and $\eta^{y_{xz} \to y_x}$ is the inverse of $\eta^{y_x \to y_{xz}}$. Since $(\eta^{x_{yz} \to x_y} \otimes \eta^{y_{xz} \to y_x}) \circ (\eta^{x_y \to x_{yz}} \otimes \eta^{y_x \to y_{xz}}) = \mathrm{id}^{xy \to xy}$, and, similarly, $(\eta^{x_y \to x_{yz}} \otimes \eta^{y_x \to y_{xz}}) \circ (\eta^{x_{yz} \to x_y} \otimes \eta^{y_{xz} \to y_x}) = \mathrm{id}^{xy_z \to xy_z}$, then $(\eta^{x_{yz} \to x_y} \otimes \eta^{y_{xz} \to y_x})$ is the inverse of $\eta^{xy \to xy_z}$, and therefore $\eta^{x_{yz} \to x_y} \boxtimes \eta^{y_{xz} \to y_x} = \eta^{xy_z \to xy}$. □

For the next property, we ask ourselves what happens when we apply two promotions in sequence, that is, first, we promote x to x_y, and then, x_y to x_{yz}. It is reasonable to expect that such a composition is equivalent to promoting x with yz directly. This is indeed the case.

Proposition 6.2.17: Sequence of promotions

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$. Then, $\eta^{x_y \to x_{yz}} \circ \eta^{x \to x_y} = \eta^{x \to x_{yz}}$, and $\eta^{x_y \to x} \circ \eta^{x_{yz} \to x_y} = \eta^{x_{yz} \to x}$.

Proof. As usual, whenever $\mathrm{ord}(x) > 0$, we write $x = (a \to b)$. Similarly, we have $y = (c \to d)$, when $\mathrm{ord}(y) > 0$ and $z = (e \to f)$ when $\mathrm{ord}(z) > 0$. Moreover, we prove each statement by induction on $\mathrm{ord}(xyz)$.

Base step: Trivial, all maps are identities. *Induction step:* We must split the proof in different cases. We first compare $\mathrm{ord}(x)$ with $\mathrm{ord}(yz)$, and then, for each case, we split the proof again by comparing $\mathrm{ord}(y)$ with $\mathrm{ord}(z)$. Moreover, Let $\gamma, \delta, \epsilon, \phi$ be types such that $yz = (\gamma\epsilon \to \delta\phi)$ as in Lemma 6.1.12.

- $\mathrm{ord}(x) = \mathrm{ord}(yz)$:
  - $\mathrm{ord}(y) = \mathrm{ord}(z)$:

$$\begin{aligned}\eta^{x \to x_{yz}}(\Gamma^x) &= \eta^{b \to b_{\delta\phi}} \circ \Gamma^{a \to b} \circ \eta^{a_{\gamma\epsilon} \to a} \\ &= \eta^{b \to b_{df}} \circ \Gamma^{a \to b} \circ \eta^{a_{ce} \to a} \\ \text{I.H.} &= \eta^{b_d \to b_{df}} \circ \eta^{b \to b_d} \circ \Gamma^{a \to b} \circ \eta^{a_c \to a} \circ \eta^{a_{ce} \to a_c} \\ &= \eta^{b_d \to b_{df}} \circ \eta^{x \to x_y}(\Gamma^x) \circ \eta^{a_{ce} \to a_c} \\ &= \eta^{x_y \to x_{yz}}(\eta^{x \to x_y}(\Gamma^x)).\end{aligned} \tag{6.58}$$

- ord(y) < ord(z):

$$\begin{aligned}
\eta^{x\to x_{yz}}(\Gamma^x) &= \eta^{b\to b_{\delta\phi}} \circ \Gamma^{a\to b} \circ \eta^{a_{\gamma\epsilon}\to a} \\
&= \eta^{b\to b_{yf}} \circ \Gamma^{a\to b} \circ \eta^{a_e\to a} \\
\text{I.H.} &= \eta^{b_y\to b_{yf}} \circ \eta^{b\to b_y} \circ \Gamma^{a\to b} \circ \eta^{a_e\to a} \\
&= \eta^{b_y\to b_{df}} \circ \eta^{x\to x_y}(\Gamma^x) \circ \eta^{a_e\to a} \\
&= \eta^{x_y\to x_{yz}}(\eta^{x\to x_y}(\Gamma^x)).
\end{aligned} \tag{6.59}$$

- ord(y) > ord(z):

$$\begin{aligned}
\eta^{x\to x_{yz}}(\Gamma^x) &= \eta^{b\to b_{\delta\phi}} \circ \Gamma^{a\to b} \circ \eta^{a_{\gamma\epsilon}\to a} \\
&= \eta^{b\to b_{dz}} \circ \Gamma^{a\to b} \circ \eta^{a_c\to a} \\
\text{I.H.} &= \eta^{b_d\to b_{dz}} \circ \eta^{b\to b_d} \circ \Gamma^{a\to b} \circ \eta^{a_c\to a} \\
&= \eta^{b_d\to b_{dz}} \circ \eta^{x\to x_y}(\Gamma^x) \\
&= \eta^{x_y\to x_{yz}}(\eta^{x\to x_y}(\Gamma^x)).
\end{aligned} \tag{6.60}$$

- ord(x) < ord(yz) (Here we use Lemma 6.2.6 to move promotion functions in and out of the ket):
  - ord(y) = ord(z):

$$\begin{aligned}
\eta^{x\to x_{yz}}(\Gamma^x) &= |\eta^{x\to x_{\delta\phi}}(\Gamma^x)\rangle^{I\to x_{\delta\phi}} \circ \eta^{I_{\gamma\epsilon}\to I} \\
&= |\eta^{x\to x_{df}}(\Gamma^x)\rangle^{I\to x_{\delta\phi}} \circ \eta^{I_{ce}\to I} \\
\text{I.H.} &= |\eta^{x_d\to x_{df}} \circ \eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_{df}} \circ \eta^{I_c\to I} \circ \eta^{I_{ce}\to I_c} \\
&= \eta^{x_d\to x_{df}} \circ |\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \circ \eta^{I_{ce}\to I_c} \\
&= \eta^{x_d\to x_{df}} \circ \eta^{x\to x_y}(\Gamma^x) \circ \eta^{I_{ce}\to I_c} \\
&= \eta^{x_y\to x_{yz}}(\eta^{x\to x_y}(\Gamma^x))
\end{aligned} \tag{6.61}$$

  - ord(y) < ord(z):

$$\begin{aligned}
\eta^{x\to x_{yz}}(\Gamma^x) &= |\eta^{x\to x_{\delta\phi}}(\Gamma^x)\rangle^{I\to x_{\delta\phi}} \circ \eta^{I_{\gamma\epsilon}\to I} \\
&= |\eta^{x\to x_{yf}}(\Gamma^x)\rangle^{I\to x_{yf}} \circ \eta^{I_e\to I} \\
\text{I.H.} &= |\eta^{x_y\to x_{yf}} \circ \eta^{x\to x_y}(\Gamma^x)\rangle^{I\to x_{yf}} \circ \eta^{I_e\to I} \\
&= \eta^{x_y\to x_{yz}}(\eta^{x\to x_y}(\Gamma^x))
\end{aligned} \tag{6.62}$$

  - ord(y) > ord(z):

$$\begin{aligned}
\eta^{x\to x_{yz}}(\Gamma^x) &= |\eta^{x\to x_{\delta\phi}}(\Gamma^x)\rangle^{I\to x_{\delta\phi}} \circ \eta^{I_{\gamma\epsilon}\to I} \\
&= |\eta^{x\to x_{dz}}(\Gamma^x)\rangle^{I\to x_{dz}} \circ \eta^{I_c\to I} \\
\text{I.H.} &= |\eta^{x_d\to x_{dz}} \circ \eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_{dz}} \circ \eta^{I_c\to I} \\
&= \eta^{x_d\to x_{dz}} \circ |\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \\
&= \eta^{x_d\to x_{dz}} \circ \eta^{x\to x_y}(\Gamma^x) \\
&= \eta^{x_y\to x_{yz}}(\eta^{x\to x_y}(\Gamma^x))
\end{aligned} \tag{6.63}$$

- $\mathrm{ord}(x) > \mathrm{ord}(yz)$

$$\begin{aligned} \eta^{x\to x_{yz}}(\Gamma^x) &= \eta^{b\to b_{yz}} \circ \Gamma^{a\to b} \\ \text{I.H.} &= \eta^{b_y\to b_{yz}} \circ \eta^{b\to b_y} \Gamma^{a\to b} \\ &= \eta^{b_y\to b_{yz}} \circ \eta^{x\to x_y}(\Gamma^x) \\ &= \eta^{x_y\to x_{yz}}(\eta^{x\to x_z}(\Gamma^x)) \end{aligned} \tag{6.64}$$

To conclude the proof, just notice that $\eta^{x_y\to x}$ is the inverse of $\eta^{x\to x_y}$ and $\eta^{x_{yz}\to x_y}$ is the inverse of $\eta^{x_y\to x_{yz}}$. As a consequence $\eta^{x_y\to x} \circ \eta^{x_{yz}\to x_y}$ is the inverse of $\eta^{x_y\to x_{yz}} \circ \eta^{x\to x_y}$. The latter is equal to $\eta^{x\to x_{yz}}$, therefore, $\eta^{x_y\to x} \circ \eta^{x_{yz}\to x_y}$ is the inverse of $\eta^{x\to x_{yz}}$ and thus $\eta^{x_y\to x} \circ \eta^{x_{yz}\to x_y} = \eta^{x_{yz}\to x}$. □

From this proposition, one of the most important properties of the promotion functions concerning the tensor product follows. Consider $\Gamma^x \boxtimes \Delta^y$. This is by definition equal to $\eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y)$. Assume now that $z \in \mathrm{Types}_{\mathcal{A}}$ is such that $z \prec y$. The product $\eta^{x\to x_z}(\Gamma^x) \boxtimes \Delta^y$ is by definition $\eta^{x_z\to x_{yz}} \circ \eta^{x\to x_z}(\Gamma^x) \boxtimes \eta^{y\to y_{xz}}(\Delta^y) = \eta^{x\to x_{yz}}(\Gamma^x) \boxtimes \eta^{y\to y_{xz}}(\Delta^y)$. However, since $z \prec y$, we have that $y_{xz} = y_z$ and $x_{yz} = x_y$. Therefore, $\eta^{x\to x_z}(\Gamma^x) \boxtimes \Delta^y = \eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y) = \Gamma^x \boxtimes \Delta^y$. The same reasoning applies to the promotion on the right-hand side with a type $z \prec x$. This implies that we can promote a side of an elementary product with any type as long as it precedes the type of the other side. This is the content of the following corollary.

Corollary 6.2.18: Partial promotion

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$ be such that $z \prec y$ and let $\Gamma^x \in L(x)$, and $\Delta^y \in L(y)$. Then, $\Gamma^x \boxtimes \Delta^y = \eta^{x\to x_z}(\Gamma^x) \boxtimes \Delta^y$ and $\Delta^y \boxtimes \Gamma^x = \Delta^y \boxtimes \eta^{x\to x_z}(\Gamma^x)$.

Another way of stating the same result is the principle of partial demotion. Let $x, \tilde{x}, y, z$ be such that $z \prec y$ and $x = \tilde{x}_z$, then the previous Corollary implies $\eta^{\tilde{x}_z\to\tilde{x}}(\Gamma^{\tilde{x}_z}) \boxtimes \Delta^y = \eta^{\tilde{x}\to\tilde{x}_z} \circ \eta^{\tilde{x}_z\to\tilde{x}}(\Gamma^{\tilde{x}_z}) \boxtimes \Delta^y = \Gamma^x \boxtimes \Delta^y$.

Corollary 6.2.19: Partial demotion

Let $x, \tilde{x}, y, z \in \mathrm{Types}_{\mathcal{A}}$ be such that $z \prec y$ and $x = \tilde{x}_z$. Let $\Gamma^x \in L(x)$, and $\Delta^y \in L(y)$. Then, $\Gamma^x \boxtimes \Delta^y = \eta^{\tilde{x}_z\to\tilde{x}}(\Gamma^x) \boxtimes \Delta^y$ and $\Delta^y \boxtimes \Gamma^x = \Delta^y \boxtimes \eta^{\tilde{x}_z\to\tilde{x}}(\Gamma^x)$.

The third easy consequence of Proposition 6.2.17 is a simplification of the promotion of the product of types (Proposition 6.2.16) in the case of types of equal order.

Lemma 6.2.20: Promotion of types of the same order and parallel composition with the identity

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$, such that $\mathrm{ord}(xy) = \mathrm{ord}(z)$. Then,

- $\eta^{x\to x_y} \boxtimes \mathrm{id}^{z\to z} = \eta^{xz\to xz_y} = \mathrm{id}^{x\to x} \boxtimes \eta^{z\to z_y}$,
- $\eta^{x_y\to x} \boxtimes \mathrm{id}^{z\to z} = \eta^{xz_y\to xz} = \mathrm{id}^{x\to x} \boxtimes \eta^{z_y\to z}$.

Proof. By definition of the product of maps of different types, we have that

$$\begin{aligned}
\eta^{x\to x_y} \boxtimes \mathrm{id}^{z\to z} &= \eta^{(x\to x_y)\to(x\to x_y)_{(z\to z)}}(\eta^{x\to x_y}) \otimes \eta^{(z\to z)\to(z\to z)_{x\to x_y}}(\mathrm{id}^{z\to z}) \\
&= (\eta^{x_y\to x_{yz}} \circ \eta^{x\to x_y} \circ \eta^{x_z\to x}) \otimes (\eta^{z\to z_{xy}} \circ \mathrm{id}^{z\to z} \circ \eta^{z_x\to z}) \\
&= (\eta^{x_y\to x_{yz}} \circ \eta^{x\to x_y} \circ \eta^{x_z\to x}) \otimes (\eta^{z\to z_{xy}} \circ \eta^{z_x\to z}) \\
\text{Proposition 6.2.17} &= (\eta^{x\to x_{yz}} \circ \eta^{x_z\to x}) \otimes (\eta^{z\to z_{xy}} \circ \eta^{z_x\to z}) \\
\text{Proposition 6.2.17} &= (\eta^{x_z\to x_{yz}} \circ \eta^{x\to x_z} \circ \eta^{x_z\to x}) \otimes (\eta^{z_x\to z_{xy}} \circ \eta^{z\to z_x} \circ \eta^{z_x\to z}) \\
&= \eta^{x_z\to x_{yz}} \otimes \eta^{z_x\to z_{xy}} \\
&= \eta^{x_z\to x_{yz}} \boxtimes \eta^{z_x\to z_{xy}} \\
\text{Proposition 6.2.16} &= \eta^{xz\to xz_y} .
\end{aligned} \tag{6.65}$$

The proof of the second equality is identical.

For the demotion map, we have

$$\begin{aligned}
\eta^{xz_y\to xz} &= (\eta^{xz\to xz_z})^{-1} \\
&= (\eta^{x\to x_y} \boxtimes \mathrm{id}^{z\to z})^{-1} \\
&= [(\eta^{x_y\to x_{yz}} \circ \eta^{x\to x_y} \circ \eta^{x_z\to x}) \otimes (\eta^{z\to z_{xy}} \circ \mathrm{id}^{z\to z} \circ \eta^{z_x\to z})]^{-1} \\
&= (\eta^{x\to x_z} \circ \eta^{x_y\to x} \circ \eta^{x_{yz}\to x_y}) \otimes (\eta^{z\to z_x} \circ \mathrm{id}^{z\to z} \circ \eta^{z\to z_{xy}}) \\
&= \eta^{x_y\to x} \boxtimes \mathrm{id}^{z\to z} .
\end{aligned} \tag{6.66}$$

Again, the proof of the second equality is identical. □

We have multiple times expressed the idea that the promotion function is the most 'natural' way to promote an operator. In a sense, it is the closest thing to an identity function, considering that the input and output spaces are not isomorphic. In the following Lemma, we see a justification for this intuition.

Lemma 6.2.21: Parallel composition with identity

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(xy) = \mathrm{ord}(z)$, and let $\Gamma^{x\to y} \in L(x \to y)$. Then, $\Gamma^{x\to y} \boxtimes \eta^{z\to z_y} = \Gamma^{x\to y} \boxtimes \mathrm{id}^{z\to z} = \Gamma^{x\to y} \boxtimes \eta^{z_x\to z}$.

Proof.

$$\begin{aligned}
\Gamma^{x\to y} \boxtimes \eta^{z\to z_y} &= (\eta^{y\to y_{z_y}} \circ \Gamma^{x\to y} \circ \eta^{z_x\to z}) \otimes (\eta^{z_y\to z_y} \circ \eta^{z\to z_y} \circ \eta^{z_x\to z}) \\
\text{Lemma 6.2.12} &= (\eta^{y\to y_z} \circ \Gamma^{x\to y} \circ \eta^{z_x\to z}) \otimes (\eta^{z_y\to z_y} \circ \eta^{z\to z_y} \circ \eta^{z_x\to z}) \\
\text{Lemma 6.2.11} &= (\eta^{y\to y_z} \circ \Gamma^{x\to y} \circ \eta^{z_x\to z}) \otimes (\eta^{z\to z_y} \circ \eta^{z_x\to z}) \\
&= (\eta^{y\to y_z} \circ \Gamma^{x\to y} \circ \eta^{z_x\to z}) \otimes (\eta^{z\to z_y} \circ \mathrm{id}^{z\to z} \circ \eta^{z_x\to z}) \\
&= \Gamma^{x\to y} \boxtimes \mathrm{id}^{z\to z} .
\end{aligned} \tag{6.67}$$

The same steps can be used to prove the equality of the demotion function. □

An analogous statement is valid for the case in which the promotion operator, the demotion operator, and the identity are on the left.

Next on the list of desirable properties of the promotion functions is unitarity. Indeed, we are working with Hilbert spaces, and the inner product plays an important role. Therefore, it is reasonable for a ‘natural’ promotion function to preserve the inner product.

Theorem 6.2.22: Promotion maps are unitary

Let $x, y \in \text{Types}_{\mathcal{A}}$. Then, $(\eta^{x \to x_y})$ is unitary.

Proof. To prove the statement, we show that $\eta^{x \to x_y}$ preserves the inner product. We begin by proving that $|\cdot\rangle$ is unitary. Note that $\{ 1 \}$ is a orthonormal basis of $L(1)$, thus, by definition, of the inner product

$$\begin{aligned} (|\Gamma^x\rangle^{I \to x}, |\Delta^x\rangle^{I \to x})^{I \to x} = (|\Gamma^x\rangle^{I \to x}(1), |\Delta^x\rangle^{I \to x}(1))_x \\ (\Gamma^x, \Delta^x)_x. \end{aligned} \tag{6.68}$$

This shows that $|\cdot\rangle$ preserves the inner product, and therefore it is unitary.

We are now ready to prove that $\eta^{x \to x_y}$ is unitary, and we do so by induction. If $\text{ord}(xy) = 0$, then the promotion function is the identity map, which is unitary. Let $\text{ord}(xy) > 0$ and assume that the statement is true for all types up to order $\text{ord}(xy) - 1$.

- If $\text{ord}(x) = \text{ord}(y)$, then $x = (a \to b)$, $y = (c \to d)$, and

  $$(\eta^{x \to x_y}(\Gamma^x), \eta^{y \to y_x}(\Delta^x))_{x_y} = (\eta^{b \to b_d} \circ \Gamma^{a \to b} \circ \eta^{a_c \to a}, \eta^{b \to b_d} \circ \Delta^{a \to b} \circ \eta^{a_c \to a}). \tag{6.69}$$

  By the induction hypothesis, $\eta^{b \to b_d}$ is unitary; therefore, when we move it from the left to the right-hand side of the inner product, it becomes $\eta^{b_d \to b}$, and it cancels out with the other $\eta^{b \to b_d}$. With that, the inner product above becomes $(\Gamma^{a \to b} \circ \eta^{a_c \to a}, \Delta^{a \to b} \circ \eta^{a_c \to a})_{(}a_c \to b)$. Let $\{ u_i^{a_c} \}_i$ be a orthonormal basis of $L(a_c)$. Observe that, since $\eta^{a_c \to a}$ is unitary by induction hypothesis, $\{ u_i^a = \eta^{a_c \to a}(u_i^{a_c}) \}_i$ is a basis of $L(a)$. Therefore,

  $$\begin{aligned} &(\Gamma^{a \to b} \circ \eta^{a_c \to a}, \Delta^{a \to b} \circ \eta^{a_c \to a})_{(}a_c \to b) \\ &\quad = \sum_i (\Gamma^{a \to b} \circ \eta^{a_c \to a}(u_i^{a_c}), \Delta^{a \to b} \circ \eta^{a_c \to a}(u^{a_c^i}))_b \\ &\quad = \sum_i (\Gamma^{a \to b}(u_i^a), \Delta^{a \to b}(u^{a^i}))_b \\ &\quad = (\Gamma^{a \to b}, \Delta^{a \to b})_{(}x). \end{aligned} \tag{6.70}$$

  This shows that $\eta^{x \to x_y}$ preserves the inner product and therefore is unitary.

- $\text{ord}(x) > \text{ord}(y)$ and $\text{ord}(x) < \text{ord}(y)$: In the proof of the previous case, what we have shown is that pre-composition and post-composition with a unitary function preserve the inner product. Indeed, we did not use any property of the promotion and demotion maps of lower order, except the fact that they were unitary. Since in the cases $\text{ord}(x) > \text{ord}(y)$ and $\text{ord}(x) < \text{ord}(y)$, $\eta^{x \to x_y}$ can be again decomposed as pre- and post-composition with unitary functions, such as lower order promotion and demotion functions and kets, it preserves the inner product as well.

□

An immediate corollary of the previous theorem is that the inner product splits for elementary tensors.

Corollary 6.2.23: Inner product of elementary tensors

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^x, \Gamma'_x \in L(x)$, $\Delta^y, \Delta'_y \in \mathrm{Types}_{\mathcal{A}}$. Then, $(\Gamma^x \boxtimes \Delta^y, \Gamma'_x \boxtimes \Delta'_y)_{xy} = (\Gamma^x, \Gamma'_x)_x(\Delta^y, \Delta'_y)_y$.

Proof. This is a direct consequence of Theorem 6.2.22 and the fact that the inner product splits in the case of the tensor product.

$$\begin{aligned}(\Gamma^x \boxtimes \Delta^y, \Gamma'_x \boxtimes \Delta'_y)_{xy} &= (\eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y), \eta^{x\to x_y}(\Gamma'_x) \otimes \eta^{y\to y_x}(\Delta'_y))_{xy} \\ &= (\eta^{x\to x_y}(\Gamma^x), \eta^{x\to x_y}(\Gamma'_x))_{x_y}(\eta^{y\to y_x}(\Delta^y)\eta^{y\to y_x}(\Delta'_y))_{y_x} \\ &= (\Gamma^x, \Gamma'_x)_x(\Delta^y, \Delta'_y)_y.\end{aligned} \tag{6.71}$$

□

We conclude this Section by noting that the ket function is a special case of the promotion function.

Lemma 6.2.24: Ket vs promotion

Let $x \in \mathrm{Types}_{\mathcal{A}}$. Then $|\cdot\rangle^{I\to x} = \eta^{x\to(I\to x)}(\cdot)$.

Proof. By definition of promotion function and Lemma 6.2.11:

$$\begin{aligned}\eta^{x\to(I\to x)}(\cdot) &= \eta^{x\to x_x} \circ |\cdot\rangle^{I\to x} \circ \eta^{I\to I} \\ &= \mathrm{id}^{x\to x} \circ |\cdot\rangle^{I\to x} \circ \mathrm{id}^{I\to I} \\ &= |\cdot\rangle^{I\to x}.\end{aligned} \tag{6.72}$$

□

6.2.3 Properties of the Square Product

In this subsection, we will present the properties of the square product. As the reader can imagine, the properties of the square tensor product mimic the properties of the standard tensor product: It is associative, bilinear, and a unit exists. These results, while relevant, are not the key results of this subsection. We will show what happens when one can sequentially compose or evaluate elementary square tensors, and these results will be fundamental in all subsequent proofs.

We begin by showing that $\boxtimes$ is associative.

Proposition 6.2.25: Associativity of $\boxtimes$

Let $x, y, z \in \mathrm{Types}_{\mathcal{A}}$ and let $\Gamma^x \in L(x), \Delta^y \in L(y), \Theta^z \in L(z)$. Then $(\Gamma^x \boxtimes \Delta^y) \boxtimes \Theta^z = \Gamma^x \boxtimes (\Delta^y \boxtimes \Theta^z)$.

Proof.

$$\begin{aligned} (\Gamma^x \boxtimes \Delta^y) \boxtimes \Theta^z &= \eta^{xy\to xy_z}(\Gamma^x \boxtimes \Delta^y) \otimes \eta^{z\to z_{xy}}(\Theta^z) \\ &= \eta^{xy\to xy_z}[\eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y)] \otimes \eta^{z\to z_{xy}}(\Theta^z) \\ \text{Proposition 6.2.16} &= (\eta^{x_y\to x_{yz}} \otimes \eta^{y_x\to y_{xz}})[\eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y\to y_x}(\Delta^y)] \otimes \eta^{z\to z_{xy}}(\Theta^z) \\ &= [\eta^{x_y\to x_{yz}} \circ \eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y_x\to y_{xz}} \circ \eta^{y\to y_x}(\Delta^y)] \otimes \eta^{z\to z_{xy}}(\Theta^z) \\ \text{Proposition 6.2.17} &= [\eta^{x\to x_{yz}}(\Gamma^x) \otimes \eta^{y\to y_{xz}}(\Delta^y)] \otimes \eta^{z\to z_{xy}}(\Theta^z) \\ &= \eta^{x\to x_{yz}}(\Gamma^x) \otimes \eta^{y\to y_{xz}}(\Delta^y) \otimes \eta^{z\to z_{xy}}(\Theta^z) \end{aligned} \tag{6.73}$$

With the same steps, one obtains

$$\Gamma^x \boxtimes (\Delta^y \boxtimes \Theta^z) = \eta^{x\to x_{yz}}(\Gamma^x) \otimes \eta^{y\to y_{xz}}(\Delta^y) \otimes \eta^{z\to z_{xy}}(\Theta^z). \tag{6.74}$$

This concludes the proof. □

We continue by showing that 1 is the unit for ⊠, as 1 is the unit of ⊗.

Proposition 6.2.26: Unit of ⊠

Let $x \in \text{Types}_{\mathcal{A}}$ and $\Gamma^x \in L(x)$. Then, $1 \boxtimes M = M = M \boxtimes 1$.

Proof. We prove it by induction on $\text{ord}(x)$. In the base step, x is elementary, and $1 \boxtimes M = 1 \otimes M = M = M \otimes 1 = M \boxtimes 1$.

Inductive step: Let $x = (a \to b)$ and let $u^a \in L(a)$.

$$\begin{aligned} (1 \boxtimes \Gamma^x)(u^a) &= (\eta^{I\to I_x}(1) \otimes \eta^{x\to x_I}(\Gamma^x))(u^a) \\ \text{I.H.} &= (\eta^{I\to I_x}(1) \otimes \eta^{x\to x_I}(\Gamma^x))(1 \boxtimes u^a) \\ &= (\eta^{I\to I_b} \circ |1\rangle^{I\to I} \circ \eta^{I_a\to I}) \otimes \Gamma^x)(\eta^{I\to I_a}(1) \otimes u^a) \\ &= \eta^{I\to I_b} \circ |1\rangle^{I\to I} \circ \eta^{I_a\to I} \circ \eta^{I\to I_a}(1) \otimes \Gamma^{a\to b}(u^a) \\ &= \eta^{I\to I_b}(1) \otimes \Gamma^{a\to b}(u^a) \\ &= 1 \boxtimes \Gamma^{a\to b}(u^a) \\ \text{I.H} &= \Gamma^{a\to b}(u^a). \end{aligned} \tag{6.75}$$

The proof of $\Gamma^x = \Gamma^x \boxtimes 1$ is identical. □

With the properties above, we are ready to prove what happens when elementary tensors are evaluated at an input. Consider $\Gamma^x \boxtimes \Delta^y \in L(xy)$. The first question that we should ask ourselves is what the input of such a function is. If $\text{ord}(xy) = 0$, then both types are elementary, and $\Gamma^x \boxtimes \Delta^y$ is a function associated with a type of order 0. This means that in our type system, there is no map of a lower order that we can consider as input. Therefore, we will restrict our analysis to the case of $\text{ord}(xy) > 0$. As usual, if $\text{ord}(x) > 0$, we write $x = (a \to b)$ and if $\text{ord}(y) > 0$ we write $y = (c \to d)$. With this notation,

$$\Gamma^x \boxtimes \Delta^y \in L(xy) = \begin{cases} L(ac \to bd) & \text{if } \text{ord}(x) = \text{ord}(y), \\ L(a \to by) & \text{if } \text{ord}(x) > \text{ord}(y), \\ L(c \to xd) & \text{if } \text{ord}(x) < \text{ord}(y). \end{cases} \tag{6.76}$$

Therefore, the input of $\Gamma^x \boxtimes \Delta^y$ depends on the relative order of x and y. This tells us that the first thing to check when writing an expression that includes the evaluation of an elementary tensor $\Gamma^x \boxtimes \Delta^y$ at its input is if the type of the input matches the left type of xy. For example, if $\mathrm{ord}(x) = \mathrm{ord}(y)$, we want the type of the input to be ac, and it is reasonable to expect that $\Gamma^{a\to b} \boxtimes \Delta^{c\to d} @ u^a \boxtimes v^c = \Gamma^{a\to b}(u^a) \boxtimes \Delta^{c\to d}(v^c)$. Similarly, if $\mathrm{ord}(x) > 0$, then the type of the input must be a, and it is reasonable to expect $\Gamma^{a\to b} \boxtimes \Delta^y @ u^a = \Gamma^{a\to b}(u^a) \boxtimes \Delta^y$. Lastly, if $\mathrm{ord}(x) < \mathrm{ord}(y)$, the type of the input map must be c, and it is reasonable to expect that $\Gamma^x \boxtimes \Delta^{c\to d} @ v^c = \Gamma^x \boxtimes \Delta^{c\to d}(v^c)$. The following theorem states that when the types match, these reasonable expectations are met.

