Quantum resource theories from the viewpoint of statistical decision theory

PhD thesis

ZHOU Wenbin

Department of mathematical informatics
School of informatics
Nagoya University

Supervisor: Prof. Dr. Francesco BUSCEMI

Nagoya, Japan

Submission date: 2021.01

"The object of life is not to be on the side of the majority, but to escape finding oneself in the ranks of the insane." — Marcus Aurelius, Meditations.

Abstract

Quantum phenomena are treated as resources for various real-world applications in quantum technologies. Comparing resources aims to discover the capability of quantum systems in information processing tasks.

In this thesis, we first studied quantum measurement and quantum incompatibility as resources. We showed that through a resource-theoretic approach, it is possible to compare a family of quantum measurements in terms of the programmable measurement device (PMD), of which users can have temporal freedom in issuing programs to control the device. The temporal setting leads to the necessity of having quantum memory in the PMDs, which bridges the connection of quantum incompatibility and quantum memory. A complete set of convertibility conditions for programmable devices is derived based on quantum state discrimination with post-measurement information game. This game scenario can be utilized as a tool to certify a genuine PMD or a device with genuine quantum memory. As a byproduct, we derived sufficient and necessary conditions for the convertibility between single POVMs (a special case of PMD) through the task of minimum-error state discrimination game.

We then studied general quantum resources through pure mathematical languages by representing quantum resources as complex density matrices and formulate a general resource theory with F-morphisms as the restricted transformation. The convertibility between resources is characterized by whether there exists F-morphisms between complex density matrics or not. The core idea from a resource-theoretic viewpoint is that it is possible to quantify resources without referring to the maximal ones. Moreover, it is enough to compare some of the entropic quantities of resources to ensure the existence of F-morphisms between them. With our resource-theoretic frame, it is possible to derive a sufficient and necessary condition of the transformation between an arbitrary resource and the maximal resource or a reference resource.

The resource-theoretic approach is essentially the core idea of the statistical comparison theory which was developed in mathematical statistics mainly by Blackwell in the 1950s. The main picture of this thesis is inspired and developed from the viewpoint of statistical comparison theory, then extended into quantum statistics based on a game-theoretic approach.

Acknowledgments

Studying quantum information science has brought me immeasurable pleasure. In the process of understanding quantum information and pursuing a Ph.D degree, I have received considerable support and help from several people, without their help I would never have made it this far.

First, I would like to thank my supervisor, Francesco Buscemi, for his guidance, teaching, and sharing in the science world during my whole Ph.D course. He has helped me build my mathematical foundations for quantum information science and has guided me through a pathway of understanding and pursuing science. He has also introduced me to many of his colleagues in the community, which whom I was able to engage in meaningful conversations and discussions about science. I highly appreciate his consistent support. During my Ph.D journey, I also had the great opportunity to work with Professor Eric Chitambar at the University of Illinois Urbana-Champaign for six months. He kindly welcomed me in his laboratory and guided me in understanding quantum incompatibility and quantum resource theories. As for my process of understanding quantum resource theories and quantum other things, I also want to thank Bartosz Regula, Ryuji Takagi, Máté Farkas, Edgar A. Aguilar, Ziwen Liu, Xiao Yuan, and Qi Zhao for their insightful discussions, corrections on the manuscripts and idea sharing. I had another great opportunity to visit Peng Cheng Laboratory in Shenzhen, hospitalized by professor Hayashi Masahito and his lab members, Kun Wang, Seunghoan Song, Yuuya Yoshida, and Ziyu Liu. I appreciate their organization and hospitality. In my early state of the Ph.D program, I had an internship opportunity in NTT basic research lab and Dr. Go Kato was my supervisor, I thank him for the guidance in understanding Shor's algorithm.

I would also like to give special thanks to Huan-Yu Ku and Shin-Liang Chen, who have been discussing with me constantly during my Ph.D journey on various topics in quantum information science. In our department of mathematical informatics, I want to thank Mrs. Mizue Kato, the secretary of our department, who has been very supportive in my paperwork and financial issues.

Financially, my Ph.D journey has been supported by RWDC leading program in school of informatics, Nagoya University. I thank their efforts for supporting my research activities.

Lastly, I would like to thank my parents from the bottom of my heart for their support, encouragement and unconditional love.

List of Figures

Contents

List of Papers

Paper 1

Francesco Buscemi, Eric Chitambar and Wenbin Zhou. Complete Resource Theory of Quantum Incompatibility as Quantum Programmability. *Phys. Rev. Lett.*, 124:120401, March 2020.

Paper 2

Wenbin Zhou and Francesco Buscemi. General state transitions with exact resource morphisms: a unified resource-theoretic approach. *Journal of Physics A: Mathematical and Theoretical*, 53(44):445303, October 2020.

Paper 3 (under review)

Shin-Liang Chen, Huan-Yu Ku, Wenbin Zhou, Jordi Tura, Yueh-Nan Chen. Robust self-testing of steerable quantum assemblages and its applications on device-independent quantum certification. *Pre-print: arXiv:2002.02823*,March 2020.

Chapter 1

Introduction to quantum information science

1.1 Research background

Quantum information science is an interdisciplinary field that involves information theory and quantum mechanics. Information theory was built in the 1950s by Claude Shannon, the father of the information age. In his well-known paper [1], he discovered that all messages with whatever length and meaning were all essentially reducible to the same element: bits, and he also showed how one could compress and encode those bits to transmit information with flawless accuracy. Quantum mechanics is the framework to describe the fundamentals of physical nature at the level of the microscopic-scale world, i.e., the quantum world, including the level of atoms, electrons, and photons. The combination of understanding quantum mechanics and the fundamentals of information theory has inspired the emergence of quantum technologies, which have been developed exponentially and rapidly in recent years. We have seen a trend in the industry that major technology companies, such as Google, IBM, Microsoft, NEC, Toshiba, Alibaba, and Huawei, are building their own quantum devices based on quantum technologies which are potentially powerful in solving typical optimization problems. Thus the trend of research in quantum information science has become evident.

To dive into the quantum information science, it is necessary to first define what is quantum information. To understand the concept, one must begin from the classical information in Shannon's framework, in which "information" is a mathematical term defined quantitatively (entropy) to quantify how uncertain a random variable (i.e, information source) is[2]. The unit of the information is the classical bit, which can be engineered physically. A bit can be represented by the state of a transistor–either ON or OFF of transistors or an electrical charge in a microscopic capacitor, among others. The carriers for classical information follow the classical information theory. Analogically, quantum information is carried in quantum particles, such as photons, and electrons. In this context, Von Neumann provides a technical definition of entropy similar to that of Shannon[3], called the Von Neu-

mann entropy. Von Neumann's definition differs from classical information in that the carrier–quantum system–that is used to build the bit follows the framework of quantum mechanics. Hence, device engineers must ensure that the carrier for the classical information process does not get into quantum effects that follow quantum mechanics. As for the quantum information process, the quantum effect must be considered. A suitable example that terminologically represents quantum information is the entanglement between different physical systems, which is a physical phenomenon named by Erwin Schrödinger in a reply letter to Einstein, Podolsky, and Rosen's EPR paper [4] in the 1930s. While Einstein was not convinced by the idea of entanglement, it was originally proved theoretically after his death by John Bell [5] in 1964 and first demonstrated in Aspect's experiment [6] in 1981. Since then entanglement has been exploited as one of the most useful quantum resources for building quantum technologies. Loosely speaking, quantum information is generally represented as information that can be encoded in the quantum system, which can be any physical phenomena that follows quantum mechanics but cannot be simulated by classical mechanics alone.

From the industrial viewpoint, the study of quantum information science aims to build and manipulate quantum systems to process quantum information of some practical tasks that solve real-world problems, such as building a quantum computer to solve complex optimization problems, and creating quantum protocols for cryptography, among many others. An example that demonstrated the importance of the development of quantum technology was the introduction of Shor's algorithm [7], which could factor exceptionally large numbers rapidly, and efficiently through manipulation of quantum systems that are built in the quantum computer. The success of the demonstration has been a thread for current cryptography technology that is mainly built based on the difficulty of factoring large numbers. Thus, building quantum cryptography has become increasingly necessary so that cryptography system can be adapted to the quantum technology era.

However, from a theoretical viewpoint, the study of quantum information science aims to understand and explore the fundamentals of the information process within quantum systems to discover new possibilities in the information processing tasks that so far have been impossible for traditional methods. Therefore, such an investigation could lead to the discovery and even the transposition of the limits of nature, which in turn could be beneficial for the industry in the process of building quantum devices. This thesis investigates the fundamentals of some quantum properties in quantum information science.

1.2 Introduction to quantum resource theory

Driven by their needs on Earth, several societies have discovered resources provided by nature that are needed to build technologies to improve our modern life. We naturally divide our nature into micro and macro-worlds by different viewing angles, since using a single standard to understand both dimensions could lead

to conflicts. Especially in the micro-world, some phenomena are very counter-intuitive and difficult to understand if the laws of the macro-world are applied to them. Over the past centuries, researchers have been striving to understand the micro-world, where quantum mechanics plays a key role. Based on research and observation of the micro-world, many innovative technologies that benefit modern societies could be built. In this context, the micro-world is here called quantum world, which is backed by quantum theories.

In the quantum world, a recently developed general framework called quantum resource theory has provided a mathematical tool to address issues in quantum information processing. Not only does this tool generalize properties of quantum systems in the information processing[8, 9], but it also indicates limitations and possibilities in some certain information processing tasks [10, 11]. In the framework of quantum resource theory, these non-classical phenomena are called resource, such as superposition, entanglement, coherence, and measurement incompatibility, while those that are not resources are called the free resource. In general, free resources are those that can be simulated without using quantum resources. Another important ingredient of quantum resource theory is the "restricted transformation" among resources. Restricted transformation cannot create resources alone. As in the theory of quantum entanglement, local operation and classical communication, as a restricted transformation, does not create the resource–entanglement–by itself.

Restricted transformations among resources lead to ordering between them. Given two resources, how one could know which one has higher ordering is the core issue in any quantum resource theory. The ordering turns out to be a pre-ordering, which is always the case in quantum resource theories in which not all resources are comparable,i.e., there does not exist a restricted transformation between some resources. One way to examine this problem is through statistical comparison between resources[12], which enables characterization of the pre-ordering between statistical models. Statistical comparisons investigate the statistical properties of statistical models through payoffs in a game-theoretic approach. Statistical models can be built from quantum resources[12]. Any resources can be compared under a statistical model (a game) to make decisions on which one has the highest ordering. The lowest ordering of resources in the ordering ranking are proved to be free resources.

Observing the statistical property of the quantum system as a resource allows one to study how powerful it is and how we can make the most use of it, as well as the limitation of the resources. For example, quantum entanglement is a powerful resource that classical physics cannot simulate and is the core for building quantum computers and quantum communication technologies. Observing the statistical property of quantum entanglement provides a tool for understanding its limitations. The more resources it can hold and the more it can be controlled, the more powerful the resource will be in some information processing tasks.

In this thesis, we reformulate the framework of quantum resource theory in a specific setting and in general setting and we apply this framework to some physical settings, and rederive some of the previous results. We first study quantum

resource theory from the viewpoint of Blackwell and Shannon, and develop a specific resource theory in addressing quantum measurement. Then we generalize the theory in a framework of quantum resource theory with various applications in other fields of physics.

Chapter 2

Preliminaries

2.1 Preliminaries of quantum information theory

Different from classical information processing, quantum information processing relies on the quantum system, which is a physical system typically at the atomic scale where quantum mechanical effects appear. Examples of them are physical systems of atoms, electrons, photons, and so on. The characteristics of a physical system are determined by physical quantities, such as position, momentum, energy, (spin) angular momentum of a particle, the polarization of a photon, etc. We describe quantum systems as **quantum states** and we describe physical quantities of quantum systems as **observables**. Measuring quantities of observables from a quantum state are called **quantum measurement**. In what follows, we will describe mathematically, what are quantum state, quantum observable, and quantum measurement.

Before we move into the description of them, we need some notations and mathematical concepts. Most of the following contents about the description of the quantum system are taken from the book [13].

2.2 Mathematical notation

2.2.1 Dirac notation

In this sector, we introduce the Dirac Notation. Dirac notation is a very useful and reasonabel tool for describing the theory of quantum mechanics. Through out the thesis, we restrict ourself to the finite complex Euclidean space $\mathbb{C}^d$ with an arbitrary dimmension d. A Dirac notation called ket vector $|\cdot\rangle$ is used to describe a column vector of $\mathbb{C}^d$. For example, we represent an element in $\mathbb{C}^3$ as $|\Psi\rangle := \Big(1, 2, 3+2i\Big)^{\mathrm{T}} \in \mathbb{C}^3$. A complex conjugate of the column vector $|\Psi\rangle$, called bra vector, is denoted by $\langle\Psi| := \Big(1, 2, 3-2i\Big)$. We normally use $\dagger$ to denote the complex conjugate symbol, so we see that, $\langle\Psi| = |\Psi\rangle^{\dagger}$. The inner product between two vectors, say $|\Phi\rangle := \Big(a_1, a_2, ..., a_d\Big)^{\mathrm{T}}$ and $|\Psi\rangle := \Big(b_1, b_2, ..., b_d\Big)^{\mathrm{T}}$, is denoted by $\langle\Phi|\Psi\rangle := \sum_i \overline{a}_i b_i$. With

inner product denoted, we use notation $\|\Psi\| := \sqrt{\langle\Psi|\Psi\rangle}$ to represent the norm of a column vector $|\Psi\rangle$. In addition, we consider another matrix product by swapping the order of a ket and bra. A product of ket vector, $d \times 1$ matrix, and a bra vector, $1 \times d$ matrix in this order is a $d \times d$ matrix. For example for $|\Phi\rangle = (a_1, a_2)^{\mathrm{T}}$ and $|\Psi\rangle = (b_1, b_2)^{\mathrm{T}} \in \mathbb{C}^2$, we have,

$$|\Phi\rangle\langle\Psi| = \begin{pmatrix} a_1\bar{b}_1 & a_1\bar{b}_2 \\ a_2\bar{b}_1 & a_2\bar{b}_2 \end{pmatrix}$$

In addtion, there is a very basic but useful calculation rule that can be applied to Dirac notation introduced above. This rule will make calculation easy for later usage. We list it as follows:

$$|\Phi\rangle\langle\Psi||\xi\rangle = \langle\Psi|\xi\rangle|\Phi\rangle$$

where we could truncate the symbols and move around the number $\langle\Psi|\xi\rangle$ around anywhere inside the term that is made of combination of Dirac notations without changing the whole intepretation of this term.

2.2.2 Hilbert space

In this sector, we introduce the mathematical definition of Hilbert space. For this thesis, we only focus on a finite-dimensional Hilbert space. A finite dimensional complex inner product space is called a Hilbert space. A linear space V is called a complex inner product space if V has an operation $V \ni |\Psi\rangle, |\Phi\rangle \mapsto \langle\Psi|\Phi\rangle \in \mathbb{C}$ with the following three properties:

1. positivity: $\langle\Phi|\Phi\rangle \geqslant 0$ with equality holds if and only if $|\Phi\rangle = 0$;
2. Hermitian symmetry: $\langle\Psi|\Phi\rangle = \overline{\langle\Phi|\Psi\rangle}$;
3. linearity: $\langle\Psi|a\Phi + b\Upsilon\rangle = a\langle\Psi|\Phi\rangle + b\langle\Psi|\Upsilon\rangle$. [1]

From the definition of inner product space, we see that the dimension is the key property that defines a Hilbert space. It is easy to prove the following statement.

Proposition 1. *All Hilbert spaces with the same dimension are isomorphic.*

By isomorphic, we mean that there exists a bijection between any two Hilbert spaces with the same dimension such that the linear structure and their inner product are preserved. So mathematically, all Hilbert spaces with the same dimension are equivalent to each other, we only care about the dimension of a Hilbert space. Assume an arbitrary Hilbert space $\mathcal{H}$ with dimension d, we write,

$$\mathcal{H} \cong \mathbb{C}^d. \tag{2.1}$$

[1] Here we write $a|\Phi\rangle + b|\Upsilon\rangle = |a\Phi + b\Upsilon\rangle$.

2.2.3 Linear operators

We talk about linear operator on a Hilbert space. Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, a map $A : \mathcal{H} \to \mathcal{K}$ is called a linear operator if it satisfies the linearity condition:

$$A(a|\Psi\rangle + b|\Phi\rangle) = aA|\Psi\rangle + bA|\Phi\rangle \quad \forall|\psi\rangle, |\Phi\rangle \in \mathcal{H}, a, b \in \mathbb{C}. \tag{2.2}$$

Let $\{|\Psi_i\rangle\}_{i=1}^{d_1}$ and $\{|\Phi_j\rangle\}_{j=1}^{d_2}$ be orthonormal basis of $\mathcal{H}$,$\mathcal{K}$ and a matrix with elements $a_{i,j} = \langle\Phi_i|A|\Psi_j\rangle$ is called the representation matrix of A. Hence, by fixing bases, one obtains an equivalent matrix representation as linear operators. In what follows, we use $\mathcal{L}(\mathcal{H}, \mathcal{K})$ to represent linear operators from $\mathcal{H}$ to $\mathcal{K}$ and $\mathcal{L}(\mathcal{H}, \mathcal{H})$ is abbreviated as $\mathcal{L}(\mathcal{H})$. Suppose $A \in \mathcal{L}(\mathcal{H}, \mathcal{K})$, we denote the adjoint operator of A as $A^\dagger$ [2] such that $\langle\Phi|A|\Psi\rangle = \langle A^\dagger\Phi||\Psi\rangle$. Notice that the matrix representation of $A^\dagger$ is a conjugate transpose of the matrix representation of A.

Two very important classes of operators are semi-definite positive operator and Hermitian operator. A semi-definite positive operator $A \in \mathcal{L}(\mathcal{H})$ is denoted as $A \geqslant 0$, such that $\langle\Psi|A|\Psi\rangle \geqslant 0, \forall|\Psi\rangle \in \mathcal{H}$. A hermitian operator $A \in \mathcal{L}(\mathcal{H})$ satisfies $A = A^\dagger$. It is easy to prove that eigenvalues of hermitian operator are real numbers. For any Hermitian operators, the following proposition is useful in later use.

Proposition 2 (Spectral decomposition)**.** *Any hermitian operator A can be written in the form*

$$A = \sum_a aP_a$$

where P_a is the eigen-projection of the eigenvalue a of A satisfying $P_aP_b = \delta_{ab}P_a$.