Theorem 6.2.27: Evaluation of elementary tensors

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{ord}(xy) > 0$, and let $\Gamma^x \in L(x)$, $\Delta^y \in L(y)$. Let $a, b, c, d \in \mathrm{Types}_{\mathcal{A}}$ be such that $x = (a \to b)$ if $\mathrm{ord}(x) > 0$ and $y = (c \to d)$ if $\mathrm{ord}(y) > 0$.

- If $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $\Gamma^{a\to b} \boxtimes \Delta^{c\to d} @ u^a \boxtimes v^c = \Gamma^{a\to b}(u^a) \boxtimes \Gamma^{c\to d}(v^c)$ for all $u^a \in L(a)$ and $v^c \in L(c)$.
- If $\mathrm{ord}(x) < \mathrm{ord}(y)$, then $\Gamma^x \boxtimes \Delta^{c\to d} @ v^c = \Gamma^x \boxtimes \Delta^{c\to d}(v^c)$ for all $v^c \in L(c)$.
- If $\mathrm{ord}(x) > \mathrm{ord}(y)$, then $\Gamma^{a\to b} \boxtimes \Delta^y @ u^a = \Gamma^{a\to b}(u^a) \boxtimes \Delta^y$ for all $u^a \in L(a)$.

Proof. We start with the case $\mathrm{ord}(x) = \mathrm{ord}(y)$. The proof follows from the definition of the product of maps.

$$\begin{aligned}
\Gamma^{a\to b} \boxtimes \Delta^{c\to d} @ u^a \boxtimes v^c &= (\eta^{b\to b_d} \circ \Gamma^{a\to b} \circ \eta^{a_c\to a}) \otimes (\eta^{d\to d_b} \circ \Delta^{c\to d} \circ \eta^{c_a\to c}) \\
&\qquad @ \eta^{a\to a_c}(u^a) \otimes \eta^{c\to c_a}(v^c) \\
&= \eta^{b\to b_d}(\Gamma^{a\to b}(u^a)) \otimes \eta^{d\to d_b}(\Delta^{c\to d}(v^c)) \\
&= \Gamma^{a\to b}(u^a) \boxtimes \Delta^{c\to d}(v^c).
\end{aligned} \tag{6.77}$$

In the case $\mathrm{ord}(x) < \mathrm{ord}(y)$, we use the definition of the product of maps and Proposition 6.2.26.

$$\begin{aligned}
\Gamma^x \boxtimes \Delta^{c\to d} @ v^c &= \Gamma^x \boxtimes \Delta^{c\to d} @ 1 \boxtimes v^c \\
&= |\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \otimes \eta^{d\to d_x} \circ \Delta^{c\to d} \\
&\qquad @ \eta^{I\to I_c}(1) \otimes \eta^{c\to c_I}(v^c) \\
&= |\eta^{x\to x_d}(\Gamma^x)\rangle^{I\to x_d}(1) \otimes \eta^{d\to d_x} \circ \Delta^{c\to d}(v^c) \\
&= \eta^{x\to x_d}(\Gamma^x) \otimes \eta^{d\to d_x}(\Delta^{c\to d}(v^c)) \\
&= \Gamma^x \boxtimes \Delta^{c\to d}(v^c).
\end{aligned} \tag{6.78}$$

The proof of the case $\mathrm{ord}(x) > \mathrm{ord}(y)$ is identical. □

Together with Proposition 6.2.16, Lemma 6.2.24, and Corollary 6.2.23, the theorem above implies the following corollary.

Corollary 6.2.28: Product of ket and bra

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^x \in L(x)$, $\Gamma^y \in L(y)$. Then,

1. $|\Gamma^x \boxtimes \Delta^y\rangle^{I\to xy} = |\Gamma^x\rangle^{I\to x} \boxtimes |\Delta^y\rangle^{I\to y}$,
2. $\langle\Gamma^x \boxtimes \Delta^y|^{xy\to I} = \langle\Gamma^x|^{x\to I} \boxtimes \langle\Delta^y|^{y\to I}$.

Proof. The first statement is a trivial consequence of Proposition 6.2.16 and Lemma 6.2.24. To prove the second statement, it is enough to show that the two functions coincide on elementary tensors $u^x \boxtimes v^y \in L(xy)$:

$$\begin{aligned} \langle\Gamma^x \boxtimes \Delta^y|^{xy\to I} \, @u^x \boxtimes u^y &= (u^x \boxtimes v^y, \Gamma^x \boxtimes \Delta^y)_{xy} \\ \text{Corollary 6.2.23} &= (u^x, |\Gamma^x)_x (v^y, \Delta^y)_y \\ &= \langle\Gamma^x|^{x\to I}\,(u^x) \boxtimes \langle\Delta^y|^{y\to I}\,(u^y) \\ \text{Theorem 6.2.27} &= \langle\Gamma^x|^{x\to I} \boxtimes \langle\Delta^y|^{y\to I} \, @u^x \boxtimes v^y. \end{aligned} \tag{6.79}$$

□

The next result that we want to show is what happens when we take the square product of two sequential compositions. That is, consider the function

$$(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e}). \tag{6.80}$$

Similarly to what happens for the tensor product, we would expect that the map above is equal to

$$(\Delta^{b\to c} \boxtimes \Xi^{e\to f}) \circ (\Gamma^{a\to b} \boxtimes \Theta^{d\to e}). \tag{6.81}$$

However, there are three problems with this expectation. First, there is no guarantee that the output type of $\Gamma^{a\to b} \boxtimes \Theta^{d\to e}$ is equal to the input type of $\Delta^{b\to c} \boxtimes \Xi^{e\to f}$. If that is not the case, then the expression in Eq. (6.81) has no meaning. Second, the input type of the expression in Eq. (6.80) may not be equal to the input type of the expression in Eq. (6.81). Again, if this is not the case, the two expressions cannot be equal. Third, the output type of the expression in Eq. (6.80) may not be equal to the output type of the expression in Eq. (6.81). Once more, if the two types are different, then the two expressions cannot be equal.

As a consequence, before checking if two expressions are equal, we should check if the types are consistent. In our case, the consistency conditions are

1. output type of $\Gamma^{a\to b} \boxtimes \Theta^{d\to e}$ equals input type of $\Delta^{b\to c} \boxtimes \Xi^{e\to f}$,
2. input type of $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e})$ equals input type of $(\Delta^{b\to c} \boxtimes \Xi^{e\to f}) \circ (\Gamma^{a\to b} \boxtimes \Theta^{d\to e})$,
3. output type of $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e})$ equals output type of $(\Delta^{b\to c} \boxtimes \Xi^{e\to f}) \circ (\Gamma^{a\to b} \boxtimes \Theta^{d\to e})$.

To simplify these requirements, we promote every operator appearing on one side of the tensor product with the type of the operator appearing on the other side. For example, in the case $\Gamma^{a\to b} \boxtimes \Theta^{d\to e}$ we define

$$\begin{aligned} \tilde{\Gamma}^{\alpha\to\beta} &= \eta^{(a\to b)\to(a\to b)_{d\to e}}(\Gamma^{a\to b}), \\ \tilde{\Theta}^{\delta\to\epsilon} &= \eta^{(d\to e)\to(d\to e)_{(a\to b)}}(\Theta^{d\to e}), \end{aligned} \tag{6.82}$$

where $\alpha, \beta, \delta, \epsilon$ are abstract types that describe the left and right types of $\tilde{\Gamma}$ and $\tilde{\Delta}$. These types will vary depending on the relative orders of a, b, d, e. Note that, $\Gamma^{a\to b} \boxtimes \Theta^{d\to e} = \tilde{\Gamma}^{\alpha\to\beta} \otimes \tilde{\Theta}^{\delta\to\epsilon} \in L(\alpha\delta \to \beta\epsilon)$, because $\alpha \to \beta$ has the same structure of $\delta \to \epsilon$, Similarly, we define

$$\begin{aligned}
\tilde{\Delta}^{\beta'\to\gamma} &= \eta^{(b\to c)\to(b\to c)_{(e\to f)}}(\Delta^{b\to c}), \\
\tilde{\Xi}^{\epsilon'\to\phi} &= \eta^{(e\to f)\to(e\to f)_{(b\to c)}}(\Xi^{e\to f}), \\
\widetilde{(\Delta \circ \Gamma)}^{\alpha'\to\gamma'} &= \eta^{(a\to c)\to(a\to c)_{(d\to f)}}(\Delta^{b\to c} \circ \Gamma^{a\to b}), \\
\widetilde{(\Xi \circ \Theta)}^{\delta'\to\phi'} &= \eta^{(d\to f)\to(d\to f)_{a\to c}}(\Xi^{e\to f} \circ \Theta^{d\to e}).
\end{aligned} \tag{6.83}$$

With this notation,

$$\Delta^{b\to c} \boxtimes \Xi^{e\to f} = \tilde{\Delta}^{\beta'\to\gamma} \otimes \tilde{\Xi}^{\epsilon'\to\phi} \in L(\beta'\epsilon' \to \gamma\phi), \tag{6.84}$$

and

$$\begin{aligned}
(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e}) = \widetilde{(\Delta \circ \Gamma)}^{\alpha'\to\gamma'} \otimes \widetilde{(\Xi \circ \Theta)}^{\delta'\to\phi'} \\
\in L(\alpha'\delta' \to \gamma'\phi').
\end{aligned} \tag{6.85}$$

We can rewrite the consistency conditions, as conditions on $\alpha, \beta \ldots$:

1. $\beta\epsilon = \beta'\epsilon'$,
2. $\alpha\delta = \alpha'\delta'$,
3. $\gamma\phi = \gamma'\phi'$.

Therefore, the expressions in Eq. (6.80) and Eq. (6.81) can be equal only if the consistency conditions above are satisfied. We are going to show that if the conditions are satisfied, then the two expressions are equal, as it was reasonable to expect. Before doing that, we need to show that we can drop the 'prime' symbol. The proof is quite long and, in some instances, repetitive. However, it is very instructive to read it to understand the power and robustness of the type formalism.

Lemma 6.2.29: Preliminary results for Theorem 6.2.30

Let $a, b, c, d, e, f \in \mathrm{Types}_{\mathcal{A}}$,and let $\alpha, \beta \in \mathrm{Types}_{\mathcal{A}}$ be such that $(\alpha \to \beta) = (a \to b)_{(d\to e)}$. Similarly, let $(\beta' \to \gamma) = (b \to c)_{(e\to f)}$, $(\delta \to \epsilon) = (d \to e)_{a\to b}$, and $(\epsilon' \to \phi) = (e \to f)_{(b\to c)}$. Furthermore, let $(\alpha' \to \gamma') = (a \to c)_{(d\to f)}$, and $(\delta' \to \phi') = (d \to f)_{(a\to c)}$. If $\beta\epsilon = \beta'\epsilon'$, $\alpha\delta = \alpha'\delta'$, and $\gamma\phi = \gamma'\phi'$, then $\beta' = \beta$, $\epsilon' = \epsilon$, $\alpha = \alpha'$, $\delta = \delta'$, $\gamma = \gamma'$, and $\phi = \phi'$.

Proof. In the first part of this lemma, we compute all the types introduced in the statement of this lemma.

$$(\alpha \to \beta) = (a \to b)_{(d\to e)} = \begin{cases} (I_d \to (a \to b)_e) & \text{if } \operatorname{ord}(ab) < \operatorname{ord}(de), \\ (a_d \to b_e) & \text{if } \operatorname{ord}(ab) = \operatorname{ord}(de), \\ (a \to b_{(d\to e)}) & \text{if } \operatorname{ord}(ab) > \operatorname{ord}(de). \end{cases} \tag{6.86}$$

In all the cases above, the type on the left of the main arrow is α, and the type on the right is β. Similarly, we compute all the other types.

$$(\beta' \to \gamma) = (b \to c)_{(e\to f)} = \begin{cases} (I_e \to (b \to c)_f) & \text{if } \operatorname{ord}(bc) < \operatorname{ord}(ef), \\ (b_e \to c_f) & \text{if } \operatorname{ord}(bc) = \operatorname{ord}(ef), \\ (b \to c_{(e\to f)}) & \text{if } \operatorname{ord}(bc) > \operatorname{ord}(ef). \end{cases} \tag{6.87}$$

$$(\delta \to \epsilon) = (d \to e)_{(a\to b)} = \begin{cases} (I_a \to (d \to e)_b) & \text{if } \operatorname{ord}(de) < \operatorname{ord}(ab), \\ (d_a \to e_b) & \text{if } \operatorname{ord}(de) = \operatorname{ord}(ab), \\ (d \to e_{(a\to b)}) & \text{if } \operatorname{ord}(de) > \operatorname{ord}(ab). \end{cases} \tag{6.88}$$

$$(\epsilon' \to \phi) = (e \to f)_{(b\to c)} = \begin{cases} (I_b \to (e \to f)_c) & \text{if } \operatorname{ord}(ef) < \operatorname{ord}(bc), \\ (e_b \to f_c) & \text{if } \operatorname{ord}(ef) = \operatorname{ord}(bc), \\ (e \to f_{(b\to c)}) & \text{if } \operatorname{ord}(ef) > \operatorname{ord}(bc). \end{cases} \tag{6.89}$$

$$(\alpha' \to \gamma') = (a \to c)_{(d\to f)} = \begin{cases} (I_d \to (a \to c)_f) & \text{if } \operatorname{ord}(ac) < \operatorname{ord}(df), \\ (a_d \to c_f) & \text{if } \operatorname{ord}(ac) = \operatorname{ord}(df), \\ (a \to c_{(d\to f)}) & \text{if } \operatorname{ord}(ac) > \operatorname{ord}(df). \end{cases} \tag{6.90}$$

$$(\delta' \to \phi') = (d \to f)_{(a\to c)} = \begin{cases} (I_a \to (d \to f)_c) & \text{if } \operatorname{ord}(df) < \operatorname{ord}(ac), \\ (d_a \to f_c) & \text{if } \operatorname{ord}(df) = \operatorname{ord}(ac), \\ (d \to f_{(a\to c)}) & \text{if } \operatorname{ord}(df) > \operatorname{ord}(ac). \end{cases} \tag{6.91}$$

Let us assume that $\beta\epsilon = \beta'\epsilon'$. From Eq. (6.86) and Eq.(6.88), we obtain

$$\beta\epsilon = \begin{cases} (a \to b)_e e_{(a\to b)} & \text{if } \operatorname{ord}(ab) < \operatorname{ord}(de), \\ b_e e_b & \text{if } \operatorname{ord}(ab) = \operatorname{ord}(de), \\ b_{(d\to e)}(d \to e)_b & \text{if } \operatorname{ord}(ab) > \operatorname{ord}(de). \end{cases} \tag{6.92}$$

We use Eq. (6.87) and Eq. (6.89) to compute $\beta'\epsilon'$

$$\beta'\epsilon' = \begin{cases} I_e e & \text{if } \operatorname{ord}(bc) < \operatorname{ord}(ef), \\ b_e e_b & \text{if } \operatorname{ord}(bc) = \operatorname{ord}(ef), \\ bI_b & \text{if } \operatorname{ord}(bc) > \operatorname{ord}(ef). \end{cases} \tag{6.93}$$

We have nine possible cases to consider, depending on the relative orders of $\operatorname{ord}(ab)$, $\operatorname{ord}(de)$, $\operatorname{ord}(ef)$, and $\operatorname{ord}(bc)$. With the assumption $\beta\epsilon = \beta'\epsilon'$, the cases are:

- $I_e e = (a \to b)_e e_{a\to b}$, which becomes $e = (a \to b)e$. Thanks to Proposition 6.1.24, we know that this implies $(a \to b) \prec e$ and $(a \to b) = I_{(a\to b)}$. As a consequence, $e_{(a\to b)} = e$, by definition of type inclusion, thus $\epsilon = \epsilon'$. Moreover, $(a \to b)_e = I_{e(a\to b)} = I_{e_{a\to b}} = I_e$ (Lemma 6.1.20), which implies $\beta = \beta'$.

- $I_e e = b_e e_b$, which becomes $e = be$. With the same arguments of the previous case, one gets $b = I_b$ and $b \prec e$, which in turn implies $\beta = \beta'$ and $\epsilon = \epsilon'$ (replace $(a \to b)$ with b everywhere in the argument).
- $I_e e = b_{(d\to e)}(d \to e)_b$, which becomes $e = b(d \to e)$. Observe that this case is impossible because the type $\mathrm{ord}(b(d \to e)) > \mathrm{ord}(e)$, and therefore the two types can never be equal.
- $b_e e_b = (a \to b)_e e_{a\to b}$, which becomes $be = (a \to b)e$. With the same argument used in Proposition 6.1.24, one gets $a = I_a$. Indeed, if the tree associated with a has one leaf not labelled by the empty string, then such a label would appear in the tree associated with $(a \to b)e$ in addition to all the labels associated with b and e. However, the tree associated with be only contains the labels from b and e; therefore, the types $(a \to b)e$ and be would not be equal.

$$\begin{aligned} b_e &= b_{be} = b_{(I_a\to b)e} \\ &= (b_{I_a\to b})_e = (I_a \to b)_e \\ &= (a \to b)_e \\ e_b &= e_{be} = e_{(I_a\to b)e} \\ &= (e_e)_{(I_a\to b)} = e_{(I_a\to b)} \\ &= e_{(a\to b)} \end{aligned} \tag{6.94}$$

  This proves that $\beta = \beta'$ and $\epsilon = \epsilon'$.

- $b_e e_b = b_e e_b$: Trivial.
- $b_e e_b = b_{d\to e}(d \to e)_b$. With the same argument of the case $b_e e_b = (a \to b)_e e_{a\to b}$, one proves that $d = I_d$, and $b_e = b_{d\to e}$ and $e_b = (d \to e)_b$.
- $bI_b = (a \to b)e_{(a\to b)}$, which becomes $b = (a \to b)e$. As in the case $I_e e = b_{(d\to e)}(d \to e)_b$, the order of types is different, and therefore this scenario can never happen.
- $bI_b = b_e e_b$: Same proof of the case $I_e e = b_e e_b$.
- $bI_b = b_{(d\to e)}(d \to e)_b$: Same proof of the case $I_e e = (a \to b)_e e_{(a\to b)}$.

We covered all possible cases and showed that $\beta = \beta'$ and $\epsilon = \epsilon'$.

Now, we turn our attention to the second statement. We compute $\alpha\delta$ using Eq. (6.86) and Eq. (6.88).

$$\alpha\delta = \begin{cases} I_d d & \text{if } \mathrm{ord}(ab) < \mathrm{ord}(de), \\ a_d d_a & \text{if } \mathrm{ord}(ab) = \mathrm{ord}(de), \\ a I_a & \text{if } \mathrm{ord}(ab) > \mathrm{ord}(de). \end{cases} \tag{6.95}$$

We compute $\alpha'\delta'$ using Eq. (6.90) and Eq. (6.91).

$$\alpha'\delta' = \begin{cases} I_d d & \text{if } \mathrm{ord}(ac) < \mathrm{ord}(df), \\ a_d d_a & \text{if } \mathrm{ord}(ac) = \mathrm{ord}(df), \\ a I_a & \text{if } \mathrm{ord}(ac) > \mathrm{ord}(df). \end{cases} \tag{6.96}$$

In this case, the possible distinct cases are as follows:

- $I_d d = I_d d$: Trivial.
- $I_d d = a_d d_a$: Same proof of $I_e e = b_e e_b$.
- $I_d d = aI_a$, which becomes $d = a$. Note that, this happens when $\mathrm{ord}(ab) < \mathrm{ord}(de)$ and $\mathrm{ord}(ac) > \mathrm{ord}(df)$, or when $\mathrm{ord}(ab) > \mathrm{ord}(de)$ and $\mathrm{ord}(ac) < \mathrm{ord}(df)$. The first scenario, corresponds to the cases $I_e e = (a \to b)_e e_{a\to b}$, $b_e e_b = (a \to b)_e e_{a\to b}$, and $bI_b = (a \to b)e_{(a\to b)}$ of the proof for $\beta = \beta'$ and $\epsilon = \epsilon'$. In all these cases, whenever the equality of types is possible, it implies $a = I_a$. Therefore, $I_d = I_a = a$, and $d = a = I_a$, which prove $\alpha = \alpha'$ and $\delta = \delta'$. The second scenario corresponds to the cases $I_e e = b_{(d\to e)}(d \to e)_b$, $b_e e_b = b_{d\to e}(d \to e)_b$, and $bI_b = b_{(d\to e)}(d \to e)_b$. As before, in all these cases $d = I_d$, and therefore $d = I_d = I_a$, and $I_d = d = a$.
- $a_d d_a = a_d d_a$: Trivial.
- $a_d d_a = aI_a$: Same proof of the case $bI_b = b_e e_b$.
- $aI_a = aI_a$: Trivial.

This proves $\alpha = \alpha'$ and $\delta = \delta'$.

Lastly, we compute $\gamma\phi$ using Eq. (6.87) and Eq. (6.89).

$$\gamma\phi = \begin{cases} (b \to c)_f f_{(b\to c)} & \text{if } \mathrm{ord}(bc) < \mathrm{ord}(ef), \\ c_f f_c & \text{if } \mathrm{ord}(bc) = \mathrm{ord}(ef), \\ c_{(e\to f)}(e \to f)_c & \text{if } \mathrm{ord}(bc) > \mathrm{ord}(ef). \end{cases} \tag{6.97}$$

We compute $\gamma'\phi'$ using Eq. (6.90) and Eq. (6.91).

$$\gamma'\phi' = \begin{cases} (a \to c)_f f_{(a\to c)} & \text{if } \mathrm{ord}(ac) < \mathrm{ord}(df), \\ c_f f_c & \text{if } \mathrm{ord}(ac) = \mathrm{ord}(df), \\ c_{(d\to f)}(d \to f)_c & \text{if } \mathrm{ord}(ac) > \mathrm{ord}(df). \end{cases} \tag{6.98}$$

The possible cases are the following.

- $(b \to c)_f f_{(b\to c)} = (a \to c)_f f_{a\to c}$, which becomes $(b \to c)f = (a \to c)f$. In this case, $\mathrm{ord}(bc) < \mathrm{ord}(ef)$, which corresponds to the first three cases of the proof for $\beta = \beta'$, $\epsilon = \epsilon'$, and in those cases $b = I_b$. This also implies that $a = I_a$, otherwise the tree associated with $(a \to c)f$ would have labels that are different from those in the tree associated with $(I_b \to c)f$, which contradicts our hypothesis.

$$\begin{aligned} (b \to c)_f &= (I_b \to c)_f = c_{(I_b\to c)f} \\ &= c_{(I_a\to c)f} = (I_a \to c)_f \\ &= (a \to c)_f, \\ f_{(b\to c)} &= f_{(b\to c)f} = f_{(a\to c)f} \\ &= f_{(a\to c)}. \end{aligned} \tag{6.99}$$

This proves $\gamma = \gamma'$ and $\phi = \phi'$.

- $(b \to c)_f f_{(b\to c)} = c_f f_c$: Same proof of $b_e e_b = (a \to b)_e e_{(a\to b)}$.
- $(b \to c)_f f_{(b\to c)} = c_{(d\to f)}(d \to f)_c$: As in the case $(b \to c)_f f_{(b\to c)} = (a \to c)_f f_{a\to c}$, we have $b = I_b$, which implies $d = I_d$.

$$\begin{aligned}(b \to c)_f &= (I_b \to c)_f = c_{(I_b\to c)f} = c_{c(d\to f)}\\ &= c_{(d\to f)},\\ f_{(b\to c)} &= f_{(b\to c)f} = f_{c(d\to f)} = f_{(I_d\to f)c} = (I_d \to f)_c\\ &= (d \to f)_c.\end{aligned} \tag{6.100}$$

- $c_f f_c = (a \to c)_f f_{(a\to c)}$: Same proof of $b_e e_b = (a \to b)_e e_{a\to b}$.
- $c_f f_c = c_f f_c$: Trivial.
- $c_f f_c = c_{d\to f}(d \to f)_c$: Same proof of $b_e e_b = b_{d\to e}(d \to e)_b$.
- $c_{(e\to f)}(e \to f)_c = (a \to c)_f f_{(a\to c)}$. In this case, $\mathrm{ord}(bc) > \mathrm{ord}(ef)$, which corresponds to the last three cases of the proof for $\beta = \beta'$ and $\epsilon = \epsilon'$. All those cases $e = I_e$. As discussed multiple times, this implies $a = I_a$. From this point on, the proof is the same as the case $(b \to c)_f f_{(b\to c)} = c_{(d\to f)}(d \to f)_c$.
- $c_{(e\to f)}(e \to f)_c = c_f f_c$: Same proof of the case $b_e e_b = b_{d\to e}(d \to e)_b$.
- $c_{(e\to f)}(e \to f)_c = c_{(d\to f)}(d \to f)_c$. Once again, $e = I_e$, which implies $d = I_d$.

$$\begin{aligned}c_{(e\to f)} &= c_{c(e\to f)} = c_{c(d\to f)}\\ &= c_{d\to f},\\ (e \to f)_c &= (I_e \to f)_c = f_{(I_e\to f)c} = f_{c(I_e\to f)}\\ &= f_{c(I_d\to f)} = f_{(I_d\to f)c} = (I_d \to f)c\\ &= (d \to f)_c.\end{aligned} \tag{6.101}$$

This concludes the proof. □

We are now ready to prove that Eq. (6.80) and Eq. (6.81) whenever the types involved satisfy the consistency conditions.

Theorem 6.2.30: Sequential composition and elementary tensors

Let $a, b, c, d, e, f \in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^{a\to b} \in L(a \to b)$, $\Delta^{b\to c} \in L(b \to c)$, $\Theta^{d\to e} \in L(d \to e)$, and $\Xi^{e\to f} \in L(e \to f)$. If $(\Delta^{b\to c} \boxtimes \Xi^{e\to f})$ is composable with $(\Gamma^{a\to b} \boxtimes \Theta^{d\to e})$ and the type of $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e})$ is equal to the type of $(\Delta^{b\to c} \boxtimes \Xi^{e\to f}) \circ (\Gamma^{a\to b} \boxtimes \Theta^{d\to e})$, that is, if the hypotheses of Lemma 6.2.29 are true, then,

$$(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes (\Xi^{e\to f} \circ \Theta^{d\to e}) = (\Delta^{b\to c} \boxtimes \Xi^{e\to f}) \circ (\Gamma^{a\to b} \boxtimes \Theta^{d\to e}). \tag{6.102}$$

Proof. Let α, β, γ, δ, ϵ, and ϕ be defined as in Lemma 6.2.29. We define

$$\begin{aligned}
\tilde{\Gamma}^{\alpha\to\beta} &= \eta^{(a\to b)\to(a\to b)_{d\to e}}(\Gamma^{a\to b}),\\
\tilde{\Theta}^{\delta\to\epsilon} &= \eta^{(d\to e)\to(d\to e)_{(a\to b)}}(\Theta^{d\to e}),\\
\tilde{\Delta}^{\beta\to\gamma} &= \eta^{(b\to c)\to(b\to c)_{(e\to f)}}(\Delta^{b\to c}),\\
\tilde{\Xi}^{\epsilon\to\phi} &= \eta^{(e\to f)\to(e\to f)_{(b\to c)}}(\Xi^{e\to f}),\\
\widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} &= \eta^{(a\to c)\to(a\to c)_{(d\to f)}}(\Delta^{b\to c}\circ\Gamma^{a\to b}),\\
\widetilde{(\Xi\circ\Theta)}^{\delta\to\phi} &= \eta^{(d\to f)\to(d\to f)_{a\to c}}(\Xi^{e\to f}\circ\Theta^{d\to e}).
\end{aligned} \tag{6.103}$$

Note that these definitions are the same as Eq. (6.82) and Eq. (6.83), considering that we can drop the 'prime' symbol as a consequence of Lemma 6.2.29. As before, we have that

$$\begin{aligned}
&(\Delta^{b\to c}\circ\Gamma^{a\to b})\boxtimes(\Xi^{e\to f}\circ\Theta^{d\to e}) = \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma}\otimes\widetilde{(\Xi\circ\Theta)}^{\delta\to\phi},\\
&\Delta^{b\to c}\boxtimes\Xi^{e\to f} = \tilde{\Delta}^{\beta\to\gamma}\otimes\tilde{\Xi}^{\epsilon\to\phi},\\
&\Gamma^{a\to b}\boxtimes\Theta^{d\to e} = \tilde{\Gamma}^{\alpha\to\beta}\otimes\tilde{\Theta}^{\delta\to\epsilon}.
\end{aligned} \tag{6.104}$$

Moreover, note that by definition $\mathrm{struct}(\alpha) = \mathrm{struct}(\delta)$, $\mathrm{struct}(\beta) = \mathrm{struct}(\epsilon)$, and $\mathrm{struct}(\gamma) = \mathrm{struct}(\phi)$. To show that Eq. (6.102) holds, it is enough to show that it holds for elementary products $u^\alpha\boxtimes v^\delta$. The left-hand side gives

$$\begin{aligned}
&(\Delta^{b\to c}\circ\Gamma^{a\to b})\boxtimes(\Xi^{e\to f}\circ\Theta^{d\to e})@u^\alpha\boxtimes v^\delta\\
&\quad= \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma}\otimes\widetilde{(\Xi\circ\Theta)}^{\delta\to\phi}@u^\alpha\otimes v^\delta\\
&\quad= (\widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma}@u^\alpha)\otimes(\widetilde{(\Xi\circ\Theta)}^{\delta\to\phi}@v^\delta).
\end{aligned} \tag{6.105}$$

The right-hand side gives

$$\begin{aligned}
&(\Delta^{b\to c}\boxtimes\Xi^{e\to f})\circ(\Gamma^{a\to b}\boxtimes\Theta^{d\to e})@u^\alpha\boxtimes v^\delta\\
&\quad= (\tilde{\Delta}^{\beta\to\gamma}\otimes\tilde{\Xi}^{\epsilon\to\phi})\circ(\tilde{\Gamma}^{\alpha\to\beta}\otimes\tilde{\Theta}^{\delta\to\epsilon})@u^\alpha\otimes v^\delta\\
&\quad= \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta}(u^\alpha)\otimes\tilde{\Xi}^{\epsilon\to\phi}\circ\tilde{\Theta}^{\delta\to\epsilon}(v^\delta).
\end{aligned} \tag{6.106}$$

Therefore, it suffices to show that $\widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} = \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta}$ and $\widetilde{(\Xi\circ\Theta)}^{\delta\to\phi} = \tilde{\Xi}^{\epsilon\to\phi}\circ\tilde{\Theta}^{\delta\to\epsilon}$. We prove the former; the proof of the latter is identical.

We start with showing that $\tilde{\Gamma}^{\alpha\to\beta} = \eta^{(a\to b)\to(a\to b)_{\alpha\to\beta}}(\Gamma^{a\to b})$. Indeed, $(a\to b)_{d\to e} = (a\to b)_{(a\to b)(d\to e)} = (a\to b)_{(a\to b)(d\to e)} = (a\to b)_{\alpha\to\beta}$. Therefore, $\tilde{\Gamma}^{\alpha\to\beta} = \eta^{(a\to b)\to(a\to b)_{(d\to e)}}(\Gamma^{a\to b}) = \eta^{(a\to b)\to(a\to b)_{(\alpha\to\beta)}}(\Gamma^{a\to b})$. Similarly, $\tilde{\Delta}^{\beta\to\gamma} = \eta^{(b\to c)\to(b\to c)_{\beta\to\gamma}}$ and $\widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} = \eta^{(a\to c)\to(a\to c)_{\alpha\to\gamma}}(\Delta^{b\to c}\circ\Gamma^{a\to b})$.

Therefore, since $\mathrm{ord}(ab)\le\mathrm{ord}(\alpha\beta)$ we have,

$$\tilde{\Gamma}^{\alpha\to\beta} = \begin{cases}\eta^{b\to b_\beta}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} & \text{if } \mathrm{ord}(ab) = \mathrm{ord}(\alpha\beta)\\ \eta^{(a\to b)\to(a\to b)_\beta}\circ\left|\Gamma^{a\to b}\right\rangle^{I\to(a\to b)}\circ\eta^{I_\alpha\to I} & \text{if } \mathrm{ord}(ab) < \mathrm{ord}(\alpha\beta)\end{cases} \tag{6.107}$$

Observe that, $\mathrm{ord}(ab) < \mathrm{ord}(\alpha\beta)$ if and only if $\mathrm{ord}(ab) < \mathrm{ord}(de)$. Tracing this scenario back to the corresponding scenarios in the proof of Lemma 6.2.29 (see discussion for the case $I_d d = anI_a$), we have that

$a = I_a$, and therefore $(a \to b) = b_{a\to b}$. This implies that

$$\begin{aligned}
&\eta^{(a\to b)\to(a\to b)_\beta} \circ \left|\Gamma^{a\to b}\right\rangle^{I\to(a\to b)} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b_{a\to b}\to b_{(a\to b)\beta}} \circ \left|\Gamma^{b_{a\to b}}\right\rangle^{I\to b_(a\to b)} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b_{a\to b}\to b_{(a\to b)\beta}} \circ \eta^{b\to b_{a\to b}} \circ \eta^{b_{a\to b}\to b} \circ \left|\Gamma^{b_{a\to b}}\right\rangle^{I\to b_(a\to b)} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b\to b_{(a\to b)\beta}} \circ \left|\eta^{b_{a\to b}\to b}(\Gamma^{b_{a\to b}})\right\rangle^{I\to b} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b\to b_\beta} \circ \left|\eta^{b_{a\to b}\to b}(\Gamma^{b_{a\to b}})\right\rangle^{I\to b} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b\to b_\beta} \circ \left|\Gamma^{I_a\to b} \circ \eta^{I\to I_a}(1)\right\rangle^{I\to b} \circ \eta^{I_\alpha\to I} \\
&\quad= \eta^{b\to b_\beta} \circ \Gamma^{I_a\to b} \circ \eta^{I\to I_a} \circ \eta^{I_\alpha\to I}.
\end{aligned} \tag{6.108}$$

Therefore,

$$\tilde{\Gamma}^{\alpha\to\beta} = \begin{cases} \eta^{b\to b_\beta} \circ \Gamma^{a\to b} \circ \eta^{a_\alpha\to a} & \text{if } \operatorname{ord}(ab) = \operatorname{ord}(\alpha\beta) \\ \eta^{b\to b_\beta} \circ \Gamma^{I_a\to b} \circ \eta^{I\to I_a} \circ \eta^{I_\alpha\to I} & \text{if } \operatorname{ord}(ab) < \operatorname{ord}(\alpha\beta) \end{cases} \tag{6.109}$$

The expression for $\tilde{\Delta}^{\beta\to\gamma}$ is

$$\tilde{\Delta}^{\beta\to\gamma} = \begin{cases} \eta^{c\to c_\gamma} \circ \Delta^{b\to c} \circ \eta^{b_\beta\to b} & \text{if } \operatorname{ord}(bc) = \operatorname{ord}(\beta\gamma) \\ \eta^{(b\to c)\to(b\to c)_\gamma} \circ \left|\Delta^{b\to c}\right\rangle^{I\to(b\to c)} \circ \eta^{I_\beta\to I} & \text{if } \operatorname{ord}(bc) < \operatorname{ord}(\beta\gamma) \end{cases} \tag{6.110}$$

Here, the proof of Lemma 6.2.29, implies that $b = I_b$ and $\beta = I_\beta$ if $\operatorname{ord}(bc) < \operatorname{ord}(\beta\gamma)$ (equivalent to $\operatorname{ord}(bc) < \operatorname{ord}(ef)$ which corresponds to the first three cases of the proof $\beta = \beta'$ and $\epsilon = \epsilon'$). Just by changing the letters in Eq. (6.108), one obtains

$$\eta^{(b\to c)\to(b\to c)_\gamma} \circ \left|\Delta^{b\to c}\right\rangle^{I\to(b\to c)} \circ \eta^{I_\beta\to I} = \eta^{c\to c_\gamma} \circ \Delta^{I_b\to c} \circ \eta^{I\to I_b} \circ \eta^{I_\beta\to I}. \tag{6.111}$$

Moreover, in the three cases relevant to the scenario, one is impossible, in another case, $b \prec \beta$ (follows straightforwardly from $b \prec e$ and $\beta = I_e$), and similarly, in the last $(a \to b) \prec \beta$. Note that, in this last case we have that $\beta_b = \beta_{(a\to b)b} = \beta \operatorname{triv}(a \to b) \operatorname{triv}(b) = \beta \operatorname{triv}(a \to b) = \beta_{a\to b} = \beta$, therefore $b \prec \beta$ in all the possible scenarios.

Now, considering that $\beta = I_\beta$, $b = I_b$ and $b \prec \beta$, we get

$$\begin{aligned}
&\eta^{(b\to c)\to(b\to c)_\gamma} \circ \left|\Delta^{b\to c}\right\rangle^{I\to(b\to c)} \circ \eta^{I_\beta\to I} \\
&\quad= \eta^{c\to c_\gamma} \circ \Delta^{I_b\to c} \circ \eta^{I_{b\beta}\to I_b} \circ \eta^{I_b\to I_{b\beta}} \circ \eta^{I\to I_b} \circ \eta^{I_\beta\to I} \\
&\quad= \eta^{c\to c_\gamma} \circ \Delta^{I_b\to c} \circ \eta^{I_{b\beta}\to I_b} \circ \eta^{I\to I_{b\beta}} \circ \eta^{I_b\beta\to I} \\
&\quad= \eta^{c\to c_\gamma} \circ \Delta^{I_b\to c} \circ \eta^{I_{b\beta}\to I_b} \\
&\quad= \eta^{c\to c_\gamma} \circ \Delta^{b\to c} \circ \eta^{b_\beta\to b}.
\end{aligned} \tag{6.112}$$

Thus,

$$\tilde{\Delta}^{\beta\to\gamma} = \eta^{c\to c_\gamma} \circ \Delta^{b\to c} \circ \eta^{b_\beta\to b} \tag{6.113}$$

regardless of the relative order of types.

Lastly, one obtains

$$\widetilde{\Delta\circ\Gamma}^{\alpha\to\gamma} = \begin{cases} \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} & \text{if } \operatorname{ord}(ac) = \operatorname{ord}(\alpha\gamma) \\ \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} & \text{if } \operatorname{ord}(ac) < \operatorname{ord}(\alpha\gamma) \end{cases} \tag{6.114}$$

by noticing that in the case $\operatorname{ord}(ac) < \operatorname{ord}(\alpha\gamma)$ one gets $I_a = a$ and following all the steps of Eq. (6.108).

This leaves us with four different cases.

- $\operatorname{ord}(ab) = \operatorname{ord}(\alpha\beta)$ and $\operatorname{ord}(ac) = \operatorname{ord}(\alpha\gamma)$:

$$\begin{aligned} \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\eta^{b_\beta\to b}\circ\eta^{b\to b_\beta}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta} \end{aligned} \tag{6.115}$$

- $\operatorname{ord}(ab) < \operatorname{ord}(\alpha\beta)$ and $\operatorname{ord}(ac) < \operatorname{ord}(\alpha\gamma)$:

$$\begin{aligned} \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} \\ &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\eta^{b_\beta\to b}\circ\eta^{b\to b_\beta}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta}. \end{aligned} \tag{6.116}$$

- $\operatorname{ord}(ab) < \operatorname{ord}(\alpha\beta)$ and $\operatorname{ord}(ac) = \operatorname{ord}(\alpha\gamma)$: We have already discussed that if $\operatorname{ord}(ab) < \operatorname{ord}(\alpha\beta)$ then $a = I_a$. Moreover, from Eq. (6.95) and the first three cases of the proof of $\alpha = \alpha'$, $\delta = \delta'$ in Lemma 6.2.29, it immediately follows that $a \prec \alpha$. With that, we have:

$$\begin{aligned} \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\eta^{b_\beta\to b}\circ\eta^{b\to b_\beta}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\eta^{b\to b_\beta}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\eta^{b\to b_\beta}\circ\Gamma^{I_a\to b}\circ\eta^{I_{a\alpha}\to I_a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\eta^{b\to b_\beta}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_a\to I}\circ\eta^{I_{a\alpha}\to I_a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\eta^{b\to b_\beta}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_{a\alpha}\to I} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\eta^{b\to b_\beta}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta}. \end{aligned} \tag{6.117}$$

- $\operatorname{ord}(ab) = \operatorname{ord}(\alpha\beta)$ and $\operatorname{ord}(ac) < \operatorname{ord}(\alpha\gamma)$: In this case too, $a = I_a$ and $a \prec \alpha$ (equivalent to the previous case, compare Eq. (6.95) and Eq. (6.96)). Therefore, using the same tricks to prove $\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} = \eta^{a_\alpha\to a}$, one obtains

$$\begin{aligned} \widetilde{(\Delta\circ\Gamma)}^{\alpha\to\gamma} &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{I_a\to b}\circ\eta^{I\to I_a}\circ\eta^{I_\alpha\to I} \\ &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \eta^{c\to c_\gamma}\circ\Delta^{b\to c}\circ\eta^{b_\beta\to b}\circ\eta^{b\to b_\beta}\circ\Gamma^{a\to b}\circ\eta^{a_\alpha\to a} \\ &= \tilde{\Delta}^{\beta\to\gamma}\circ\tilde{\Gamma}^{\alpha\to\beta}. \end{aligned} \tag{6.118}$$

All the cases have been considered; thus, the proof is completed. □

Next, we show that analogously to the tensor product, the square product is bilinear.

Proposition 6.2.31: Bilinearity

The product $\boxtimes$ is bilinear.

Proof. Observe that for every $x, y \in \mathrm{Types}_{\mathcal{A}}$ and $\Gamma^x \in L(x)$, $\Delta^y \in L(y)$ we have that $\Gamma^x \boxtimes \Delta^y = \eta^{x\to x_y}(\Gamma^x) \otimes \eta^{y \to y_x}(\Delta^y)$. Since the promotion functions are linear and $\otimes$ is bilinear, it immediately follows that $\boxtimes$ is bilinear. □

We conclude this Section with two handy properties of the square product. The first is that the promotion of a function, $\eta^{x\to x_y}(\Gamma^x)$, can be expressed as the square product of the function with the promotion of 1.

Lemma 6.2.32: Parallel composition with promotions of 1

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^x \in L(x)$. Then, $\eta^{x\to x_y}(\Gamma^x) = \Gamma^x \boxtimes \eta^{I \to I_y}(1)$.

Proof.

$$\begin{aligned}\eta^{x\to x_y}(\Gamma^x) &= \eta^{x\to x_y}(\Gamma^x) \boxtimes 1\\ &= \eta^{x\to x_y}(\Gamma^x) \boxtimes \eta^{I\to I_{xy}}(1)\\ &= \Gamma^x \boxtimes \eta^{I_y \to I_{xy}} \circ \eta^{I\to I_y}(1)\\ &= \Gamma^x \boxtimes \eta^{I\to I_y}(1).\end{aligned} \tag{6.119}$$

□

The second is that in the case of functions from or to the trivial space $L(I)$, one can switch the tensor product with sequential composition.

Lemma 6.2.33: Parallel composition of maps from with I as input or output

Let $a, b \in \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{ord}(a) = \mathrm{ord}(b)$. Let $\Gamma^{a\to I} \in L(a \to I)$, and $\Delta^{I\to b} \in L(I \to b)$. Then, $\Gamma^{a\to I} \boxtimes \Delta^{I\to b} = \Delta^{I\to b} \circ \Gamma^{a\to I} = \Delta^{I\to b} \boxtimes \Gamma^{a\to I}$.

Proof. All three maps have type $(a \to b)$. We show that their action coincide for every $u^a \in L(a)$:

$$\begin{aligned}&\Delta^{I\to b} \circ \Gamma^{a\to I}(u^a) = \Delta^{I\to b}(\Gamma^{a\to I}(u^a)) = \Delta^{I\to b}(1)\Gamma^{a\to I}(u^a),\\ &(\Delta^{I\to b} \boxtimes \Gamma^{a\to I})(u^a) = (\Delta^{I\to b} \boxtimes \Gamma^{a\to I})(1 \boxtimes u^a)\\ &\qquad = \Delta^{I\to b}(1) \boxtimes \Gamma^{a\to I}(u^a) = \Delta^{I\to b}(1)\Gamma^{a\to I}(u^a),\\ &(\Gamma^{a\to I} \boxtimes \Delta^{I\to b})(u^a) = (\Gamma^{a\to I} \boxtimes \Delta^{I\to b})(u^a \boxtimes 1)\\ &\qquad = \Gamma^{a\to I}(u^a) \boxtimes \Delta^{I\to b}(1) = \Delta^{I\to b}(1)\Gamma^{a\to I}(u^a).\end{aligned} \tag{6.120}$$

□

6.2.4 Swap Operators

An important operator in product spaces is the swap operator, that is, the operator that swaps the order of elements in an elementary tensor: $u \otimes v \to v \otimes u$. In this subsection, we present the maps that play the same role in the case of the square product.

Definition 6.2.34: Swap operators

Let $x, y \in \text{Types}_{\mathcal{A}}$. Then, $\sigma^{xy\to yx}$ is the map defined on elementary tensor $u^x \boxtimes v^y \in L(xy)$ as $\sigma^{xy\to yx}(u^x \boxtimes v^y) = v^y \boxtimes u^x$ and extended by linearity to all elements in $L(xy)$.

We first observe that this map is well-defined. Let $\left\{ u^x_i \boxtimes v^y_\alpha \right\}_{i,\alpha}$ and $\left\{ \tilde{u}^x_i \boxtimes \tilde{v}^y_\alpha \right\}_{i,\alpha}$ be two basis for $L(xy)$. Let U_{ij} and $V_{\alpha\beta}$ be the matrices for the change of basis. That is, $\tilde{u}^x_i = \sum_j U_{ij} u^x_j$ and $\tilde{v}^y_\alpha = \sum_\beta V_{\alpha\beta} v^y_\beta$. Let $\Gamma^{xy} = \sum_{i\alpha} G_{i\alpha} u^x_i \boxtimes v^y_\alpha = \sum_{i\alpha} \tilde{G}_{i\alpha} \tilde{u}^x_i \boxtimes \tilde{v}^y_\alpha$. Observe that

$$\begin{aligned}\sum_{i\alpha} \tilde{G}_{i\alpha} \tilde{u}^x_i \boxtimes \tilde{v}^y_\alpha &= \sum_{i\alpha j\beta} \tilde{G}_{i\alpha} U_{ij} u^x_j \boxtimes V_{\alpha\beta} v^y_\beta \\ &= \sum_{j\beta} (\sum_{i\alpha} \tilde{G}_{i\alpha} U_{ij} V_{\alpha\beta}) u^x_j \boxtimes v^y_\beta .\end{aligned} \tag{6.121}$$

Therefore, $G_{j\beta} = \sum_{i\alpha} \tilde{G}_{i\alpha} U_{ij} V_{\alpha\beta}$. If we compute the output of the swap map for the two different decompositions of Γ^{xy}, we obtain,

$$\begin{aligned}\sigma^{x,y}(\sum_{i\alpha} \tilde{G}_{i\alpha} \tilde{u}^x_i \boxtimes \tilde{v}^y_\alpha) &= \sum_{i\alpha} \tilde{G}_{i\alpha} \tilde{v}^y_\alpha \boxtimes \tilde{u}^x_i \\ \sigma^{x,y}(\sum_{i\alpha} G_{i\alpha} u^x_i \boxtimes v^y_\alpha) &= \sum_{i\alpha} G_{i\alpha} v^y_\alpha \boxtimes u^x_i \\ &= \sum_{i\alpha} \sum_{j\beta} \tilde{G}_{j\beta} U_{ji} V_{\beta\alpha} v^y_\alpha \boxtimes u^x_i \\ &= \sum_{j\beta} \tilde{G}_{j\beta} \sum_{\alpha} V_{\beta\alpha} v^y_\alpha \boxtimes \sum_i U_{ji} u^x_i \\ &= \sum_{j\beta} \tilde{G}_{j\beta} \tilde{v}^y_\beta \boxtimes \tilde{u}^x_j .\end{aligned} \tag{6.122}$$

In both cases, the outputs are the same; therefore, $\sigma^{xy\to yx}$ is well defined.

It follows immediately from the definition of the swap operators that swapping twice an elementary tensor is the same as doing nothing.

Proposition 6.2.35: Inverse of swap operators

Let $x, y \in \text{Types}_{\mathcal{A}}$. Then, $\sigma^{yx\to xy}$ is the inverse of $\sigma^{xy\to yx}$.

Since we prove most of our results by induction, it is helpful to have a recursive definition of the swap operator. This is the content of the following Lemma.

Lemma 6.2.36: Inductive definition of the swap operator

Let $x, y \in \mathrm{Types}_{\mathfrak{A}}$ be such that $\mathrm{ord}(xy) > 0$, let $x = (a \to b)$ whenever $\mathrm{ord}(x) > 0$, $y = (c \to d)$ whenever $\mathrm{ord}(y) > 0$, and let $\Gamma^{xy} \in L(xy)$.

1. If $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $\sigma^{xy \to yx}(\Gamma^{xy}) = \sigma^{bd \to db} \circ \Gamma^{ac \to bd} \circ \sigma^{ca \to ac}$.
2. If $\mathrm{ord}(x) < \mathrm{ord}(y)$, then $\sigma^{xy \to yx}(\Gamma^{xy}) = \sigma^{xd \to dx} \circ \Gamma^{c \to xd}$.
3. If $\mathrm{ord}(x) > \mathrm{ord}(y)$, then $\sigma^{xy \to yx}(\Gamma^{xy}) = \sigma^{by \to yb} \circ \Gamma^{a \to by}$.

Proof. For the proof, it is enough to show that the expressions above coincide on the elementary tensor $\Gamma^x \boxtimes \Delta^y \in L(xy)$. If so, then the two functions are the same by linearity.

1. Let $v^c \boxtimes u^a \in L(ca)$.

$$\begin{aligned}\sigma^{xy \to yx}(\Gamma^x \boxtimes \Delta^y) @ v^c \boxtimes u^a &= \Delta^{c \to d} \boxtimes \Gamma^{a \to b} @ v^c \boxtimes u^a \\ &= \Delta^{c \to d}(v^c) \boxtimes \Gamma^{a \to b}(u^a).\end{aligned} \tag{6.123}$$

   Similarly,

$$\begin{aligned}&\sigma^{bd \to db} \circ (\Gamma^x \boxtimes \Delta^y) \circ \sigma^{ca \to ac} @ v^c \boxtimes u^a \\ &= \sigma^{bd \to db} \circ (\Gamma^{a \to b} \boxtimes \Delta^{c \to d}) @ u^a \boxtimes v^c \\ &= \sigma^{bd \to db} @ (\Gamma^{a \to b}(u^a) \boxtimes \Delta^{c \to d}(v^c)) \\ &= \Delta^{c \to d}(v^c) \boxtimes \Gamma^{a \to b}(u^a).\end{aligned} \tag{6.124}$$

2. Let $v^c \in L(c)$.

$$\begin{aligned}\sigma^{xy \to yx}(\Gamma^x \boxtimes \Delta^y) @ v^c &= \Delta^{c \to d} \boxtimes \Gamma^x @ v^c \\ &= \Delta^{c \to d}(v^c) \boxtimes \Gamma^x.\end{aligned} \tag{6.125}$$

   Similarly,

$$\begin{aligned}\sigma^{xd \to dx} \circ (\Gamma^x \boxtimes \Delta^y) @ v^c &= \sigma^{xd \to dx}(\Gamma^x \boxtimes \Delta^{c \to d}(v^c)) \\ &= \Delta^{c \to d}(v^c) \boxtimes \Gamma^x.\end{aligned} \tag{6.126}$$

3. Analogous to the previous case.

□

From the Lemma above, one deduces the following equalities for the swap of elementary tensors.

Corollary 6.2.37

Let $x, y \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{ord}(xy) > 0$, let $x = (a \to b)$ whenever $\mathrm{ord}(x) > 0$, $y = (c \to d)$ whenever $\mathrm{ord}(y) > 0$, and let $\Gamma^x \in L(x)$, $\Delta^y \in L(y)$.

1. If $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $\sigma^{bd \to db} \circ \Gamma^x \boxtimes \Delta^y = \Delta^y \boxtimes \Gamma^x \circ \sigma^{ac \to ca}$.
2. If $\mathrm{ord}(x) < \mathrm{ord}(y)$, then $\sigma^{xd \to dx} \circ \Gamma^x \boxtimes \Delta^y = \Delta^y \boxtimes \Gamma^x$.
3. If $\mathrm{ord}(x) > \mathrm{ord}(y)$, then $\sigma^{by \to yb} \circ \Gamma^x \boxtimes \Delta^y = \Delta^y \boxtimes \Gamma^x$.

Moreover, the three cases of Lemma 6.2.36 can be merged into a single abstract expression.

Corollary 6.2.38

Let $x, y \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{ord}(xy) > 0$, let $x = (a \to b)$ whenever $\mathrm{ord}(x) > 0$, $y = (c \to d)$ whenever $\mathrm{ord}(y) > 0$, and let $\Gamma^{xy} \in L(xy)$. Then, $\sigma^{xy \to yx} = \sigma^{\beta\delta \to \delta\beta} \circ \Gamma^{\alpha\gamma \to \beta\delta} \circ \sigma^{\gamma\alpha \to \alpha\gamma}$, where

$$\alpha = \begin{cases} a & \text{if } \mathrm{ord}(x) \geq \mathrm{ord}(y) \\ I & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y) \end{cases} \qquad \beta = \begin{cases} b & \text{if } \mathrm{ord}(x) \geq \mathrm{ord}(y) \\ x & \text{if } \mathrm{ord}(x) < \mathrm{ord}(y) \end{cases}$$
$$\gamma = \begin{cases} c & \text{if } \mathrm{ord}(x) \leq \mathrm{ord}(y) \\ I & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y) \end{cases} \qquad \delta = \begin{cases} d & \text{if } \mathrm{ord}(x) \leq \mathrm{ord}(y) \\ y & \text{if } \mathrm{ord}(x) > \mathrm{ord}(y) \end{cases} \tag{6.127}$$

Proof. The only thing to show is that for any type z, $\sigma^{zI \to Iz} = \mathrm{id}^{z \to z} = \sigma^{Iz \to zI}$. This is trivial: $\sigma^{zI \to Iz}(1 \boxtimes u^z) = u^z \boxtimes 1 = u^z = 1 \boxtimes u^z$ for all $u^z \in L(z)$. This shows that $\mathrm{id}^{z \to z} = \sigma^{Iz \to zI}$. The proof of $\sigma^{zI \to Iz} = \mathrm{id}^{z \to z}$ is analogous. With this property at hand, one can precompose the functions in the second and third case of Lemma 6.2.36 with the proper swap operator and obtain the desired result. □

6.2.5 Hermitian Preserving Maps

In this and the following Subsection, we describe two essential subsets of $L(x)$, which are the generalizations of the set of Hermitian-preserving channels and completely positive-preserving channels. These two subsets are very relevant in quantum mechanics. Indeed, quantum states are Hermitian and positive semidefinite. Quantum channels are Hermitian-preserving and completely positive-preserving. Quantum superchannels preserve Hermitian-preserving maps and completely preserve completely positive-preserving maps. This kind of pattern is well-suited for a recursive definition, which is the bread and butter of our type system. In this Subsection, we focus on the generalization of Hermitian-preserving, while in the following Subsection, we focus on the generalization of completely positive-preserving.

Definition 6.2.39: Hermitian preserving

Let $x \in \mathrm{Types}_{\mathcal{A}}$. The set of Hermitian-preserving maps of type x, denoted as $H(x)$ is

- The set of Hermitian maps on the Hilbert space $\mathcal{H}_E$, if $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$.
- The set $H(x) = \left\{ \Gamma^{a\to b} \mid \Gamma^{a\to b}(u^a) \in H(b), \forall u^a \in H(a) \right\}$, if $x = (a \to b)$, and $a, b \in \mathrm{Types}_{\mathcal{A}}$.

From the definition, it follows that $H(x)$ is a real vector space for all $x \in \mathrm{Types}_{\mathcal{A}}$. Indeed, this is the case for elementary types, and the recursive structure of the definition guarantees that $H(x)$ is a real vector space if both $H(a)$ and $H(b)$ are, when $x = (a \to b)$. Moreover, by induction, one shows that the (real) dimension of $H(x)$ is simply the product of the square of the elementary types labelling the leaves of the tree associated with x. That is $|H(x)| = \prod_i |E_i|^2$ (compare with Lemma 6.2.2).

Of more interest is the fact that the complex inner product defined on $L(x)$ becomes a real inner product on $H(x)$, as in the case of the Hilbert-Schmidt inner product of Hermitian matrices.

Proposition 6.2.40: Inner product on $H(x)$

Let $x \in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^x, \Delta^x \in H(x)$, then $(\Gamma^x, \Delta^x)_x \in \mathbb{R}$, where $(\cdot,\cdot)_x$ is the inner product defined in Definition 6.2.3. Moreover, there is an orthonormal basis of $L(x)$ that is a basis of $H(x)$.

Proof. This is the first instance of an inductive proof where it is easier to prove two statements simultaneously because we can use the inductive hypothesis of both statements in the inductive step. We proceed by induction on the order of types, as usual. The statements are true for elementary types. These are well-known results about Hermitian matrices. For the inductive step, let $x \in \mathrm{Types}_{\mathcal{A}}$ be a type such that $\mathrm{ord}(x) > 0$, and let us assume that the statements are true for all types up to order $\mathrm{ord}(x) - 1$. Let $a, b \in \mathrm{Types}_{\mathcal{A}}$ be such that $x = (a \to b)$, and let $\left\{ u_i^a \right\}_i$ be the orthonormal basis of $L(a)$ with elements in $H(a)$. Starting with the definition of the inner product, we have

$$(\Gamma^x, \Delta^x)_x = \sum_i (\Delta^x(u_i^a), \Gamma^x(u_i^a))_b. \tag{6.128}$$

Now, since $\Gamma^x, \Delta^x \in H(x)$ and $u_i^a \in H(a)$ for all i, we have that $\Delta^x(u_i^a), \Gamma^x(u_i^a) \in H(b)$ for all i. By inductive hypothesis, we obtain $(\Delta^x(u_i^a), \Gamma^x(u_i^a))_b \in \mathbb{R}$ for all i, from which immediately follows $(\Gamma^x, \Delta^x)_x \in \mathbb{R}$.