Binary relations is considerably used throughout the thesis, of which preorder and partial order are of interest.

Definition 2.2.1 (Preorder and partial order)**.** *Consider some set $\mathcal{S}$ and a binary relation $\succ$ on $\mathcal{S}$. Then $\succ$ is a preorder, if it is reflexive and transitive; i.e., for all a, b and $c \in \mathcal{S}$, we have that:*

1. *$a \succ a$;*
2. *if $a \succ b$ and $b \succ c$, then $a \succ c$.*

If a preorder is also antisymmetric, that is, $a \succ b$ and $b \succ a$ implies $a = b$, then it is a partial order.

2.3 Quantum terminology and postulates

2.3.1 Quantum state

Quantum mechanics is a mathematical framework for the development of quantum physical theories. The postulates of quantum mechanics provides the ground

[2]The symbol † is called dagger, it usually describes adjoint operator in physics.

to describe in mathematical terms all quantum phenomena. It is noted that with postulates quantum mechanics becomes easy to formulate without touching the deep theories that back the postulates. In quantum physics, a quantum state is a mathematical entity that provides a probability distribution for the outcomes of each possible measurement on a quantum system.

Postulate 1: *For any quantum system, there is an associated Hilbert space $\mathcal{H}$ in a way that a physical state is represented by a unit vector of $\mathcal{H}$, the state space.*

For each quantum system, such as an electron, a photon, we assume that there is an associated Hilbert space with which all the physics on the system are described. We denote a pure quantum states as a unit vectors $|\psi\rangle \in \mathcal{H}$, such that $\langle\psi|\psi\rangle = 1$. By linearity, we notice that there exists state as a unit vector such that $|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle \in \mathcal{H}$ for some unit vectors $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ and $a, b \in \mathbb{C}$. We claim these states contain superposition, which is one of the fundamental phenomena in quantum physics. Note that superposition is a counter-intuitive concept. An interpretation of superposition state is that it is a state that coexists with another state. The Schroedinger's Cat is a famous paradox example that describes superposition.

Besides the pure state, there exists mixed state, which is an ensemble of pure states with probability distribution associated with them. The mixed states can describe those states that are not completely known. We denote $\{p(i), |\Psi_i\rangle\}$ as an ensemble of pure states, where $p(i)$ is a probability distribution. In reality, mixed states happened when dealing with a physical system that consists of many particles, so it is impossible to describe it with a pure state. A typical example of a mixed state physics is a thermal equilibrium state. Under the limitation of pure states, we redefine the quantum state through density operators. For an ensemble $\{p(i), |\Psi_i\rangle\}$, we define a density operator,

$$\rho := \sum_i p(i)|\Psi_i\rangle\langle\Psi_i|$$

It is easy to see that $\mathrm{Tr}\{\rho\} = 1$ and $\rho \geqslant 0$. Now we can modify the previous postulate as follows,

Postulate 1': *For any quantum system, there is an associated Hilbert space $\mathcal{H}$ in a way that a state is completely described by its density operator, which is a semi-definite positive operator $\rho \geqslant 0$ with trace one, acting on Hilbert space $\mathcal{H}$ of the system.*

While it is necessary to represent the quantum state through density operators, the unit vector representation could be useful in most of the cases where the probability does not play an important role.

Another good function of the density operator is to represent the classical state. Suppose we have a random variable X, and with a probability distribution $p(x)$ attached to its outcome x. We write a diagonal density operator as,

$$\rho = \sum_x p(x)|x\rangle\langle x|$$

where $\{|x\rangle\}$ forms a orthonormal basis for a Hilbert space with dimension $|X|$ and x is the label.

Remark. *We notice that quantum state can represent both classical random variable and quantum random variable. In information theory, a random variable is the information source, so we could understand the quantum state as a quantum resource such that it is not simulatable by a classical random variable in terms of information.*

2.3.2 Quantum measurement

In quantum physics, measurement is the testing or manipulation of a physical system to yield a numerical result with prediction. The prediction that quantum physics makes are in general probabilistic, which means that after measuring on a quantum system, we will get numerical results with uncertainty.

Postulate 2: *An observable of a physical quantity is represented by a Hermitian operator on $\mathcal{H}$ where the measurement outcome is one of its eigenvalues. If we measure a physical quantity $A \in \mathcal{L}(\mathcal{H})$ under a state $|\Psi\rangle$, then the probability to observe an outcome a is given by the **Born rule**, $Pr(A = a||\Psi\rangle) := \langle\Psi|P_a|\Psi\rangle$, where P_a is the eigen-projection of the eigenvalue a of A.*

There is another way to represent the measurement that is more general than observable, for details, we encourage readers to refer to standard textbooks[14]. In general, we use positive operator-valued measure (POVM) to represent quantum measurements. Mathematically, a POVM $\mathbb{P}$ is a set $\{P_a\}$ of operators on $\mathcal{H}$ with a label set $\mathcal{A}$ that is indexed by integers, such that all $P_a \in \mathcal{L}(\mathcal{H})$, i.e.,$\forall a, \mathcal{L}(\mathcal{H}) \ni P_a \geqslant 0$ and $\sum_a P_a = \mathbb{1}$. In POVM's, we only care about the statistics of measurement instead of concerning the exact value, which is a real number, that comes out from the measurement device. We normally represent the result of its labels. In what follows, we use POVM and measurement interchangeably. For a quantum state that is represented by a density operator ρ, the probability of getting a-th outcome from the measuring device is calculated as, $Pr(a) = \mathrm{Tr}\{\rho P_a\}$, which is derived by the Born rule.

Another important assumption for quantum mechanics is that we will need to describe the composite system from Hilbert spaces associated with the subsystems. The following postulate describes it in detail.

Postulate 3: *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Let S_1 and S_2 be two physical systems, and their corresponding state spaces $\mathcal{H}$ and $\mathcal{K}$, the state space of composite system $S := S_1 + S_2$ is tensor product $\mathcal{H} \otimes \mathcal{K}$. An observable of $\mathcal{H}$ represented by a Hermitian operator A on $\mathcal{H}$ is identical with the observable of S represented by $A \otimes \mathbb{1}_2$ on $\mathcal{H} \otimes \mathcal{K}$, where $\mathbb{1}_2$ is the identity operator on $\mathcal{K}$, the same goes for $\mathbb{1}_1 \otimes B$ on $\mathcal{H} \otimes \mathcal{K}$.*

2.3.3 Quantum channel

In micro-world, the evolution of particles is described by Schrodinger's equation, the following postulate gives an equivalent explanation of it. For a closed system, such as the whole universe, we are interested in the mechanics of its elements.

Postulate 4: *The time evolution of a closed quantum system is described by a unitary transformation. That is, the state ρ of the system at time t_1 is evolved to the state ρ' of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,*

$$\rho' = U\rho U^\dagger. \tag{2.3}$$

A quantum channel takes a quantum state as input and outputs another quantum state. Through the postulate, we can describe a quantum channel as follows,

Definition 2.3.1 (Stinespring representation)**.** *Let $\rho \in \mathcal{L}(\mathcal{H})$, let E be the environment system, U be the unitary operator on $\mathcal{H} \otimes E$, e_0 be an initial pure state of environment system, $e_0 \in \mathcal{L}(\mathcal{E})$, the quantum channel $\mathcal{N}$ is defined by*

$$\mathcal{N}(\rho) = Tr_E\{U(\rho \otimes e_0)U^\dagger\} \tag{2.4}$$

Hence, any quantum evolution either in an open quantum system or closed quantum system can be translated into a specific map, quantum channel, that operates on the system.

In general, we denote the quantum channel as a linear map that satisfies typical conditions, defined as follows:

Definition 2.3.2 (CPTP map as channel)**.** *Given input and output Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, a linear map $\mathcal{N} : A \mapsto B$, where $A \in \mathcal{L}(\mathcal{H}), B \in \mathcal{L}(\mathcal{K})$, is a quantum channel if and only if it is a complete positive trace preserving (CPTP) map. Let $\mathcal{R}$ be an extra quantum system with arbitrary dimensionality, let $\omega_{AR} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{R})$, then quantum channel $\mathcal{N}$ satisfies the following conditions:*

1. $A \geqslant 0 \implies B \geqslant 0$ *and* $\mathrm{Tr}\{A\} = \mathrm{Tr}\{B\}$;
2. $\omega_{AR} \geqslant 0 \implies \omega_{AB} := (\mathcal{N} \otimes id_R)(\omega_{AR}) \geqslant 0$ *and* $\mathrm{Tr}\{\omega_{AR}\} = \mathrm{Tr}\{\omega_{AB}\}$;

where id_R denotes the identity map on system R.

Equivalent to CPTP map, we can have another representation of the quantum channel, which is the Kraus representation.

Definition 2.3.3 (Kraus representation)**.** *A map $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \longrightarrow \mathcal{L}(\mathcal{H}_B)$ is a CPTP map if and only if it has a decomposition as follows:*

$$\mathcal{N}(\rho) = \sum_{l=0}^{d-1} V_l \rho V_l^\dagger, \tag{2.5}$$

where $\rho \in \mathcal{L}(\mathcal{H}_A), V_l \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ *for all* $l \in \{0, ..., d-1\}$,

$$\sum_{l=0}^{d-1} V_l^\dagger V_l = I_A, \tag{2.6}$$

and d *need not be any larger than* $dim(\mathcal{H}_A \otimes \mathcal{H}_B)$.

Among all these descriptions of quantum channels, we can prove that they are all equivalent to each other. The following theorem gives us a different method to represent a quantum channel in our interests.

Theorem 1. *The following conditions are equivalent for a linear map* $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \longrightarrow \mathcal{L}(\mathcal{H}_B)$, *where* $\mathcal{H}_A$ *and* $\mathcal{H}_B$ *represent the Hilbert space for system* A *and* B.

1. $\mathcal{N}$ *is a CPTP map.*
2. $\mathcal{N}$ *admits a Kraus representation.*
3. $\mathcal{N}$ *admits a Stinespring representation.*

2.3.4 Quantum instrument

In quantum measurement described above, we only care about the classical output statistics of the measuring device by ignoring the quantum output. When both the classical outputs and quantum outputs are of interest. We need another mathematical framework, a quantum instrument, to describe the process.

Definition 2.3.4 (Quantum instrument)**.** *Let discrete set* $x \in \mathcal{X}$ *be the outcome set of a measurement, let* $\{\mathcal{E}_x\}_{x\in\mathcal{X}}$ *denote trace non-increasing complete positive maps such that* $\sum_x \mathcal{E}_x$ *as a whole is CPTP maps. The probability of measuring a specific outcome* x *on a input quantum state* ρ *is given by,*

$$p(x|\rho) := \mathrm{Tr}\{\mathcal{E}_x(\rho)\} \tag{2.7}$$

while the quantum output of the measuring device is given by,

$$\rho_x := \frac{\mathcal{E}_x(\rho)}{p(x|\rho)}. \tag{2.8}$$

Through the quantum instrument, we can reformulate the measurement process when an after-measurement quantum state needs to be described.

Postulate 4': *A general measurement process of a quantum system is described by a quantum instrument.*

Chapter 3

Statistical model comparison

3.1 Introduction to quantum statistical model comparisons

This chapter introduces a well-known theory on comparisons of statistical models in terms of statistical decision problems in the quantum domain. Before we move into the quantum domain, we will give an intuitive understanding of what is a statistical decision problem by introducing a classical example. Suppose that an experimentalist wants to identify what an unknown object is with classifiable features, an example could be, what is the classification of bacteria? The experimentalist could take some samples of the bacteria and develop them in Petri dishes, then try to identify them through the statistics of shapes of their growth in the dishes. In this context, the experimentalist made a decision based on the observation of statistics from bacteria. Let us formulate the statistical decision problems in mathematical language.

3.2 Classical Statistics

Definition 3.2.1 (Noisy channel)**.** *We define a discrete channel to be a system consisting of an input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and a conditional probability distribution $N = \{p(y|x)\}_{x\in\mathcal{X},y\in\mathcal{Y}}$, satisfying $p(y|x) \geqslant 0$ and $\sum_{y\in\mathcal{Y}} p(y|x) = 1, \forall x \in \mathcal{X}$. We then denote the noisy channel as,*

$$N : \mathcal{X} \to \mathcal{Y}. \tag{3.1}$$

Definition 3.2.2 (Statistical model)**.** *A (finite) statistical model $\boldsymbol{m}$ is a triple $(\Theta, \mathcal{X}, \Omega)$, where Θ is a (finite) discrete parameter set $\{\theta\}_{\theta\in\Theta}$, $\mathcal{X}$ is a (finite) discrete sample set $\{x\}_{x\in\mathcal{X}}$, and Ω is a noisy channel $\Omega : \Theta \to \mathcal{X}$.*

Remark. *Intuitively, when we try to extract information from the parameter set, we assume no perfection, i.e., there will be always noise behind the information extraction. Normally, we see Θ as the set of states of nature.*

Definition 3.2.3 (Statistical decision problem)**.** *A statistical decision problem is a triple* $(\Theta, \mathcal{U}, \ell)$*, where* $\mathcal{U}$ *is a finite decision set* $\{u\}_{u\in\mathcal{U}}$*, and* $\ell : \Theta \times \mathcal{U} \to \mathbb{R}$ *is a payoff function.*

Definition 3.2.4. *The expected payoff* $\mathbb{E}$ *of a statistical model* $\boldsymbol{m} = (\Theta, \mathcal{X}, \Omega)$ *with respect to a statistical decision problem* $(\Theta, \mathcal{U}, \ell)$ *is given by the following,*

$$\mathbb{E}(\boldsymbol{m}) := \max_{d} \sum_{u,x,\theta} \ell(\theta, u) d(u|x) \omega(x|\theta) |\Theta|^{-1} \tag{3.2}$$

where $d : \mathcal{X} \to \mathcal{U}$ *is a decision strategy and the strategy set is* $\{d : d(u|x) \geqslant 0, \quad \sum_u d(u|x) = 1 \quad \forall x \in \mathcal{X}\}$*.*

Remark. *Note that if we set the payoff function to be a delta function, we end up with a guessing game in which we guess at the best strategy of the unknown parameter set* Θ*.*

Definition 3.2.5 (Statistical model comparison)**.** *Given two statistical models* $\boldsymbol{m} = (\Theta, \mathcal{X}, \Omega)$ *and* $\boldsymbol{m}' = (\Theta, \mathcal{Y}, \Omega')$*, we say that,*

$$\boldsymbol{m} \succ \boldsymbol{m}', \tag{3.3}$$

whenever for all decision problems $(\Theta, \mathcal{U}, \ell)$*,*

$$\mathbb{E}(\boldsymbol{m}) \geqslant \mathbb{E}(\boldsymbol{m}'). \tag{3.4}$$

Definition 3.2.6 (Noisy channel degrading)**.** *Given two noisy channels* $\Omega : \Theta \to \mathcal{X}$ *and* $\Omega' : \Theta \to \mathcal{Y}$ *with same initial inputs, we say* Ω *can be degraded into* Ω' *whenever there exists another noisy channel* $\phi : \mathcal{X} \to \mathcal{Y}$ *such that* $\Omega' = \Omega \circ \phi$*. This degradedness is denoted by*

$$\Omega \succ \Omega'. \tag{3.5}$$

Remark. *We notice that* Ω *is more "informative" than* Ω'*, meaning that* Ω *is less noisy than* Ω' *in terms of information extraction because we add an additional noisy channel right after* Ω*.*

A well-known theorem that connects the "informativeness" and statistical comparison is the Blackwell-Sherman-Stein (BSS, [15][16][17]) theorem.

Theorem 2 (BSS theorem)**.** *Given two statistical models* $\boldsymbol{m} = (\Theta, \mathcal{X}, \Omega)$ *and* $\boldsymbol{m}' = (\Theta, \mathcal{Y}, \Omega')$*, we have the following equivalence:*

$$\Omega \succ \Omega' \iff \boldsymbol{m} \succ \boldsymbol{m}'. \tag{3.6}$$

Remark. *We have characterized the noisy channel degradedness in terms of statistical decision game payoff's comparisons.*

3.3 Shannon Channel ordering with encoding and decoding

The BSS theorem gives us a different viewpoint of understanding the comparison between noisy channels, in which situations we only compare noisy channels with the same dimension of inputs. The possible reason is that the information parameter Θ as input in general is unknown, and we are extracting information from unknown natural statistics. So comparing noisy channels with the limitation of being unable to touching of the input fits well in the BSS scenario.

However, when it comes to noisy channels for which we have control over both input and output, we are interested in how to transform one noisy channel into another one with the control of processing both input and output of noisy channels.

Definition 3.3.1 (Channel inclusion [18]). *Given two noisy channels $K_1 : \mathcal{X} \to \mathcal{Y}$ and $K_2 : \mathcal{W} \to \mathcal{Z}$ such that the inputs, $\mathcal{X}, \mathcal{W}$ and the outputs, $\mathcal{Y}, \mathcal{Z}$ are different. We say K_1 can be transformed into K_2 whenever there exists n pairs of noisy pre-channels $R_\alpha : \mathcal{W} \to \mathcal{X}$ and post-channels $T_\alpha : \mathcal{Y} \to \mathcal{Z}$ and a probability distribution g_α with $1 \leqslant \alpha \leqslant n$, such that $K_2 = \sum_\alpha g_\alpha R_\alpha \circ K_1 \circ T_\alpha$. This transformation is denoted by*

$$K_1 \supseteq K_2 \tag{3.7}$$

if we represent the channel K_1, K_2, R_α and T_α as conditional probability distributions $p(y|x)$, $q(z|w)$, $t_\alpha(z|y)$ and $r_\alpha(x|w)$, respectively. We equivalently write the channel inclusion as,

$$q(z|w) = \sum_{\alpha,x,y} g_\alpha t_\alpha(z|y) p(y|x) r_\alpha(x|w) \tag{3.8}$$

Remark. *We denote channel transformation in terms of channel inclusion notation that is originally addressed by Shannon[18].*

Definition 3.3.2 (Pure channel). *A pure channel is defined as the one whose transition probabilities are either 0 and 1.*

Suppose a channel $\mathcal{N} : \mathcal{X} \to \mathcal{Y}$ is denoted by conditional probability distribution $p(y|x)$. We now consider transmitting a block of codewords with length n.