We now prove that there exists an orthonormal basis for $L(x)$ that is contained in $H(x)$. Let $\left\{ u_i^a \right\}_i$ and $\left\{ v_j^b \right\}_j$ be the basis for $L(a)$ and $L(b)$ that are also basis for $H(a)$ and $H(b)$, respectively. Then, $\left\{ \left|v_j^b\right\rangle^{I\to b} \circ \left\langle u_i^a\right|^{a\to I} \right\}_{i,j}$ is an orthonormal basis for $L(x)$. We show that $\left|v_j^b\right\rangle^{I\to b} \circ \left\langle u_i^a\right|^{a\to I} \in H(x)$ for all i, j. Let $u^a \in H(a)$, then

$$\left|v_j^b\right\rangle^{I\to b} \circ \left\langle u_i^a\right|^{a\to I} @ u^a = \left|v_j^b\right\rangle^{I\to b} @ (u^a, u_i^a)_a = v_j^b (u^a, u_i^a)_a. \tag{6.129}$$

By inductive hypothesis, $(u^a, u_i^a)_a \in \mathbb{R}$, which implies $v_j^b(u^a, u_i^a)_a \in H(b)$. This is true for all $u^a \in H(a)$, therefore $\left|v_j^b\right\rangle^{I\to b} \circ \left\langle u_i^a\right|^{a\to I} \in H(x)$ for all i, j. Since the elements of $\left\{ \left|v_j^b\right\rangle^{I\to b} \circ \left\langle u_i^a\right|^{a\to I} \right\}_{i,j}$ are linearly independent in the complex field, they are linearly independent in the real field. Moreover, since the (complex) dimension of $L(x)$ is equal to the (real) dimension of $H(x)$, there are enough elements in $\left\{ \left| v_j^b \right\rangle^{I\to b} \circ \left\langle u_i^a \right|^{a\to I} \right\}_{i,j}$ to span $H(x)$.

$\square$

In other words, this Proposition shows that $H(x)$ is a real Hilbert space with $(\cdot,\cdot)_x$ as the inner product.

Next on the list is to understand what happens when a H-preserving map $\Gamma^x \in H(x)$ is promoted to a map Γ^{x_y}. In the next Proposition, we show that, as one would expect, promotion and demotion maps are H-preserving, which means that the promoted map Γ^{x_y} is in $H(x_y)$.

Proposition 6.2.41: Promotion and demotion maps are H-preserving

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. Then, $\eta^{x\to x_y} \in H(x \to x_y)$ and $\eta^{x_y\to x} \in H(x_y \to x)$.

Proof. In the base step, the promotion and demotion maps are identities; therefore, they are H-preserving. For the inductive step, it is enough to show that the ket map is H-preserving. Indeed, promotion and demotion maps are the sequential composition of promotion and demotion of lower order with the ket function, and the sequential composition of H-preserving maps is H-preserving. Now, for arbitrary x, $|\cdot\rangle^{I\to x}$ maps a map Γ^x to the map $|\Gamma^x\rangle^{I\to x} \in L(I \to x)$ defined as $|\Gamma^x\rangle^{I\to x}(1) = \Gamma(x)$. Now, assume that $\Gamma^x \in H(x)$. We have to prove that $|\Gamma^x\rangle^{I\to x} \in H(I \to x)$. Note that $H(I) = \mathbb{R}$. Let $a \in \mathbb{R}$, then $|\Gamma^x\rangle^{I\to x}(a) = a\Gamma^x$. Since $H(x)$ is a real vector space, $a\Gamma^x \in H(x)$, and therefore $|\Gamma^x\rangle^{I\to x} \in H(I \to x)$. This shows that $|\cdot\rangle^{I\to x}$ is H-preserving. $\square$

Since the promotion and demotion maps are one the inverse of the other and $\eta^{x\to x_y}(H(x)) \subseteq H(x_y)$ and $\eta^{x_y\to x}(H(x_y)) \subseteq H(x)$, we have that H_x and $H(x_y)$ are isomorphic.

Corollary 6.2.42: Isomorphism of $H(x)$ and $H(x_y)$

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. Then, $H(x)$ is isomorphic to $H(x_y)$.

We have discussed promotion and demotion operators, and it is trivial that the sequential composition of H-preserving functions is H-preserving. What is left to discuss is the square product of H-preserving functions. With the next Theorem, we show that the product of H-preserving functions is H-preserving.

Theorem 6.2.43: Product of H-preserving maps

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, and let $H(x) \boxtimes H(y) := \mathrm{Span}\{\, \Gamma^x \boxtimes \Delta^y \mid \Gamma^x \in H(x), \Delta^y \in H(y) \,\}$. Then, $H(xy) = H(x) \boxtimes H(y)$.

Proof. We prove it by induction on $\mathrm{ord}(xy)$. If $\mathrm{ord}(xy) = 0$, then both $x = E$ and $y = G$ are elementary types, and the tensor product of the space of Hermitian matrices on a Hilbert space E with the space of Hermitian matrices on G is the vector space of Hermitian matrices on EG. We assume now that $\mathrm{ord}(xy) > 0$ and that the statement is true for all types up to order $\mathrm{ord}(xy) - 1$. We first show that $H(x) \boxtimes H(y) \subseteq H(xy)$.

By definition, if $\left\{ \Gamma_i^x \right\}_i$ and $\left\{ \Delta_j^y \right\}_j$ are bases for $H(x)$ and $H(y)$, respectively, then $\left\{ \Gamma_i^x \boxtimes \Delta_j^y \right\}_{i,j}$ is a basis for $H(x) \boxtimes H(y)$. We show that $\Gamma_i^x \boxtimes \Delta_j^y \in H(xy)$ for arbitrary i, j. We split the proof into three cases.

- $\mathrm{ord}(x) = \mathrm{ord}(y)$, and $x = (a \to b)$, $y = (c \to d)$. Then, $\Gamma_i^x \boxtimes \Delta_j^y \in H(ac \to bd)$. By induction hypothesis, any element of $H(ac)$ can be written as a linear combination of elementary tensors $u^a \boxtimes v^b$, with $u^a \in H(a)$ and $v^c \in H(c)$. The action of $\Gamma_i^x \boxtimes \Delta_j^y$ on elementary tensors gives $\Gamma_i^x(u^a) \boxtimes \Delta_j^y(v^c) \in H(b) \boxtimes H(d) = H(bd)$, where the last equality follows by induction hypothesis. Therefore, by linearity, $\Gamma_i^x \boxtimes \Delta_j^y @ H_{ac} \subseteq H(bd)$. This proves, that $\Gamma_i^x \boxtimes \Delta_j^y \in H(xy)$.
- $\mathrm{ord}(x) > \mathrm{ord}(y)$, and $x = (a \to b)$. Then, $\Gamma_i^x \boxtimes \Delta_j^y \in H(a \to by)$. Let $u^a \in H(a)$. By Theorem 6.2.27, we have that $\Gamma_i^x \boxtimes \Delta_j^y @ u^a = \Gamma_i^x(u^a) \boxtimes \Delta_j^y \in H(b) \boxtimes H(y) = H(by)$. As in the previous case, this implies that $\Gamma_i^x \boxtimes \Delta_j^y \in H(xy)$.
- $\mathrm{ord}(x) < \mathrm{ord}(y)$: Symmetric to the previous case.

We have shown that a basis of $H(x) \boxtimes H(y)$ is contained in $H(xy)$. Since these two real vector spaces have the same dimension, $\left\{ \Gamma_i^x \boxtimes \Delta_j^y \right\}_{i,j}$ is a basis of $H(xy)$, and therefore the two vector spaces are equal. □

Thanks to this Theorem, it is easy to see that, as in the case of quantum channels, there is no need to require that a map is completely Hermitian-preserving.

Corollary 6.2.44: Completely H-preserving maps

Let $x \in \mathrm{Types}_{\mathcal{A}}$. Then $\Gamma^x \in H(x)$ if and only if $\Gamma^x \boxtimes \mathrm{id}^{y \to y} \in H(x(y \to y))$ for all $y \in \mathrm{Types}_{\mathcal{A}}$.

Proof. Necessity follows from Theorem 6.2.43. For sufficiency, we choose $y = I_x$. Since $\Gamma^x \boxtimes \mathrm{id}^{I_x \to I_x} \in H(I_x \to x)$ and $\eta^{I \to I_x}(1) \in H(I_x)$ because promotion functions are H-preserving, we have that $\Gamma^x \boxtimes \mathrm{id}^{I_x \to I_x} @ \eta^{I \to I_x}(1) \in H(x)$. Observe that $\Gamma^x \boxtimes \mathrm{id}^{I \to I} @ \eta^{I \to I_x}(1) = \Gamma^x \boxtimes \eta^{I \to I_x}(1) = \Gamma^x \boxtimes 1 = \Gamma^x$, which implies $\Gamma^x \in H(x)$. □

6.2.6 Positive-Preserving Maps

In this subsection, we show how to extend the notion of completely positive-preserving channels to maps of arbitrary type x. The key point is that a map is *completely* positive-preserving if it is positive-preserving when tensored with an identity of the same order. For example, channels are tensored with the identity channel.

Definition 6.2.45: Completely positive-preserving maps

Let $x \in \mathrm{Types}_{\mathcal{A}}$. The set $K(x)$ of completely positive-preserving maps of type x is defined as

- if $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$, $K(x)$ is the set of positive semi-definite matrices on $\mathcal{H}_E$,
- if $x = (a \to b)$, then $K(x)$ is the set of all $\Gamma^x \in H(x)$ such that $\Gamma^x \boxtimes \mathrm{id}^{y \to y}(H(ay)) \subseteq H(by)$, for all types y such that $\mathrm{ord}(y) = \mathrm{ord}(x) - 1$.

Note that this definition ensures that the order type of $y \to y$ is the same as the order of x. Moreover, if $x = (A \to B)$, the set $K(x)$ coincides with $\mathrm{CP}(A \to B)$, the set of completely positive-preserving operators from A to B.

After we define a space, we usually study its geometry. For example, $L(x)$ is a complex Hilbert space, and $H(x)$ is a real Hilbert space. $K(x)$ is a closed, convex, and pointed cone, where pointed means that $K(x) \cap -K(x) = \{ \mathbf{0}_x \}$, and $\mathbf{0}_x$ is the map sending all inputs to the zero vector of the output space.

Proposition 6.2.46: Geometry of $K(x)$

Let $x \in \mathrm{Types}_{\mathcal{A}}$, then $K(x)$ is a convex, closed, and pointed cone.

Proof. Base step: The set of positive semidefinite matrices is a convex, closed, and pointed cone.

Inductive step: Let $\mathrm{ord}(x) > 0$ and assume for all types z up to order $\mathrm{ord}(x) - 1$, $K(z)$ is a convex, closed, and pointed cone. Let $a, b \in \mathrm{Types}_{\mathcal{A}}$ be such that $x = (a \to b)$.

1. $K(x)$ is a convex cone: Let $\{ \Gamma^x_i \}_i$ be a finite sets of maps in $K(x)$, and let $\alpha_i \in \mathbb{R}_+$. Let $y \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(y) = \mathrm{ord}(x) - 1$, and let $u^{ay} \in K(ay)$. Then,

$$\sum_i \alpha_i \Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay} = \sum_i \alpha_i (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay}), \tag{6.130}$$

   and each $\Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay} \in K(by)$. Since $H(by)$ is a convex cone we have that $\sum_i \alpha_i (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay}) \in K(by)$. This is true for all $y \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(y) = \mathrm{ord}(x) - 1$ and all $u^{ay} \in K(ay)$. Therefore, $\sum_i \alpha_i (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y}) \in K(x)$.

2. $K(x)$ is closed: Let $\{ \Gamma^x_i \}_i$ be a sequence of maps in $K(x)$ converging in the norm induced by the inner product $(\cdot, \cdot)_x$ to Γ^x. That is, $\lim_{i \to \infty} \left\| \Gamma^x - \Gamma^x_i \right\|_x = 0$. We want to show that $\Gamma^x \in K(x)$. Let $y \in \mathrm{Types}_{\mathcal{A}}$ be such that $\mathrm{ord}(y) = \mathrm{ord}(x) - 1$ and let $u^{ay} \in K(ay)$. First, we show that $\Gamma^x \boxtimes \mathrm{id}^{y \to y}$ is equal to $\lim_{i \to \infty} (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y})$. Indeed,

$$\begin{aligned} \lim_{i \to \infty} \left\| \Gamma^i_x \boxtimes \mathrm{id}^{y \to y} - \Gamma^x \boxtimes \mathrm{id}^{y \to y} \right\|_{x(y \to y)} &= \lim_{i \to \infty} \left\| (\Gamma^i_x - \Gamma^x) \boxtimes \mathrm{id}^{y \to y} \right\|_{x(y \to y)} \\ \text{Corollary 6.2.23} &= \lim_{i \to \infty} \left\| (\Gamma^i_x - \Gamma^x) \right\|_x \|\mathrm{id}^{y \to y}\|_{y \to y} \\ &= 0. \end{aligned} \tag{6.131}$$

   This implies that

$$\Gamma^x \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay} = \lim_{i \to \infty} (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y}) @ u^{ay} = \lim_{i \to \infty} \Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay}. \tag{6.132}$$

   Since $\Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay} \in K(by)$ for all i and $K(by)$ is closed by induction hypothesis, we have that $\Gamma^x \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay} = \lim_{i \to \infty} (\Gamma^x_i \boxtimes \mathrm{id}^{y \to y} \,@ u^{ay}) \in K(by)$. This, in turn, implies that $\Gamma^x \in K(x)$.

3. $K(x)$ is pointed: Let $\Gamma^x \in K(x) \cap -K(x)$. Then for any $y \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(y) = \mathrm{ord}(x) - 1$, and any u^{ay} we have that $\Gamma^x \boxtimes \mathrm{id}^{y \to y}(u^{ay}) \in K(by)$ and $-\Gamma^x \boxtimes \mathrm{id}^{y \to y}(u^{ay}) \in K(by)$. However, $K(by)$ is pointed by induction hypothesis, therefore $\Gamma^x \boxtimes \mathrm{id}^{y \to y}(u^{ay}) = \mathbf{0}_{by}$. This is true for all u^{ay}, therefore $\Gamma^x = \mathbf{0}_x$.

□

At this point, we would turn our attention to promotion and demotion maps, with the hope that they are K-preserving. That is, we would like to show that if $\Gamma^x \in K(x)$ then its promotion $\eta^{x\to x_y}(\Gamma^x)$ is in $K(x_y)$. However, if one tries to prove this result, one would notice that many properties of the square product are needed. In particular, one needs the square product of maps in K to be in K. Therefore, we must prove the result for the promotion function in conjunction with the result for the square product. While doing so, we also prove that $|\cdot\rangle$ is K-preserving, that a function in K is K-preserving, and that the swap operators are in K.

Theorem 6.2.47: Properties of K

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$.

1. If $\Gamma^x \in K(x)$, then $\eta^{x\to x_y}(\Gamma^x) \in K(x_y)$. Similarly, if $\Delta^{x_y} \in K(x_y)$, then $\eta^{x_y\to x}(\Delta^{x_y}) \in K(x)$
2. If $\Gamma^x \in K(x)$ and $\Delta^y \in K(y)$, then $\Gamma^x \boxtimes \Delta^y \in K(xy)$.
3. If $\Gamma^x \in K(x)$, then $|\Gamma^x\rangle^{I\to x} \in K(I \to x)$. Similarly, if $\Theta^{I\to x} \in K(I \to x)$, then $\Theta^{I\to x}(1) \in K(x)$.
4. If $\Gamma^x \in K(x)$, $x = (a \to b)$ and $u^a \in K(a)$, then $\Gamma^x(u^a) \in K(b)$.
5. $\sigma^{xy\to yx} \in K(xy \to yx)$.

Proof. We prove it by induction on $\mathrm{ord}(xy)$. *Base step:* Let $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$, $y = F \in \mathrm{EleTypes}_{\mathcal{A}}$.

1. $\eta^{E\to E_F} = \eta^{E_F\to E} = \mathrm{id}^{E\to E}$, which is obviously K-preserving.
2. Since $\Gamma^E, \Delta^F \geq 0$, then $\Gamma^E \otimes \Delta^F \geq 0$.
3. Let $\Gamma^E \geq 0$. For any $G \in \mathrm{EleTypes}_{\mathcal{A}}$, and $u^G \in K(G)$,

$$\begin{aligned}(\left|\Gamma^E\right\rangle^{I\to E} \boxtimes \mathrm{id}^{G\to G})(u^G) &= (\left|\Gamma^E\right\rangle^{I\to E} \boxtimes \mathrm{id}^{G\to G})(1 \boxtimes u^G)\\ &= \left|\Gamma^E\right\rangle^{I\to E}(1) \boxtimes \mathrm{id}^{G\to G}(u^G)\\ &= \Gamma^E \boxtimes u^G,\end{aligned} \tag{6.133}$$

   which, as we have just shown in point 2, is in $K(EG)$. This shows that $\left|\Gamma^E\right\rangle^{I\to E} \in K(I \to E)$.

   If we assume that $\Theta^{I\to E} \in K(I \to E)$, then $\Theta^{I\to E} \boxtimes \mathrm{id}^{I\to I}$ is K-preserving. In particular, since $1 \boxtimes 1 \in K(I \boxtimes I)$, we have that

$$\begin{aligned}\Theta^{I\to E} \boxtimes \mathrm{id}^{I\to I}(1 \boxtimes I) &= \Theta^{I\to E}(1) \boxtimes \mathrm{id}^{I\to I}(1)\\ &= \Theta^{I\to E}(1) \boxtimes 1\\ &= \Theta^{I\to E}(1).\end{aligned} \tag{6.134}$$

   Since $\Theta^{I\to E} \boxtimes \mathrm{id}^{I\to I}$ is K-preserving, then $\Theta^{I\to E}(1) \in K(E)$.

4. Not compatible with the base case.

5. $\sigma^{EG\to GE}$ is simply the swap channel, which is completely positive-preserving. Therefore $\sigma^{EG\to GE} \boxtimes \mathrm{id}^{F\to F}$ is K-preserving for any F, and thus, $\sigma^{EG\to GE} \in K(EG \to GE)$.

Inductive step: Here we assume that $\mathrm{ord}(xy) > 0$ and that the properties *1-5* are true for all types up to order $\mathrm{ord}(xy) - 1$. This implies that we can use any of the properties *1-5* for types of order up to $\mathrm{ord}(xy) - 1$ in the proof of the inductive step. Moreover, we can use a proven statement at order $\mathrm{ord}(xy)$ to prove a subsequent statement at order $\mathrm{ord}(xy)$. Note that this does not result in a circular proof if we are careful and only use proven statements. In the following, we write $x = (a \to b)$ whenever $\mathrm{ord}(x) > 0$, $y = (c \to d)$ whenever $\mathrm{ord}(y) > 0$, and $z = (e \to f)$, whenever $\mathrm{ord}(z) > 0$.

1. Promotion and demotion functions are K-preserving:

   - $\mathrm{ord}(x) = \mathrm{ord}(y)$:

     $$\begin{aligned} \eta^{x\to x_y}(\Gamma^x) \boxtimes \mathrm{id}^{z\to z} &= \eta^{x_y\to x_y(z\to z)}(\eta^{x\to x_{xy}}(\Gamma^x)) \boxtimes \mathrm{id}^{z\to z} \\ \text{Proposition 6.2.17} &= (\eta^{b\to b_{dz}} \circ \Gamma^{a\to b} \circ \eta^{a_{cz}\to a}) \boxtimes \mathrm{id}^{z\to z} \\ \text{Theorem 6.2.30} &= (\eta^{b\to b_{dz}} \boxtimes \mathrm{id}^{z\to z}) \circ (\Gamma^x \boxtimes \mathrm{id}^{z\to z}) \circ (\eta^{a_{cz}\to a} \boxtimes \mathrm{id}^{z\to z}) \\ \text{Lemma 6.2.20} &= \eta^{bz\to bz_d} \circ (\Gamma^x \boxtimes \mathrm{id}^{z\to z}) \circ \eta^{az_c\to az}. \end{aligned} \tag{6.135}$$

     By the induction hypothesis, the promotion and demotion functions appearing in the last line of the equation above are K-preserving, and so is $\Gamma^x \boxtimes \mathrm{id}^{z\to z}$ because $\Gamma^x \in K(x)$. The concatenation of K-preserving functions is K-preserving, and therefore, $\eta^{x\to x_y}(\Gamma^x) \boxtimes \mathrm{id}^{z\to z}$ is K-preserving for any z such that $\mathrm{ord}(z) = \mathrm{ord}(xy) - 1$. This implies that $\eta^{x\to x_y}(\Gamma^x) \in K(x_y)$ for any $\Gamma^x \in K(x)$, and this shows that $\eta^{x\to x_y}$ is K-preserving.

     The proof of $\eta^{x_y\to y}$ being K-preserving follows exactly the same steps: $\eta^{x_y\to x}(\Delta^{x_y}) \boxtimes \mathrm{id}^{z\to z}$ can be written as a concatenation of K-preserving function, which implies that $\eta^{x_y\to x}(\Delta^{x_y}) \in K(x)$ and that $\eta^{x\to x_y}$ is K-preserving.

   - $\mathrm{ord}(x) < \mathrm{ord}(y)$:

     $$\begin{aligned} \eta^{x\to x_y}(\Gamma^x) \boxtimes \mathrm{id}^{z\to z} &= \eta^{x_y\to x_y(z\to z)}(\eta^{x\to x_{xy}}(\Gamma^x)) \boxtimes \mathrm{id}^{z\to z} \\ &= (\eta^{x_d\to x_{dz}} |\eta^{x\to x_d}(\Gamma^a)\rangle^{I\to x_d} \circ \eta^{I_c\to I} \circ \eta^{I_{zc}\to I_c}) \boxtimes \mathrm{id}^{z\to z} \\ \text{Lemma 6.2.6} &= |\eta^{x\to x_{dz}}(\Gamma^x)\rangle^{I\to x_{dz}} \circ \eta^{I_{cz}\to I}. \end{aligned} \tag{6.136}$$

     Note that, since $\mathrm{ord}(xdz) < \mathrm{ord}(xy)$ we can use the inductive hypothesis about $|\cdot\rangle^{I\to x_{dz}}$ being K-preserving. Therefore, we can conclude that $\eta^{x\to x_y}$ is K-preserving with the same argument used in the previous case. Similarly, we can prove that $\eta^{x_y\to x}$ is K-preserving using the inductive hypothesis for the evaluation at 1.

   - $\mathrm{ord}(x) > \mathrm{ord}(y)$: the proof is analogous to the case $\mathrm{ord}(x) = \mathrm{ord}(y)$ but simpler.

2. Completely positive-preserving maps are closed under the tensor product. Let $z \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(z) = \mathrm{ord}(xy) - 1$. We want to show that $\Gamma^x \boxtimes \Delta^y \boxtimes \mathrm{id}^{z\to z}$ is K-preserving. To this end, we write $\Gamma^x \boxtimes \Delta^y \boxtimes \mathrm{id}^{z\to z}$ as a sequential composition of K-preserving maps. First, note that

$$\begin{aligned} \Gamma^x \boxtimes \Delta^y \boxtimes \mathrm{id}^{z\to z} &= \eta^{xy\to xy_{z\to z}}(\Gamma^x \boxtimes \Delta^y) \boxtimes \mathrm{id}^{z\to z} \\ &= \eta^{x\to x_{y(z\to z)}}(\Gamma^x) \boxtimes \eta^{y\to y_{x(z\to z)}}(\Delta^y) \boxtimes \mathrm{id}^{z\to z}\,. \end{aligned} \tag{6.137}$$

Since $\mathrm{ord}(xy(z \to z)) = \mathrm{ord}(xy)$, we can use the already proven fact that promotion maps of order $\mathrm{ord}(xy)$ are K-preserving. Therefore, $\tilde{\Gamma}^{x_{y(z\to z)}} \in K(x_{y(z\to z)})$ and $\tilde{\Delta}^{y_{x(z\to z)}} \in K(y_{x(z\to z)})$. For simplicity, let α be the type associated with the left subtree of $x_{y(z\to z)}$ and β the type associated with the right subtree. That is $\alpha \to \beta = x_{y(z\to z)}$. Similarly, let $\gamma \to \delta = y_{x(z\to z)}$. It is important to notice that $\mathrm{ord}(\alpha) = \mathrm{ord}(\beta) = \mathrm{ord}(\gamma) = \mathrm{ord}(\delta) = \mathrm{ord}(xy) - 1$. So far, we have that $\Gamma^x \boxtimes \Delta^y \boxtimes \mathrm{id}^{z\to z} = \tilde{\Gamma}^{\alpha\to\beta} \boxtimes \tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{z\to z}$. Using Theorem 6.2.30 and the recursive definition of swap operators in Lemma 6.2.36 we get:

$$\begin{aligned} \tilde{\Gamma}^{\alpha\to\beta} \boxtimes \tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{z\to z} &= (\mathrm{id}^{\beta\to\beta} \boxtimes \tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{z\to z}) \\ &\quad \circ (\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \mathrm{id}^{\gamma\to\gamma} \boxtimes \mathrm{id}^{z\to z}) \\ &= (\sigma^{(\gamma\beta\to\delta\beta)\to(\beta\gamma\to\beta\delta)}(\tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{\beta\to\beta}) \boxtimes \mathrm{id}^{z\to z}) \\ &\quad \circ (\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \mathrm{id}^{\gamma\to\gamma} \boxtimes \mathrm{id}^{z\to z}) \\ &= [(\sigma^{\delta\beta\to\beta\delta} \circ (\tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{\beta\to\beta}) \circ \sigma^{\beta\gamma\to\gamma\beta}) \boxtimes \mathrm{id}^{z\to z}] \\ &\quad \circ (\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \mathrm{id}^{\gamma\to\gamma} \boxtimes \mathrm{id}^{z\to z}) \\ &= (\sigma^{\delta\beta\to\beta\delta} \boxtimes \mathrm{id}^{z\to z}) \circ (\tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{\beta z\to\beta z}) \\ &\quad \circ (\sigma^{\beta\gamma\to\gamma\beta} \boxtimes \mathrm{id}^{z\to z}) \circ (\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \mathrm{id}^{\gamma z\to\gamma z}). \end{aligned} \tag{6.138}$$

We have already discussed that $\tilde{\Gamma}^{\alpha\to\beta} \in K(\alpha \to \beta)$ and $\tilde{\Delta}^{\gamma\to\delta} \in K(\gamma \to \delta)$. Therefore, $(\tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{\beta z\to\beta z})$ and $(\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \mathrm{id}^{\gamma z\to\gamma z})$ are K-preserving, since $\mathrm{ord}(\gamma z) = \mathrm{ord}(\alpha \to \beta) - 1$ and $\mathrm{ord}(\beta z) = \mathrm{ord}(\gamma \to \delta) - 1$. Moreover, by induction hypothesis $\sigma^{\delta\beta\to\beta\delta} \in K(\delta\beta \to \beta\delta)$ and $\sigma^{\beta\gamma\to\gamma\beta} \in K(\beta\gamma \to \gamma\beta)$. Again, this implies that when tensored with identity maps of the same order, as in the equation above, they are K-preserving. Since, the sequential composition of K-preserving maps is, by definition, K-preserving, then $\tilde{\Gamma}^{\alpha\to\beta} \boxtimes \tilde{\Delta}^{\gamma\to\delta} \boxtimes \mathrm{id}^{z\to z}$ is K-preserving. Thus, $\Gamma^x \boxtimes \Delta^y \boxtimes \mathrm{id}^{z\to z}$ is K preserving for any $z \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(z) = \mathrm{ord}(xy) - 1$, and therefore $\Gamma^x \boxtimes \Delta^y \in K(xy)$.

3. Ket and evaluation at 1 are K-preserving. We want to show that $|\Gamma^x\rangle^{I\to x} \boxtimes \mathrm{id}^{z\to z}$ is K-preserving when $\Gamma^x \in K(x)$ and $\mathrm{ord}(z) = \mathrm{ord}(I \to x) - 1 = \mathrm{ord}(x)$. Let $u^z \in K(z)$. Then,

$$\begin{aligned} |\Gamma^x\rangle^{I\to x} \boxtimes \mathrm{id}^{z\to z} @ u^z &= |\Gamma^x\rangle^{I\to x} \boxtimes \mathrm{id}^{z\to z} @ 1 \boxtimes u^z \\ &= |\Gamma^x\rangle^{I\to x}(1) \boxtimes \mathrm{id}^{z\to z}(u^z) \\ &= \Gamma^x \boxtimes u^z. \end{aligned} \tag{6.139}$$

Since $\mathrm{ord}(xz) = \mathrm{ord}(x) \le \mathrm{ord}(xy)$, we can use either the inductive hypothesis for property *2* or the proof in the inductive step of property *2* to conclude that $\Gamma^x \boxtimes u^z \in K(xz)$. Therefore, $|\Gamma^x\rangle^{I\to x} \boxtimes \mathrm{id}^{z\to z}$ is K-preserving, and $|\Gamma^x\rangle^{I\to x} \in K(I \to x)$.

Let us assume that $\Theta^{I\rightarrow x} \in K(I \rightarrow x)$. We want to show that $\Theta^{I\rightarrow x}(1) \in K(x)$. We know that $\Theta^{I\rightarrow x} \boxtimes \mathrm{id}^{I_x\rightarrow I_x}$ is K-preserving. Since $1 \in K(I)$, then using property *1* we have that $\eta^{I\rightarrow I_x}(1) \in K(I_x)$. Therefore, $\Theta^{I\rightarrow x} \boxtimes \mathrm{id}^{I_x\rightarrow I_x} @\eta^{I\rightarrow I_x}(1) \in K(x)$. Note that

$$\begin{aligned}
\Theta^{I\rightarrow x} \boxtimes \mathrm{id}^{I_x\rightarrow I_x} @\eta^{I\rightarrow I_x}(1) &= \Theta^{I\rightarrow x} \boxtimes \mathrm{id}^{I_x\rightarrow I_x}\ 21 \boxtimes \eta^{I\rightarrow I_x}(1) \\
&= \Theta^{I\rightarrow x}(1) \boxtimes \mathrm{id}^{I_x\rightarrow I_x}(\eta^{I\rightarrow I_x}(1)) \\
&= \Theta^{I\rightarrow x}(1) \boxtimes \eta^{I\rightarrow I_x}(1) \\
\text{Lemma-6.2.21} &= \Theta^{I\rightarrow x}(1) \boxtimes \mathrm{id}^{I\rightarrow I}(1) \\
&= \Theta^{I\rightarrow x}(1) \boxtimes 1 \\
&= \Theta^{I\rightarrow x}(1).
\end{aligned} \tag{6.140}$$

This implies that $\Theta^{I\rightarrow x}(1) \in K(x)$.