Given an input sequence, denoted by

$$x^n = (x_1 x_2, ..., x_n) \in \mathcal{X}^n \tag{3.9}$$

after transmitting the codewords through noisy channel, we obtain,

$$y^n = (y_1 y_2, ..., y_n) \in \mathcal{Y}^n \tag{3.10}$$

we denote this transmission as,

$$\mathcal{N}^n : \mathcal{X}^n \to \mathcal{Y}^n \tag{3.11}$$

and we represent this channel in the form of conditional probability as,

$$p(y^n|x^n) := \prod_{i=1}^{n} p(y_i|x_i) \tag{3.12}$$

Remark. *This definition of channels is built by the assumption that the channel is discrete, independent and identical distributed (i.i.d), or, equivalently, memoryless discrete channel.*

Definition 3.3.3 (Encoder)**.** *We call an encoding channel an **encoder** such that it transforms message $\mathcal{M} = \{1, 2, ..., M\}$ into codewords with a length of n, denoted by*

$$\mathcal{E} : \mathcal{M} \to \mathcal{X}^n \tag{3.13}$$

Definition 3.3.4 (Decoder)**.** *We call a decoding channel a **decoder** such that it transforms the output coming from the channel into original message $\mathcal{M} = \{1, 2, ..., M\}$ that was sent into the channel, denoted by*

$$\mathcal{D} : \mathcal{Y}^n \to \mathcal{M} \tag{3.14}$$

Definition 3.3.5 (Code)**.** *A code $(\mathcal{M}, n)$ for the noisy channel $\mathcal{N} : \mathcal{X} \to \mathcal{Y}$ consists of the following:*

1. *A message set $\mathcal{M} = \{1, 2, ..., M\}$;*
2. *An encoder $\mathcal{E} : \mathcal{M} \to \mathcal{X}^n$;*
3. *A decoder $\mathcal{D} : \mathcal{Y}^n \to \mathcal{M}$;*

Definition 3.3.6 (Average probability error)**.** *For a code $(\mathcal{M}, n)$ with encoder $\mathcal{E}$ and decoder $\mathcal{D}$ for the noisy channel $\mathcal{N} : \mathcal{X} \to \mathcal{Y}$, represented by conditional probility distributions $p(y|x)$, if message $\mathcal{M}$ is chosen randomly with probability distribution $p(m)$, we obtain the average probability error as a function defined as,*

$$\epsilon_n\Big(\mathcal{N}, \mathcal{M}, \mathcal{E}, \mathcal{D}, p(m)\Big) := \sum_{m \in \mathcal{M}} p(m) \sum_{y^n : \mathcal{D}(y^n) \neq m} p\Big(y^n | \mathcal{E}(m)\Big) \tag{3.15}$$

Remark. *The intuition is that n copies of noisy channel $\mathcal{N}$ is able to transmit M messages within the average probability error ϵ_n.*

In the paper [18], Shannon proved the characterization of channel inclusion in terms of coding theorem.

Theorem 3 (Shannon channel ordering)**.** *For the following statements,*

1. *$K_1 \supseteq K_2$;*
2. *There exists a code $(\mathcal{M}, n)$ with $\epsilon_n\Big(K_2, \mathcal{M}, \mathcal{E}, \mathcal{D}, p(m)\Big) \implies$ There exists a code $(\mathcal{M}, n)$ with $\epsilon'_n\Big(K_1, \mathcal{M}, \mathcal{E}', \mathcal{D}', p(m)\Big)$, such that $\epsilon'_n \leqslant \epsilon_n$;*

we have the implication:

$$1 \qquad \Longrightarrow \qquad 2 \tag{3.16}$$

Remark. *The intuition for statement 2 is that for the same code $(\mathcal{M}, n)$, noisy channel K_1 can always perform no worse than noisy channel K_2 in terms of decoding the encoded message $\mathcal{M}$ sent through the noisy channel under average probability error .*

Remark. *Note that statement 2 is only a necessary condition for the characterization of Shannon channel ordering, the sufficiency was left open.*

It is noteworthy that, in the Shannon ordering, the input and output alphabets do not need to be the same, whereas in the characterization of noisy channel degrading in Theorem 2, one must assume that the noisy channels share the same input alphabet but can have different output alphabets. In this sense, one can regard degradedness order as a special case of Shannon ordering. Since Shannon's coding theory allows for the encoding and decoding of a noisy channel, it is thus natural to allow noisy channels to be injected into both input and output for the transformation. For example, after some effort, an engineer discovers a better encoder and decoder pair than previous ones, then a transformation with both encoding and decoding is necessary. However, the characterization of the ordering for Shannon is weaker than that of BSS's ordering since the sufficiency was not true.

Understanding how much noise is contained in a noisy channel and how to transform a noisy channel into another one plays a key role in classical information communication. Similarly, in quantum information science we are interested in the comparison between the quantum noisy channels, the completely positive, trace-preserving (CPTP), as well as the transformation between them. The mathematical structure of the CPTP map covers not only the physical evolution of a quantum system but also quantum state preparation, discarding of a (sub)system, and quantum measurement. The following sections discuss quantum measurement in the form of a noisy channel. As mentioned above, similarly to quantum measurement, a classical noisy channel is essentially a platform for information extraction. The only difference lies on whether the input state is a classical or quantum state. Stimulated by classical statistics, we now move to quantum statistics.

3.4 Quantum Statistics

In the classical formulation of statistical decision problems, one seeks for the optimal decision procedure. In the quantum domain, there exists parameter space Θ and the space of decisions $\mathcal{U}$, however, each $\theta \in \Theta$ corresponds to the quantum state ρ_θ on Hilbert space $\mathcal{H}$ of the quantum system $\mathcal{S}$. A decision is to be made according to a quantum measurement on $\mathcal{S}$, and the problem is thus to find the optimum quantum measurement. Let us formulate it in mathematical languages.

Definition 3.4.1 (Quantum Statistical model)**.** *A quantum statistical model **n** is a triple $(\Theta, \mathcal{H}, \{\rho_\theta\})$, where $\mathcal{H}$ is the Hilbert space of quantum system $\mathcal{S}$, for each $\theta \in \Theta$, ρ_θ corresponds to a quantum state of quantum system $\mathcal{S}$.*

Definition 3.4.2 (Decision from POVMs)**.** *For any (finite) decision set $\mathcal{U} = \{u\}$, a POVM $\{P_u\}$ on system $\mathcal{S}$ is correspondent.*

Definition 3.4.3 (Quantum Statistical decision problem)**.** *The expected payoff $\mathbb{E}$ of a quantum statistical model $\boldsymbol{n} = (\Theta, \mathcal{H}, \{\rho_\theta\})$ with respect to a statistical decision problem $(\Theta, \mathcal{U}, \ell)$ is given by the following,*

$$\mathbb{E}(\boldsymbol{n}) := \max_{P_u} \sum_{\theta,u} \ell(\theta, u) \operatorname{Tr}\{P_u \rho_\theta\}\, p(\theta). \tag{3.17}$$

where we normally set $p(\theta) = |\Theta|^{-1}$ to be an a priori probability over the unknown parameter θ.

Remark. *We see that the state discrimination problem with minimum error is a special case of quantum statistical decision problems. Hence, we usually consider a statistical game as a special case of quantum statistical decision problems. In what follows, we will make use of this form several times.*

Definition 3.4.4 (Quantum Statistical model comparison)**.** *Given two statistical models $\boldsymbol{n} = (\Theta, \mathcal{H}, \{\rho_\theta\})$ and $\boldsymbol{n}' = (\Theta, \mathcal{H}', \{\sigma_\theta\})$, we say that,*

$$\boldsymbol{n} \succ \boldsymbol{n}', \tag{3.18}$$

whenever for all decision problems $(\Theta, \mathcal{U}, \ell)$,

$$\mathbb{E}(\boldsymbol{n}) \geqslant \mathbb{E}(\boldsymbol{n}'). \tag{3.19}$$

In the following chapters, we will make use of this idea through special case of decision games, in most cases we make use of game-theoretic approaches.

Chapter 4

Game-theoretic approach

4.1 Game-theoretic approach to quantum measurement

This chapter discusses the noisy processing of quantum measurement, and guessing games based on measurement that serve as a characterization of quantum statistic comparisons.

Quantum measurement aims to retrieve information on the state of a quantum system, which can be considered a noisy channel process in the quantum domain where the inputs will be quantum states and the outputs will be classical statistics. Natural questions to ask include how noisy a typical quantum measurement is and how to compare the noisiness of two quantum measurements? As discussed above in classical statistics,nosiness is addressed through the processing of noisy channels. As for quantum measurement, which is represented by POVM, one makes use of the operational processing of the measurement process. Here, measurement is regarded as a device with input and output that can be processed them independently or simultaneously.

4.2 POVM and its processings

There are two types of POVM processings: pre-processing and post-processing. POVM pre-processing is a quantum channel applied to the input state of the POVM, which is what we understand the evolution of measurement in the Heisenberg picture. Post-processing is a purely classical noisy channel applied to classical measurement outcomes.

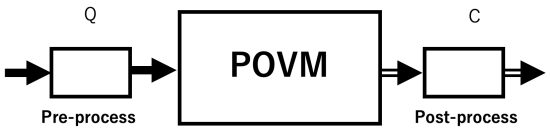

Figure 4.1: Processing of quantum measurements.

The above figure 4.2 illustrate how to process a POVM, we give the mathematical definitions as follows:

Definition 4.2.1 (Pre-processing of POVMs). *POVMs $\mathbb{P} = \{P^a\}$ and $\mathbb{Q} = \{Q^a\}$ are defined as functions that take quantum state as input and output result from same set $\mathcal{A}$. We say $\mathbb{P}$ can be transformed to $\mathbb{Q}$ whenever there exists a CPTP map $\mathcal{E}$ such that $\mathbb{Q} = \mathbb{P} \circ \mathcal{E}$. This transformation is denoted by*

$$\mathbb{P} \succ_{pre} \mathbb{Q}. \tag{4.1}$$

Remark. *Pre-processing of POVM is an unital complete positive map and it might change the dimension of the POVM but will not change the number of outcomes.*

Definition 4.2.2 (Post-processing of POVMs). *Given two POVMs $\mathbb{P} = \{P^a\}$ and $\mathbb{Q} = \{Q^b\}$ as functions with same input quantum system and different output $\mathcal{A}$ and $\mathcal{B}$. We say $\mathbb{P}$ can be transformed to $\mathbb{Q}$ whenever there exists a noisy channel $\phi : \mathcal{A} \to \mathcal{B}$ such that $\mathbb{Q} = \phi \circ \mathbb{P}$. This transformation is denoted by*

$$\mathbb{P} \succ_{post} \mathbb{Q}. \tag{4.2}$$

Remark. *Post-processing of POVM might change the number of the measurement outcomes but it will not change the dimension of POVM.*

Since it is possible to do pre- or post-processing on measurement, then there is no reason to not consider doing pre- and post-processing together, probably with some shared randomness. We will define the most general processing for the process of measurement as follows.

Definition 4.2.3 (General processing of POVMs). *Given two POVMs $\mathbb{P} = \{P_A^a\}$ on system $\mathcal{H}_A$ with outcome set $\mathcal{A}$ and $\mathbb{Q} = \{Q_B^b\}$ on system $\mathcal{H}_B$ with outcome set $\mathcal{B}$, we say $\mathbb{P}$ can be transformed to $\mathbb{Q}$ whenever there exists a CPTP map $\mathcal{E}_r$, a noisy channel ϕ^r and randomness set $\{\mu(r)\}$ such that $\mathbb{Q} = \sum_r \mu(r)\phi_r \circ \mathbb{P} \circ E_r$, denoted by*

$$\mathbb{P} \succ \mathbb{Q}. \tag{4.3}$$

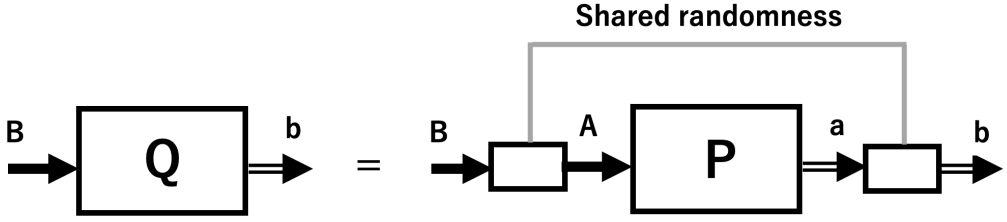

Figure 4.2: How to process POVMs: with shared randomness.

Remark. *A general process of POVM combines both pre-processing and post-processing of POVMs with shared randomness, mathematically $Q_B^b = \sum_r \mu(r) \sum_a w(b|a,r)\mathcal{E}_r^\dagger(P_A^a)$, and $w(b|a,r)$ is a conditional probability distribution.*

Definition 4.2.4 (Trivial POVMs). *A POVM $\mathbb{P} = \{P^a\}$ is trivial if its effects are proportional to identity, such that, $P^a = p(a)\mathbb{1}$, $p_a \geqslant 0 \quad \forall a$, and $\sum_a p(a) = 1$.*

Remark. *We see trivial POVM as a classical simulation of quantum POVM, in other words, it is a classical resource that will behave like quantum measurements.*

Lemma 4.2.1. *Trivial POVMs cannot be transformed into non-trivial POVMs under general processing, or we say that general processing preserve the triviality of POVMs.*

Proof. It is straightforward due to the fact that $\mathcal{E}^\dagger(\mathbb{1}) = \mathbb{1}$. ∎

Lemma 4.2.2. *General processing induces a pre-order among POVMs.*

Proof. Suppose $\mathbb{P}$, $\mathbb{Q}$,and $\mathbb{R}$ are POVMs on $\mathcal{L}(\mathcal{H})$, then $\mathbb{P} \succ \mathbb{P}$; and $\mathbb{P} \succ \mathbb{Q}, \mathbb{Q} \succ \mathbb{R} \implies \mathbb{P} \succ \mathbb{R}$. ∎

The pre-order induces a special class of POVMs, the clean POVMs. We know that in most of the cases, the transformation of POVMs is not reversible since the transformation is a noisy process. We are interested in finding a class of POVMs that are reversible.

Definition 4.2.5 (Clean POVMs[19]). *Given a quantum system $\mathcal{H}$, we denote the set of clean POVMs on $\mathcal{L}(\mathcal{H})$ in terms of pre-order as,*

$$\mathcal{G}(\mathcal{H}) := \{\mathbb{Q} \in \mathcal{L}(\mathcal{H}) : \forall \mathbb{P} \in \mathcal{L}(\mathcal{H}), \mathbb{P} \succ \mathbb{Q} \implies \mathbb{Q} \succ \mathbb{P}\}\,. \tag{4.4}$$

Remark. *Clean POVMs can be seen as those most non-noisy measurements since we can reverse the noisy process through some transformations.*

Lemma 4.2.3. *Rank one POVM is clean POVM.*

Proof. Suppose $\mathbb{Q} = \{Q^b\}$, and suppose $Q^b = |\omega_b\rangle\langle\omega_b|$, let us normalize $|\tilde{\omega_b}\rangle = \frac{1}{\sqrt{\alpha_b}}|\omega_b\rangle$, then we have the following equalities:

$$\alpha_b = \mathrm{Tr}\Big\{Q^b\Big\} = \mathrm{Tr}\Big\{Q^b|\tilde{\omega_b}\rangle\langle\tilde{\omega_b}|\Big\}$$
$$\sum_b \alpha_b = \mathrm{Tr}\{\mathbb{1}\} = d$$

Suppose a POVM $\mathbb{P}$ satisfies $\mathbb{P} \succ \mathbb{Q}$, then we will prove it is also rank one. Recall that $\mathbb{P} \succ \mathbb{Q}$ implies the following,

$$Q^b = \sum_r \mu(r) \sum_a w(b|a,r)\mathcal{N}_r^\dagger(P^a) \quad \forall b$$

then we have,

$$\begin{aligned}
\alpha_b &= \sum_r \mu(r) \sum_a w(b|a,r)\, \mathrm{Tr}\Big\{|\tilde{\omega_b}\rangle\langle\tilde{\omega_b}|\mathcal{N}_r^\dagger(P^a)\Big\} \\
&= \sum_r \mu(r) \sum_a w(b|a,r)\, \mathrm{Tr}\{\mathcal{N}_r(|\tilde{\omega_b}\rangle\langle\tilde{\omega_b}|)P^a\} \\
&\leq \sum_r \mu(r) \sum_a w(b|a,r)\lambda_M(P^a)
\end{aligned}$$

where $\lambda_M(P^a)$ is the maximal eigenvalue of P^a, and the equality holds because we know that for any state ρ, $\mathrm{Tr}\{\rho P^a\} \in [\lambda_m(P^a), \lambda_M(P^a)]$. Then we do the summation on both side on the above inequality since both sides are positive numbers,

$$\begin{aligned}\sum_b \alpha_b = d &\leq \sum_b \sum_r \mu(r) \sum_a w(b|a,r)\lambda_M(P^a) \\ &= \sum_r \mu(r) \sum_a \sum_b w(b|a,r)\lambda_M(P^a) \\ &= \sum_a \lambda_M(P^a) \\ &\leq \sum_a \mathrm{Tr}\{P^a\} = \mathrm{Tr}\{\mathbb{1}\} = d\end{aligned}$$

where the last inequality holds because the trace of P^a is the sum of all eigenvalues of P^a.
Then it is true that,

$$\sum_a \lambda_M(P^a) = \sum_a \mathrm{Tr}\{P^a\}$$

since $0 \leq \lambda_M(P^a) \leq \mathrm{Tr}\{P^a\}$ holds for all a, then

$$\lambda_M(P^a) = \mathrm{Tr}\{P^a\} \quad \forall a$$

So $\mathbb{P}$ is also a rank one POVM.
Next, we show that $\mathbb{Q} \succ \mathbb{P}$ holds also. Recall that,

$$Q^b = \sum_{a,r} \mu(r) w(b|a,r) \mathcal{N}_r^\dagger(P^a) \quad \forall b$$

Since Q^b is rank one, then the POVM $\mathcal{N}_r^\dagger(P^a)$ must also be rank one, such that,

$$\mu(r) w(b|a,r) \mathcal{N}_r^\dagger(P^a) = \beta_{a,r} Q^b$$

where $\beta_{a,r}$ are some positive numbers satisfying $\sum_{a,r} \beta_{a,r} = 1$. By applying theorem 11.2 in "clean POVM" paper, we can obtain that, there exists an unitary U_r such that,

$$\mathcal{N}_r^\dagger(P^a) = U_r P^a U_r^\dagger \quad \forall a$$

By plugging the above equation into the previous one, we can obtain the following,

$$\mu(r) w(b|a,r) U_r P^a U_r^\dagger = \beta_{a,r} Q^b$$

so,

$$\mu(r) w(b|a,r) P^a = \beta_{a,r} U_r^\dagger Q^b U_r$$

so,

$$\sum_{b,r} \mu(r) w(b|a,r) P^a = P^a = \sum_{b,r} \beta_{a,r} U_r^\dagger Q^b U_r \quad \forall a$$

Since $\beta_{a,r}$ are some positive numbers satisfying $\sum_{a,r} \beta_{a,r} = 1$, we can always decompose it as conditional probability distributions, such that $\beta_{a,r} = \mu'(r)\omega'(a|r)$, so, there exists μ' and ω' such that,

$$P^a = \sum_{b,r} \mu'(r) \omega'(a|r) U_r^\dagger Q^b U_r \quad \forall a$$

thus we have $\mathbb{Q} \succ \mathbb{P}$ which concludes the proof. ∎

4.3 Measurement incompatibility

Until now, this thesis has discussed a single POVM processing and its corresponding guessing game. When considering a set of POVMs, such as in the situation of Bell experiment [20], a family of measurements is required. Discussing family of measurement starts from the investigation of a pair of measurements, where a quantum phenomenon, called measurement incompatibility, exists as a non-classical physics phenomenon. In what follows, this thesis discusses equivalent definitions of measurement incompatibility and their processing related to guessing games.