4. Let z be a type such that $z = I_z$ and $\mathrm{ord}(z) = \mathrm{ord}(x) - 1$. We want to show that $\Gamma^{a\rightarrow b}(u^a) = \eta^{b_z\rightarrow b} \circ (\Gamma^x \boxtimes \mathrm{id}^{I_z\rightarrow I_z}) \circ \eta^{a\rightarrow a_z} @u^a$ for all u^a.

$$\begin{aligned}
\eta^{b_z\rightarrow b} \circ (\Gamma^x \boxtimes \mathrm{id}^{I_z\rightarrow I_z}) \circ \eta^{a\rightarrow a_z} @u^a &= \eta^{b_z\rightarrow b} \circ (\Gamma^x \boxtimes \mathrm{id}^{I_z\rightarrow I_z}) @\eta^{a\rightarrow a_z}(u^a) \\
\text{Lemma 6.2.32} &= \eta^{b_z\rightarrow b} \circ (\Gamma^x \boxtimes \mathrm{id}^{I_z\rightarrow I_z}) @u^a \boxtimes \eta^{I\rightarrow I_z}(1) \\
&= \eta^{b_z\rightarrow b} @(\Gamma^x(u^a) \boxtimes \eta^{I\rightarrow I_z}(1)) \\
\text{Lemma 6.2.32} &= \eta^{b_z\rightarrow b} @(\eta^{b\rightarrow b_z}(\Gamma^x(u^a))) \\
&= \Gamma^x(u^a).
\end{aligned} \tag{6.141}$$

   We have proven that the promotion function are K-preserving, and so is $\Gamma^x \boxtimes \mathrm{id}^{I_z\rightarrow I_z}$ because $\Gamma^x \in K(x)$. Since Γ^x is equal to the composition of K-preserving functions, it is K-preserving.

5. We want to show that $\sigma^{xy\rightarrow yx} \in K(x)$. That is, we want to show that for every $z \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(z) = \mathrm{ord}(xy)$, $\sigma^{xy\rightarrow yx} \boxtimes \mathrm{id}^{z\rightarrow z}$ is K-preserving. To achieve that, we first prove

$$\begin{aligned}
\sigma^{xy\rightarrow yx} \boxtimes \mathrm{id}^{z\rightarrow z}(\Gamma^{xyz}) = (\sigma^{\beta_f\delta_f\rightarrow\delta_f\beta_f} \boxtimes \mathrm{id}^{f_{\beta\delta}\rightarrow f_{\beta\delta}}) \circ \Gamma^{\alpha\gamma e\rightarrow\beta\delta f} \\
\circ (\sigma^{\gamma_e\alpha_e\rightarrow\alpha_e\gamma_e} \boxtimes \mathrm{id}^{e_{\alpha\gamma}\rightarrow e_{\alpha\gamma}})
\end{aligned} \tag{6.142}$$

   for every $\Gamma^{xyz} \in L(xyz)$, where α, β, γ, and δ are defined as in Corollary 6.2.38. Let $u^x \in L(x)$, $v^y \in L(y)$, and $w^z \in L(z)$. Let $\tilde{u}^{\alpha_e\rightarrow\beta_f} = \eta^{x\rightarrow x_z}(u^x)$ and let $\tilde{v}^{\gamma_e\rightarrow\delta_f} = \eta^{y\rightarrow y_z}(v^y)$ and $w^{e_{\alpha\gamma}\rightarrow f_{\beta\delta}} =$

$\eta^{z\to z_{xy}}(w^z)$. Then, using Theorem 6.2.30, we get

$$\begin{aligned}
&(\sigma^{\beta_f\delta_f\to\delta_f\beta_f}\boxtimes \mathrm{id}^{f_{\beta\delta}\to f_{\beta\delta}})\circ(u^x\boxtimes v^y\boxtimes w^z)\circ(\sigma^{\gamma_e\alpha_e\to\alpha_e\gamma_e}\boxtimes \mathrm{id}^{e_{\alpha\gamma}\to e_{\alpha\gamma}})\\
&=(\sigma^{\beta_f\delta_f\to\delta_f\beta_f}\boxtimes \mathrm{id}^{f_{\beta\delta}\to f_{\beta\delta}})\circ(\tilde{u}^{\alpha_e\to\beta_f}\boxtimes\tilde{v}^{\gamma_e\to\delta_f}\boxtimes\tilde{w}^{e_{\alpha\gamma}\to f_{\beta\delta}})\\
&\qquad\circ(\sigma^{\gamma_e\alpha_e\to\alpha_e\gamma_e}\boxtimes \mathrm{id}^{e_{\alpha\gamma}\to e_{\alpha\gamma}})\\
&=(\sigma^{\beta_f\delta_f\to\delta_f\beta_f}\circ(\tilde{u}^{\alpha_e\to\gamma_f}\boxtimes\tilde{v}^{\gamma_e\to\delta_f})\circ\sigma^{\gamma_e\alpha_e\to\alpha_e\gamma_e})\\
&\qquad\boxtimes(\mathrm{id}^{f_{\beta\delta}\to f_{\beta\delta}}\circ\tilde{w}^{e_{\alpha\gamma}\to f_{\beta\delta}}\circ\mathrm{id}^{e_{\alpha\gamma}\to e_{\alpha\gamma}})\\
&=\sigma^{x_zy_z\to y_zx_z}(\tilde{u}^{x_z}\boxtimes\tilde{v}^{y_z})\boxtimes\tilde{w}^{z_{xy}}\\
&=\tilde{v}^{y_z}\boxtimes\tilde{u}^{x_z}\boxtimes\tilde{w}^{z_{xy}}\\
&=v^y\boxtimes u^x\boxtimes w^z\\
&=(\sigma^{xy\to yx}\boxtimes\mathrm{id}^{z\to z})(u^x\boxtimes v^y\boxtimes w^z).
\end{aligned}\tag{6.143}$$

By linearity,

$$\begin{aligned}
(\sigma^{xy\to yx}\boxtimes\mathrm{id}^{z\to z})(\Gamma^{xyz})=(\sigma^{\beta_f\delta_f\to\delta_f\beta_f}\boxtimes \mathrm{id}^{f_{\beta\delta}\to f_{\beta\delta}})\circ\Gamma^{xyz}\\
\circ(\sigma^{\gamma_e\alpha_e\to\alpha_e\gamma_e}\boxtimes \mathrm{id}^{e_{\alpha\gamma}\to e_{\alpha\gamma}}).
\end{aligned}\tag{6.144}$$

By inductive hypothesis, the swap operators appearing on the right-hand side of Eq. (6.144) are in K, and therefore, they are K-preserving when tensored with an identity of the same order, as in our case. Moreover, if $\Gamma^{xyz}\in K(xyz)$, then as a consequence of property *4*, Γ^{xyz} is K-preserving. Since sequential composition of K-preserving maps is K-preserving, we have that $\sigma^{xy\to yx}\boxtimes\mathrm{id}^{z\to z}$ is K-preserving for all $z\in\mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(z)=\mathrm{ord}(xy)$, and therefore $\sigma^{xy\to yx}\in K(xy\to yx)$.

□

Combining the fact that promotion maps are K-preserving with Lemma 6.2.20, one immediately obtains that the promotion maps are in K.

Corollary 6.2.48: Promotion maps are in K

Let $x,y\in\mathrm{Types}_{\mathcal{A}}$. Then $\eta^{x\to x_y}\in K(x\to x_y)$ and $\eta^{x_y\to x}\in K(x_y\to x)$.

We have discussed promotion functions and parallel composition. We turn our attention to sequential composition. As in the case of quantum channels, we will prove that the sequential composition of maps in K is in K. Note that this is far less trivial than stating that the sequential composition of K-preserving maps is K-preserving because for a map to be in K, it has to be K-preserving when tensored with identities of the proper order.

Lemma 6.2.49: Sequential composition of K-maps

Let $x=(a\to b), y=(b\to c)\in\mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^{a\to b}\in K(x)$, $\Delta^{b\to c}\in K(y)$. Then, $\Delta^{b\to c}\circ\Gamma^{a\to b}\in K(a\to c)$.

Proof. We prove it by induction. In the base case, $\mathrm{ord}(x)=\mathrm{ord}(y)=1$. Let $a=A\in\mathrm{EleTypes}_{\mathcal{A}}$, $b=B\in\mathrm{EleTypes}_{\mathcal{A}}$ and $c=C\in\mathrm{EleTypes}_{\mathcal{A}}$. Then, $K(A\to B)$ is $\mathrm{CP}(A\to B)$, the set of completely positive-preserving maps from $L(A)$ to $L(B)$, and $K(B \to C)$ is $\mathrm{CP}(B \to C)$. The fact that the sequential composition of completely positive-preserving maps is completely positive-preserving is a well-known result in quantum information.

Inductive step: Let $\mathrm{ord}(xy) > 1$. We want to prove that $(\Delta^{b\to c} \cdot M_{a\to b}) \boxtimes \mathrm{id}^{z\to z}$ is K-preserving for all $z = (e \to f) \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(z) = \mathrm{ord}(ac)$.

- If $\mathrm{ord}(x) > \mathrm{ord}(y)$, then $\mathrm{ord}(a) > \mathrm{ord}(bc)$. Therefore, $\mathrm{ord}(z) = \mathrm{ord}(a)$. Now, we present a long chain of equalities to show that $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z}$ can be written as a sequence of function from which it is easier to deduce that it is K-preserving. Let $u^a \boxtimes v^z \in L(az)$. Using properties of promotion and demotion functions, together with Theorem 6.2.27 and Theorem 6.2.30 we have:

$$\begin{aligned}
&\mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ ((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ (\Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z} @ u^a \boxtimes v^z) \\
&= \mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ ((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ (\Gamma^{a\to b}(u^a) \boxtimes v^z) \\
&= \mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ ((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ \\
&\qquad ((\big|\eta^{b\to b_f}[\Gamma^{a\to b}(u^a)]\big\rangle^{I\to b_f} \circ \eta^{I_e\to I}) \boxtimes v^z) \\
&= \mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ (\big|\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \Gamma^{a\to b}(u^a)\big\rangle^{I\to c_f} \circ \eta^{I_e\to I}) \boxtimes (\eta^{f\to f_{bc}} \circ v^z) \\
&= \mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ \eta^{c\to c_z}(\Delta^{b\to c} \circ \Gamma^{a\to b}(u^a)) \boxtimes \eta^{z\to z_{bc}}(v^z) \\
&= \mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ \eta^{c\to c_z}(\Delta^{b\to c} \circ \Gamma^{a\to b}(u^a)) \boxtimes \eta^{z\to z_b}(v^z) \\
&= \eta^{c\to c_z}(\Delta^{b\to c} \circ \Gamma^{a\to b}(u^a)) \boxtimes v^z \\
&= [\Delta^{b\to c} \circ \Gamma^{a\to b}(u^a)] \boxtimes v^z \\
&= (\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z} @ u^a \boxtimes v^z.
\end{aligned} \tag{6.145}$$

  Therefore, for every $w^{az} \in K(az)$, we have

$$\begin{aligned}
&(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z} @ w^{az} = \\
&\mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ ((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ (\Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z} @ w^{az}).
\end{aligned} \tag{6.146}$$

  Now, observe that, since $\Gamma^{a\to b} \in K(a \to b)$, we have that $\Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z} @ w^{az} \in K(bz)$. By induction hypothesis, $\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}$ is in $K(b_f \to c_f)$, and, using Theorem 6.2.47 we get $((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \in K(bf \to cf_b)$. Note that $\mathrm{ord}(bf \to cf_b) \leq \mathrm{ord}(x) - 1$ and $\mathrm{ord}(bz) = \mathrm{ord}(x) - 1$. Therefore, we can use the inductive hypothesis again, to obtain that $((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ (\Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z} @ w^{az}) \in K(e \to cf_b) = K(cz_b)$. Lastly, since the identity and the promotion function are in K and therefore K-preserving, we have that $\mathrm{id}^{c_z\to c_z} \boxtimes \eta^{z_b\to z} @ ((\eta^{c\to c_f} \circ \Delta^{b\to c} \circ \eta^{b_f\to b}) \boxtimes \eta^{f\to f_{bc}}) \circ (\Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z} @ u^a \boxtimes v^z) \in K(cz)$. This shows that $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z}$ is K-preserving, and therefore $\Delta^{b\to c} \circ \Gamma^{a\to b} \in K(a \to c)$.

- If $\mathrm{ord}(x) < \mathrm{ord}(y)$, then $\mathrm{ord}(c) > \mathrm{ord}(a), \mathrm{ord}(b)$ and $\mathrm{ord}(z) = \mathrm{ord}(c)$. Once again, we use a long chain of equalities to write $\Delta^{b\to c} \circ \Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z}$ as a sequence of functions and deduce that it is

K-preserving. Let $u^a \boxtimes v^z \in L(az)$.

$$\begin{aligned}
&\Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ ((\eta^{b\to b_f} \circ \Gamma^{a\to b} \circ \eta^{a_f\to a}) \boxtimes \eta^{f\to f_{ab}}) \circ (u^a \boxtimes v^z) \\
&= \Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ ((\eta^{b\to b_f} \circ \Gamma^{a\to b} \circ \eta^{a_f\to a}) \boxtimes \eta^{f\to f_{ab}}) \circ \\
&\qquad (|\eta^{a\to a_f}(u^a)\rangle^{I\to a_f} \circ \eta^{I_e\to I} \boxtimes v^z) \\
&= \Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ (\left|\eta^{b\to b_f} \circ \Gamma^{a\to b}(u^a)\right\rangle^{I\to b_f} \circ \eta^{I_e\to I}) \boxtimes (\eta^{f\to f_{ab}} \circ v^z) \\
&= \Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ \eta^{b\to b_z}(\Gamma^{a\to b}(u^a)) \boxtimes \eta^{z\to z_{ab}}(v^z) \\
&= \Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ \Gamma^{a\to b}(u^a) \boxtimes \eta^{z\to z_a}(v^z) \\
&= \Delta^{b\to c} \circ \Gamma^{a\to b}(u^a) \boxtimes v^z \\
&= (\Delta^{b\to c} \circ \Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z}) @ u^a \boxtimes v^z.
\end{aligned} \tag{6.147}$$

With an analogous argument to the one used before, one concludes that $\Delta^{b\to c} \boxtimes \eta^{z_a\to z} @ ((\eta^{b\to b_f} \circ \Gamma^{a\to b} \circ \eta^{a_f\to a}) \boxtimes \eta^{f\to f_{ab}}) \circ w^{az} \in K(cz)$ whenever $w^{az} \in K(az)$. Therefore $\Delta^{b\to c} \circ \Gamma^{a\to b} \boxtimes \mathrm{id}^{z\to z}$ is K-preserving and $\Delta^{b\to c} \circ \Gamma^{a\to b} \in K(a \to c)$.

- If $\mathrm{ord}(x) = \mathrm{ord}(y)$, then $\mathrm{ord}(b) > \mathrm{ord}(ac)$ or (non-exclusive) $\mathrm{ord}(a) = \mathrm{ord}(c)$. Since, $\mathrm{ord}(z) = \mathrm{ord}(ac)$, we have that $\mathrm{ord}(zb) = \mathrm{ord}(abc) = \mathrm{ord}(x) - 1 = \mathrm{ord}(y - 1)$. Therefore,

$$\begin{aligned}
(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z} &= (\Delta^{b\to c} \circ \Gamma^{a\to b}) \circ (\eta^{z_b\to z} \circ \eta^{z\to z_b}) \\
&= (\Delta^{b\to c} \boxtimes \eta^{z_b\to z}) \circ (\Gamma^{a\to b} \circ \eta^{z\to z_b}).
\end{aligned} \tag{6.148}$$

By Theorem 6.2.47, both functions enclosed between round brackets are K-preserving, and therefore $(\Delta^{b\to c} \circ \Gamma^{a\to b}) \boxtimes \mathrm{id}^{z\to z}$ is K-preserving and $\Delta^{b\to c} \circ \Gamma^{a\to b} \in K(a \to c)$.

□

6.3 The Choi Isomorphism

In this last Section, we extend the Choi isomorphism from the case of channels and states to the case of maps of order $n + 1$ and maps of order n. In the literature, various approaches to the Choi isomorphism beyond channels and states have been explored. For example, in Ref. [264], the Choi isomorphism is extended to superchannels, and different options are presented. However, from a linear algebra perspective, the Choi isomorphism presents no ambiguity. If one has a linear map $f : X \to Y$, where X and Y are finite-dimensional Hilbert spaces, then the Choi isomorphism is defined following these steps.

1. Fix an orthonormal basis $\{u^i\}_i$ of X.
2. Define the Choi state Φ in $X \otimes X$ as $\Phi = \sum_i u^i \otimes u^i$.
3. The Choi isomorphism maps f in $\mathfrak{L}(X, Y)$, into $f \otimes \mathrm{id}^{Y\to Y}(\Phi) = \sum_i f(u^i) \otimes u^i \in Y \otimes X$.

Each of the following three subsections will be dedicated to one of the steps above, and we will adapt this very high-level definition of the Choi isomorphism to our type system. After that, we will derive properties of the Choi isomorphism. The most remarkable property is that the Choi isomorphism is K-preserving. This generalizes the well-known results for channels and states [66]. Lastly, we will conclude this Section by generalizing the link product.

6.3.1 Orthonormal Bases

As detailed above, the first step to define the Choi isomorphism is to fix an orthonormal basis. Here, our goal is to fix bases for all spaces $L(x)$, with $x \in \text{Types}_{\mathcal{A}}$. As usual, we will start from the smallest components, the elements of the alphabet. For all Hilbert spaces $\mathcal{H}_A$, where $A \in \mathcal{A}$, we fix an orthonormal basis $\left\{ u_i^A \right\}_i$. The orthonormal basis of $\mathcal{H}_E$, associated with an elementary type $E = A_1 \dots A_n$, with $A_i \in \mathcal{A}$, is $\left\{ u_{i_1}^{A_1} \otimes \cdots \otimes u_{i_n}^{A_n} \right\}_{i_1,\dots,i_n}$, where $\left\{ u_{i_j}^{A_j} \right\}_{i_j}$ is the orthonormal basis of A_j. Lastly, we provide the following recursive definition for the orthonormal basis of $L(x)$.

Definition 6.3.1: Fixed orthonormal basis

Let $x \in \text{Types}_{\mathcal{A}}$.

- If $x = E \in \text{EleTypes}_{\mathcal{A}}$, and $\left\{ u_j^E \right\}$ is the orthonormal basis of $\mathcal{H}_E$, then $\left\{ u_j^E (u_i^E)^\dagger \right\}_{i,j}$, is the orthonormal basis of $L(E)$.
- If $x = (a \to b)$ and $\left\{ u_i^a \right\}_i$ and v_j^b are the orthonormal bases of $L(a)$ and $L(b)$, then $\left\{ \left| v_j^b \middle\rangle\!\middle\langle u_i^a \right|^{a \to b} \right\}_{i,j}$, is the orthonormal basis of $L(x)$.

Since promotion and demotion functions play a key role in the formalism for types, a natural question is what the promotion of an element of a fixed orthonormal basis is. When we introduced promotion functions, we stressed the concept that they represent the most natural isomorphism between a space $L(x)$ and the space of promoted functions $L(x_y)$. Therefore, it comes as no surprise that promotion and demotion maps preserve the fixed orthonormal bases.

Proposition 6.3.2: Promotion preserves orthonormal bases

Let $x, y \in \text{Types}_{\mathcal{A}}$, then $\eta^{x \to x_y}$ maps the fixed orthonormal basis of $L(x)$ to the fixed orthonormal basis of $L(x_y)$

Proof. $L(x)$ and $L(x_y)$ have the same dimension and $\eta^{x \to x_y}$ is an isomorphism. It is enough to show that for every u^x in the basis of $L(x)$, $\eta^{x \to x_y}(u^x)$ is an element of the basis of $L(x_y)$. We prove it by induction on $\text{ord}(xy)$. If $\text{ord}(xy) = 0$, then $x_y = x$, $L(x) = L(x_y)$ and $\eta^{x \to x_y} = \text{id}^{x \to x}$. Therefore, each element of the basis is mapped to itself.

Assume now $\text{ord}(xy) > 0$. As usual, we write $x = (a \to b)$ whenever $\text{ord}(x) > 0$ and $y = (c \to d)$ whenever $\text{ord}(y) > 0$. Moreover, in the first case, $u^x = |w^b\rangle\langle u^a|^{a \to b}$ for some v^a in the fixed orthonormal basis of $L(a)$ and w^b in the fixed orthonormal basis of $L(b)$.

- $\mathrm{ord}(x) = \mathrm{ord}(y)$:

  $$\begin{aligned} \eta^{x\to x_y}(u^x) &= \eta^{b\to b_d} \circ \left|w^b\right\rangle^{I\to b} \circ \langle v^a|^{a\to I} \circ \eta^{a_c\to a} \\ &= \left|\eta^{b\to b_d}(w^b)\right\rangle^{I\to b_d} \circ \langle \eta^{a\to a_c}(v^a)|^{a_c\to I} . \end{aligned} \tag{6.149}$$

  By induction hypothesis, $\eta^{b\to b_d}(w^b)$ is an element of the orthonormal basis of $L(b_d)$ and $\eta^{a\to a_c}(v^a)$ of $L(a_c)$. Therefore, by definition, $|\eta^{b\to b_d}(w^b)\rangle\langle\eta^{a\to a_c}(v^a)|^{a_c\to b_d}$ is an element of the orthonormal basis of $L(x_y)$.

- $\mathrm{ord}(y) > \mathrm{ord}(x)$:

  $$\eta^{x\to x_y}(u^x) = |\eta^{x\to x_d}(u^x)\rangle^{I\to x_d} \circ \eta^{I_c\to I}. \tag{6.150}$$

  By induction hypothesis $\eta^{x\to x_d}(u^x)$ is an element of the fixed orthonormal basis of $L(x_d)$. Moreover, by induction hypothesis again, $\eta^{I_c\to I}$ maps the single element in the orthonormal basis of I_c, to the single element in the orthonormal basis of I, that is, $\eta^{I_c\to I} = |1\rangle^{I\to I} \circ \left\langle v^{I_c}\right|^{I_c\to I}$. However, $|1\rangle^{I\to I}$ is just the identity map on I, therefore $\eta^{I_c\to I} = \left\langle v^{I_c}\right|^{I_c\to I}$. As a consequence, $|\eta^{x\to x_d}(u^x)\rangle\langle v^{I_c}|^{I_c\to x_d}$ is an element of the orthonormal basis of $L(x_y)$.

- $\mathrm{ord}(x) > \mathrm{ord}(y)$:

  $$\begin{aligned} \eta^{x\to x_y}(u^x) &= \eta^{b\to b_y} \circ \left|w^b\right\rangle^{I\to b} \circ \langle v^a|^{a\to I} \\ &= \left|\eta^{b\to b_y}(w^b)\right\rangle^{I\to b_y} \circ \langle v^a|^{a\to I} . \end{aligned} \tag{6.151}$$

  As before, by induction hypothesis, $\eta^{b\to b_y}(w^b)$ is an element of the orthonormal basis of $L(x_b)$, and therefore $\left|\eta^{b\to b_y}(w^b)\right\rangle^{I\to b_y} \circ \langle v^a|^{a\to I}$ is an element of the orthonormal basis of $L(x_y)$.

□

With the results regarding promotion and demotion maps, we can turn our attention to the orthonormal bases of product spaces. We show that the fixed orthonormal basis for a product space $L(xy)$ is composed of elementary tensors of elements of the bases of the $L(x)$ and $L(y)$. Note that this is not a new construction: xy is a type, and therefore, the fixed orthonormal basis of $L(xy)$ is defined as in Definition 6.3.1.

Proposition 6.3.3: orthonormal basis of product space

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$, and let $\left\{ u_i^x \right\}_i$, and $\left\{ v_j^y \right\}_j$ be the fixed orthonormal bases of $L(x)$ and $L(y)$. Then, $\left\{ u_i^x \boxtimes v_j^y \right\}_{i,j}$ is the orthonormal basis of $L(xy)$.

Proof. The statement is true by definition for $\mathrm{ord}(xy) = 0$. Let $\mathrm{ord}(xy) > 0$. Then, according to Lemma 6.1.12, we can write $xy = (\alpha\gamma \to \beta\delta)$. In addition, note that $x_y = (\alpha \to \beta)$, and $y = (\gamma \to \delta)$. By induction hypothesis, the basis of $L(\alpha\gamma)$ is $\left\{ u_l^\alpha \boxtimes v_m^\gamma \right\}_{l,m}$ and the basis of $L(\beta\delta)$ is $\left\{ u_p^\beta \boxtimes v_q^\delta \right\}_{pq}$. By construction, the basis of $L(xy)$ is $\left\{ \left| u_p^\beta \boxtimes v_q^\delta \rangle\langle u_l^\alpha \boxtimes v_m^\gamma \right|^{\alpha\beta\to\gamma\delta} \right\}_{l,m,p,q} = \left\{ \left| u_p^\beta \rangle\langle u_l^\alpha \right|^{\alpha\to\gamma} \boxtimes |v_q^\delta\rangle\langle v_m^\gamma|^{\beta\to\delta} \right\}_{l,m,p,q}$. We have used Theorem 6.2.30 and Corollary 6.2.28 in the equality. Now, observe that each $|u_p^\beta\rangle\langle u_l^\alpha|^{\alpha\to\gamma}$ is an element of the basis of $L(x_y)$, therefore, by Proposition 6.3.2 it is equal to $\eta^{x\to x_y}(u_i^x)$ for some i and every element of the basis of $L(x_y)$ is the image of a different element of the basis of $L(x)$. Similarly, $|v_q^\delta\rangle\!\langle v_m^\gamma|^{\beta\to\delta} = \eta^{y\to y_x}(v_j^y)$ for some j. Therefore, the basis of $L(xy)$ is $\left\{ \eta^{x\to x_y}(u_i^x) \boxtimes \eta^{y\to y_x}(v_j^x) \right\}_{i,j} = \left\{ u_i^x \boxtimes v_j^y \right\}_{i,j}$. □

6.3.2 The Choi state

The second step for defining the Choi isomorphism, as detailed at the beginning of this Section, is to define the Choi state. To simplify later proofs, we provide a recursive definition of the Choi state Φ^{xx}.

Definition 6.3.4: Choi state

Let $x \in \mathrm{Types}_{\mathcal{A}}$. The Choi state Φ^{xx} is recursively defined as follows.

- If $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$, then $\Phi^{EE} = \sum_i u_i^E \otimes u_i^E$, where $\left\{ u_i^E \right\}_i$ is the fixed orthonormal basis of $L(E)$.
- If $x = (a \to b)$, then $\Phi^{xx} = |\Phi^{bb}\rangle\!\langle\Phi^{aa}|^{aa\to bb}$.

This definition looks quite different from what we stated at the beginning of this Section and from what a reader would expect. Indeed, the Choi state is usually defined as the sum of the elements of the fixed orthonormal basis of the Hilbert space tensored with their copies. We show in the following Proposition that the two definitions are equivalent.

Proposition 6.3.5: Alternative definition of the Choi state

Let $x \in \mathrm{Types}_{\mathcal{A}}$. Then, $\Phi^{xx} = \sum_i u_i^x \otimes u_i^x$, where $\left\{ u_i^x \right\}_i$ is the fixed orthonormal basis of $L(x)$.

Proof. The base case is true by definition. For the inductive case, we have:

$$\begin{aligned} \Phi^{xx} &= |\Phi^{bb}\rangle\!\langle\Phi^{aa}|^{aa\to bb} \\ \text{I.H.} &= \sum_{j,k} |u_j^b \otimes u_j^b\rangle\!\langle u_k^a \otimes u_k^a|^{aa\to bb} \\ &= \sum_{j,k} |u_j^b\rangle\!\langle v_k^a|^{a\to b} \otimes |u_j^b\rangle\!\langle v_k^a|^{a\to b} \\ &= \sum_i u_i^x \otimes u_i^x, \end{aligned} \tag{6.152}$$

because $|u_j^b\rangle\!\langle v_k^a|^{a\to b}$ is by definition an element of the orthonormal basis of $L(x)$. □

As we have seen in Proposition 6.3.2, the promotion and demotion functions preserve the fixed orthonormal bases. It is, therefore, a straightforward consequence of the Proposition above that they preserve the Choi state as well.

Corollary 6.3.6: Promotion of the Choi state

Let $x, y \in \mathrm{Types}_{\mathcal{A}}$. Then, $\eta^{xx\to xx_y}(\Phi^{xx}) = \Phi^{x_y x_y}$.

Proof. This follows from Proposition 6.2.16.

$$\begin{aligned}\eta^{xx\to xx_y}(\Phi^{xx}) &= (\eta^{x\to x_y}\otimes\eta^{x\to x_y})(\sum_i u_i^x\otimes u_i^x)\\ &= \sum_i \eta^{x\to x_y}(u_i^x)\otimes\eta^{x\to x_y}(u_i^x)\\ \text{Proposition 6.3.2} &= \sum_i u_i^{x_y}\otimes u_i^{x_y}\\ &= \Phi^{x_y x_y}.\end{aligned} \tag{6.153}$$

□

We conclude this subsection with an expression for the Choi state associated with a product of types. This expression generalizes the one presented in Section 5.4 for bipartite systems, where we showed that $\Phi^{ABA'B'} = (\mathcal{I}^A\otimes\mathcal{S}^{A'B\to BA'}\otimes\mathcal{I}^{B'})(\Phi^{AA'}\otimes\Phi^{BB'})$. The natural way to generalize this equation for a generic product type xy is

$$\Phi^{xyxy} = \mathrm{id}^{x_y\to x_y}\boxtimes\sigma^{xy\to yx}\boxtimes\mathrm{id}^{y_x\to y_x} @ \Phi^{xx}\boxtimes\Phi^{yy}. \tag{6.154}$$

Observe that we have used the identities on the promoted spaces $L(x_y)$ and $L(y_x)$ to ensure that $\mathrm{id}^{x_y\to x_y}\boxtimes\sigma^{xy\to yx}\boxtimes\mathrm{id}^{y_x\to y_x}\in L(xxyy)\to L(xyxy)$. In the following Proposition, we prove that the expression in Eq. (6.154) is indeed the Choi state of Φ^{xyxy}.