In fact, there are several equivalent definitions of measurement incompatibility, and some are more intuitive than others.

4.3.1 Equivalent definition of measurement incompatibility

Measurement incompatibility is usually considered in the case of a pair of measurements. Suppose we have two measurements $\mathbb{M} := \{M(i)\}$ and $\mathbb{N} := \{N(j)\}$ with a finite discrete outcome set, where, for simplicity, i and j label the i-th and j-th outcomes of the measurements, and $\mathbb{M}$ and $\mathbb{N}$ are said to be compatible whenever there exists a third (joint) measurement $\mathbb{S} = \{S(i,j)\}$ with two indexes outcome set such that, $M(i) = \sum_j S(i,j)$ and $N(j) = \sum_i S(i,j)$ for all i and j. And they are called incompatible if they are not compatible. In such a case, one calls $\mathbb{M}$ and $\mathbb{N}$ jointly measurable under the joint measurement $\mathbb{S}$.

The above definition tells that if two measurements are marginals of a third measurement, then they are compatible.

Lemma 4.3.1. *Two POVMs $\mathbb{M}_1$ and $\mathbb{M}_2$ are joint measurable if and only if there exists another POVM $\mathbb{G}$ such that both $\mathbb{M}_1$ and $\mathbb{M}_2$ are marginals of $\mathbb{G}$.*

Lemma 4.3.2. *Suppose $\{M_1(a_i), M_2(b_j), \cdots\}$ is a set of measurements with $a_i \in \Omega_1$, $b_j \in \Omega_2$ and the rest, then they are joint measurable if and only if there exists a "mother" measurement $\{S(x_k)\}$ with $x_k \in \Omega$ and a set of conditional probabilities $\{\mu^1_{x_k}(a_i), \mu^2_{x_k}(b_j), \cdots\}$, such that,*

$$M_1(a_i) = \sum_k \mu^1_{x_k}(a_i) S(x_k), \quad M_2(b_j) = \sum_k \mu^2_{x_k}(b_j) S(x_k), \quad \cdots, \tag{4.5}$$

where i, j and k denote ith, jth and kth respectively.

In addition to the above definition, we can have an equivalent definition as follows:

Definition 4.3.1 (General definition of compatibility)**.** *Let $\mathcal{A}$ be a finite and discrete label set for a family of measurements $\{\mathbb{M}_x\}_{x\in\mathcal{X}}$, denoting x-th measurement as $\{M(a|x)\}$. If there exists another measurement $\{G(\lambda)\}$ and a conditional probability distribution $p(a|x,\lambda)$, such that, for all a and x,*

$$M(a|x) = \sum_\lambda p(a|x,\lambda) G(\lambda) \tag{4.6}$$

then we call $\{\mathbb{M}_x\}_{x\in\mathcal{X}}$ *compatible.*

The above definition provides a way to represent a set of measurements as an operational device with quantum and classical input and classical output. Without loss of generality, we assume the classical input as independent of the quantum input and as serving as a programmable side to control what is inside the device. As for the compatible measurements, we can represent the device as per Figure 4.3 below.

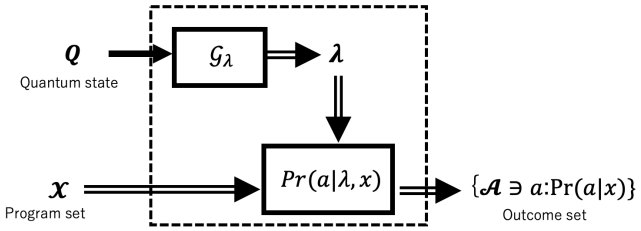

Figure 4.3: The inner structure of a compatible measurement as a device.

In the compatible measurement described in the picture above, we can see that inside the box there is one "mother" measurement and post-processing of the measurement outcome. However, when we are given a black box, how could we know whether it is compatible? Studying the compatibility of measurement then becomes a way to analyze the structure of a device. We denote this device as a programmable measurement devices(PMD). Any PMD that is equivalent to compatible measurements are denoted by $simple$ PMDs. Being programmable means that this device can behave as an incompatible measurement. Figure 4.4 shows the structure of a PMD as a black box without knowing anything inside. This model enables the investigation of the programmability of the PMD.

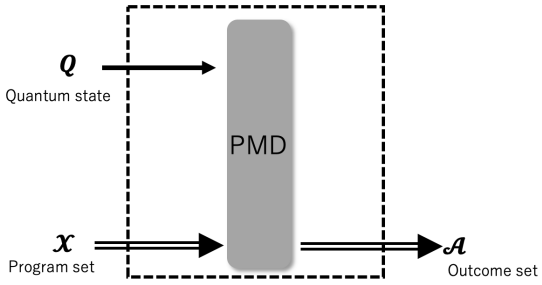

Figure 4.4: A Programmable measurement device as a black box.

Definition 4.3.2 (PMD)**.** *We define PMD as* $cq \to c$ *channel with two inputs (quantum and classical) that are always assumed to be separate systems; thus, we allow a time lag between the two inputs. An arbitrary PMD is denoted as* $\mathcal{M} := M^Q(a|x)$*, where* Q

denotes the input quantum system, $x \in \mathcal{X}$ represents the program set and $a \in \mathcal{A}$ represents the output set.

Notice that, as the program set and quantum system input are independent of each other, we could inject some time lag for the program set. The injection of time lag induces a fact that the quantum input system must be stored in a quantum memory until the program sets to come. If no additional quantum memory is allowed, then without any quantum memory inside the PMD black box, we are assured that this PMD is just a $simple$ PMD. Accordingly, the following proposition characterizes the PMD in terms of quantum memory.

Proposition 3. *A PMD box is incompatible if and only if it has quantum memory inside.*

Remark. *We note that quantum memory plays a key role in the programmability of PMDs. We regard quantum memory as a resource that is not free in the sense that we assume it cannot be simulated by classical physics.*

4.3.2 Transformation of PMDs

We treat a PMD that is equivalent to the compatible measurement as a useless resource and the others as not useful resources. When given two PMDs and we do not know whether they are compatible or incompatible, we are interested in processing between PMDs.

Definition 4.3.3. *Given two PMDs $\mathcal{M} := M^Q(a|x)$ and $\mathcal{N} := N^{Q'}(b|y)$, with $a \in \mathcal{A}, b \in \mathcal{B}, x \in \mathcal{X}$ and $y \in \mathcal{Y}$, we say $M^Q(a|x)$ can be processed into $N^{Q'}(b|y)$ whenever there exists quantum noisy channel to process the quantum input from system Q' to system Q, and noisy channel to process classical output up to some shared randomness, denoted by*

$$\mathcal{M} \succ \mathcal{N}, \tag{4.7}$$

mathematically, we equivalently define it as,

$$N^{Q'}(b|y) = \sum_r \mu(r) \sum_a q(b|a,x,i,y,r)p(x|i,y,r)(\mathcal{E}_{i|r}^{Q'\to Q})^\dagger[M^Q(a|x)] \tag{4.8}$$

where $\mu(r)$ is a probability distribution, $\{\mathcal{E}_{i|r}^{Q'\to Q}\}$ is a family of quantum instruments labeled by r, with classical outcome i, and $q(b|a,x,i,y,r), p(x|i,y,r)$ are conditional probability distributions that represent classical noisy channels.

Lemma 1. *All simple devices are equivalent to each other, that is, given any two simple devices $M^Q(a|x)$ and $N^{Q'}(b|y)$, possibly defined on different Hilbert spaces $\mathcal{H}^Q$ and $\mathcal{H}^{Q'}$, both relations hold:*

$$M^Q(a|x) \succ N^{Q'}(b|y) \quad \textit{and} \quad N^{Q'}(b|y) \succ M^Q(a|x)\ .$$

Proof. For any two simple PMDs. Let us denote by $I^Q(a|x)$ the trivial PMD, i.e. the PMD with alphabets $\mathcal{A} = \mathcal{X} = \{0\}$ and Hilbert space $\mathcal{H}^Q = \mathbb{C}$. Clearly the trivial

PMD can be attained from any other using the free operations. Showing that the converse is true will complete the proof of the lemma. Using the trivial PMD as the input PMD in Eq. (4.8), we verify that the instruments $\{\mathcal{E}_{i|r}^{Q'\to Q}\}$ are, in fact, POVMs $E^{Q'}(i|r)$: this is so because $\dim \mathcal{H}^Q = 1$. Since these POVMs can be freely chosen, all devices of the form

$$\begin{aligned} N^{Q'}(b|y) &= \sum_r \mu(r) \sum_{i,j} q(b|j,r)p(j|i,y,r)\; E^{Q'}(i|r) \\ &= \sum_r \sum_i p'(b|i,y,r)\; E^{Q'}(i,r)\;, \end{aligned}$$

can be obtained from the trivial PMD, where $E^{Q'}(i,r) := \mu(r)E^{Q'}(i|r)$ is considered now as a POVM with two outcome indices. Since the above coincides with the definition of simple PMDs, the desired conclusion is reached. ∎

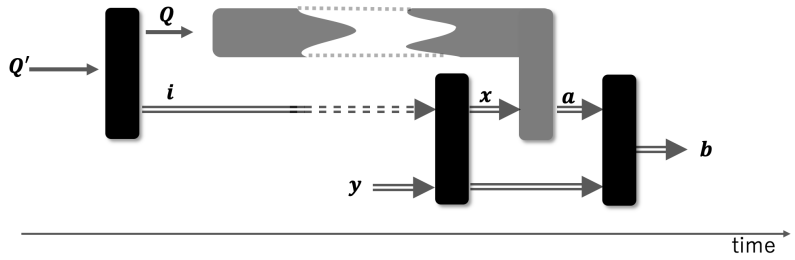

Figure 4.5: PMD transformation under noisy channels

As depicted in Figure 4.5, transforming a PMD into another is under an allowed process, in which we are only allowed to conduct the processing of the measurement device up to some shared randomness. In fact, the time evolution scenario could be transferred into a space-like scenario with equivalence. In the space-like scenario, we could treat an event after some time as a faraway spatial event, from which we are not allowed to communicate to get rid of time travel in its equivalent time evolution scenario. Figure 4.6 shows the process in the space-like scenario.

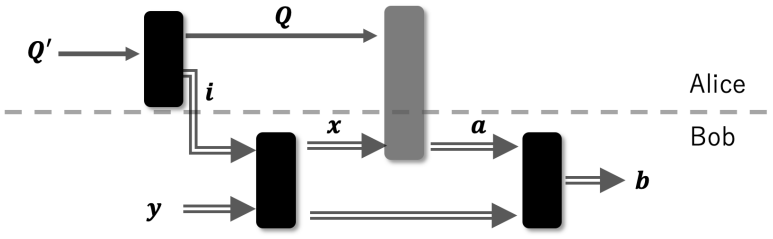

Figure 4.6: PMD transformation in space-like scenario

In this scenario, two parties are separate from each other and no communication is allowed between them. In this framework, the transformation process is a

one-way LOCC. In addition, by the structure of this space-like scenario, we obtain the equivalence between transformation and one-way LOCC.

Proposition 4. *Given two PMDs $\mathcal{M} := M^{Q}(a|x)$ and $\mathcal{N} := N^{Q'}(b|y)$, with $a \in \mathcal{A}, b \in \mathcal{B}, x \in \mathcal{X}$ and $y \in \mathcal{Y}$, $\mathcal{M} \succ \mathcal{N}$ if and only if $\mathcal{M}$ can be transformed into $\mathcal{N}$ by one-way LOCC.*

Proof. It is obvious from Fig 4.5 above that $M^{Q}(a|x) \succ N^{Q'}(b|y)$ implies an implementation by one-way LOCC. Conversely, every one-way LOCC protocol from Alice to Bob consists here of (i) a one-way LOCC pre-processing, (ii) local side channels that are quantum for Alice and classical for Bob, and (iii) one-way LOCC post-processing. Since Alice receives no output from the PMD, any local post-processing and forward communication she performs can be included in her pre-processing. What remains is exactly as depicted in Fig 4.6. ∎

Remark. *We find the equivalence between the time evolution and space-like scenario.*

4.4 Game-theoretic approaches

The major task in quantum computation and quantum information is to show that quantum resources can conduct information processing tasks that would be impossible or much more difficult using only classical resources. Among the most fundamental quantum resources, quantum entanglement stands out. With quantum theory backing behind, entanglement can shape our intuition in understanding nature. In an intuitive world, the idea of "local realism", is natural, which means that faraway events can't influence each other faster than the speed of light (what is known as locality), and the properties of objects have a definite value even if we do not measure them (what is known as realism). However, in 1964, John S. Bell in [20] showed in his now-famous bell theorem that with quantum entanglement, the local realism can be violated through what the so-called "bell game" or "bell experiment" and that this violation can be tested through the game. This game scenario approach has currently served as a useful tool to certify the existence of quantum entanglement.

In fact, game-theoretical approaches are also essential in certifying quantum resources. The following section introduces guessing games for certifying specific types of POVMs.

At the beginning of this thesis, we introduced general quantum statistical decision problems. At the interest of POVMs, this thesis focuses on minimum error state distinguishing games (guessing games), which is a special case of general quantum statistical decision problem. Given a POVM $\mathbb{P} = \{P^a\}$, we want to know how good it is for a guessing game. Suppose we are given an index set $x \in \mathcal{X}$ and we encode them into quantum states, we then input the quantum states into the measurement, being allowed to process the measurement device to obtain a better result. We are interested in finding the best strategy that gives us the maximum guessing probability for the index x. Figure 4.7 shows the structure of this guessing game.

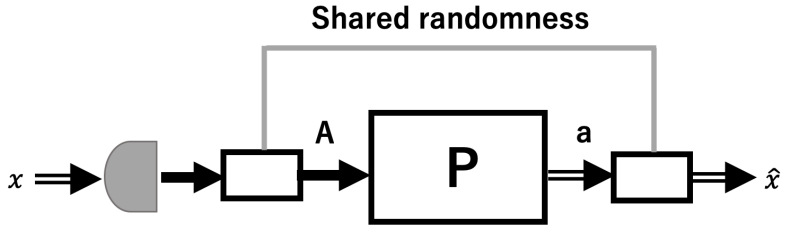

Figure 4.7: Guessing the index set through a measurement device: minimum-error state discrimination game

As shown in Figure 4.7, we see that the orange boxes are seen as processing of the measuring device up to some shared randomness, and the maximum guessing strategy is to find the suitable processings. Mathematically, we phrase the game payoff (the guessing probability) as follows,

$$P_{\text{guess}}(\{\rho^x\},\mathbb{P}) := \max_{\mu,w,\{\mathcal{N}_r\}} \sum_r \mu(r) \sum_{b,x} w(x|b,r)\,\mathrm{Tr}\Big\{Q_B^b\,\mathcal{N}_r(\rho^x)\Big\}\ . \tag{4.9}$$

where $\{\rho^x\}$ is the encoded ensemble of quantum states, $\mu(r)$ is the shared randomness, $w(x|b,r)$ is conditional probability distribution, and $\{\mathcal{N}_r\}$ is an instrument.

When dealing with a set of POVMs, we introduce the so-called "post information guessing games", which consist of the following components: (i) the referee picks one pair $(w,z) \in \mathcal{W}\times\mathcal{Z}$ at random according to the distribution $p(w,z)$, (ii) at time t_0, the normalized quantum state $p(w,z)^{-1}\rho^R_{w,z}$ is sent to the player followed, after some finite time at time t_1, by the index w, and (iii) the player attempts to maximize the probability of correctly guessing the value z using the given PMD $\mathcal{M} := M^Q(a|x)$ and any processings of the PMD. In this game, the label w is interpreted as the "post information" since it is imported into the program register of the PMD *after* the quantum state, and it cannot be used in any pre-processing of the PMD.

When playing guessing games with post information, certain processing strategies will lead to greater success probabilities in guessing z. In particular, if the referee's questions $\rho^R_{w,z}$ are encoded on a quantum system that is different from the quantum input of the PMD $M^Q(a|x)$, then the player must conduct some kind of quantum pre-processing of R into Q, represented without loss of generality by a quantum instrument $\{\mathcal{E}_i^{R\to Q}\}$. The optimum success probability over all strategies is thus given by

$$\begin{aligned} P_{\text{guess}}(M^Q(a|x);\rho^R_{w,z}) := \max_{\mu,q,p,\mathcal{E}} \sum_{w,z,r,i,x,a} & \mu(r)q(z|a,w,i,r)p(x|w,i,r) \\ & \times \mathrm{Tr}\Big\{\mathcal{E}^{R\to Q}_{i|r}(\rho^R_{w,z})\ M^Q(a|x)\Big\}\ , \end{aligned} \tag{4.10}$$

where the probability distribution $\mu(r)$ is included to describe mixed strategies, i.e. those in which a different strategy, labeled by r, is chosen at random.

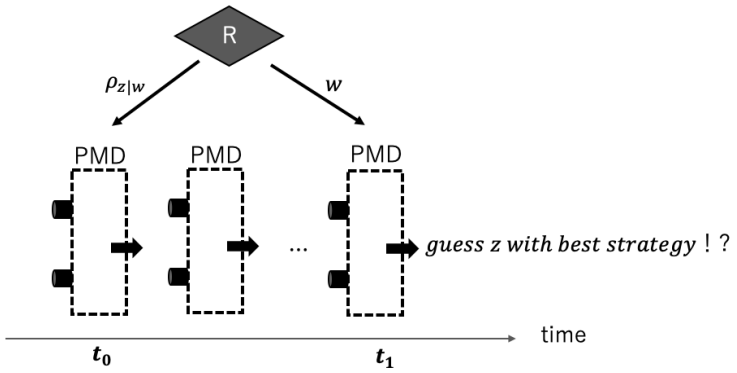

Figure 4.8: Post-information guessing game.

In the picture above, we note the time evolution plays an important role in this game.

4.4.1 Game-theoretical approach to POVMs

The set of all correlated pre and post-processing constitutes a closed and convex set. This is easy to see due to the inclusion of shared randomness. Given a POVM on $\mathcal{H}_A$ with output alphabet $\mathcal{A} = \{a\}$, let us denote by $\mathfrak{M}$ any transformation that takes it into a POVM on $\mathcal{H}_B$ with output alphabet $\mathcal{B} = \{b\}$.

Theorem 4. *Given two POVMs $\mathbb{P} = \{P_A^a\}$ on system $\mathcal{H}_A$ and $\mathbb{Q} = \{Q_B^b\}$ on system $\mathcal{H}_B$, the relation*

$$\mathbb{P} \succ \mathbb{Q} \tag{4.11}$$

holds if and only if, for any ensemble $\{\rho_R^x\}$, such that $\sum_x \mathrm{Tr}\{\rho_R^x\} = 1$,

$$P_{\mathrm{guess}}(\{\rho_R^x\}, \mathbb{P}) \geqslant P_{\mathrm{guess}}(\{\rho_R^x\}, \mathbb{Q}) \ , \tag{4.12}$$

Remark. *In what follows, we will always highlight the original theorems that are derived in our published papers with gray background.*

Proof. The direction (4.11) $\implies$ (4.12) is trivial: if the POVM $\mathbb{P}$ can be used to simulate the POVM $\mathbb{Q}$, then the former cannot be worse than the latter in any discrimination task. We hence need to prove only the converse.

Fix a basis of self-adjoint operators $\{X_B^j\}$. Then, relation (4.11) is equivalent to the following:

$$\mathrm{Tr}\Big\{Q_B^b \ X_B^j\Big\} = \mathrm{Tr}\left\{\left\{\sum_r \mu(r) \sum_a w(b|a,r) \mathcal{N}_r^\dagger(P_A^a)\right\} \ X_B^j\right\} \ , \qquad \forall b, \forall j \ .$$

Denote by $\vec{s}(\mathbb{Q})$ the vector whose entries are the $|\mathcal{B}| \times |\mathcal{J}|$ real numbers above, and by $\vec{r}(\mathbb{P}, \mathfrak{M})$ the same vector on the right-hand side. Consider also the set of all such vectors that can be obtained from POVM $\mathbb{P}$ by varying the pre/post-processing $\mathfrak{M}$; denote such set by

$$\mathcal{S}(\mathbb{P}) := \{\vec{r}(\mathbb{P}, \mathfrak{M}) : \mathfrak{M}\} \ .$$

Such a set is closed and convex because closed and convex is the set of all transformations $\mathfrak{M}$. Hence, we can say that relation (4.11) is equivalent to

$$\vec{s}(\mathbb{Q}) \in \mathcal{S}(\mathbb{P}) \ ,$$

namely (as an application of the separation theorem)

$$\vec{s}(\mathbb{Q}) \cdot \vec{c} \leqslant \max_{\vec{r} \in \mathcal{S}(\mathbb{P})} \vec{r} \cdot \vec{c} \ , \qquad \forall \vec{c} \ .$$

Denoting by Y_B^b the self-adjoint operators obtained as $Y_B^b := \sum_j c(b,j) X_B^j$, we have that relation (4.11) is equivalent to

$$\sum_b \mathrm{Tr}\Big\{Q_B^b \ Y_B^b\Big\} \leqslant \max_{\mathfrak{M}} \sum_b \mathrm{Tr}\Big\{\mathfrak{M}(\mathbb{P})_B^b \ Y_B^b\Big\} \ , \qquad \forall \{Y_B^b\} \ .$$

We now shift and rescale the operators Y_B^b to $\rho_B^b := \frac{Y_B^b + C}{\sum_b \mathrm{Tr}\left\{Y_B^b + C\right\}} \geqslant 0$, so that the ρ_B^b form an ensemble. This can always be done by choosing the constant operator C large enough. Then, by noticing that $\sum_b \mathrm{Tr}\Big\{Q_B^b \ C\Big\}$ does not depend on the POVM $\mathbb{Q}$, we can massage the above equation arriving at the following conclusion: relation (4.11) is equivalent to

$$\sum_b \mathrm{Tr}\Big\{Q_B^b \ \rho_B^b\Big\} \leqslant \max_{\mathfrak{M}} \sum_b \mathrm{Tr}\Big\{\mathfrak{M}(\mathbb{P})_B^b \ \rho_B^b\Big\} \ , \qquad \forall \{\rho_B^b\} \ .$$

But now, a *sufficient* condition for relation (4.11) is that

$$\max_{\mathfrak{M}} \sum_b \mathrm{Tr}\Big\{\mathfrak{M}(\mathbb{Q})_B^b \ \rho_B^b\Big\} \leqslant \max_{\mathfrak{M}} \sum_b \mathrm{Tr}\Big\{\mathfrak{M}(\mathbb{P})_B^b \ \rho_B^b\Big\} \ , \qquad \forall \{\rho_B^b\} \ .$$

∎

A straightforward corollary from Theorem 4 can used as a theoretical too to determine if a POVM is not a trivial one.

Corollary 1. *A POVM $\mathbb{P}$ is not a trivial one if and only if there exists a guessing game such that,*

$$P_{\mathrm{guess}}(\{\rho_R^x\}, \mathbb{P}) > 0. \tag{4.13}$$

4.4.2 Game-theoretic approach to quantum incompatibility

The shared randomness of processing PMD shows that the set of PMDs forms a convex set. Accordingly, we can derive the following theorem that characterizes the PMD transformations in terms of quantum games.

Theorem 5. *Given two PMDs $\mathcal{M} := M^Q(a|x)$ and $\mathcal{N} := N^{Q'}(b|y)$, with $a \in \mathcal{A}, b \in \mathcal{B}, x \in \mathcal{X}$ and $y \in \mathcal{Y}$, then for all guessing games wih post-information $\{\rho^R_{w,z} : \omega \in \mathcal{W}, z \in \mathcal{Z}\}$, the following equivalence holds,*

$$\mathcal{M} \succ \mathcal{N} \iff P_{\text{guess}}(M^Q(a|x); \rho^R_{w,z}) \geqslant P_{\text{guess}}(N^{Q'}(b|y); \rho^R_{w,z}). \tag{4.14}$$

For simplicity, It is equivalent to consider only guessing games with $\mathcal{H}^R = \mathcal{H}^{Q'}$, $\mathcal{W} = \mathcal{Y}$, and $\mathcal{Z} = \mathcal{B}$.

Proof. For the sake of notation, we will denote the processing of a PMD $M^Q(a|x)$ as prescribed in Eq. (4.8) simply by

$$[\mathcal{T}(M)](b|y) \ .$$

In particular, the set of all allowed mappings of PMDs with input Hilbert space $\mathcal{H}^Q$, input alphabet $\mathcal{X}$, and output alphabet $\mathcal{A}$, into PMDs with input Hilbert space $\mathcal{H}^{Q'}$, input alphabet $\mathcal{Y}$, and output alphabet $\mathcal{B}$, will be denoted by $\mathscr{T}$:

$$\mathscr{T} := \{\mathcal{T} : \text{PMD}(\mathcal{H}^Q, \mathcal{X}, \mathcal{A}) \to \text{PMD}(\mathcal{H}^{Q'}, \mathcal{Y}, \mathcal{B})\} \ .$$

A crucial observation is that the set $\mathscr{T}$ is convex due to the presence of shared randomness (represented by the probability distribution $\mu(r)$ in Eq (4.8) .

The implication (a) $\implies$ (b) is trivial: since processings of the form (4.8) are always allowed when playing guessing games with post-information, as prescribed in Eq. (4.10)), if PMD $M^Q(a|x)$ can simulate $N^{Q'}(b|y)$, then any strategy that can be reached from the latter can be reached also from the former. Hence, we only need to prove explicitly the implication (b) $\implies$ (a).

We begin by noticing that condition (a) is equivalent to the existence of a mapping $\mathcal{T}$ of the form (4.8) such that

$$[\mathcal{T}(M)]^{Q'}(b|y) = N^{Q'}(b|y), \qquad \forall b, \forall y \ . \tag{4.15}$$

Let us fix a basis of self-adjoint operators $\{X^{Q'}_j : j \in \mathcal{J}\}$. Then, relation (4.15) is equivalent to the following:

$$\text{Tr}\Big\{N^{Q'}(b|y)\ X^{Q'}_j\Big\} = \text{Tr}\Big\{[\mathcal{T}(M)]^{Q'}(b|y)\ X^{Q'}_j\Big\} \ , \qquad \forall b, \forall y, \forall j \ .$$

Denote by $\vec{s}(N)$ the vector whose entries are the $|\mathcal{B}| \times |\mathcal{Y}| \times |\mathcal{J}|$ real numbers above, and by $\vec{r}(M, \mathcal{T})$ the same vector on the right-hand side.

Let us consider now the set of all such vectors that can be obtained from PMD $M^Q(a|x)$ by varying the processing $\mathcal{T}$ in $\mathscr{T}$; denote such set by

$$\mathscr{S}(M) := \{\vec{r}(M, \mathcal{T}) : \mathcal{T} \in \mathscr{T}\} \ .$$

Such a set is closed and convex because closed and convex is the set of all transformations $\mathscr{T}$. Hence, we can say that relation (4.15) is equivalent to

$$\vec{s}(N) \in \mathscr{S}(M) \ ,$$

that is, by applying the separation theorem for convex sets,

$$\vec{s}(N) \cdot \vec{c} \leqslant \max_{\vec{r} \in \mathscr{S}(M)} \vec{r} \cdot \vec{c} \ , \qquad \forall \vec{c} \in \mathbb{R}^{|\mathcal{B}| \times |\mathcal{Y}| \times |\mathcal{J}|} \ .$$

Denoting by $Y_{b,y}^{Q'}$ the self-adjoint operators obtained as $Y_{b,y}^{Q'} := \sum_j c(b,y,j) X_j^{Q'}$, we have that relation (4.15) is equivalent to

$$\sum_{b,y} \mathrm{Tr}\Big\{ N^{Q'}(b|y) \, Y_{b,y}^{Q'} \Big\} \leqslant \max_{\mathcal{T} \in \mathscr{T}} \sum_{b,y} \mathrm{Tr}\Big\{ [\mathcal{T}(M)]^{Q'}(b|y) \, Y_{b,y}^{Q'} \Big\} \ , \forall \{ Y_{b,y}^{Q'} : \text{self-adjoint} \} \ .$$

We now shift and rescale the self-adjoint operators $Y_{b,y}^{Q'}$ to $\rho_{b,y}^{Q'} := \frac{Y_{b,y}^{Q'} + C}{\sum_{b,y} \mathrm{Tr}\Big\{ Y_{b,y}^{Q'} + C \Big\}} \geqslant 0$, so that the $\rho_{b,y}^{Q'}$ form an ensemble. This can always be done by choosing the constant operator C large enough. Then, by noticing that $\sum_{b,y} \mathrm{Tr}\Big\{ N^{Q'}(b|y) \, C \Big\} = |\mathcal{Y}| \, \mathrm{Tr}\{C\}$ does not depend on the particular PMD $N^{Q'}(b|y)$, we can rewrite the above equation arriving at the following conclusion: condition (4.15) is equivalent to

$$\sum_{b,y} \mathrm{Tr}\Big\{ N^{Q'}(b|y) \, \rho_{b,y}^{Q'} \Big\} \leqslant \max_{\mathcal{T} \in \mathscr{T}} \sum_{b,y} \mathrm{Tr}\Big\{ [\mathcal{T}(M)]^{Q'}(b|y) \, \rho_{b,y}^{Q'} \Big\} \ , \forall \text{ ensembles } \{ \rho_{b,y}^{Q'} : b \in \mathcal{B}, y \in \mathcal{Y} \} \ .$$

Comparing the above relation with the expression (4.10) of the optimal guessing probability in guessing games with post-information, we recognize that the above equation means that, for any guessing game with post-information $\{\rho_{b,y}^{Q'} : b \in \mathcal{B}, y \in \mathcal{Y}\}$, it holds that

$$\sum_{b,y} \mathrm{Tr}\Big\{ N^{Q'}(b|y) \, \rho_{b,y}^{Q'} \Big\} \leqslant P_{\mathrm{guess}}(M^{Q}(a|x); \rho_{b,y}^{Q'}) \ . \tag{4.16}$$

But then, a *sufficient* condition for relation (4.15) is that

$$P_{\mathrm{guess}}(N^{Q'}(b|y); \rho_{b,y}^{Q'}) \leqslant P_{\mathrm{guess}}(M^{Q}(a|x); \rho_{b,y}^{Q'}) \ ,$$

for all guessing game with post-information $\{\rho_{b,y}^{Q'} : b \in \mathcal{B}, y \in \mathcal{Y}\}$. ∎

Simply by noticing that it is impossible to turn a simple PMD into an incompatible one using free operations, we obtain as a corollary that quantum incompatibility can always be witnessed using a suitable guessing game with post-information.

Corollary 2. *A PMD $M^Q(a|x)$ is incompatible, if and only if there exists an ensemble $\{\rho^Q_{x,a} : x \in \mathcal{X}, a \in \mathcal{A}\}$ such that*

$$\sum_{a,x} \mathrm{Tr}\Big\{ M^Q(a|x)\ \rho^Q_{x,a} \Big\} > P^{\textit{simple}}_{\mathrm{guess}}(\rho^Q_{x,a})\ ,$$

where $P^{\textit{simple}}_{\mathrm{guess}}(\rho^Q_{x,a})$ is defined as the optimum guessing probability achievable with simple PMDs.

Proof. First, we notice that, for any guessing game with post-information, the optimum guessing probability is the same *for all* simple PMDs. This is a direct consequence of Theorem 5 and Lemma 1.

Then, the statement is proved by contradiction. Suppose that, for all guessing game with post-information $\{\rho^Q_{x,a} : x \in \mathcal{X}, a \in \mathcal{A}\}$, the opposite relation holds, that is

$$\sum_{a,x} \mathrm{Tr}\Big\{ M^Q(a|x)\ \rho^Q_{x,a} \Big\} \leqslant P^{\mathrm{simple}}_{\mathrm{guess}}(\rho^Q_{x,a})\ .$$

But then, by means of Eq. (4.16) in the proof above, one would conclude that it is possible to obtain $M^Q(a|x)$ by acting with a free operation on a simple PMD, in contradiction with the fact that $M^Q(a|x)$ is incompatible. ∎

Chapter 5

Convex optimization

5.1 Convex optimization

In optimization theory, the duality principle provides a different perspective to optimization problems. Usually, the term "dual problem" refers to the Lagrangian dual problem. The solution of the dual problem provides a lower bound to the solution of the primal (minimization) problem. However, if strong duality holds, the solutions of primal and dual problems are equal. The following contents refer to the paper [9, 21].

5.1.1 The underlying intuition

This section introduces a mathematical tool that is useful in solving problems not only in this thesis but also in many other research setting. In most cases in which quantum information theory plays a key role in some information processing tasks, convex optimization is of particular interest. This is because quantum objects [1] are formulated as mathematical elements such that their set is convex. In the case of quantum states, it is known that a set of bipartite quantum states forms a convex set with a subset of separable states. Similarly, one notices that both a set of measurements and a set of quantum channels are in the form of convex set with a subset of useless elements. Hence, it is indeed crucial to investigate the convexity of quantum objects is indeed necessary.

Figure 5.1 shows that, given a state ρ in the convex set of states, whether it belongs to the subset of separable states is of our interest. It is known that quantum entangled state serves as a powerful resource in many quantum information processing tasks. Then, a natural question to ask is: how much resource does a quantum entangled state contain? Or how close the state ρ is to the subset of separable states that we consider null resource. A common quantifier is defined as follows: given an object ρ, what is the least amount of mixing $p \in [0, 1]$ with another object σ such that $(1-p)\rho + p\sigma$ belongs to the subset of separable states. This

[1] Here quantum objects can represent any quantum phenomena statistics, rather that only being limited to the quantum state itself

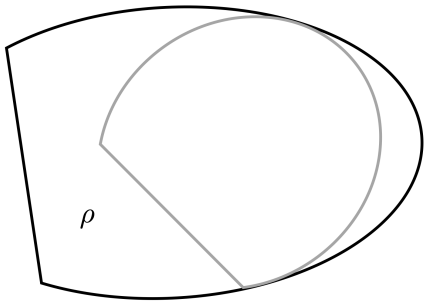

Figure 5.1: Convex set of quantum states and its subset of separable states.

is core idea behind the robustness of a resource, which can be understood as the robustness of the resource contained in ρ to noise in the form of mixing an arbitrary object σ.

5.2 Basics of semi-definite programming

Given the Hermitian operators A, B and a hermiticity-preserving linear map $\Phi(\cdot)$, we consider an optimization problem in the following form.

$$\begin{array}{ll} \text{maximize} & \langle A, X\rangle \\ \text{subject to} & \Phi(X) = B \\ & X \geqslant 0 \end{array} \tag{5.1}$$

where the notation $\langle\cdot,\cdot\rangle$ is defined as the inner product, such that $\langle A, B\rangle := \mathrm{Tr}\{AB\}$. This problem is to maximize the real linear function $\mathrm{Tr}\{AX\}$, over the subset of positive semi-definite operator X which satisfies the constraint $\Phi(X) = B$. This problem is called the semi-definite programming (SDP) problem. The set of operators X that satisfy the constraints in 5.1 are called primal feasible. The maximal value of $\langle A, X\rangle$ over the primal feasible set, denoted by α, is called primal optimal value.