Proposition 6.3.7: Choi state of product of types

Let $x, y\in\mathrm{Types}_{\mathcal{A}}$. Then,

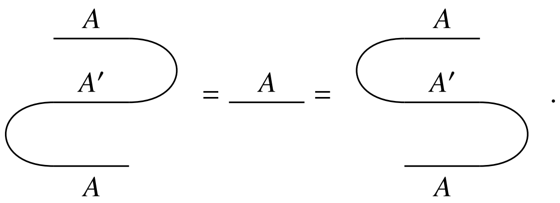
$$\Phi^{xyxy} = \mathrm{id}^{x_y\to x_y}\boxtimes\sigma^{xy\to yx}\boxtimes\mathrm{id}^{y_x\to y_x} @ \Phi^{xx}\boxtimes\Phi^{yy}. \tag{6.155}$$

Proof. We start from the right-hand side of the equation. First, we write $\Phi^{xx}\boxtimes\Phi^{yy}$ as $\sum_{i,j} u_i^x\boxtimes u_i^x\boxtimes u_j^y\boxtimes u_j^y$. This expression is equal to $\sum_{i,j} u_i^{x_y}\boxtimes u_i^x\boxtimes u_j^y\boxtimes u_j^{y_x}$, where we have promoted the first and the last elements of the elementary tensors. Therefore,

$$\begin{aligned}&\mathrm{id}^{x_y\to x_y}\boxtimes\sigma^{xy\to yx}\boxtimes\mathrm{id}^{y_x\to y_x} @ \Phi^{xx}\boxtimes\Phi^{yy}\\ &\quad= \sum_{i,j}\mathrm{id}^{x_y\to x_y}\boxtimes\sigma^{xy\to yx}\boxtimes\mathrm{id}^{y_x\to y_x} @ u_i^{x_y}\boxtimes u_i^x\boxtimes u_j^y\boxtimes u_j^{y_x}\\ &\quad= \sum_{i,j} u_i^{x_y}\boxtimes u_j^y\boxtimes u_i^x\boxtimes u_j^{y_x}\\ &\quad= \sum_{i,j} u_i^x\boxtimes u_j^y\boxtimes u_i^x\boxtimes u_j^y.\end{aligned} \tag{6.156}$$

Now, note that $u_i^x\boxtimes u_j^y$ is an element of the basis of $L(xy)$, therefore, by Proposition 6.3.5, one obtains $\sum_{i,j} u_i^x\boxtimes u_j^y\boxtimes u_i^x\boxtimes u_j^y = \Phi^{xyxy}$. □

We conclude this Section by showing that as the Choi state $\Phi^{AA'}$ satisfies the 'snake equations' (see Section 5.4), so does the Choi state Φ^{xx} for all types x. As a reminder, the snake equations are



(6.157)

To translate these into equations for Φ^{xx}, note that the application of the Choi effect sends any bipartite state $M_{AA'}$ to $\mathrm{Tr}_{AA'}[\Phi^{AA'}M_{AA'}] = (M_{AA'}, \Phi^{AA'})_{AA'} = \left\langle \Phi^{AA'} \right|^{AA'} @ M_{AA'}$. With this notation, the first equation consists of sequentially composing $\left\langle \Phi^{AA'} \right|^{AA'} \otimes \mathcal{I}^A$ with $\mathcal{I}^A \otimes \left| \Phi^{AA'} \right\rangle^{AA'}$, where we have promoted the state $\Phi^{AA'}$ to the preparation channel $\left| \Phi^{AA'} \right\rangle^{AA'}$. If one wants to write this expression for a generic type x, one gets

$$(\langle \Phi^{xx}|^{xx\to I} \boxtimes \mathrm{id}^{x\to x}) \circ (\mathrm{id}^{x\to x} \boxtimes |\Phi^{xx}\rangle^{I\to xx}). \tag{6.158}$$

In the following Lemma, we show that this and the expression corresponding to the other diagram in Eq. (6.157) are equal to the identity.

Lemma 6.3.8: Snake equations

Let $x \in \mathrm{Types}_{\mathcal{A}}$. Then,

$$\begin{aligned} (\mathrm{id}^{x\to x} \boxtimes \langle \Phi^{xx}|^{xx\to I}) \circ (|\Phi^{xx}\rangle^{I\to xx} \boxtimes \mathrm{id}^{x\to x}) &= \mathrm{id}^{x\to x}, \\ (\langle \Phi^{xx}|^{xx\to I} \boxtimes \mathrm{id}^{x\to x}) \circ (\mathrm{id}^{x\to x} \boxtimes |\Phi^{xx}\rangle^{I\to xx}) &= \mathrm{id}^{x\to x}. \end{aligned} \tag{6.159}$$

Proof. Using Proposition 6.3.5 one gets:

$$\begin{aligned} (\mathrm{id}^{x\to x} \boxtimes \langle \Phi^{x} x|^{xx\to I}) \circ (|\Phi^{xx}\rangle^{I\to xx} \boxtimes \mathrm{id}^{x\to x}) &= \sum_{i,j} (\left|u_i^x\right\rangle^{I\to x} \boxtimes \left\langle u_j^x \middle| u_i^x \right\rangle^{I\to I}) \boxtimes \left\langle u_j^x \right|^{x\to I} \\ &= \sum_i \left|u_i^x\right\rangle^{I\to x} \boxtimes \left\langle u_i^x\right|^{x\to I} \\ \text{Lemma 6.2.33} &= \sum_i |u_i^x\rangle\langle u_i^x|^{x\to x} \\ &= \mathrm{id}^{x\to x}. \end{aligned} \tag{6.160}$$

The last equality is just the decomposition of the identity on an orthonormal basis. The proof of the second equation in the Lemma is identical to the proof of the first. □

6.3.3 The Choi Map

We have extensively discussed the Choi isomorphism for channels in Chapter 5. The Choi matrix associated with the channel $\mathcal{M}^{A\to B}$ is $(\mathcal{M}^{A\to B} \otimes \mathcal{I}^{A'})(\Phi^{AA'})$. In the language of types, we can say that the Choi isomorphism maps a linear function in $L(x)$, where $x = (A \to B)$, into a function in $L(BA)$. Extending this concept to a map Γ^x associated with any type $x = (a \to b)$, we would define the output of the Choi isomorphism as $\Gamma^{a\to b}\boxtimes \mathrm{id}^{a\to a} @\Phi^{aa}$. However, there is no guarantee that the input type of $\Gamma^{a\to b}\boxtimes \mathrm{id}^{a\to a}$ matches the type of Φ^{aa}. Indeed, $(a\to b)|(a\to a)$ is either $aa\to ba$ or $a\to b(a\to a)$, depending on the relative order of a and b. To fix this, we use the identity on the promoted space a_b. This guarantees that $\Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b}\in L(aa_b\to ba)$. To match the input type of this map, we use the promoted version of Φ^{aa}, that is, $\Phi^{a_b a_b}$.

Definition 6.3.9: The Choi map

Let $x=(a\to b)\in \mathrm{Types}_{\mathcal{A}}$, and let $\Gamma^{x}\in L(x)$. The Choi map $C^{x\to ba}$ is defined as

$$C^{x\to ba}(\Gamma^{x})=\Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b} @\Phi^{a_b a_b}. \tag{6.161}$$

It is handy to express the Choi map with respect to the fixed orthonormal basis. This will simplify many proofs later on. Indeed, a standard strategy that we will use is to prove a statement for an element of the fixed orthonormal basis and then extend it by linearity to all vectors in the space.

Lemma 6.3.10: Equivalent expression for the Choi map

Let $x=(a\to b)\in \mathrm{Types}_{\mathcal{A}}$. $C^{x\to ba}(\Gamma^{x})=\sum_i \Gamma^{a\to b}(u_i^a)\boxtimes u_i^a$, where $\left\{\, u_i^a \,\right\}_i$ is the fixed orthonormal basis associated with $L(a)$.

Proof. The proof is straightforward. It follows directly from the definition of the tensor product, promotion and demotion functions, Theorem 6.2.27. First, we express the Choi state as a sum of elementary products of elements of the orthonormal basis.

$$\Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b} @\Phi^{a_b a_b}=\sum_i \Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b} @u_i^{a_b}\boxtimes u_i^{a_b}. \tag{6.162}$$

Then, we observe that $\Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b} = \eta^{(a\to b)\to(a\to b)_{(a\to a)}}(\Gamma^{a\to b})\boxtimes \mathrm{id}^{a_b\to a_b\to a_b\to a_b}$, which we can write as $\eta^{b\to b_a}\circ\Gamma^{a\to b}\circ\eta^{a_b\to b}\boxtimes \mathrm{id}^{a_b\to a_b}$. Therefore,

$$\begin{aligned}\Gamma^{a\to b}\boxtimes \mathrm{id}^{a_b\to a_b} @\Phi^{a_b a_b} &= \sum_i \eta^{b\to b_a}\circ\Gamma^{a\to b}\circ\eta^{a_b\to b}\boxtimes \mathrm{id}^{a_b\to a_b} @u_i^{a_b}\boxtimes u_i^{a_b}\\ \text{Theorem 6.2.27} &= \sum_i \eta^{b\to b_a}\circ\Gamma^{a\to b}\circ\eta^{a_b\to b}(u_i^{a_b})\boxtimes \mathrm{id}^{a_b\to a_b}(u_i^{a_b})\\ &= \sum_i \eta^{b\to b_a}\circ\Gamma^{a\to b}(u_i^{a})\boxtimes u_i^{a_b}.\end{aligned} \tag{6.163}$$

Observe that $\eta^{b\to b_a}\circ\Gamma^{a\to b}(u_i^a)\boxtimes u_i^{a_b}=\eta^{b\to b_a}(\Gamma^{a\to b}(u_i^a)\boxtimes\eta^{a\to a_b}(u_i^a)$ which is by definition $\Gamma^{a\to b}(u_i^a)\boxtimes u_i^a$. This shows that $C^{x\to ba}(\Gamma^{x})=\sum_i \Gamma^{a\to b}(u_i^a)\boxtimes u_i^a$. ☐

Defining the Choi map is only half of the process of defining the Choi isomorphism. We also need to find the inverse map. We use the inverse map of the Choi isomorphism for channels as a guide. If $M_{BA'}$ is a bipartite matrix, the channel $\mathcal{M}^{A\to B}$ associated with $M_{BA'}$ acts on a state ρ^A as

$$\mathcal{M}^{A\to B}(\rho^A)=\mathrm{Tr}_{A'A}[(\mathbb{1}^B\otimes\Phi^{A'A})(M_{A'A}\otimes\rho^A)]. \tag{6.164}$$

It is not immediate to translate this expression into our type system because we do not have a notion of partial tracing. However, it is immediate to show that $\mathcal{M}^{A\to B}(\rho^A) = (\mathcal{I}^B \otimes \left\langle \Phi^{A'A} \right|)(M_{BA'} \otimes \rho^A)$. Indeed, $\left\langle \Phi^{A'A} \right|$ is the map that sends a matrix $N_{A'A}$ to $\mathrm{Tr}_{A'A}(\Phi^{A'A} N_{A'A})$. Now, it is easier to rewrite this expression for generic types a and b, a function $\Gamma^{ba} \in L(ba)$, and a test vector $u^a \in L(a)$. A first attempt gives

$$\mathrm{id}^{b\to b} \boxtimes \langle \Phi^{aa}|^{aa\to I} @ \Gamma^{ba} \boxtimes u^a. \tag{6.165}$$

However, there is no guarantee that the input of $\mathrm{id}^{b\to b} \boxtimes \langle \Phi^{aa}|^{aa\to I}$ is baa, as we need. To fix this, we use the standard method of promoting types.

$$\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I} @ \Gamma^{ba} \boxtimes u^a. \tag{6.166}$$

Now the input type of $\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I}$ is baa as we need. However, the output type is b_a, and not b as we would like in analogy to the case of channels and states. Therefore, the last step is to reduce b_a to b.

$$\eta^{b_a\to b}(\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I} @ \Gamma^{ba} \boxtimes u^a) \tag{6.167}$$

This is our candidate for the inverse of the Choi map.

Definition 6.3.11: Inverse Choi map

Let $x = (a \to b) \in \mathrm{Types}_{\mathcal{A}}$. The inverse Choi map $\mathcal{IC}^{ba\to x}$ is defined as follows. For every $\Gamma^{ba} \in L(ba)$, $\mathcal{IC}^{ba\to x}(\Gamma^{ba})$ is the map in $L(x)$ that acts on $u^a \in L(a)$ as

$$\mathcal{IC}^{ba\to x}(\Gamma^{ba}) @ u^a = \eta^{b_a\to b}(\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I} @ \Gamma^{ba} \boxtimes u^a). \tag{6.168}$$

In the following Theorem, we show that the map defined in Definition 6.3.11 is indeed the inverse of the Choi map.

Theorem 6.3.12: The Choi isomorphism

Let $x = (a \to b)$, then $C^{x\to ba}$ is an isomorphism, and $\mathcal{IC}^{ba\to x}$ is the inverse of $C^{x\to ba}$.

Proof. We show that $\mathcal{IC}^{ba\to x}$ is both left and right inverse of $C^{x\to ba}$.

Let $\Gamma^{ba} \in L(ba)$. We want to show that $C^{x\to ba}(\mathcal{IC}^{ba\to x}(\Gamma^{ba})) = \Gamma^{ba}$. We start by writing the Choi map in terms of the fixed orthonormal basis.

$$C^{x\to ba}(\mathcal{IC}^{ba\to x}(\Gamma^{ba})) = \sum_i [\mathcal{IC}^{ba\to x}(\Gamma^{ba}) @ u_i^a] \boxtimes u_i^a. \tag{6.169}$$

Now, we use the definition of $\mathcal{IC}^{ba\to x}$ to get

$$\sum_i [\eta^{b_a\to b}(\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I} @ \Gamma^{ba} \boxtimes u_i^a)] \boxtimes u_i^a. \tag{6.170}$$

By definition of the square product, we can promote the term on the left of $\boxtimes$ with $\eta^{b\to b_a}$. However $\eta^{b\to b_a} \circ \eta^{b_a\to b} = \mathrm{id}^{b_a\to b_a}$. Therefore, the expression above becomes

$$\sum_i (\mathrm{id}^{b_a\to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b\to I} @ \Gamma^{ba} \boxtimes u_i^a) \boxtimes u_i^a. \tag{6.171}$$

Now, we observe that $\Gamma^{ba} \boxtimes u_i^{a} = \Gamma^{ba} \boxtimes u_i^{a_b}$, where $u_i^{a_b}$ is the promotion of u_i^{a}. Moreover, thanks to Corollary 6.2.28, we can write $\langle \Phi^{a_b a_b}|^{a_b a_b \to I}$ as $\sum_j \left\langle u_j^{a_b} \right|^{a_b \to I} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I}$. Thus, Eq. (6.171) becomes,

$$\sum_{i,j} (\mathrm{id}^{b_a \to b_a} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} @\Gamma^{ba} \boxtimes u_i^{a_b}) \boxtimes u_i^{a}. \tag{6.172}$$

Theorem 6.2.27 allows us to split the action of the terms between round brackets and obtain

$$(\mathrm{id}^{b_a \to b_a} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} @\Gamma^{ba}) \boxtimes (\left\langle u_j^{a_b} \right|^{a_b \to I} @ \boxtimes u_i^{a_b}). \tag{6.173}$$

Since $\{ u_i^{a_b} \}_i$ is an orthonormal basis, the term on the right is $\delta_{i,j}$. As a consequence, we can write the equation above as

$$\sum_{i} (\mathrm{id}^{b_a \to b_a} \boxtimes \langle u_i^{a_b} |^{a_b \to I} @\Gamma^{ba}) \boxtimes u_i^{a}. \tag{6.174}$$

Now, we use the fact that $L(ba) = L(b_a) \boxtimes L(a_b)$ to write Γ^{ba} as a linear combination of elementary tensors, that is $\Gamma^{ba} = \sum_{m,n} G_{m,n} u_m^{b_a} \boxtimes u_n^{a_b}$. With that, we obtain

$$\sum_{i,m,n} (\mathrm{id}^{b_a \to b_a} \boxtimes \langle u_i^{a_b} |^{a_b \to I} @G_{m,n} u_m^{b_a} \boxtimes u_n^{a_b}) \boxtimes u_i^{a}. \tag{6.175}$$

Once again, we split the action of the square product, and we observe that $\langle u_i^{a_b} |^{a_b \to I} @u_n^{a_b} = \delta_{i,n}$. Therefore, Eq. (6.171) becomes $\sum_{m,n} G_{m,n} u_m^{b_a} \boxtimes u_n^{a} = \Gamma^{ba}$. This shows $C^{x \to ba}(\mathcal{I}C^{ba \to x}(\Gamma^{ba})) = \Gamma^{ba}$.

Let now $\Gamma^{x} \in L(x)$, and let u_k^{a} be an element of the orthonormal basis of $L(a)$. We want to show that $\mathcal{I}C^{ba \to x}(C^{x \to ba}(\Gamma^{x}))@u_k^{a} = \Gamma^{x}(u_k^{a})$. Using the same concepts that we used in the proof of the previous case, we get:

$$\begin{aligned}
&\mathcal{I}C^{ba \to x}(C^{x \to ba}(\Gamma^{x}))@u_k^{a} = \eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @C^{x \to ba}(\Gamma^{x}) \boxtimes u_k^{a}) \\
&= \sum_{i} \eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @\Gamma^{x}(u_i^{a}) \boxtimes u_i^{a} \boxtimes u_k^{a}) \\
&= \sum_{i} \eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @\eta^{b \to b_a}(\Gamma^{x}(u_i^{a})) \boxtimes u_i^{a_b} \boxtimes u_k^{a_b}) \\
&= \sum_{i,j} \eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} @\eta^{b \to b_a}(\Gamma^{x}(u_i^{a})) \boxtimes u_i^{a_b} \boxtimes u_k^{a_b}) \\
&= \sum_{i,j} \eta^{b_a \to b}(\eta^{b \to b_a}(\Gamma^{x}(u_i^{a})) \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} (u_i^{a_b}) \boxtimes \left\langle u_j^{a_b} \right|^{a_b \to I} (u_k^{a_b})) \\
&= \eta^{b_a \to b}(\eta^{b \to b_a}(\Gamma^{x}(u_k^{a}))) \\
&= \Gamma^{x}(u_k^{a}).
\end{aligned} \tag{6.176}$$

This is true for every element of the fixed orthonormal basis of $L(a)$, and therefore, by linearity, we have that $\mathcal{I}C^{ba \to x}(C^{x \to ba}(\Gamma^{x}))@v^{a} = \Gamma^{x}(v^{a})$ for every $v^{a} \in L(a)$. This shows that $\Gamma^{x} = \mathcal{I}C^{ba \to x}(C^{x \to ba}(\Gamma^{x}))$.

□

6.3.4 Properties of the Choi Isomorphism

In the previous Subsections, we were able to define the Choi isomorphism associated with any type $x = (a \to b)$, generalizing the well-known isomorphism from maps in $\mathfrak{L}(A \to B)$. However, what makes the Choi isomorphism so interesting is the easy translation of properties of a map in $\mathfrak{L}(A \to B)$ to properties of the corresponding matrix on BA. For example, a channel $\mathcal{M}^{A\to B}$ is Hermitian-preserving if and only if its Choi matrix M_{BA} is Hermitian and completely positive-preserving if its Choi matrix is positive semidefinite. Here, we show that these properties are true for maps of type $x = (a \to b)$, generalizing what happens in the case of channels.

We begin by demonstrating that the Choi map preserves the inner product, and therefore it is unitary.

Proposition 6.3.13: The Choi map is unitary

Let $x \in ty$, $\Gamma^x, \Delta^x \in L(x)$. Then, $(\Gamma^x, \Delta^x)_x = (C^{x\to ba}(\Gamma^x), C^{x\to ba}(\Delta^x))_{ba}$.

Proof. We write the inner product $(C^{x\to ba}(\Gamma^x), C^{x\to ba}(\Delta^x))_{ba}$ expanding the Choi isomorphism in terms of the fixed orthonormal basis of $L(ba)$.

$$(C^{x\to ba}(\Gamma^x), C^{x\to ba}(\Delta^x))_{ba} = \sum_{i,j} (\Gamma^x(u_i^a) \boxtimes u_i^a, \Delta^x(u_j^a) \boxtimes u_j^a). \tag{6.177}$$

Using Corollary 6.2.23, we can split the inner product of the tensor space into the product of two inner products.

$$\sum_{i,j} (\Gamma^x(u_i^a) \boxtimes u_i^a, \Delta^x(u_j^a) \boxtimes u_j^a) = \sum_{i,j} (\Gamma^x(u_i^a), \Delta^x(u_j^a))_b (u_i^a, u_j^a)_a. \tag{6.178}$$

The second inner product is $\delta_{i,j}$. Therefore the expression above becomes $\sum_i (\Gamma^x(u_i^a), \Delta^x(u_i^a))_b$. To conclude the proof, we notice that this expression is the definition of the inner product $(\Gamma^x, \Delta^x)_x$. □

We are ready to prove the main Theorem of this subsection, the one where we show that a map Γ^x, where $x = (a \to b)$, is in $K(x)$ if and only if $C^{x\to ba}(\Gamma^x) \in K(ba)$. Note that if $x = (A \to B)$, this reduces to the known results about channels. While proving these statements, we will prove the other four properties. This simplifies the proof because it allows us to use more inductive hypotheses.

Theorem 6.3.14: The Choi isomorphism is K-preserving

Let $x \in \mathrm{Types}_{\mathcal{A}}$.

1. $H(x)$ has a basis composed of elements of $K(x)$.
2. $\Phi^{xx} \in K(xx)$.
3. If $x = (a \to b)$, $C^{x\to ba}$ and $\mathcal{I}C^{ba\to x}$ are K-preserving.
4. $K(x)$ is self dual, i.e., $K^*(x) = K(x)$, where $K^* = \{\, \Gamma^x \in H(x) \mid (\Delta^x, \Gamma^x)_x \geq 0, \forall \Delta^x \in K(x) \,\}$.
5. $\Gamma^x \in K(x)$ implies $\langle \Gamma^x|^{x\to I} \in K(I \to x)$.

Proof. This is one of the cases in which we prove multiple statements at the same time by induction to be able to use the inductive hypothesis for all the statements. Since there is no base case for statement *3*, we cannot use its inductive hypothesis. Moreover, we can use already proven statements to prove the subsequent statement.

Base case: The base case for statements *1*, *2* and *4* follows from properties of positive semidefinite matrices. For statement *5*, let $x = E \in \mathrm{EleTypes}_{\mathcal{A}}$, and let $F \in \mathrm{EleTypes}_{\mathcal{A}}$, we have to show that $\langle \Gamma^x|^{x \to I} \boxtimes \mathrm{id}^{F \to F}$ is K-preserving. Let $\Delta^{EF} \in K(EF)$, that is, K_{EF} is a positive semidefinite matrix. Note that, for $\Theta^F \in K(F)$ we have

$$\begin{aligned}(\Theta^F, (\left\langle \Gamma^E \right|^{E \to I} \boxtimes \mathrm{id}^{F \to F})(\Delta^{EF}))_F &= ((\left|\Gamma^E\right\rangle^{I \to E} \boxtimes \mathrm{id}^{F \to F})(\Theta^F), \Delta^{EF})_{EF} \\ &= (\Gamma^E \otimes \Theta^F, \Delta^{EF})_{EF} \\ &\geq 0,\end{aligned} \tag{6.179}$$

because the inner product of two positive semidefinite matrices is positive. By self duality, we have $(\left\langle \Gamma^E \right|^{E \to I} \boxtimes \mathrm{id}^{F \to F})(\Delta^{EF}) \in K(F)$, and therefore $\left\langle \Gamma^E \right|^{E \to I} \boxtimes \mathrm{id}^{F \to F}$ is K-preserving.

Inductive step:

1. By induction hypothesis *1* there exists a basis of $H(a)$ such that $\left\{ u_i^a \right\}_i \subseteq K(a)$ and similarly, a basis of $H(b)$ such that $\left\{ u_j^b \right\}_j \subseteq K(b)$. The set $\left\{ \left| u_i^b \middle\rangle\!\middle\langle u_j^a \right|^{a \to b} \right\}^{i,j}$ is a basis of $H(x)$. Moreover, $\left|u_i^b\right\rangle^{I \to b} \in K(I \to b)$ because of Theorem 6.2.47 and $\left\langle u_i^a\right|^{a \to I} \in K(a \to I)$ by inductive hypothesis *5*. The sequential composition of maps in K is in K (Lemma 6.2.49) and therefore $|u_i^b\rangle\!\langle u_j^a|^{a \to b} \in K(x)$.

2. By inductive hypothesis *2*, $\Phi^{aa} \in K(aa)$ and $\Phi^{bb} \in K(bb)$. Therefore, by inductive hypothesis *5* and Lemma 6.2.49 $\Phi^{xx} = |\Phi^{bb}\rangle\!\langle\Phi^{aa}|^x \in K(xx)$.

3. Let $\Gamma^x \in K(x)$. Thus, $\Gamma^x \boxtimes \mathrm{id}^{a_b \to a_b}$ is K-preserving. Since by induction hypothesis *2* $\Phi^{a_b a_b} \in K(a_b a_b)$, we have $C^{x \to ba}(\Gamma^x) = \Gamma^x \boxtimes \mathrm{id}^{a_b \to a_b} @ \Phi^{a_b a_b} \in K(ba)$. Now, let $\Gamma^{ba} \in K(ba)$. We have to show that $\mathcal{I}C^{ba \to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y \to y}$ is K-preserving, where $\mathrm{ord}(y) = \mathrm{ord}(ab) = \mathrm{ord}(x-1)$. In the next chain of equalities, we have the goal to write $\mathcal{I}C^{ba \to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y \to y}$ as the sequential composition of K-preserving functions. As usual, we use an element $u_i^a \boxtimes u_j^y$ of the fixed orthonormal basis of $L(ay)$ to prove that two functions are equal.

$$\begin{aligned}&\mathcal{I}C^{ba \to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y \to y} @ u_i^a \boxtimes u_j^y = [\mathcal{I}C^{ba \to x}(\Gamma^{ba}) @ u_i^a] \boxtimes \mathrm{id}^{y \to y}(u_j^y) \\ &= [\eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @ \Gamma^{ba} \boxtimes u_i^a)] \boxtimes \mathrm{id}^{y \to y}(u_j^y) \\ &= [\eta^{b_a \to b}(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @ \Gamma^{ba} \boxtimes u_i^a)] \boxtimes \eta^{y_a \to y} \circ \eta^{y \to y_a}(u_j^y) \\ &= (\eta^{b_a \to b} \boxtimes \eta^{y_a \to y}) @ [(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @ \Gamma^{ba} \boxtimes u_i^a) \boxtimes \eta^{y \to y_a}(u_j^y)] \\ &= \eta^{b y_a \to b y} @ [(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @ \Gamma^{ba} \boxtimes u_i^a) \boxtimes \eta^{y \to y_a}(u_j^y)] \\ &= \eta^{b y_a \to b y} @ [(\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} @ \Gamma^{ba} \boxtimes u_i^a) \boxtimes u_j^y] \\ &= \eta^{b y_a \to b y} @ (\mathrm{id}^{b_a \to b_a} \boxtimes \langle \Phi^{a_b a_b}|^{a_b a_b \to I} \boxtimes \mathrm{id}^{y \to y} @ \Gamma^{ba} \boxtimes u_i^a \boxtimes u_j^y).\end{aligned} \tag{6.180}$$

By linearity, we have that for every $\Delta^{ay} \in K(ay)$

$$\begin{aligned} &\mathcal{I}C^{ba\to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y\to y} @\Delta^{ay} \\ &= \eta^{by_a\to by} @(\mathrm{id}^{b_a\to b_a} \boxtimes \langle\Phi^{a_b a_b}|^{a_b a_b\to I} \boxtimes \mathrm{id}^{y\to y} @\Gamma^{ba} \boxtimes \Delta^{ay}). \end{aligned} \tag{6.181}$$

We have shown that the product of maps in K is in K. Moreover, promotion and identity maps are K-preserving and by induction hypothesis *2* and *5* so $\langle\Phi^{a_b a_b}|^{a_b a_b\to I}$ is K-preserving as well. Therefore, $\mathcal{I}C^{ba\to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y\to y} @\Delta^{ay} \in K(by)$, which shows that $\mathcal{I}C^{ba\to x}(\Gamma^{ba}) \boxtimes \mathrm{id}^{y\to y}$ is K-preserving and thus $\mathcal{I}C^{ba\to x}(\Gamma^{ba}) \in K(x)$.

4. Let $\Gamma^x \in K(x)$. By the proof of the previous point, $\Gamma^x \in K(x)$ if and only if $C^{x\to ba}(\Gamma^x) \in K(ba)$. Now, using the inductive hypothesis *4*, we have that $C^{x\to ba}(\Gamma^x) \in K(ba)$ if and only if $(\Delta^{ba}, C^{x\to ba}(\Gamma^x))_{ba} \geq 0$ for all $\Delta^{ba} \in K(ba)$. Now, since the Choi map is an isomorphism between $K(ba)$ and $K(x)$, as for the previous point, the above statement is equivalent to $(C^{x\to ba}(\Delta^x), C^{x\to ba}(\Gamma^x))_{ba} \geq 0$ for all $\Delta^x \in K(x)$. Lastly, since the Choi map is an isometry (Proposition 6.3.13), the previous statement becomes $(\Delta^x, \Gamma^x)_x \geq 0$ for all $\Delta^x \in K(x)$, which implies $\Gamma^x \in K_*$.