For every SDP program, we introduce Lagrange multipliers for each constrain in 5.1. Let us introduce the Hermitian operators Y and Z as the Lagrange multipliers associated to the first and second constraints respectively in 5.1. The Lagrangian is defined as,

$$\begin{aligned} \mathcal{L} &= \langle A, X\rangle + \langle Y, B - \Phi(X)\rangle + \langle Z, X\rangle \\ &= \langle A - \Phi^\dagger(Y) + Z, X\rangle + \langle Y, B\rangle \end{aligned} \tag{5.2}$$

From the above structure of the Lagrangian, we see that if $Z \geqslant 0$, $\langle A, X\rangle \leqslant \mathcal{L}$ for all primal feasible operators because $\langle Z, X\rangle \geqslant 0$. Thus, the Lagrangian upper bounds the primal optimal value α. Moreover, if the condition $A - \Phi^\dagger(Y) + Z = 0$ holds, then the Lagrangian $\mathcal{L}$ is independent of X and equal to $\langle Y, B\rangle$. Thus we can achieve the best possible upper bound α by minimizing $\mathcal{L}$ over the Lagrange

multipliers, subject to the condition $A - \Phi^\dagger(Y) + Z = 0$ and $Z \geqslant 0$. We can write the dual problem,

$$\begin{array}{ll} \text{minimize} & \langle Y, B\rangle \\ \text{subject to} & \Phi^\dagger(Y) \geqslant A \end{array} \tag{5.3}$$

Similarly, we have the dual optimal value, denoted by β, such that we always have $\alpha \leqslant \beta$, which is called weak duality. However, in most cases, we have $\alpha = \beta$, which is called strong duality. The condition for strong duality to hold is that either the primal or the dual problem is strictly feasible, that one can find either a positive-definite $X > 0$ such that $\Phi(X) = B$, or a Y such that $\Phi^\dagger(Y) - A > 0$. In this way, we turn a primal problem into a dual problem from a different perspective.

Note that we only considered the simplest case of a single optimization variable, with a single equality constraint. This can easily extend to complicated cases with multiple optimization variables and multiple inequality and equality constraints.

5.3 Robustness of POVM with relation to a guessing game

The following definition of robustness quantifies how much usefulness a POVM holds concerning trivial POVMs, i.e., the robustness of POVM.

Definition 5.3.1 (Robustness of POVM)**.** *Robustness of POVM is defined to capture how tolerant a POVM is to mixing before it becomes a trivial POVM. The robustness of a POVM $\mathbb{P} = \{P^a\}$ is defined as*

$$\mathfrak{R}(\mathbb{P}) = \min\left\{r \geqslant 0 : \frac{P^a + rQ^a}{1+r} \propto \mathbb{1}, Q^a \propto \mathbb{1}\right\} ,$$

where $\propto \mathbb{1}$ means that the effect of POVM is proportional to $\mathbb{1}$, which indicates a trivial POVM (classically simulatable).

The following subsection presents the connection of robustness of POVM to the guessing game through convex optimization analysis.

5.3.1 Convex optimization of robustness of POVM

Following the idea in [22], we rewrite the robustness of POVM $\mathbb{P} = \{P^a\}$. We denote by $R(\mathbb{P})$ as follows:

$$\begin{aligned} \mathfrak{R}(\mathbb{P}) = \min \quad & r \\ \text{s.t.} & \frac{P^a + rQ^a}{1+r} = q(a)\mathbb{1}, \quad \forall a \\ & Q^a \geq 0 \quad \forall a \\ & \sum_a Q^a = \mathbb{1}. \end{aligned} \tag{5.4}$$

Note that the POVM $\{Q^a\}$ above can be defined as any POVMs. However, by solving Q^a, the above constraints can be equally reformulated as a SDP as follows:

$$\begin{aligned} \mathfrak{R}(\mathbb{P}) = \min & \sum_a \tilde{q}(a) - 1 \\ \text{s.t.} \quad & \tilde{q}(a)\mathbb{1} \geq P^a, \quad \forall a \end{aligned} \tag{5.5}$$

Where we define $\tilde{q}(a) := (1+r)q(a)$. For later usage we denote the optimal $q^*(a)$ for the above SDP, then we have $(1+\mathfrak{R}(\mathbb{P}))q^*(a)\mathbb{1} \geq P^a$ for all a.
Next, we write the dual of the above SDP. Let's introduce the dual variables positive semi-definite ρ^a, then the associated Lagrangian is as follows:

$$\begin{aligned} \mathcal{L} &= \sum_a \tilde{q}(a) - 1 - \sum_a \mathrm{Tr}\{\rho^a[(\tilde{q}(a)\mathbb{1} - P^a)]\} \\ &= \sum_a \mathrm{Tr}\{\rho^a P^a\} - 1 + \sum_a \tilde{q}(a)(1 - \mathrm{Tr}\{\rho^a\}) \end{aligned} \tag{5.6}$$

We thus ensure that the defined Lagrangian upper bounds the primal object function whenever the primal constraints are satisfied. By making the constraints $\mathrm{Tr}\{\rho_a\} = 1$ for all a, this Lagrangian is independent of the primal constraints. We can thus write the dual SDP as:

$$\begin{aligned} \mathfrak{R}(\mathbb{P}) = \max & \sum_a \mathrm{Tr}\{\rho^a P^a\} - 1 \\ \text{s.t.} \quad & \rho^a \geq 0, \quad , \mathrm{Tr}\{\rho^a\} = 1 \quad \forall a \end{aligned} \tag{5.7}$$

In the following, we connect the robustness to our guessing games. Recall that for any POVM $\mathbb{P} = \{P^a\}$ on system $\mathcal{H}_A$, its guessing probability $P_{\text{guess}}(\mathcal{E}, \mathbb{P})$ for an ensemble $\mathcal{E} := \{\rho^x\}$ is defined as:

$$\max_{\mu,w,\{\mathcal{N}_r\}} \sum_r \mu(r) \sum_{a,x} w(x|a,r)\, \mathrm{Tr}\{P^a\, \mathcal{N}_r(\rho_R^x)\} \ . \tag{5.8}$$

Where $\sum_x \mathrm{Tr}\{\rho^x\} = 1$. Let us denote $p(x) = \mathrm{Tr}\{\rho^x\}$, $p^*(x) = \max_x p(x)$.
Let us first define the free POVMs the POVM whose elements are all proportional to the identity. Suppose $P_a = \alpha_a \mathbb{1}$, we have,

$$P_{\text{guess}}(\mathcal{E}, \mathbb{P}) = \max_{\mu,w,\{\mathcal{N}_r\}} \sum_r \mu(r) \sum_{a,x} w(x|a,r)\, \mathrm{Tr}\Big\{\mathcal{N}_r^\dagger(\alpha_a \mathbb{1})\rho_R^x\Big\} \tag{5.9}$$

$$= \max_{\mu,w} \sum_r \mu(r) \sum_{a,x} w(x|a,r)\alpha_a\, \mathrm{Tr}\{\rho_R^x\} \tag{5.10}$$

$$= \max_{\mu,w} \sum_r \mu(r) \sum_{a,x} w(x|a,r)\alpha_a p(x) \tag{5.11}$$

$$\leq \max_{\mu,w} \sum_r \mu(r) p^*(x) \sum_{a,x} w(x|a,r)\alpha_a \tag{5.12}$$

$$= p^*(x) =: P_{\text{guess}}^{\text{free}}(\mathcal{E}) \tag{5.13}$$

Where we denote $P^{\text{free}}_{\text{guess}}(\mathcal{E})$ as the guessing probability by using trivial POVMs, and $P_{\text{guess}}(\{\rho^x_R\}, \mathbb{P})$ achieve its maximum $p^*(x)$ bound by letting $w(x|a,r) = \delta_{a,x'}$ and x' is the place where $p^*(x) := \max_x p(x)$.
It is evident that all free POVMs give a constant guessing probability for any ensemble $\mathcal{E} := \{\rho^x\}$, which is the $p^*(x) := \max_x p(x)$.

We can now state the following theorem, in which we make the connection of quantification of POVM as a resource to the advantage for which we compare to useless resources, the trivial POVM.

Theorem 6. *The robustness of a POVM $\mathbb{P}$,denote by $\mathfrak{R}(\mathbb{P})$, satisfy the following:*

$$1 + \mathfrak{R}(\mathbb{P}) = \max_{\mathcal{E}} \frac{P_{\text{guess}}(\mathcal{E}, \mathbb{P})}{P^{\textit{free}}_{\text{guess}}(\mathcal{E})} \tag{5.14}$$

where $\mathcal{E} := \{\rho^x\}$ and $\sum_x \text{Tr}\{\rho^x\} = 1$.

Proof. We first prove the following:

$$1 + \mathfrak{R}(\mathbb{P}) \geq \frac{P_{\text{guess}}(\mathcal{E}, \mathbb{P})}{P^{\text{free}}_{\text{guess}}(\mathcal{E})} \qquad \forall \mathcal{E} \tag{5.15}$$

We first assume that $P_{\text{guess}}(\mathcal{E}, \mathbb{P})$ achieve its maximal optimum through $\mu^*, w^*, \mathcal{N}^*$, i.e.,

$$P_{\text{guess}}(\mathcal{E}, \mathbb{P}) = \sum_r \mu^*(r) \sum_{a,x} w^*(x|a,r) \,\text{Tr}\{P^a \,\mathcal{N}^*_r(\rho^x_R)\} \tag{5.16}$$

Recall that $[1 + R(\mathbb{P})]q^*(a)\mathbb{1} \geq P_a$ since we denoted the optimal $q^*(a)$ as the constraint for the solution of primal SDP. Then we have,

$$\begin{aligned}
P_{\text{guess}}(\mathcal{E}, \mathbb{P}) &\leq \sum_r \mu^*(r) \sum_{a,x} w^*(x|a,r)[1 + \mathfrak{R}(\mathbb{P})]q^*(a) \,\text{Tr}\{\mathbb{1}\mathcal{N}^*_r(\rho^x_R)\} && (5.17)\\
&\leq \sum_r \mu^*(r) \sum_{a,x} w^*(x|a,r)[1 + \mathfrak{R}(\mathbb{P})]q^*(a)p(x) && (5.18)\\
&\leq \sum_r \mu^*(r) \sum_{a,x} w^*(x|a,r)[1 + \mathfrak{R}(\mathbb{P})]q^*(a)p^*(x) && (5.19)\\
&= [1 + \mathfrak{R}(\mathbb{P})]p^*(x) \sum_r \mu^*(r) \sum_{a,x} w^*(x|a,r)q^*(a) && (5.20)\\
&= [1 + \mathfrak{R}(\mathbb{P})]p^*(x) && (5.21)\\
&= [1 + \mathfrak{R}(\mathbb{P})]P^{\text{free}}_{\text{guess}}(\mathcal{E}) && (5.22)
\end{aligned}$$

The first inequality holds because we plug in the inequality $[1 + R(\mathbb{P})]q^*(a)\mathbb{1} \geq P^a$ into the trace. The second inequality holds because $\mathcal{N}^*$ is trace preserving. the

third inequality holds by replacing the maximal $p^*(x)$ with $p(x)$ for all x.
Next, we prove the following direction,

$$1 + \mathfrak{R}(\mathbb{P}) \leq \max_{\mathcal{E}} \frac{P_{\text{guess}}(\mathcal{E}, \mathbb{P})}{P^{\text{free}}_{\text{guess}}(\mathcal{E})} \tag{5.23}$$

Suppose we have the ensemble $\mathcal{E}^* = \{\rho^x\}$, such that $p(x) = \text{Tr}\{\rho^x\} = \frac{1}{|\mathcal{X}|}$, i.e, uniform probability distribution. Hence $p^*(x) := \max_x p(x) = \frac{1}{|\mathcal{X}|}$.
Suppose we have the optimal $\{\rho^{a*}\}$ for the dual SDP defined above and we get $1 + R(\mathbb{P}) = \sum_a \text{Tr}\{\rho^{a*} P^a\}$, where we restrict $\rho^{a*} \geq 0$ and $\text{Tr}\{\rho^{a*}\} = 1$. Instead of proving the above inequality, we prove the following,

$$1 + \mathfrak{R}(\mathbb{P}) \leq \frac{P_{\text{guess}}(\mathcal{E}^*, \mathbb{P})}{P^{\text{free}}_{\text{guess}}(\mathcal{E}^*)} \tag{5.24}$$

After the above setting, we can manipulate our objective equation as follows,

$$P_{\text{guess}}(\mathcal{E}, \mathbb{P}) = \max_{\mu, w, \{\mathcal{N}_r\}} \sum_r \mu(r) \sum_{a,x} w(x|a,r) \, \text{Tr}\{P^a \, \mathcal{N}_r(\rho^x_R)\} \tag{5.25}$$

$$= \max_{\mu, w, \{\mathcal{N}_r\}} \sum_a \text{Tr}\left\{ P^a \left\{ \frac{1}{|\mathcal{X}|} \sum_{r,x} \mu(r) w(x|a,r) \mathcal{N}_r(|\mathcal{X}| \rho^x_R) \right\} \right\} \tag{5.26}$$

Since we can always find a CPTP map to transfer between any two states (an extreme example is the discard and prepare map, which is indeed a CPTP), we are able to define the following:

$$\rho^{a*} = \sum_{r,x} \mu^*(r) w^*(x|a,r) \mathcal{N}^*_r(|\mathcal{X}| \rho^x_R) \tag{5.27}$$

where $\mu^*, w^*, \mathcal{N}^*$ serve as the sub-optimal constraints for $P_{\text{guess}}(\mathcal{E}, \mathbb{P})$ and we can verify that it indeed satisfies $\rho^{a*} \geq 0$ and $\text{Tr}\{\rho^{a*}\} = 1$. Then, we conclude our proof with the following:

$$P_{\text{guess}}(\mathcal{E}, \mathbb{P}) \geqslant \sum_a \text{Tr}\left\{ \frac{1}{|\mathcal{X}|} \rho^{a*} P^a \right\} = [1 + \mathfrak{R}(\mathbb{P})] p^*(x) = P^{\text{free}}_{\text{guess}}(\mathcal{E})[1 + \mathfrak{R}(\mathbb{P})] \tag{5.28}$$

∎

5.4 Robustness of PMD with relation to post-information guessing game

Recall the definition of simple PMDs, which is the equivalence of a family of compatible measurements.

Definition 5.4.1 (Simple PMDs, *alias* Compatible POVMs)**.** *A PMD $M^Q(a|x)$ is called simple if its constituting POVMs can be written as*

$$M^Q(a|x) = \sum_{i \in \mathcal{I}} p(a|i,x) \tilde{M}^Q(i), \tag{5.29}$$

where the $\tilde{M}^Q(i)$ are elements of a single POVM (sometime referred to as the "mother" POVM), and $p(a|i,x)$ is a conditional probability distribution.

As noted previously, any convex mixing of simple PMDs can be directly incorporated into the "mother" POVM. We now describe this in a bit more detail. Suppose that $M^Q(a|x)$ admits a decomposition of the form

$$M^Q(a|x) = \sum_r \mu(r) \sum_{i\in\mathcal{I}} p(a|i,x,r)\tilde{M}^Q(i|r),$$

where $\mu(r)$ is a probability distribution and $\tilde{M}^Q(i|r)$ is now a *family* of POVMs indexed by the shared random index r. Then, simply by noticing that $\mu(r)\tilde{M}^Q(i|r)$ is itself a normalized two-outcome indexed POVM, it is possible to conclude that Definition 5.4.1 is fully general and no further random variables are needed.

5.4.1 Convex optimization of robustness of PMDs

This section proves the connection of robustness and the guessing game scenario. We show that the generalized robustness of PMD is an exact quantifier for the advantage in some guessing games. Suppose we are given a PMD, $\{M(a|x) : a \in \mathcal{A}, x \in \mathcal{X}\}$ on the Hilbert space $\mathcal{H}^Q$, and an ensemble $\{\rho_{a,x}\}$. According to the theorem in main text, it is possible to restrict all $\rho_{a,x} \in \mathcal{L}(\mathcal{H}^Q) \quad \forall a,x$.

At first, we define the set of *simple* PMDs as,

$$\begin{aligned}\mathcal{F}_{Q,\mathcal{A},\mathcal{X}} &= \{\{M(a|x)\}_{a,x} : \exists \text{ POVM } \{\widetilde{M}^Q(i)\}, p(a|x,i), \text{s.t. } M^Q(a|x) \\ &\quad = \sum_{i\in\mathcal{I}} p(a|i,x)\widetilde{M}^Q(i) \quad \forall a,x\},\end{aligned} \tag{5.30}$$

Definition (5.29) identifies $\mathcal{F}_{Q,\mathcal{A},\mathcal{X}}$ as a collection of POVM familes that is convex and closed. Since it is possible to fix $Q,\mathcal{A},\mathcal{X}$, in what follows we ignore the subscripts. We define $\mathbb{M} := \{M(a|x)\}_{a,x}$ and in what follows we use the same font style to represent PMDs. We denote by $\mathcal{Z}$ the set of general PMDs, that is to say $\mathcal{F} \subseteq \mathcal{Z}$.

For usefullness, we define a real vector space $\mathcal{V}$ as,

$$\mathcal{V} := \left\{ \mathbb{V} = \begin{pmatrix} V_1 \\ \vdots \\ V_d \end{pmatrix} : V_i = V_i^\dagger \quad \forall i \right\} \tag{5.31}$$

on $\mathbb{R}_+$, while we define its inner product as $\langle \mathbb{A}|\mathbb{B}\rangle := \sum_i \langle A_i, B_i\rangle = \langle \mathbb{B}|\mathbb{A}\rangle$, and the notation $\langle\cdot,\cdot\rangle$ is defined as the inner product, such that $\langle A,B\rangle := \mathrm{Tr}\{AB\}$, and $d = |\mathcal{X}||\mathcal{A}|$. Note that each element in $\mathcal{Z}$ corresponds to a unique vector in $\mathcal{V}$.

Let us first define the convex cone generated by *simple* PMDs as,

$$\mathcal{C} := \{c\mathbb{W} : c \in \mathbb{R}_+, \mathbb{W} \in \mathcal{F}\}, \tag{5.32}$$

as well as its dual,

$$\mathcal{C}^* := \{\mathbb{E} \in \mathcal{V} : \langle \mathbb{E}|\mathbb{F}\rangle \geq 0, \forall \mathbb{F} \in \mathcal{C}\}. \tag{5.33}$$

We then define the generalized robustness of PMD $\mathbb{M}$ with respect to $\mathcal{F}$,

$$\mathfrak{R}(\mathbb{M}) := \min\{r \in \mathbb{R}_+ : \mathbb{M} + r\mathbb{N} \in \mathcal{C}, \quad \mathbb{N} \in \mathcal{Z}\} \tag{5.34}$$

where $\mathbb{M} + r\mathbb{N} \in \mathcal{C}$ is equivalent to the fact that the family of $\{M(a|x) + rN(a|x)\}$ considered as a vector defined in (5.31) is in set $\mathcal{C}$.

In order to see the connection between robustness and guessing game, let's us define $\mathbb{N}' := r\mathbb{N}$, i.e., $N'(a|x) = rN(a|x) \forall a, x$, and define $\mathbb{N}' \succ 0$ the same fashion as, $N'(a|x) \geq 0 \quad \forall a, x$, then we rewrite the definition (5.34) as a conic form problem (which we call primal problem) with generalized inequality $\succ$, i.e., given $\mathbb{M}$, we want:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \mathbb{M} + \mathbb{N}' \in \mathcal{C} \\ & \mathbb{N}' \succ 0, \\ & \sum_a N'(a|x) = \lambda \mathbb{1}, \quad \forall x. \end{array} \tag{5.35}$$

Introducing hermitian operators γ_x as Lagrange multiplies, we can write the Lagrangian with respect to $\mathbb{M}$ as

$$\mathcal{L}(\lambda, \mathbb{N}', \mathbb{A}, \mathbb{B}, \{\gamma_x\}) = \lambda - \langle \mathbb{M} + \mathbb{N}'|\mathbb{A}\rangle - \langle \mathbb{N}'|\mathbb{B}\rangle - \sum_x \left\langle \lambda \mathbb{1} - \sum_a N'(a|x), \gamma_x \right\rangle \tag{5.36}$$

$$= -\langle \mathbb{M}|\mathbb{A}\rangle + \lambda(1 - \sum_x \mathrm{Tr}\{\gamma_x\}) + \sum_{a,x} \left\langle N'(a|x), \gamma_x - \beta_{a,x} - \alpha_{a,x} \right\rangle. \tag{5.37}$$

where the dual variables satisfy $\mathbb{A} \in \mathcal{C}^*$, $\mathbb{B} \succ 0$, and the elements of $\mathbb{A}$ and $\mathbb{B}$ are $\{\alpha_{a,x}\}$ and $\{\beta_{a,x}\}$ respectively. Then we write the dual function as,

$$g(\mathbb{A}, \mathbb{B}, \{\gamma_x\}) = \min_{\lambda, \mathbb{N}'} \mathcal{L}(\lambda, \mathbb{N}', \mathbb{A}, \mathbb{B}, \{\gamma_x\}) \tag{5.38}$$

$$= -\langle \mathbb{M}|\mathbb{A}\rangle + \min_{\lambda, \mathbb{N}'} \left(\lambda(1 - \sum_x \mathrm{Tr}\{\gamma_x\}) + \sum_{a,x} \left\langle N'(a|x), \gamma_x - \beta_{a,x} - \alpha_{a,x} \right\rangle \right) \tag{5.39}$$

since g is linear function and a linear function is bounded below only when it is identical zero. Thus, $g = -\infty$ (trivial bound), except only when the following two conditions hold,

$$\begin{cases} \sum_x \mathrm{Tr}\{\gamma_x\} = 1 \\ \gamma_x - \beta_{a,x} - \alpha_{a,x} = 0 \quad \forall a, x, \end{cases}$$

in which cases, $g(\mathbb{A}, \mathbb{B}, \{\gamma_x\}) = -\langle \mathbb{M}|\mathbb{A}\rangle$. So we can write the dual problem to define the upper bound of dual fucntion as follows,

$$\begin{array}{ll} \text{maximize} & -\langle \mathbb{M}|\mathbb{A}\rangle \\ \text{subject to} & \mathbb{A} \in \mathcal{C}^* \\ & \mathbb{B} \succ 0, \\ & \gamma_x - \beta_{a,x} - \alpha_{a,x} = 0 \quad \forall a, x, \\ & \sum_x \mathrm{Tr}\{\gamma_x\} = 1, \qquad \gamma_x = \gamma_x^\dagger. \end{array} \tag{5.40}$$

we can get rid of the dual variable $\mathbb{B}$ by combining the second and third constriant as the condition $\gamma_x - \alpha_{a,x} \geq 0 \quad \forall a, x$, becasue $\mathbb{B}$ is only the constraint of dual variables, then the above problem reduces to,

$$\begin{array}{ll} \text{maximize} & -\langle \mathbb{M}|\mathbb{A}\rangle \\ \text{subject to} & \mathbb{A} \in \mathcal{C}^* \\ & \gamma_x - \alpha_{a,x} \geq 0 \qquad \forall a, x, \\ & \sum_x \mathrm{Tr}\{\gamma_x\} = 1, \quad \gamma_x = \gamma_x^\dagger. \end{array} \tag{5.41}$$

Define a new variable $\mathbb{W}$, such that its element $\omega_{a,x} := \gamma_x - \alpha_{a,x}$, and we see that,

$$-\langle \mathbb{M}|\mathbb{A}\rangle = \langle \mathbb{M}|\mathbb{W}\rangle - 1 = \sum_{a,x} \left\langle M(a|x), \omega_{a,x} \right\rangle - 1,$$

then we can rewrite the dual problem as,

$$\begin{array}{ll} \text{maximize} & \langle \mathbb{M}|\mathbb{W}\rangle - 1 \\ \text{subject to} & \mathbb{A} \in \mathcal{C}^* \\ & \mathbb{W} \succ 0 \\ & \sum_x \mathrm{Tr}\{\gamma_x\} = 1, \quad \gamma_x = \gamma_x^\dagger. \end{array} \tag{5.42}$$

where $\mathbb{W} \succ 0$ is equivalent to, $\gamma_x - \alpha_{a,x} \geq 0 \forall a, x$. To see that the strong duality holds, that is to say, the optimal value of the dual is equal to the optimal value of the primal problem, let's choose $\alpha_{a,x} = \frac{1}{2|\mathcal{X}|\,\mathrm{Tr}\{\mathbb{1}\}}\mathbb{1}, \forall a, x$, i.e., $\mathbb{A} \succ 0$ (thus $\mathbb{A}$ is in the interior of $\mathcal{C}^*$), and $\gamma_x := 2\alpha_{a,x}$, we then see that $\gamma_x - \alpha_{a,x} = \alpha_{a,x} > 0 \quad \forall a, x$ and $\sum_x \mathrm{Tr}\{\gamma_x\} = 1$. These choices can be noticed to strictly satisfy the conditions (5.42). So Slater's theorem ensures that the strong duality holds.

With the above techniques, we are able to connect the robustness of PMD to post information guessing games that was introduced previously.

Theorem 7. *For any PMD, with its robustness related to guessing games, we have,*

$$1 + \mathfrak{R}(\{M(a|x)\}_{a,x}) = \max_{\{\rho_{a,x}\}} \frac{P_{\text{guess}}(M(a|x); \rho_{a,x})}{P_{\text{guess}}^{\textit{simple}}(\rho_{a,x})}\,, \tag{5.43}$$

Proof. We first show the right hand side is smaller than or equal to the left hand side for all possible ensembles, then we show a special choosing ensemble satisfies that the right hand side is greater than or equal to the left hand side, which can be seen as the optimal ensemble.

According to the definition of the general robustness of $\{M(a|x)\}_{a,x}$, one can write $M(a|x) = (1+r)F(a|x) - rN(a|x)$ for some $\mathbb{F} \in \mathcal{F}$ with elements as $F(a|x)$, where $r = \mathfrak{R}(\{M(a|x)\}_{a,x})$. By using the same notations as shown in Theorem 5, and according to Lemma 1, we obtain,

$$P_{\text{guess}}^{\text{simple}}(\rho_{a,x}) = \max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \rho_{a,x}, \right\rangle \tag{5.44}$$

and,

$$\begin{aligned}
P_{\text{guess}}(M(a|x);\rho_{a,x}) &= \max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(M)(a|x), \rho_{a,x} \right\rangle && (5.45)\\
&= \sum_{a,x} \left\langle \mathcal{T}^*(M)(a|x), \rho_{a,x} \right\rangle && (5.46)\\
&= (1+r)\sum_{a,x} \left\langle \mathcal{T}^*(F)(a|x), \rho_{a,x} \right\rangle - r\sum_{a,x} \left\langle \mathcal{T}^*(N)(a|x), \rho_{a,x} \right\rangle && (5.47)\\
&\leqslant (1+r)\max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \rho_{a,x} \right\rangle && (5.48)\\
&= (1+r)P_{\text{guess}}^{\text{simple}}(\rho_{a,x}), && (5.49)
\end{aligned}$$

where the third equality holds because the optimized $\mathcal{T}^*$ is linear.

Next we choose an ensemble $\rho_{a,x} = \omega_{a,x}$ (up to normalization constriant of the ensemble) satisfying the constraint in where we consider the set of optimal $\{\omega_{a,x}\}$ appear in the dual problem (5.42), under this ensemble, we obtain,

$$\begin{aligned}
\frac{P_{\text{guess}}(M(a|x);\rho_{a,x})}{P_{\text{guess}}^{\text{simple}}(\rho_{a,x})} &= \frac{\max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(M)(a|x), \omega_{a,x} \right\rangle}{\max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \omega_{a,x} \right\rangle} && (5.50)\\
&\geqslant \frac{\sum_{a,x} \left\langle M(a|x), \omega_{a,x} \right\rangle}{\max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \omega_{a,x} \right\rangle} && (5.51)\\
&\geqslant \frac{\sum_{a,x} \left\langle M(a|x), \omega_{a,x} \right\rangle}{\max_{\mathcal{T}\in\mathscr{T}} \sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \gamma_{x} \right\rangle} && (5.52)\\
&= \frac{\sum_{a,x} \left\langle M(a|x), \omega_{a,x} \right\rangle}{\sum_{x} \left\langle \mathbb{1}, \gamma_{x} \right\rangle} && (5.53)\\
&= 1 + \mathfrak{R}(\{M(a|x)\}_{a,x}) && (5.54)
\end{aligned}$$

where the first inequality holds because of maximization over all possible $\mathcal{T}$, the second inequality holds because of the constraint $\gamma_x - \omega_{a,x} = \alpha_{a,x}$ and $\mathbb{A} \in \mathcal{C}^*$ in (5.42), which brings the fact that $\sum_{a,x} \left\langle \mathcal{T}(F)(a|x), \gamma_x - \omega_{a,x} \right\rangle \geqslant 0$, and the last equality holds because we have that $\sum_a \mathcal{T}(F)(a|x) = \mathbb{1}$ and also $\sum_x \mathrm{Tr}\{\gamma_x\} = 1$, which concludes the proof. ∎

Chapter 6

Convex quantum resource theory

6.1 Framework

This chapter generalizes the discussions of previous chapters and introduces general resource theory. In Chapter 3, we studied noisy channel transformation from the viewpoint of statistical model comparisons. For classical noisy channel comparisons, we introduced the transformation among noisy channels through statistical decision games at BSS Theorem (2). Then in Chapter 4, we apply the same idea to the quantum domain and extend it to quantum measurement comparison. The comparison of quantum measurement was then characterized through statistical guessing games at Theorem (4) for a single measurement case and Theorem (5) for a family of measurement cases. In all these object comparisons, an important concept is the transformation map that transfers one object into another. Now let us list all the transformation maps in the following:

1. Post-processing of noisy channel in BSS ordering;
2. Pre- and post-processing of noisy channel in Shannon ordering;
3. Pre- and post-processing in POVM ordering;
4. Pre- and post-processing in PMD ordering;

We notice that in the above transformation, each has its own physical or informational interpretations when applied to some specific cases. Moreover, by introducing a transformation map between both classical and quantum noisy channels[1], we have established a preorder among the noisy channels. However, we are not interested in the class of lower bound of the ordering in the classical domain, but in the quantum domain, the class of lower bound of ordering and its beyond become important. The reason lies in the core difference between quantum and classical physics. Quantum physics, comprises resources that classical physics cannot simulate, which gives some advantage for quantum resources in

[1]We say that quantum measurement is a special case of a quantum noisy channel.

some certain information processing tasks. The disadvantages from classical resources when compared to quantum resources result in the class of object that lies in the lower bound of the ordering. Understanding the advantage of the quantum resource is one of the most important tasks in quantum information science.

Another important feature of the transformation of noisy channels in the quantum domain is that it preserves the class of the lower bound of the ordering it establishes. This is natural because it will not create the quantum resource from the classically simulatable operations. For example, in the case of POVM ordering, pre-and post-processing transformation will not generate a non-trivial POVM from trivial POVMs, and in the case of PMD ordering, pre-and post-processing transformation will not generate a non-simple PMD from simple PMDs.

Chapter 5 introduced the concept of robustness of quantum resource, which is a quantifier of the resource. Quantifying quantum resources become natural when ordering of quantum resource is critical. The motivations for the quantum resource lie in the fact that these classically simulatable resources should be quantified as zero, while those that are non-classically simulatable resources should be quantified positively. The preorder of quantum resource implies that there is not only one class of quantification of the resource. In what follows, we will introduce, besides robustness, other quantifiers of quantum resources.

We are now ready to introduce the following postulates for convex quantum resource theory.

Postulate 1: *Quantum resource forms a closed and convex set with a subset of free resource.*

Postulate 2: *There exist restricted transformations between resources that preserve the null resource.*

6.2 Underlying motivation of general quantum resource theory

Quantum and classical information theories can be viewed as theories of interconversion between various resources[23], which can be quantum or classical, static or dynamic, noisy or noiseless. An example in the classical domain for dynamical resource is Shannons channel coding theorem [1], where the task is to transform noisy classical channels into noiseless channels for arbitrary message transmission, under the restriction that no side communication is allowed between sender and receiver. Another example in the quantum domain but for static resources is entanglement concentration[24], where the task is to concentrate n pairs of particles in identical partly entangled pure states into a smaller number of maximally entangled pairs of particles, under the restriction that only local operations are allowed. With interest in physics, the same idea can be applied to the field of thermodynamics, where there is interest in the possibility of transforming one thermal state

into another under the constrained process. The restricted process is the one that preserves the thermal equilibrium states and the states that break the equilibrium are called static resources [25]. A generalized theory could be derived as resource theory, which would make it possible to explain fundamentally the nature from the viewpoint of resource. In addition, we see these core elements in information processing tasks as resources. In this chapter, resource theory is investigated with the largest restricted transformations, i.e., those morphisms that preserve the set of free resource.

The following section formulates the theory in mathematical languages.

6.3 Preliminaries

6.3.1 Mathematical notations and settings

We denote by $\mathcal{D}(\mathbb{C}^m)$ the set of all m-by-m complex density matrices ρ, i.e., $\rho \geqslant 0$ and $\mathrm{Tr}\{\rho\} = 1$, which are used here to represent quantum states of m-dimensional quantum systems. Within $\mathcal{D}(\mathbb{C}^m)$, we identify a non-empty closed convex subset $\mathcal{F}$ as the set of "free states". The closure and the convexity of $\mathcal{F}$ is crucial in various steps of our proofs, for example, when invoking the closure under convex mixtures, or when applying a variant of the minimax theorem that requires convex domain. Here and throughout this work, *resource morphisms* (or more precisely $\mathcal{F}$-morphisms[2]) are defined as completely positive, trace-preserving (CPTP) linear maps $\mathcal{E} : \mathcal{D}(\mathbb{C}^m) \to \mathcal{D}(\mathbb{C}^m)$ such that $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{F}$. More generally, one may consider CPTP maps that change the dimension of the system, for example, $\mathcal{E} : \mathcal{D}(\mathbb{C}^m) \to \mathcal{D}(\mathbb{C}^n)$. Also in this case, whenever the *output* free set $\mathcal{F}'$ is also specified, it is possible to define a notion of resource morphisms by the condition that $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{F}'$. However in what follows, for the sake of readability, we will restrict ourselves to the case of equal input and output dimensions and $\mathcal{F} = \mathcal{F}'$, keeping in mind however that all the results we derive can be straightforwardly extended to the general case. We will go back to the more general setting, with different input and output systems, in Section 6.5 when discussing various applications like the tasks of resource dilution and distillation.

Definition 6.3.1 (Resourcefulness Preorder)**.** *Given two density matrices $\rho, \sigma \in \mathcal{D}(\mathbb{C}^m)$, we write $\rho \succ_\epsilon \sigma$ whenever there exists a resource morphism $\mathcal{E}$ such that $\frac{1}{2}\left\|\sigma - \mathcal{E}(\rho)\right\|_1 \leqslant \epsilon$, where $\left\|X\right\|_1 := \mathrm{Tr}\left\{\sqrt{X^\dagger X}\right\}$ denotes the trace-norm. In particular, we write $\rho \succeq \sigma$ whenever $\rho \succ_{\epsilon=0} \sigma$.*

In the above definition, the preorder $\succ_\epsilon$ has been introduced with respect to the distance induced by the trace-norm, although it is possible to use any other

[2]Here we prefer the term "resource morphisms" to the more common "free operations" because it reminds the fact that the foundational concept, in the geometric approach in which we are working, is the free set $\mathcal{F}$, not the set of allowed transformations, which are just defined as all those that map $\mathcal{F}$ into itself.

well-behaved distance measure between density matrices (like the fidelity, for example) without substantially changing the results [26, 27]. In any case, an important thing to notice in Definition 6.3.1 is that, even though errors are allowed in the state transformation, we always require the constraint $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{F}$ to be *strictly* satisfied.

The resourcefulness preorder naturally lead us to define a maximally resourceful element as follows:

Definition 6.3.2 (Maximally resourceful element)**.** *An element $\alpha \in \mathcal{D}(\mathbb{C}^m)$ is said to be maximally resourceful if $\alpha \succ \sigma$ for any $\sigma \in \mathcal{D}(\mathbb{C}^m)$.*

Given a general resource theory, an important question is to whether the theory possesses maximally resourceful elements or not. In the following we will consider sufficient conditions for their existence. However, we recall that our main results do not rely in any way on the existence of maximally resourceful elements.

6.3.2 Information-theoretic divergences

In what follows, for any operator $\rho \in \mathcal{D}(\mathbb{C}^m)$ we denote by Π_ρ the orthogonal projector onto its support (i.e., the orthogonal complement of its kernel). Moreover, for any $\epsilon \in [0,1]$, we denote by $\mathsf{B}^\epsilon(\rho)$ the set of operators $\{\rho' \in \mathcal{D}(\mathbb{C}^m) : \|\rho - \rho'\|_1 \leqslant 2\epsilon\}$ and by $\mathsf{P}^\epsilon(\rho)$ the set of operators $\{P : 0 \leqslant P \leqslant \mathbb{1} \text{ and } \mathrm{Tr}\{\rho P\} \geqslant 1-\epsilon\}$. All logarithms are taken in base 2.