5. Let $\Gamma^x \in K(x)$, to show that $\langle\Gamma^x|^{x\to I} \in K(x \to I)$ we have to show that $\langle\Gamma^x|^{x\to I} \boxtimes \mathrm{id}^{y\to y}$ is K-preserving for all $y \in \mathrm{Types}_{\mathcal{A}}$ such that $\mathrm{ord}(y) = \mathrm{ord}(x)$. Let $\Delta^{xy} \in K(xy)$, we have to show that $\langle\Gamma^x|^{x\to I} \boxtimes \mathrm{id}^{y\to y} @\Delta^{xy} \in K(y)$. Since, $\mathrm{ord}(y) = \mathrm{ord}(x)$, $K(y)$ is self dual and it is equivalent to show that $(\Theta^y, \langle\Gamma^x|^{x\to I} \boxtimes \mathrm{id}^{y\to y} @\Delta^{xy})_y \geq 0$ for all $\Theta^y \in K(y)$.

$$\begin{aligned} (\Theta^y, (\langle\Gamma^x|^{x\to I} \boxtimes \mathrm{id}^{y\to y})(\Delta^{xy}))_y &= ((|\Gamma^x\rangle^{I\to x} \boxtimes \mathrm{id}^{y\to y})(\Theta^y), \Delta^{xy})_{xy} \\ &= (\Gamma^x \otimes \Theta^y, \Delta^{xy})_{xy} \\ &\geq 0, \end{aligned} \tag{6.182}$$

because $\Gamma^x \otimes \Theta^y, \Delta^{xy} \in K(xy)$ and $K(xy)$ is self dual.

□

This long Theorem gave us much information about the geometry of the cone K. First, it says that it is generating, that is, every element $\Gamma^x \in H(x)$ can be written as $\Gamma^x = \Delta^x - \Theta^x$, with $\Delta^x, \Theta^x \in K(x)$. Moreover, the cone is self-dual, which is quite a valuable property. For example, it becomes easier to write the dual conic linear programs based on these cones.

An easy corollary of Theorem 6.3.14 is the fact that the Choi isomorphism is H-preserving as well.

Corollary 6.3.15: The Choi isomorphism is H-preserving

Let $x = (a \to b) \in \mathrm{Types}_{\mathcal{A}}$. Then, $C^{x\to ba}$ and $\mathcal{I}C^{ba\to x}$ are H-preserving.

Proof. Let $\Gamma^x \in H(x)$. Since $K(x)$ is generating, we can write $\Gamma^x = \Delta^x - \Theta^x$, with $\Delta^x, \Theta^x \in K(x)$. By linearity, $C^{x\to ba}(\Gamma^x) = C^{x\to ba}(\Delta^x) - C^{x\to ba}(\Theta^x)$. Since the Choi isomorphism is K-preserving, both $C^{x\to ba}(\Delta^x)$ and $C^{x\to ba}(\Theta^x)$ are in $K(ba) \subseteq H(ba)$. This implies that $C^{x\to ba}(\Gamma^x)$ is a real combination of elements of $H(ba)$, and therefore, it is in $H(ba)$. This proves that $C^{x\to ba}$ is H-preserving. The proof that $\mathcal{I}C^{ba\to x}$ is H-preserving is analogous. □

Note that if $x = (A \to B)$, the above Lemma states that a channel is H-preserving if and only if its Choi matrix is Hermitian, reproducing the known results for channels.

Thanks to the Choi isomorphism, one can reduce a map Γ^x of order n to a map of order $n-1$. This process can be iterated until the map reaches order 0. We name the matrix obtained with such iteration 'the Choi matrix' of Γ^x. Thanks to the results about the Choi isomorphism being both K-preserving and H-preserving, one immediately concludes that a map is in K if and only if its Choi matrix is positive semidefinite, and a matrix is in H if and only if its Choi matrix is Hermitian.

6.3.5 The Link Product

Now, we have a way of converting maps of type $x = (a \to b)$ to maps of type ba, which is of lower order. A natural question is whether we can also convert operations between maps. For example, in the case of channels, if $\mathcal{M}^{A\to B}$ and $\mathcal{N}^{B\to C}$ are channels with Choi matrices $M_{BA'}$ and $N_{CB'}$, then the Choi matrix of $\mathcal{N}^{B\to C}\circ\mathcal{M}^{A\to B}$ is $N_{CB'} * M_{BA'} = \mathrm{Tr}_{BB'}[(\mathbb{1}^C \otimes \Phi^{B'B} \otimes \mathbb{1}^{A'})(N_{CB'} \otimes M_{BA'})]$. As we have discussed already several times, to transform this into an equation for maps of arbitrary types, we can think of $\mathrm{Tr}_{BB'}[\Phi^{B'B}\cdot]$ as the action of $\langle\Phi^{B'B}\big|^{B'B}$. A first attempt is to write the link product between two maps Γ^{ba} and Δ^{cb} as

$$\mathrm{id}^{c\to c} \boxtimes \langle\Phi^{bb}\big|^{bb\to I} \boxtimes \mathrm{id}^{a\to a} @ \Delta^{cb} \boxtimes \Gamma^{ba}. \tag{6.183}$$

The reader who is now familiar with the type system would immediately realize that the input type of $\mathrm{id}^{c\to c} \boxtimes \langle\Phi^{bb}\big|^{bb\to I} \boxtimes \mathrm{id}^{a\to a}$ may not be $cbba$, as is the type of $\Delta^{cb} \boxtimes \Gamma^{ba}$. The standard quick fix is to promote all the types:

$$\mathrm{id}^{c_{ba}\to c_{ba}} \boxtimes \langle\Phi^{b_{ac}b_{ac}}\big|^{b_{ac}b_{ac}\to I} \boxtimes \mathrm{id}^{a_{bc}\to a_{bc}} @ \Delta^{cb} \boxtimes \Gamma^{ba}. \tag{6.184}$$

The expression above fixes the issue of the input type but raises another issue. Now, the output type is $c_b a_b$, and not just ca, as we would want in analogy to the link product of matrices. The solution is to use $\eta^{c_{ab}\to c}$ and $\eta^{a_{bc}\to a}$ instead of $\mathrm{id}^{c_{ab}\to c_{ab}}$ and $\mathrm{id}^{a_{bc}\to a_{bc}}$. In this way, the extra b is removed from the output type, and the input type is still $cbba$. Therefore, our candidate for the link product is the following.

Definition 6.3.16: Link product

Let $a, b, c \in ty$, $\Gamma^{ba} \in L(ba)$, and $\Delta^{cb} \in L(cb)$. The link product '$*$': $L(cb) \times L(ba) \to L(ca)$ is defined as follows:

$$\Delta^{cb} * \Gamma^{ba} = \eta^{c_{ba}\to c} \boxtimes \langle\Phi^{b_{ac}b_{ac}}\big|^{b_{ac}b_{ac}\to I} \boxtimes \eta^{a_{bc}\to a} @ \Delta^{cb} \boxtimes \Gamma^{ba}. \tag{6.185}$$

The link product defined above has the property that we were looking for. It is the operation that corresponds to the sequential composition of maps.

Theorem 6.3.17: Link product and sequential composition

Let $x = (a \to b) \in \mathrm{Types}_{\mathcal{A}}$, $y = (b \to c) \in \mathrm{Types}_{\mathcal{A}}$. Let $\Gamma^x \in L(x)$, $\Delta^y \in L(y)$, $\Theta^{ba} \in L(ba)$, $\Xi^{cb} \in L(cb)$. Then,

1. $C^{(a\to c)\to ca}(\Delta^{b\to c} \circ \Gamma^{a\to b}) = C^{y\to cb}(\Delta^y) * C^{x\to ba}(\Gamma^a)$,
2. $\mathcal{I}C^{ca\to(a\to c)}(\Xi^{cb} * \Theta^{ba}) = \mathcal{I}C^{cb\to y}(\Xi^{cb}) \circ \mathcal{I}C^{ba\to x}(\Theta^{ba})$.

Proof. First, we observe that the two properties are equivalent because C is an isomorphism. For example, assume property *1* is true. To prove property *2*, we notice that for every Ξ^{cb} there exists Δ^y such that $\Xi^{cb} = C^{y\to cb}(\Delta^y)$. Similarly, $\Theta^{ba} = C^{x\to ba}(\Gamma^x)$. Therefore, $\mathcal{I}C^{ca\to(a\to c)}(\Xi^{cb} * \Theta^{ba}) = \mathcal{I}C^{ca\to(a\to c)}(C^{y\to cb}(\Delta^y) * C^{x\to ba}(\Gamma^x))$. Now, using property *1*, the last expression becomes $\mathcal{I}C^{ca\to(a\to c)}(C^{(a\to c)\to ca}(\Delta^{b\to c}\circ\Gamma^{a\to b}))$, which is equal to $\Delta^{b\to c}\circ\Gamma^{a\to b} = \mathcal{I}C^{cb\to y}(\Xi^{cb})\circ\mathcal{I}C^{ba\to x}(\Theta^{ba})$. Therefore, $\mathcal{I}C^{ca\to(a\to c)}(\Xi^{cb} * \Theta^{ba}) = \mathcal{I}C^{cb\to y}(\Xi^{cb}) \circ \mathcal{I}C^{ba\to x}(\Theta^{ba})$. The proof of *2* ⇒ *1* is identical. It is then enough to prove either of the two properties, and the other will follow. We choose to prove property *1*.

We start from the right-hand side, and we expand the Choi map in terms of the orthonormal bases:

$$C^{y\to cb}(\Delta^y) * C^{x\to ba}(\Gamma^a) = \sum_{i,j} (\Delta^y(u_j^b) \boxtimes u_j^b) * (\Gamma^x(u_i^a) \boxtimes u_i^a). \tag{6.186}$$

Next, we rewrite the expression above using the definition of the link product.

$$\sum_{i,j} \eta^{c_{ba}\to c} \boxtimes \left\langle \Phi^{b_{ac}b_{ac}} \right|^{b_{ac}b_{ac}\to I} \boxtimes \eta^{a_{bc}\to a} @ \Delta^y(u_j^b) \boxtimes u_j^b \boxtimes \Gamma^x(u_i^a) \boxtimes u_i^a. \tag{6.187}$$

Note that the definition of the square product gives $\Delta^y(u_j^b) \boxtimes u_j^b \boxtimes \Gamma^x(u_i^a) \boxtimes u_i^a = \eta^{c\to c_{ba}}(\Delta^y(u_j^b)) \boxtimes \eta^{b\to b_{ac}}(u_j^b) \boxtimes \eta^{b\to b_{ca}}(\Gamma^x(u_i^a)) \boxtimes \eta^{a\to a_{bc}}(u_i^a)$. Using this and Theorem 6.2.27, Eq. (6.187) becomes

$$\begin{aligned}
&\sum_{i,j} [\eta^{c_{ba}\to c} @ \eta^{c\to c_{ba}}(\Delta^y(u_j^b))] \boxtimes [\left\langle \Phi^{b_{ac}b_{ac}} \right|^{b_{ac}b_{ac}\to I} @ \eta^{b\to b_{ac}}(u_j^b) \\
&\qquad \boxtimes \eta^{b\to b_{ca}}(\Gamma^x(u_i^a))] \boxtimes [\eta^{a_{bc}\to a} @ \eta^{a\to a_{bc}}(u_i^a)] \\
&= \sum_{i,j} \Delta^y(u_j^b) \boxtimes [\left\langle \Phi^{b_{ac}b_{ac}} \right|^{b_{ac}b_{ac}\to I} @ \eta^{b\to b_{ac}}(u_j^b) \boxtimes \eta^{b\to b_{ca}}(\Gamma^x(u_i^a))] \boxtimes u_i^a.
\end{aligned} \tag{6.188}$$

Now, we focus on the term between square brackets. Observe that, Corollary (6.2.28) allows us to write $\left\langle \Phi^{b_{ac}b_{ac}} \right|^{b_{ac}b_{ac}\to I} = \sum_k \left\langle u_k^{b_{ac}} \right|^{b_{ac}\to I} \boxtimes \left\langle u_k^{b_{ac}} \right|^{b_{ac}\to I}$. With Lemma 6.2.6, we can go one step further and obtain $\left\langle \Phi^{b_{ac}b_{ac}} \right|^{b_{ac}b_{ac}\to I} = \sum_k (\left\langle u_k^b \right|^{b\to I} \circ \eta^{b_{ac}\to b}) \boxtimes (\left\langle u_k^b \right|^{b\to I} \circ \eta^{b_{ac}\to b})$. Using this expression and Theorem 6.2.27, the term inside square brackets in Eq. (6.188) is equal to $\sum_k (\left\langle u_k^b \right|^{b\to I} @ u_j^b) \boxtimes (\left\langle u_k^b \right|^{b\to I} @ \Gamma^x(u_i^a)) = \left\langle u_j^b \right|^{b\to I} @ \Gamma^x(u_i^a)$. Observe that this is a number; therefore, by linearity of Δ^y, we can write the last expression Eq. (6.188) as

$$\begin{aligned}
&\sum_{i,j} \left[\left\langle u_j^b \right|^{b\to I} @ \Gamma^x(u_i^a) \right] \Delta^y(u_j^b)) \boxtimes u_i^a \\
&= \sum_i \Delta^y \left(\sum_j u_j^b \left[\left\langle u_j^b \right|^{b\to I} @ \Gamma^x(u_i^a) \right] \right) \boxtimes u_i^a.
\end{aligned} \tag{6.189}$$

The scalar multiplication of u_j^b with the number $\left\langle u_j^b \right|^{b\to I} @\Gamma^x(u_i^a)$ can be rewritten as $\left| u_j^b \right\rangle^{I\to b} \circ \left\langle u_j^b \right|^{b\to I} @\Gamma^x(u_i^a)$. Thus, the input of Δ^y becomes

$$\begin{aligned}\sum_j u_j^b \left[\left\langle u_j^b \right|^{b\to I} @\Gamma^x(u_i^a) \right] &= \sum_j \left| u_j^b \right\rangle^{I\to b} \circ \left\langle u_j^b \right|^{b\to I} @\Gamma^x \\ &= \mathrm{id}^{b\to b} @\Gamma^x(u_i^a) \\ &= \Gamma^x(u_i^a).\end{aligned} \tag{6.190}$$

Following the chain of equalities, we have shown that

$$C^{y\to cb}(\Delta^y) * C^{x\to ba}(\Gamma^a) = \sum_i \Delta^y(\Gamma^x(u_i^a)) \boxtimes u_i^a = \sum_i \Delta^y \circ \Gamma^x(u_i^a) \boxtimes u_i^a, \tag{6.191}$$

and the right-hand side is by definition equal to $C^{(a\to c)\to ca}(\Delta^y \circ \Gamma^x)$. □

Chapter 7

Conclusion and Future Directions

This Thesis investigates the properties of quantum resource theories. After a brief introduction to the framework of quantum resource theories in Chapter 2, we focused on the resource theory of entanglement in Chapter 3. The first important results concern conversion distances between pure states. With Definition 3.1.1, we introduce a new conversion distance, $T_\star$, which we prove to be equivalent to the conversion distance based on the trace distance, as defined in Definition 2.1.3. Most importantly, with Theorem 3.1.6, we derive a closed-form expression for such distance. It is worth noticing that these results were possible thanks to the characterization of pure-to-mixed state conversion in LOCC (Proposition 3.1.2), which extends the results already known for pure-to-pure state LOCC-conversions, whether deterministic or probabilistic [26, 100].

With the closed formula for $T_\star$, we were able to characterize universal embezzling families (Definition 3.2.1 and Refs. [52, 53]) in terms of a simple optimization problem stated in Theorem 3.2.4. Thanks to that, we were able to determine that the embezzling family originally proposed by van Dam and Hayden [52] is unique under reasonable assumptions, in the sense that any other embezzling family has to be similar to the van Dam and Hayden family in the asymptotic limit. For specific families of states, we were also able to compute their star conversion distance to maximally entangled states, which showed that these families are universally embezzling if only if they are exactly the van Dam and Hayden family (see Figure 3.1). It is worth noting that the original embezzling protocol proposed by van Dam and Hayden only requires local operations, while we also allow for classical communication. However, classical communication does not appear to provide a significant advantage; in both cases, embezzling families must be similar to the van Dam and Hayden families. This suggests a direction for future work, namely, to determine whether LOCC embezzlement implies LO embezzlement (the other direction is trivial) and, therefore, to investigate if classical communication is relevant in entanglement embezzlement or not.

A different future direction involves extending the results about approximate conversion and entanglement embezzlement to other resource theories that share some of the properties of entanglement theory. Indeed, embezzlement has also been studied in other resource theories, including non-uniformity [265, 266], coherence [47, 267–269], and athermality [33, 174, 265, 270, 271]. Fundamental limits for embezzlement have been proved in Ref. [272]. All these resource theories are based on majorization (which was the crucial tool in defining and finding a closed formula for $T_\star$) or relative majorization. It is worth noting that, so far, research on entanglement embezzlement has focused solely on families composed of pure states. Another future research direction is the development of a more comprehensive theory of embezzlement that includes families of mixed states.

The next set of results presented in this Thesis concerns the resource theory of quantum thermodynamics, specifically the tasks of cooling and heating. The first step was to define what it means to cool and heat a quantum system. We opted for a definition compatible with the zeroth law of thermodynamics. We interpret cooling and heating as converting a system initially at equilibrium with a bath at temperature T to the state that would be at equilibrium with a bath at temperature $\tilde{T}$. With this definition, we derived expressions for the maximal and minimal temperatures to which a system can be heated and cooled, using the resourcefulness of a given state out of thermal equilibrium with the bath and closed thermal operations (Theorems 4.1.2 and 4.1.3). These quantities can be easily computed numerically for systems of any dimension and analytically for qubits. Moreover, we showed that quasi-classical state convertibility is completely characterized by the ability to heat and cool qubits. That is, a quasi-classical state can be converted into another with closed thermal operations if and only if the first state cools and heats qubits to lower and higher temperatures than the second does.

Still, in the resource theory of thermodynamics, we approached cooling and heating from a different perspective. Suppose that an agent wants to cool or a qubit, originally at a temperature T to a temperature $\tilde{T}$ using another state out of thermal equilibrium. Which qubits, identified by their energy gap, can be cooled or heated to the temperature $\tilde{T}$? We demonstrated that the ability to cool or heat more qubits to the target temperature $\tilde{T}$ also characterizes quasi-classical conversion. That is, a quasi-classical resource can be converted into another with closed thermal operations if and only if the first resource can cool or heat to the target temperature $\tilde{T}$ at least as many different qubits as the second can. Note that all of these results provide only fundamental bounds (similar to Landauer's principle [273]), which are very far from what can be achieved with current technologies, even though thermal machines in the quantum limit aiming to approach these bounds are under active investigation [189–196].

To facilitate the experimental realization of optimal cooling and heating protocols, a future direction of this work is to identify the thermal operation (or sequence of thermal operations) that achieves optimal cooling or heating. Indeed, all the results of Chapter 4 only guarantee the existence of a closed thermal operation that achieves optimal cooling or heating. However, they do not indicate what this operation is. A first step in this direction was taken by Orr [4], who found the GPO that achieves the optimal cooling of a qubit using another qubit as a resource. Another research direction is to extend these results to states that are not quasi-classical. Unfortunately, this is a particularly challenging task because a characterization for state conversion under closed thermal operation is still missing for states that are not quasi-classical.

In Chapter 5, we identified a key property of the resource theories of magic states, imaginarity, SEP, NPT, and non-negativity of quantum amplitudes: a channel is free if and only if its suitably renormalized Choi matrix is a free state. With that property in mind, we presented a new method for constructing a resource theory from the set of free states based on the Choi isomorphism, and we derived which conditions on the free states are necessary and sufficient for constructing a mathematically consistent Choi-defined resource theory. Our interest is driven by the significant advantages that such a property guarantees. These advantages stem from the fact that every question in Choi-defined resource theories can be naturally framed as a question involving only free states, which often have a well-understood structure. Indeed, we introduce several resource quantifiers that can all be computed with CLPs if the set of free states is convex and closed.

Moreover, we demonstrated that in any CDRT, free operations coincide with CRNG operations. This provides a new constructive way to define CRNG operations when the set of free states satisfies the conditions of Theorem 5.2.5.

A natural future direction is to explore Choi-defined dynamical resource theories [68, 274, 275], i.e., resource theories for channels, where a superchannel [65, 264] is free if and only if its Choi matrix is the Choi matrix of a free channel. Notably, this property has already been considered in the resource theories of dynamical entanglement [64] and magic channels [70], leading to interesting results. The characterization of Choi-defined dynamical resource theories is a particularly challenging problem because the techniques employed in this work cannot be easily translated into superchannels. Indeed, there are some subtleties in the definition of the Choi matrix of a superchannel. For example, in Ref. [264], several different Choi matrices are constructed from the same superchannel, which differ in the order in which systems are considered. As in the case presented in this article, an ordering of systems is mandatory if one wants to construct resource theories from Choi matrices, as inputs and outputs play different roles, especially if we want such a construction to be generalizable, without ambiguity, to even higher-order maps [80, 276].

Chapter 6 is the first step in the direction of higher-order Choi-defined resource theories. The goal of this Chapter is to introduce a formalism to describe maps of any order and give a definition of the Choi isomorphism that can be used for maps of any order. We reach the goal thanks to the type formalism [80]. We proved several properties of types, with a focus on those relevant to the parallel and sequential composition of higher-order maps. Then, we described the spaces of maps labelled with a given type and provided formulas to compute the map associated with parallel and sequential composition of maps of any order. We then characterized subspaces that are important for quantum theory, providing a generalization to higher-order maps of Hermitian-preserving and completely positive-preserving maps. Lastly, we defined the Choi isomorphism for higher-order maps, and we showed that it preserves Hermitian-preserving and completely positive-preserving higher-order maps (a generalization of the well-known results for channels and states [66]). Since the Choi isomorphism lowers the order of the map, one can use it to deduce properties of higher-order maps from maps of lower order, such as matrices.

The next step is to investigate the generalization of trace-preserving maps, a concept associated with determinism, to higher orders and determine how trace-preserving higher-order maps transform under the Choi isomorphism. Another direction is to determine which higher-order quantum maps are *physically implementable*, that is, maps that can be simulated with a sequence of channels with memories, known as quantum networks [66, 221, 246]. In general, this is not a one-to-one correspondence. Therefore, one could also ask the opposite question. That is, which higher-order quantum maps can one simulate with a given quantum network? This research direction would provide new insights into maps that cannot be simulated with a quantum circuit, such as the quantum switch [73–79].

Bibliography

[1] E. Zanoni, T. Theurer, and G. Gour, Complete characterization of entanglement embezzlement, Quantum **8**, 1368 (2024).

[2] T. Theurer, E. Zanoni, C. M. Scandolo, and G. Gour, Thermodynamic state convertibility is determined by qubit cooling and heating, New J. Phys. **25**, 123017 (2023).

[3] E. Zanoni and C. M. Scandolo, Choi-defined resource theories, Phys. Rev. A **111**, 062407 (2025).

[4] M. Orr, Qubit cooling in the resource theory of thermodynamics (2025).

[5] E. Schrödinger, Discussion of probability relations between separated systems, Math. Proc. Camb. Philos. Soc. **31**, 555 (1935).

[6] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).

[7] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz **1**, 195 (1964).

[8] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[9] S. J. Freedman and J. F. Clauser, Experimental test of local hidden-variable theories, Phys. Rev. Lett. **28**, 938 (1972).

[10] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[12] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, Vol. 1 (IEEE Bangalore, 1984) pp. 175–179.

[13] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, Theor. Comput. Sci. **560**, 7 (2014), theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[14] A. K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[15] C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum cryptography without Bell's theorem, Phys. Rev. Lett. **68**, 557 (1992).

[16] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett **70**, 1895 (1993).

[17] C. H. Bennett and S. J. Wiesner, Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states, Phys. Rev. Lett. **69**, 2881 (1992).

[18] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, Proc. R. Soc. A **439**, 553 (1992).

[19] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994) pp. 124–134.

[20] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96 (Association for Computing Machinery, New York, NY, USA, 1996) p. 212–219.

[21] D. Gottesman, The Heisenberg representation of quantum computers, in *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. P. Corney, R. Delbourgo, and P. D. Jarvis (International Press, Cambridge, MA, 1999) pp. 32–43, arXiv:quant-ph/9807006 .

[22] S. Bravyi and A. Kitaev, Universal quantum computation with ideal clifford gates and noisy ancillas, Phys. Rev. A **71**, 022316 (2005).

[23] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).

[24] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of noisy entanglement and faithful teleportation via noisy channels, Phys. Rev. Lett. **76**, 722 (1996).

[25] H.-K. Lo and S. Popescu, Concentrating entanglement by local actions: Beyond mean values, Phys. Rev. A **63**, 022301 (2001).

[26] M. A. Nielsen, Conditions for a class of entanglement transformations, Phys. Rev. Lett. **83**, 436 (1999).

[27] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[28] G. C. Wick, A. S. Wightman, and E. P. Wigner, The intrinsic parity of elementary particles, Phys. Rev. **88**, 101 (1952).

[29] Y. Aharonov and L. Susskind, Charge superselection rule, Phys. Rev. **155**, 1428 (1967).

[30] T. Rudolph and B. C. Sanders, Requirement of optical coherence for continuous-variable quantum teleportation, Phys. Rev. Lett. **87**, 077903 (2001).

[31] B. C. Sanders, S. D. Bartlett, T. Rudolph, and P. L. Knight, Photon-number superselection and the entangled coherent-state representation, Phys. Rev. A **68**, 042329 (2003).

[32] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reference frames, superselection rules, and quantum information, Rev. Mod. Phys. **79**, 555 (2007).

[33] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource theory of quantum states out of thermal equilibrium, Phys. Rev. Lett. **111**, 250404 (2013).

[34] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nat. Commun. **4**, 2059 (2013).

[35] M. Lostaglio, An introductory review of the resource theory approach to thermodynamics, Rep. Prog. Phys. **82**, 114001 (2019).

[36] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, Negative quasi-probability as a resource for quantum computation, New J. Phys. **14**, 113011 (2012).

[37] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, The resource theory of stabilizer quantum computation, New J. Phys. **16**, 013009 (2014).

[38] M. Howard and E. Campbell, Application of a resource theory for magic states to fault-tolerant quantum computing, Phys. Rev. Lett. **118**, 090501 (2017).

[39] M. Ahmadi, H. B. Dang, G. Gour, and B. C. Sanders, Quantification and manipulation of magic states, Phys. Rev. A **97**, 062332 (2018).

[40] J. R. Seddon and E. T. Campbell, Quantifying magic for multi-qubit operations, Proc. R. Soc. A: Math. Phys. Eng. Sci. **475**, 20190251 (2019).

[41] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Resource theory of imaginarity: Quantification and state conversion, Phys. Rev. A **103**, 032401 (2021).

[42] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Operational resource theory of imaginarity, Phys. Rev. Lett. **126**, 090401 (2021).

[43] A. Hickey and G. Gour, Quantifying the imaginarity of quantum mechanics, J. Phys. A **51**, 414009 (2018).

[44] G. Gour and R. W. Spekkens, The resource theory of quantum reference frames: manipulations and monotones, New J. Phys. **10**, 033023 (2008).

[45] I. Marvian and R. W. Spekkens, The theory of manipulations of pure state asymmetry: I. basic tools, equivalence classes and single copy transformations, New J. Phys. **15**, 033001 (2013).

[46] I. Marvian and R. W. Spekkens, How to quantify coherence: Distinguishing speakable and unspeakable notions, Phys. Rev. A **94**, 052324 (2016).

[47] T. Baumgratz, M. Cramer, and M. B. Plenio, Quantifying coherence, Phys. Rev. Lett. **113**, 140401 (2014).

[48] J. Åberg, Quantifying superposition (2006), arXiv:quant-ph/0612146 [quant-ph] .

[49] A. Streltsov, G. Adesso, and M. B. Plenio, Colloquium: Quantum coherence as a resource, Rev. Mod. Phys. **89**, 041003 (2017).

[50] A. Winter and D. Yang, Operational resource theory of coherence, Phys. Rev. Lett. **116**, 120404 (2016).

[51] H. Zhu, Z. Ma, Z. Cao, S.-M. Fei, and V. Vedral, Operational one-to-one mapping between coherence and entanglement measures, Phys. Rev. A **96**, 032316 (2017).

[52] W. van Dam and P. Hayden, Universal entanglement transformations without communication, Phys. Rev. A **67**, 060302 (2003).

[53] D. Leung and B. Wang, Characteristics of universal embezzling families, Phys. Rev. A **90**, 042331 (2014).

[54] E. M. Rains, Entanglement purification via separable superoperators (1998), arXiv:quant-ph/9707002 [quant-ph] .

[55] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, Entangling operations and their implementation using a small amount of entanglement, Phys. Rev. Lett. **86**, 544 (2001).

[56] A. W. Harrow and M. A. Nielsen, Robustness of quantum gates in the presence of noise, Phys. Rev. A **68**, 012308 (2003).

[57] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[58] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[59] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature?, Phys. Rev. Lett. **80**, 5239 (1998).

[60] E. M. Rains, Bound on distillable entanglement, Phys. Rev. A **60**, 179 (1999).

[61] K. Audenaert, M. B. Plenio, and J. Eisert, Entanglement cost under positive-partial-transpose-preserving operations, Phys. Rev. Lett. **90**, 027901 (2003).

[62] S. Ishizaka and M. B. Plenio, Multiparticle entanglement manipulation under positive partial transpose preserving operations, Phys. Rev. A **71**, 052303 (2005).

[63] W. Matthews and A. Winter, Pure-state transformations and catalysis under operations that completely preserve positivity of partial transpose, Phys. Rev. A **78**, 012317 (2008).

[64] G. Gour and C. M. Scandolo, Dynamical entanglement, Phys. Rev. Lett. **125**, 180505 (2020).

[65] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, EPL **83**, 30004 (2008).

[66] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[67] G. Gour and C. M. Scandolo, Entanglement of a bipartite channel, Phys. Rev. A **103**, 062422 (2021).

[68] G. Gour and C. M. Scandolo, Dynamical resources (2020), arXiv:2101.01552 [quant-ph] .

[69] G. Saxena, E. Chitambar, and G. Gour, Dynamical resource theory of quantum coherence, Phys. Rev. Res. **2**, 023298 (2020).

[70] G. Saxena and G. Gour, Quantifying multiqubit magic channels with completely stabilizer-preserving operations, Phys. Rev. A **106**, 042422 (2022).

[71] B. Regula and R. Takagi, Fundamental limitations on distillation of quantum channel resources, Nat. Commun. **12**, 4411 (2021).

[72] M. Hayashi and H. Yamasaki, Generalized quantum stein's lemma and second law of quantum resource theories (2025), arXiv:2408.02722 [quant-ph] .

[73] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).

[74] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, Indefinite causal order enables perfect quantum communication with zero capacity channels, New J. Phys. **23**, 033039 (2021).

[75] M. M. Taddei, J. Cariñe, D. Martínez, T. García, N. Guerrero, A. A. Abbott, M. Araújo, C. Branciard, E. S. Gómez, S. P. Walborn, L. Aolita, and G. Lima, Computational advantage from the quantum superposition of multiple temporal orders of photonic gates, PRX Quantum **2**, 010320 (2021).

[76] A. Feix, M. Araújo, and Č. Brukner, Quantum superposition of the order of parties as a communication resource, Phys. Rev. A **92**, 052326 (2015).

[77] M. Araújo, P. A. Guérin, and Ä. Baumeler, Quantum computation with indefinite causal structures, Phys. Rev. A **96**, 052315 (2017).

[78] K. Simonov, M. Caleffi, J. Illiano, J. Romero, and A. S. Cacciapuoti, Universal quantum computation via superposed orders of single-qubit gates (2024), arXiv:2311.13654 [quant-ph] .

[79] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301 (2012).

[80] A. Bisio and P. Perinotti, Theoretical framework for higher-order quantum theory, Proc. R. Soc. A **475**, 20180706 (2019).

[81] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).

[82] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).

[83] G. Gour, *Quantum Resource Theories* (Cambridge University Press, Cambridge, 2025).

[84] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[85] B. Coecke, T. Fritz, and R. W. Spekkens, A mathematical theory of resources, Inf. Comput. **250**, 59 (2016).

[86] F. Albarelli, M. G. Genoni, M. G. A. Paris, and A. Ferraro, Resource theory of quantum non-gaussianity and wigner negativity, Phys. Rev. A **98**, 052350 (2018).

[87] L. Lami, B. Regula, X. Wang, R. Nichols, A. Winter, and G. Adesso, Gaussian quantum resource theories, Phys. Rev. A **98**, 022335 (2018).

[88] L. Lami, R. Takagi, and G. Adesso, Assisted concentration of gaussian resources, Phys. Rev. A **101**, 052305 (2020).

[89] L. Lami, B. Regula, R. Takagi, and G. Ferrari, Framework for resource quantification in infinite-dimensional general probabilistic theories, Phys. Rev. A **103**, 032424 (2021).

[90] B. Regula, L. Lami, G. Ferrari, and R. Takagi, Operational quantification of continuous-variable quantum resources, Phys. Rev. Lett. **126**, 110403 (2021).

[91] G. Vidal, D. Jonathan, and M. A. Nielsen, Approximate transformations and robust manipulation of bipartite pure-state entanglement, Phys. Rev. A **62**, 012304 (2000).

[92] A. Uhlmann, The “transition probability” in the state space of a *-algebra, Rep. Math. Phys. **9**, 273 (1976).

[93] R. Jozsa, Fidelity for mixed quantum states, J. Mod. Opt **41**, 2315 (1994).

[94] C. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Trans. Inf. Theory **45**, 1216 (1999).

[95] M. Tomamichel, R. Colbeck, and R. Renner, Duality between smooth min- and max-entropies, IEEE Trans. Inf. Theory **56**, 4674 (2010).

[96] G. Gour and M. Tomamichel, Optimal extensions of resource measures and their applications, Phys. Rev. A **102**, 062401 (2020).

[97] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability, Phys. Rev. Res. **1**, 033170 (2019).

[98] G. Gour, Role of quantum coherence in thermodynamics, PRX Quantum **3**, 040323 (2022).

[99] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer Series in Statistics (Springer, 2011).

[100] D. Jonathan and M. B. Plenio, Minimal conditions for local pure-state entanglement manipulation, Phys. Rev. Lett. **83**, 1455 (1999).

[101] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).

[102] A. Lenard, Thermodynamical proof of the Gibbs formula for elementary quantum systems, J. Stat. Phys **19**, 575 (1978).

[103] A. Riera, C. Gogolin, and J. Eisert, Thermalization in nature and on a quantum computer, Phys. Rev. Lett. **108**, 080402 (2012).

[104] J. Uffink, Bluff your way in the second law of thermodynamics, Stud. Hist. Philos. Sci. B **32**, 305 (2001).

[105] H. R. Brown and J. Uffink, The origins of time-asymmetry in thermodynamics: The minus first law, Stud. Hist. Philos. Sci. B **32**, 525 (2001), the Conceptual Foundations of Statistical Physics.

[106] A. Y. Kitaev, Quantum computations: algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[107] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law, Int. J. Theor. Phys **39**, 2717 (2000).

[108] F. Buscemi and G. Gour, Quantum relative Lorenz curves, Phys. Rev. A **95**, 012110 (2017).

[109] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, Quantum majorization and a complete set of entropic conditions for quantum thermodynamics, Nat. Commun. **9**, 5352 (2018).

[110] C. T. Chubb, M. Tomamichel, and K. Korzekwa, Beyond the thermodynamic limit: finite-size corrections to state interconversion rates, Quantum **2**, 108 (2018).

[111] M. Horodecki, J. Oppenheim, and C. Sparaciari, Extremal distributions under approximate majorization, J. Phys. A **51**, 305301 (2018).

[112] M. P. Müller, Correlating thermal machines and the second law at the nanoscale, Phys. Rev. X **8**, 041051 (2018).

[113] K. Korzekwa, C. T. Chubb, and M. Tomamichel, Avoiding irreversibility: Engineering resonant conversions of quantum resources, Phys. Rev. Lett. **122**, 110403 (2019).

[114] F. vom Ende, Strict positivity and d-majorization, Linear Multilinear A. **70**, 4023 (2022).

[115] U. Singh, S. Das, and N. J. Cerf, Partial order on passive states and Hoffman majorization in quantum thermodynamics, Phys. Rev. Res. **3**, 033091 (2021).

[116] D. Blackwell, Equivalent comparisons of experiments, Ann. Math. Stat. **24**, 265 (1953).

[117] A. Uhlmann, Ordnungsstrukturen im Raume der gemischten Zustände und die H-Theoreme von Felderhof-van Kampen, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat. R **27**, 213 (1978).

[118] E. Ruch, R. Schranner, and T. H. Seligman, The mixing distance, J. Chem. Phys **69**, 386 (1978).

[119] J. M. Renes, Relative submajorization and its use in quantum resource theories, J. Math. Phys **57**, 122202 (2016).

[120] H. Joe, Majorization and divergence, J. Math. Anal. Appl. **148**, 287 (1990).

[121] G. Dahl, Matrix majorization, Linear Algebra Its Appl. **288**, 53 (1999).

[122] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying entanglement, Phys. Rev. Lett. **78**, 2275 (1997).

[123] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, Exact and asymptotic measures of multipartite pure-state entanglement, Phys. Rev. A **63**, 012307 (2000).

[124] M. B. Plenio and S. Virmani, An introduction to entanglement measures, Quantum Inf. Comput. **7**, 1 (2007).

[125] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[126] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Entanglement-assisted classical capacity of noisy quantum channels, Phys. Rev. Lett. **83**, 3081 (1999).

[127] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem, IEEE Trans. Inf. Theory **48**, 2637 (2006).

[128] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels, IEEE Trans. Inf. Theory **60**, 2926 (2014).

[129] S. Popescu, Bell's inequalities versus teleportation: What is nonlocality?, Phys. Rev. Lett. **72**, 797 (1994).

[130] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Phys. Rev. A **60**, 1888 (1999).

[131] M. Berta, M. Christandl, and R. Renner, The quantum reverse Shannon theorem based on one-shot information theory, Commun. Math. Phys. **306**, 579 (2011).

[132] R. Cleve, L. Liu, and V. I. Paulsen, Perfect embezzlement of entanglement, J. Math. Phys. **58**, 012204 (2017).

[133] D. Jonathan and M. B. Plenio, Entanglement-assisted local manipulation of pure quantum states, Phys. Rev. Lett. **83**, 3566 (1999).

[134] L. Lami, B. Regula, and A. Streltsov, No-go theorem for entanglement distillation using catalysis, Phys. Rev. A **109**, L050401 (2024).

[135] R. Ganardi, T. V. Kondra, and A. Streltsov, Catalytic and asymptotic equivalence for quantum entanglement, Phys. Rev. Lett. **133**, 250201 (2024).

[136] J. Son and N. H. Y. Ng, A hierarchy of thermal processes collapses under catalysis, Quantum Sci. Technol. **10**, 015011 (2024).

[137] S. H. Lie and N. H. Y. Ng, Catalysis always degrades external quantum correlations, Phys. Rev. A **108**, 012417 (2023).

[138] L. v. Luijk, R. F. Werner, and H. Wilming, Covariant catalysis requires correlations and good quantum reference frames degrade little, Quantum **7**, 1166 (2023).

[139] C. Datta, R. Ganardi, T. V. Kondra, and A. Streltsov, Is there a finite complete set of monotones in any quantum resource theory?, Phys. Rev. Lett. **130**, 240204 (2023).

[140] J. Son and N. H. Y. Ng, Catalysis in action via elementary thermal operations, New J. Phys. **26**, 033029 (2024).

[141] P. Lipka-Bartosik, M. Perarnau-Llobet, and N. Brunner, Operational definition of the temperature of a quantum state, Phys. Rev. Lett. **130**, 040401 (2023).

[142] K. Korzekwa and M. Lostaglio, Optimizing thermalization, Phys. Rev. Lett. **129**, 040602 (2022).

[143] C. Datta, T. V. Kondra, M. Miller, and A. Streltsov, Entanglement catalysis for quantum states and noisy channels, Quantum **8**, 1290 (2024).

[144] P. Lipka-Bartosik and P. Skrzypczyk, Catalytic quantum teleportation, Phys. Rev. Lett. **127**, 080502 (2021).

[145] T. V. Kondra, C. Datta, and A. Streltsov, Catalytic transformations of pure entangled states, Phys. Rev. Lett. **127**, 150503 (2021).

[146] M. Karvonen, Neither contextuality nor nonlocality admits catalysts, Phys. Rev. Lett. **127**, 160402 (2021).

[147] H. Wilming, Entropy and reversible catalysis, Phys. Rev. Lett. **127**, 260402 (2021).

[148] S. H. Lie and H. Jeong, Randomness for quantum channels: Genericity of catalysis and quantum advantage of uniformness, Phys. Rev. Res. **3**, 013218 (2021).

[149] P. Lipka-Bartosik and P. Skrzypczyk, All states are universal catalysts in quantum thermodynamics, Phys. Rev. X **11**, 011061 (2021).

[150] N. Shiraishi and T. Sagawa, Quantum thermodynamics of correlated-catalytic state conversion at small scale, Phys. Rev. Lett. **126**, 150502 (2021).

[151] P. Lipka-Bartosik, H. Wilming, and N. H. Y. Ng, Catalysis in quantum information theory, Rev. Mod. Phys. **96**, 025005 (2024).

[152] C. Datta, T. Varun Kondra, M. Miller, and A. Streltsov, Catalysis of entanglement and other quantum resources, Rep. Prog. Phys. **86**, 116002 (2023).

[153] O. Regev and T. Vidick, Elementary proofs of Grothendieck theorems for completely bounded norms, J. Operat. Theor. **71**, 491 (2014).

[154] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques, Boll. Soc. Mat. São-Paulo **8**, 1 (1953), reprinted in Resenhas **4**, 401 (1996).

[155] G. Pisier, Grothendieck's theorem, past and present, Bull. Amer. Math. Soc. **49**, 237 (2012).

[156] D. Leung, B. Toner, and J. Watrous, Coherent state exchange in multi-prover quantum interactive proof systems, Chic. J. Theoret. Comput. Sci. **2013**, 11 (2013).

[157] I. Dinur, D. Steurer, and T. Vidick, A parallel repetition theorem for entangled projection games, Comput. Complex. **24**, 201 (2015).

[158] O. Regev and T. Vidick, Quantum XOR games, ACM Trans. Comput. Theory **7**, 15 (2015).

[159] Z. Ji, D. Leung, and T. Vidick, A three-player coherent state embezzlement game, Quantum **4**, 349 (2020).

[160] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys **1**, 231 (1969).

[161] A. S. Holevo, Statistical decision theory for quantum systems, J. Multivariate Anal. **3**, 337 (1973).

[162] G. Vidal, Entanglement of pure states for a single copy, Phys. Rev. Lett. **83**, 1046 (1999).

[163] G. Chiribella and C. M. Scandolo, Entanglement and thermodynamics in general probabilistic theories, New J. Phys. **17**, 103027 (2015).

[164] A. W. Roberts and D. E. Varberg, *Convex Functions*, edited by P. A. Smith and S. Ellenberg, Pure and Applied Mathemathics; A Series of Monographs and Textbooks, Vol. 57 (ACADEMIC PRESS New York and London, 1973).

[165] J. von Neumann, Some matrix-inequalities and metrization of matrix-space, Tomsk. Univ. Rev. **1**, 286 (1937), in *John von Neumann Collected Works*, Vol. IV, edited by A. H. Taub (Pergamon, 1962) pp. 205–218.

[166] L. Mirsky, A trace inequality of John von Neumann, Monatsh. Math **79**, 303 (1975).

[167] T. S. Ferguson, *Mathematical statistics: A decision theoretic approach*, Probability and Mathematical Statistics: A Series of Monographs and Textbooks (Academic Press, 1967).

[168] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms* (MIT press, 2022).

[169] R. A. Wijsman, A useful inequality on ratios of integrals, with application to maximum likelihood estimation, J. Am. Stat. Assoc **80**, 472 (1985).

[170] W. Rudin, *Real and Complex Analysis*, Higher Mathematics Series (McGraw-Hill Education, 1987).

[171] J. Åberg, Truly work-like work extraction via a single-shot analysis, Nat. Commun. **4**, 1925 (2013).

[172] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nat. Commun. **5**, 4185 (2014).

[173] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, A measure of majorization emerging from single-shot statistical mechanics, New J. Phys. **17**, 073001 (2015).

[174] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, The second laws of quantum thermodynamics, Proc. Natl. Acad. Sci. U.S.A. **112**, 3275 (2015).

[175] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Quantum coherence, time-translation symmetry, and thermodynamics, Phys. Rev. X **5**, 021001 (2015).

[176] M. Lostaglio, D. Jennings, and T. Rudolph, Description of quantum coherence in thermodynamic processes requires constraints beyond free energy, Nat. Commun. **6**, 6383 (2015).

[177] J. Goold, M. Huber, A. Riera, L. d. Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, J. Phys. A **49**, 143001 (2016).

[178] K. Korzekwa, *Coherence, thermodynamics and uncertainty relations*, Ph.D. thesis, Imperial College London (2016).

[179] L. Masanes and J. Oppenheim, A general derivation and quantification of the third law of thermodynamics, Nat. Commun. **8**, 14538 (2017).

[180] J. Scharlau and M. P. Mueller, Quantum Horn's lemma, finite heat baths, and the third law of thermodynamics, Quantum **2**, 54 (2018).

[181] P. Faist, M. Berta, and F. Brandão, Thermodynamic capacity of quantum processes, Phys. Rev. Lett. **122**, 200601 (2019).

[182] P. Faist, M. Berta, and F. G. S. L. Brandao, Thermodynamic implementations of quantum processes, Commun. Math. Phys. **384**, 1709 (2021).

[183] M. Lostaglio, Á. M. Alhambra, and C. Perry, Elementary Thermal Operations, Quantum **2**, 52 (2018).

[184] P. Mazurek and M. Horodecki, Decomposability and convex structure of thermal processes, New J. Phys. **20**, 053040 (2018).

[185] P. Mazurek, Thermal processes and state achievability, Phys. Rev. A **99**, 042110 (2019).

[186] A. d. O. Junior, J. Czartowski, K. Życzkowski, and K. Korzekwa, Geometric structure of thermal cones, Phys. Rev. E **106**, 064109 (2022).

[187] F. vom Ende and G. Dirr, The d-majorization polytope, Linear Algebra Its Appl. **649**, 152 (2022).

[188] F. vom Ende and E. Malvetti, The thermomajorization polytope and its degeneracies, Entropy **26**, 10.3390/e26020106 (2024).

[189] H. E. D. Scovil and E. O. Schulz-DuBois, Three-level masers as heat engines, Phys. Rev. Lett. **2**, 262 (1959).

[190] L. P. Faucheux, L. S. Bourdieu, P. D. Kaplan, and A. J. Libchaber, Optical thermal ratchet, Phys. Rev. Lett. **74**, 1504 (1995).

[191] M. O. Scully, Quantum afterburner: Improving the efficiency of an ideal heat engine, Phys. Rev. Lett. **88**, 050602 (2002).

[192] J. Baugh, O. Moussa, C. A. Ryan, A. Nayak, and R. Laflamme, Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance, Nature **438**, 470 (2005).

[193] R. Uzdin, A. Levy, and R. Kosloff, Equivalence of quantum heat machines, and quantum-thermodynamic signatures, Phys. Rev. X **5**, 031044 (2015).

[194] R. Silva, G. Manzano, P. Skrzypczyk, and N. Brunner, Performance of autonomous quantum thermal machines: Hilbert space dimension as a thermodynamical resource, Phys. Rev. E **94**, 032120 (2016).

[195] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, Experimental demonstration of quantum effects in the operation of microscopic heat engines, Phys. Rev. Lett. **122**, 110601 (2019).

[196] M. L. Bera, M. Lewenstein, and M. N. Bera, Attaining carnot efficiency with quantum and nanoscale heat engines, npj Quantum Inf. **7**, 31 (2021).

[197] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics*, Lecture Notes in Physics, Vol. 784 (Springer Berlin, Heidelberg, 2009).

[198] F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds., *Thermodynamics in the Quantum Regime*, Fundamental Theories of Physics, Vol. 195 (Springer Cham, 2018).

[199] N. Yunger Halpern and D. T. Limmer, Fundamental limitations on photoisomerization from thermodynamic resource theories, Phys. Rev. A **101**, 042116 (2020).

[200] G. Spaventa, S. F. Huelga, and M. B. Plenio, Capacity of non-Markovianity to boost the efficiency of molecular switches, Phys. Rev. A **105**, 012420 (2022).

[201] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, Unifying paradigms of quantum refrigeration: A universal and attainable bound on cooling, Phys. Rev. Lett. **123**, 170605 (2019).

[202] L. J. Schulman and U. V. Vazirani, Molecular scale heat engines and scalable quantum computation, in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC ’99 (Association for Computing Machinery, New York, NY, USA, 1999) pp. 322–329.

[203] P. O. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen, Algorithmic cooling and scalable nmr quantum computers, Proc. Natl. Acad. Sci. U.S.A. **99**, 3388 (2002).

[204] J. M. Fernandez, S. Lloyd, T. Mor, and V. Roychowdhury, Algorithmic cooling of spins: A practicable method for increasing polarization, International Journal of Quantum Information **02**, 461 (2004).

[205] P. Taranto, F. Bakhshinezhad, P. Schüttelkopf, F. Clivaz, and M. Huber, Exponential improvement for quantum cooling through finite-memory effects, Phys. Rev. Appl. **14**, 054005 (2020).

[206] O. Svelto, *Principles of Lasers* (Springer New York, NY, 2010).

[207] M. Kardar, *Statistical Physics of Particles* (Cambridge University Press, 2007).

[208] F. vom Ende, Which bath hamiltonians matter for thermal operations?, J. Math. Phys **63**, 112202 (2022).

[209] G. Gour, Quantum resource theories in the single-shot regime, Phys. Rev. A **95**, 062314 (2017).

[210] D. Rosset, F. Buscemi, and Y.-C. Liang, Resource theory of quantum memories and their faithful verification with minimal assumptions, Phys. Rev. X **8**, 021033 (2018).

[211] P. Skrzypczyk and N. Linden, Robustness of measurement, discrimination games, and accessible information, Phys. Rev. Lett. **122**, 140403 (2019).

[212] R. Takagi and B. Regula, General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks, Phys. Rev. X **9**, 031053 (2019).

[213] H. Wilming and R. Gallego, Third law of thermodynamics as a single inequality, Phys. Rev. X **7**, 041033 (2017).

[214] T. Gonda and R. W. Spekkens, Monotones in general resource theories, Compositionality **5**, 7 (2023).

[215] C. M. Scandolo, G. Gour, and B. C. Sanders, Covariant influences for finite discrete dynamical systems, Phys. Rev. E **107**, 014203 (2023).

[216] R. Cockett, I. J. Geng, C. M. Scandolo, and P. V. Srinivasan, Extending resource monotones using kan extensions, Electronic Proceedings in Theoretical Computer Science **380**, 203 (2023).

[217] N. Johnston and J. Sikora, Completely positive completely positive maps (and a resource theory for non-negativity of quantum amplitudes), Linear Algebra Its Appl. **653**, 395 (2022).

[218] A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics, Vol. 54 (American Mathematical Society, Providence, 2002).

[219] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*, 2053-2563 (IOP Publishing, Bristol, 2023).

[220] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Its Appl **10**, 285 (1975).

[221] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum networks: general theory and applications, Acta Phys. Slovaca **61**, 273 (2011).

[222] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Quantum nonlocality without entanglement, Phys. Rev. A **59**, 1070 (1999).

[223] E. Chitambar, W. Cui, and H.-K. Lo, Increasing entanglement monotones by separable operations, Phys. Rev. Lett. **108**, 240504 (2012).

[224] P. Faist, J. Oppenheim, and R. Renner, Gibbs-preserving maps outperform thermal operations in the quantum regime, New J. Phys. **17**, 043003 (2015).

[225] E. Kaur, S. Das, M. M. Wilde, and A. Winter, Resource theory of unextendibility and nonasymptotic quantum capacity, Phys. Rev. A **104**, 022401 (2021).

[226] L. Pankowski, F. G. S. L. Brandão, M. Horodecki, and G. Smith, Entanglement distillation by extendible maps, Quantum Inf. Comput. **13**, 751 (2013).

[227] I. Marvian, Coherence distillation machines are impossible in quantum thermodynamics, Nat. Commun. **11**, 25 (2020).

[228] K.-D. Wu, T. V. Kondra, C. M. Scandolo, S. Rana, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Resource theory of imaginarity in distributed scenarios, Commun. Phys. **7**, 171 (2024).

[229] G. Vidal, Entanglement monotones, J. Mod. Opt **47**, 355 (2000).

[230] N. Datta, Min- and max-relative entropies and a new entanglement monotone, IEEE Trans. Inf. Theory **55**, 2816 (2009).

[231] T. Cooney, M. Mosonyi, and M. M. Wilde, Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication, Commun. Math. Phys **344**, 797 (2016).

[232] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Approaches for approximate additivity of the holevo information of quantum channels, Phys. Rev. A **97**, 012332 (2018).

[233] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, Lett. Math. Phys. **110**, 2277 (2020).

[234] K. Fang, O. Fawzi, R. Renner, and D. Sutter, Chain rule for the quantum relative entropy, Phys. Rev. Lett. **124**, 100501 (2020).

[235] L. 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 (Association for Computing Machinery, New York, NY, USA, 2003) pp. 10–19.

[236] M. Hayashi and Y. Ouyang, Tight Cramér-Rao type bounds for multiparameter quantum metrology through conic programming, Quantum **7**, 1094 (2023).

[237] B. Coecke and A. Kissinger, The compositional structure of multipartite quantum entanglement, in *Automata, Languages and Programming*, edited by S. Abramsky, C. Gavoille, C. Kirchner, F. Meyer auf der Heide, and P. G. Spirakis (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010) pp. 297–308.

[238] B. Coecke, C. Heunen, and A. Kissinger, Categories of quantum and classical channels, Quantum Inf. Process. **15**, 5179–5209 (2016).

[239] B. Coecke, R. Duncan, A. Kissinger, and Q. Wang, Generalised compositional theories and diagrammatic reasoning, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht, 2016) pp. 309–366.

[240] A. Kissinger and S. Uijlen, Picturing indefinite causal structure, in *Proceedings 13th International Conference on Quantum Physics and Logic, Glasgow, Scotland, 6-10 June 2016*, Electronic Proceedings in Theoretical Computer Science, Vol. 236, edited by R. Duncan and C. Heunen (Open Publishing Association, 2017) pp. 87–94.

[241] A. Kissinger and S. Uijlen, A categorical semantics for causal structure, Log. Methods Comput. Sci. **15**, 3 (2019).

[242] J. H. Selby, C. M. Scandolo, and B. Coecke, Reconstructing quantum theory from diagrammatic postulates, Quantum **5**, 445 (2021).

[243] B. Coecke and A. Kissinger, *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press, 2017).

[244] B. Coecke, Quantum picturalism, Contemp. Phys. **51**, 59 (2010).

[245] B. Coecke and É. Paquette, Categories for the practising physicist, in *New Structures for Physics*, edited by B. Coecke (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011) pp. 173–286.

[246] M. Wilson and G. Chiribella, A mathematical framework for transformations of physical processes (2023), arXiv:2204.04319 [quant-ph] .

[247] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite causal order in a quantum switch, Phys. Rev. Lett. **121**, 090503 (2018).

[248] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).

[249] Č. Brukner, Quantum causality, Nat. Phys. **10**, 259 (2014).

[250] O. Oreshkov, Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics, Quantum **3**, 206 (2019).

[251] E. Castro-Ruiz, F. Giacomini, and Č. Brukner, Dynamics of quantum causal structures, Phys. Rev. X **8**, 011047 (2018).

[252] G. Chiribella and Z. Liu, Quantum operations with indefinite time direction, Commun. Phys. **5**, 190 (2022).

[253] G. M. D'Ariano, F. Manessi, and P. Perinotti, Determinism without causality, Physica Scripta **2014**, 014013 (2014).

[254] M. Wilson and G. Chiribella, Causality in higher order process theories, Electron. Proc. Theor. Comput. Sci. **343**, 265 (2021a).

[255] S. Milz, J. Bavaresco, and G. Chiribella, Resource theory of causal connection, Quantum **6**, 788 (2022).

[256] D. Jia and F. Costa, Causal order as a resource for quantum communication, Phys. Rev. A **100**, 052319 (2019).

[257] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, Phys. Rev. A **97**, 012127 (2018).

[258] I. A. Luchnikov, S. V. Vintskevich, D. A. Grigoriev, and S. N. Filippov, Machine learning non-Markovian quantum dynamics, Phys. Rev. Lett. **124**, 140502 (2020).

[259] G. Zambon and G. Adesso, Quantum processes as thermodynamic resources: The role of non-Markovianity, Phys. Rev. Lett. **134**, 200401 (2025).

[260] G. D. Berk, A. J. P. Garner, B. Yadin, K. Modi, and F. A. Pollock, Resource theories of multi-time processes: A window into quantum non-Markovianity, Quantum **5**, 435 (2021).

[261] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation (3rd Edition)* (Addison-Wesley Longman Publishing Co., Inc., USA, 2006).

[262] K. Rosen, *Discrete Mathematics and Its Applications, Eighth Edition* (McGraw-Hill Education, 2019).

[263] J. Hefford and M. Wilson, A profunctorial semantics for quantum supermaps, in *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS ’24 (Association for Computing Machinery, New York, NY, USA, 2024).

[264] G. Gour, Comparison of quantum channels by superchannels, IEEE Trans. Inf. Theory **65**, 5880 (2019).

[265] N. H. Y. Ng, L. Mančinska, C. Cirstoiu, J. Eisert, and S. Wehner, Limits to catalysis in quantum thermodynamics, New J. Phys. **17**, 085004 (2015).

[266] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Yunger Halpern, The resource theory of informational nonequilibrium in thermodynamics, Phys. Rep. **583**, 1 (2015).

[267] S. Du, Z. Bai, and Y. Guo, Conditions for coherence transformations under incoherent operations, Phys. Rev. A **91**, 052120 (2015).

[268] S. Chen, X. Zhang, Y. Zhou, and Q. Zhao, One-shot coherence distillation with catalysts, Phys. Rev. A **100**, 042323 (2019).

[269] R. Takagi and N. Shiraishi, Correlation in catalysts enables arbitrary manipulation of quantum coherence, Phys. Rev. Lett. **128**, 240501 (2022).

[270] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nat. Commun. **4**, 2059 (2013).

[271] P. Lipka-Bartosik and P. Skrzypczyk, All states are universal catalysts in quantum thermodynamics, Phys. Rev. X **11**, 011061 (2021).

[272] R. Rubboli and M. Tomamichel, Fundamental limits on correlated catalytic state transformations, Phys. Rev. Lett. **129**, 120506 (2022).

[273] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev **5**, 183 (1961).

[274] Z.-W. Liu and A. Winter, Resource theories of quantum channels and the universal role of resource erasure (2019), arXiv:1904.04201 [quant-ph] .

[275] Y. Liu and X. Yuan, Operational resource theory of quantum channels, Phys. Rev. Res. **2**, 012035 (2020).

[276] P. Perinotti, Causal structures and the classification of higher order quantum computations, in *Time in Physics*, edited by R. Renner and S. Stupar (Springer International Publishing, Cham, 2017) pp. 103–127.