Definition 6.3.3 (Relative entropies)**.** *Given two density matrices $\rho, \sigma \in \mathcal{D}(\mathbb{C}^m)$, we define*

1. *the Umegaki relative entropy [28]:*

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}\{\rho\,(\log\rho - \log\sigma)\}\ , & \text{if } \Pi_\sigma \geqslant \Pi_\rho\ , \\ +\infty\ , & \text{otherwise}\ ; \end{cases} \tag{6.1}$$

2. *the hypothesis testing relative entropy [29]: for any $\epsilon \in [0,1]$*

$$D_h^\epsilon(\rho\|\sigma) := -\log \min_{P \in \mathsf{P}^\epsilon(\rho)} \mathrm{Tr}\{\sigma\, P\}\ , \tag{6.2}$$

with the convention $-\log 0 := +\infty$; for $\epsilon = 0$, one recovers the min-divergence, defined as [30]

$$D_{\min}(\rho\|\sigma) := -\log \mathrm{Tr}\{\sigma\, \Pi_\rho\}\ ; \tag{6.3}$$

3. *the max-divergence [30]:*

$$D_{\max}(\rho\|\sigma) := \begin{cases} \log\min\{\lambda \in \mathbb{R} : \lambda\sigma - \rho \geqslant 0\}\ , & \text{if } \Pi_\sigma \geqslant \Pi_\rho\ , \\ +\infty\ , & \text{otherwise}\ ; \end{cases} \tag{6.4}$$

in this case we also define a "smoothed" version as follows: for any $\epsilon \in [0,1]$,

$$D_{\max}^\epsilon(\rho\|\sigma) := \inf_{\rho' \in \mathsf{B}^\epsilon(\rho)} D_{\max}(\rho'\|\sigma)\ . \tag{6.5}$$

A crucial property satisfied by all these divergences is the monotonicity under CPTP linear maps, that is, for example, $D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma) \leqslant D(\rho\|\sigma)$ and analogously for the others. In a resource theory characterized by the set of free states $\mathcal{F}$, we also introduce the following:

Definition 6.3.4 (Max-divergence of resourcefulness)**.** *Given two density matrices $\rho, \sigma \in \mathcal{D}(\mathbb{C}^m)$ and a non-empty closed convex subset $\mathcal{F} \subseteq \mathcal{D}(\mathbb{C}^m)$, the max-divergence relative to $\mathcal{F}$ is defined as*

$$D_{\max,\mathcal{F}}(\rho\|\sigma) := \log \inf \left\{ \lambda \in \mathbb{R} : \frac{\lambda\sigma - \rho}{\lambda - 1} \in \mathcal{F} \right\} , \tag{6.6}$$

with the convention that $\inf \varnothing = +\infty$. Its "smoothed" version is defined in analogy with (6.5)*, that is*

$$D^{\epsilon}_{\max,\mathcal{F}}(\rho\|\sigma) := \inf_{\rho' \in \mathsf{B}^{\epsilon}(\rho)} D_{\max,\mathcal{F}}(\rho'\|\sigma) . \tag{6.7}$$

We notice that if $\mathcal{F} = \mathcal{D}(\mathbb{C}^m)$, then $D_{\max,\mathcal{F}}(\rho\|\sigma) = D_{\max}(\rho\|\sigma)$, but in general $D_{\max,\mathcal{F}}(\rho\|\sigma) \geqslant D_{\max}(\rho\|\sigma)$. Moreover, while $D_{\max,\mathcal{F}}$ may fail to be monotonic under general CPTP maps, it is monotonic under the action of resource morphisms, that is, CPTP maps that map $\mathcal{F}$ into itself. This can be easily seen by noticing that, if for some λ, $(\lambda - 1)^{-1}(\lambda\sigma - \rho)$ is in $\mathcal{F}$, then, for any resource morphism $\mathcal{E}$, also $(\lambda - 1)^{-1}\mathcal{E}(\lambda\sigma - \rho) = (\lambda - 1)^{-1}(\lambda\mathcal{E}(\sigma) - \mathcal{E}(\rho))$ is automatically in $\mathcal{F}$, so that $D_{\max,\mathcal{F}}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$ cannot be larger than $D_{\max,\mathcal{F}}(\rho\|\sigma)$.

6.3.3 Resource monotones

We say that a function $f : \mathcal{D}(\mathbb{C}^m) \to [0, +\infty]$ constitutes a *resource monotone* if it achieves its global minimum on all elements of $\mathcal{F}$, and it does not increase under the action of resource morphisms, i.e., $f(\rho) \geqslant f(\mathcal{E}(\rho))$ for any resource morphism $\mathcal{E}$. More properties can be demanded (and are indeed desirable) in order to fruitfully work with concrete examples of resource monotones. The information-theoretic divergences introduced above can be used to introduce resource monotones that inherit many useful properties from the parent divergence. In our construction, the following quantities play a central role [31].

Definition 6.3.5 (Entropic Resource Monotones)**.** *Given a non-empty closed convex set $\mathcal{F} \subseteq \mathcal{D}(\mathbb{C}^m)$, for any density matrix $\rho \in \mathcal{D}(\mathbb{C}^m)$ and any $\epsilon \in [0, 1]$, we define the following quantities:*

1. $\mathfrak{D}(\rho) := \inf_{\omega\in\mathcal{F}} D(\rho\|\omega)$*;*
2. $\mathfrak{D}^{\epsilon}_{h}(\rho) := -\log \max_{\omega\in\mathcal{F}} \min_{P\in\mathsf{P}^{\epsilon}(\rho)} \mathrm{Tr}\{P\,\omega\}$*, with the convention* $-\log 0 := +\infty$*;*
3. $\mathfrak{D}^{\epsilon}_{\max}(\rho) := \inf_{\omega\in\mathcal{F}} D^{\epsilon}_{\max}(\rho\|\omega)$*;*
4. $\mathfrak{D}^{\epsilon}_{\max,\mathcal{F}}(\rho) := \inf_{\omega\in\mathcal{F}} D^{\epsilon}_{\max,\mathcal{F}}(\rho\|\omega)$.

In the case $\epsilon = 0$, we simply remove the superscript; the only exception is $\mathfrak{D}_h^{\epsilon=0}(\rho)$, for which we will use the special notation $\mathfrak{D}_{\min}(\rho)$.

The above quantities are all well-behaved resource monotones. This fact is a direct consequence of the monotonicity of the parent divergences under the action of resource morphisms.

Definition 6.3.6 (Free fraction and generalized free fraction)**.** *Given a non-empty closed convex free set $\mathcal{F} \subseteq \mathcal{D}(\mathbb{C}^m)$, the free fraction of a density matrix $\rho \in \mathcal{D}(\mathbb{C}^m)$ is defined by the formula*

$$\mathfrak{F}(\rho) := \max\{p \in [0,1] : \exists\, \omega \in \mathcal{F} \text{ s.t. } p\rho + (1-p)\omega \in \mathcal{F}\} \ . \tag{6.8}$$

When mixing with general $\omega \in \mathcal{D}(\mathbb{C}^m)$ instead of $\omega \in \mathcal{F}$ is allowed, one obtains the generalized free fraction, defined as

$$\mathfrak{F}_g(\rho) := \max\{p \in [0,1] : \exists\, \omega \in \mathcal{D}(\mathbb{C}^m) \text{ s.t. } p\rho + (1-p)\omega \in \mathcal{F}\} \ . \tag{6.9}$$

The free fraction and the generalized free fraction are related to the *robustness* $\mathfrak{R}(\rho)$ [32] and the *generalized robustness* $\mathfrak{R}_g(\rho)$ [33], respectively, through the relations $\mathfrak{F}(\rho)^{-1} = 1 + \mathfrak{R}(\rho)$ and $\mathfrak{F}_g(\rho)^{-1} = 1 + \mathfrak{R}_g(\rho)$, and they are both directly related with the entropic resource monotones in Definition 6.3.5 as follows:

$$-\log \mathfrak{F}(\rho) = \log(1 + \mathfrak{R}(\rho)) = \mathfrak{D}_{\max,\mathcal{F}}(\rho) \ , \tag{6.10}$$

$$-\log \mathfrak{F}_g(\rho) = \log(1 + \mathfrak{R}_g(\rho)) = \mathfrak{D}_{\max}(\rho) \ , \tag{6.11}$$

with the convention $-\log 0 := +\infty$. In particular, we have that $\mathfrak{D}_{\max}(\rho)$ coincides with the *generalized logarithmic robustness* of [30, 34], while $\mathfrak{D}_{\max,\mathcal{F}}(\rho)$ coincides with the *logarithmic robustness* of [35]. For this reason, in what follows, when speaking of $\mathfrak{D}_{\max}(\rho)$ (respectively, $\mathfrak{D}_{\max,\mathcal{F}}(\rho)$) we will follow the mainstream convention and call it "generalized log-robustness" (respectively, "log-robustness") even though, depending on the context, it would be more appropriate to use the term we introduced above, that is, "generalized log-free fraction" (respectively, "log-free fraction").

Remark. *All resource monotones introduced above would still be well-defined monotones even if the class of resource morphisms were enlarged to comprise also positive, but not completely-positive, linear maps. This seems no coincidence, since at the single-shot level, where no rule for composing system is given yet, there really is no compelling mathematical reason to limit the discussion to CPTP linear maps only. This is a common feature of various problems in quantum statistics, in particular quantum decision theory, where the theory becomes simpler if one works with quantum statistical morphism (which may violate complete positivity) and introduce CPTP maps as special cases, rather than starting from the beginning with fully blown CPTP maps [12]. Here we do not investigate further into this point, and simply justify the assumption of CPTP-ness on practical grounds.*

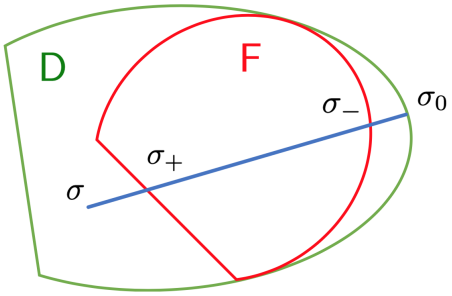

Figure 6.1: Convex set of quantum states and its subset of separable states.

Figure 6.2: Geometric intuition for the generalized free fraction introduced in Definition 6.3.6. Here σ_0 denotes the optimized density matrix, which is able to achieve, by means of convex mixing, the generalized free fraction of σ.

6.3.4 Optimal convex decompositions

Our main results rely on the following construction, whose intuitive picture is given in Fig. 6.2 below.

Given $\sigma \in \mathcal{D}(\mathbb{C}^m)$, assuming $\sigma \notin \mathcal{F}$, let us fix a convex decomposition achieving its generalized free fraction and write it as

$$\sigma_+ = \mathfrak{F}_g(\sigma)\sigma + [1 - \mathfrak{F}_g(\sigma)]\sigma_0 \, . \tag{6.12}$$

In the above equation, due to the optimality of $\mathfrak{F}_g$, $\sigma_+ \in \mathcal{F}$ lies on the border of $\mathcal{F}$, while σ_0 lies on the border of $\mathcal{D}(\mathbb{C}^m)$, as depicted in Fig. 6.2. The above decomposition includes the situation in which $\mathfrak{F}_g(\sigma) = 0$, that is, $\sigma_+ = \sigma_0$. For any decomposition as in (6.12), another free state σ_- can be uniquely defined using the max-divergence of resourcefulness (Definition 6.3.4) as follows:

$$\sigma_- := \frac{2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)}\sigma_+ - \sigma}{2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)} - 1} \tag{6.13}$$

$$= \left[\frac{\mathfrak{F}_g(\sigma)\, 2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)} - 1}{2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)} - 1}\right]\sigma + \left[1 - \frac{\mathfrak{F}_g(\sigma)\, 2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)} - 1}{2^{D_{\max,\mathcal{F}}(\sigma\|\sigma_+)} - 1}\right]\sigma_0 \, , \tag{6.14}$$

whenever $D_{\max,\mathcal{F}}(\sigma\|\sigma_+) < +\infty$, or $\sigma_- := \sigma_+$ otherwise. In order to derive (6.14) we just plugged (6.12) into (6.13) and rearranged terms. Notice that since $\sigma \notin \mathcal{F}$, we have $\sigma \neq \sigma_+$ and $D_{\max,\mathcal{F}}(\sigma\|\sigma_+) > 0$. It is easy to check that, by construction, σ_-, as σ_+, lies on the intersection between the border of $\mathcal{F}$ and the segment joining σ with σ_0. Our main results will originate from a careful evaluation of the relative distances between these four points in state space.

6.4 Resource-theoretic approach to state transitions

In this section, we state and prove the main results of this paper. Firstly we derive, for any finite-dimensional resource theory in which the set of free states is

non-empty closed and convex, sufficient conditions for the existence of a resource morphism between any two states, given in terms of resource monotones. Such conditions are formulated so to allow, in general, non-zero errors in the state transition, while the operation implementing the transition is an exact resource morphism.

Theorem 8. *Let us arbitrarily fix two states, $\rho, \sigma \in \mathcal{D}(\mathbb{C}^m)$, and two values $\epsilon_1, \epsilon_2 \in [0,1]$. Let us moreover choose $\tilde{\sigma} \in \mathsf{B}^{\epsilon_2}(\sigma)$ and $\tilde{\sigma}_+ \in \mathcal{F}$ so that $D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+) = \mathfrak{D}^{\epsilon_2}_{\max}(\sigma)$.*

1. *If $\mathfrak{D}^{\epsilon_1}_h(\rho) = +\infty$, then $\rho \succ_{\epsilon_1} \sigma$.*
2. *If $\mathfrak{D}^{\epsilon_2}_{\max}(\sigma) = 0$, then $\rho \succ_{\epsilon_2} \sigma$.*
3. *If $\mathfrak{D}^{\epsilon_1}_h(\rho) < +\infty$ and $\mathfrak{D}^{\epsilon_2}_{\max}(\sigma) > 0$, then*
   (a) *either $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) < +\infty$; in such a case, $\rho \succ_{\epsilon_1+\epsilon_2} \sigma$ if the following two conditions simultaneously hold:*

$$\mathfrak{D}^{\epsilon_1}_h(\rho) \geqslant \mathfrak{D}^{\epsilon_2}_{\max}(\sigma) \tag{6.15}$$

and

$$2^{-\max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)} \geqslant \frac{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) - \mathfrak{D}^{\epsilon_1}_h(\rho)} - 1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1} \; ; \tag{6.16}$$

   (b) *or $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) = +\infty$; in such a case, $\rho \succ_{\epsilon_1+\epsilon_2} \sigma$ if condition* (6.15) *above holds together with*

$$\max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega) = \min_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega) \; . \tag{6.17}$$

Remark. *As discussed in Section 6.3.4, the assumption $\mathfrak{D}^{\epsilon_2}_{\max}(\sigma) > 0$ in case (iii.a) guarantees that also $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) > 0$, so that the denominator appearing in the right-hand side of* (6.16) *is strictly greater than zero. Also, since $\mathfrak{D}^{\epsilon_1}_h(\rho) \geqslant -\log(1-\epsilon_1)$ independently of ρ, the parameter ϵ_1 can be modulated so to compensate, to some extent, eventual lack of resource in the initial state.*

Condition (6.17) is stronger than condition (6.16), in the sense that if the former is satisfied, the latter is also satisfied. This is because, by multiplying both sides by $2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1 > 0$ (see preceding remark), condition (6.16) becomes

$$2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) - \max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)} - 2^{-\max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)} \geqslant 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) - \min_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)} - 1 \; ,$$

and this, if $\max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega) = \min_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)$, becomes equivalent to

$$2^{-\max_{\omega\in\mathcal{F}} D^{\epsilon_1}_h(\rho\|\omega)} \leqslant 1 \; ,$$

which is always trivially satisfied, due to the non-negativity of the hypothesis testing relative entropy. In other words, we have shown the following:

Corollary 3. *Given a state $\rho \in \mathcal{D}(\mathbb{C}^m)$, suppose that $\max_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega) = \min_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega)$. Then, for any σ,*

$$\mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max}^{\epsilon_2}(\sigma) \qquad \Longrightarrow \qquad \rho \succ_{(\epsilon_1+\epsilon_2)} \sigma \ .$$

Corollary 3, for rank-one ρ and $\epsilon_1 = 0$, recovers Theorem 2 in Ref. [31].

Proof of Theorem 8. Case (i) is easily proved as follows. The condition $\mathfrak{D}_h^{\epsilon_1}(\rho) = +\infty$ guarantees the existence of an operator $P \in \mathsf{P}^{\epsilon_1}(\rho)$ such that $\mathrm{Tr}\{P\,\omega\} = 0$ for all $\omega \in \mathcal{F}$. Hence, by constructing a CPTP map as follows:

$$\mathcal{E}(\cdot) := \mathrm{Tr}\{P\,\cdot\}\,\sigma + \mathrm{Tr}\{(\mathbb{1} - P)\,\cdot\}\,\varphi \ ,$$

where φ is an arbitrarily fixed element of $\mathcal{F}$, we see that $\mathcal{E}$ maps all free states to φ, so that $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{F}$, while $\|\mathcal{E}(\rho) - \sigma\|_1 \leqslant 2(1 - \mathrm{Tr}\{P\rho\}) \leqslant 2\epsilon_1$.

Case (ii) follows trivially from the fact that condition $\mathfrak{D}_{\max}^{\epsilon_2}(\sigma) = 0$ guarantees the existence of at least one free state which is ϵ_2-close to σ. Hence, the sought resource morphism is trivially given by the CPTP map that prepares any one such states.

Now we move on to case (iii). We begin by looking at condition (6.15), which is the same in both (iii.a) and (iii.b), and rewrite it as follows

$$-\log \mathrm{Tr}\{P^*\,\omega^*\} \geqslant D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+) \ , \tag{6.18}$$

where

- the operators $P^* \in \mathcal{P}^{\epsilon_1}(\rho)$ and $\omega^* \in \mathcal{F}$ are chosen to satisfy:

$$\begin{aligned} \mathrm{Tr}\{P^*\,\omega^*\} &= 2^{-\mathfrak{D}_h^{\epsilon_1}(\rho)} && (6.19)\\ &:= \max_{\omega\in\mathcal{F}} \min_{P\in\mathsf{P}^{\epsilon_1}(\rho)} \mathrm{Tr}\{P\,\omega\} && (6.20)\\ &= \min_{P\in\mathsf{P}^{\epsilon_1}(\rho)} \max_{\omega\in\mathcal{F}} \mathrm{Tr}\{P\,\omega\} \ ; && (6.21) \end{aligned}$$

  the equality in the third line follows from the minimax theorem, for example, in Kakutani's formulation [36, 37], whose hypotheses are satisfied since both optimizations range over convex sets and the functional to be optimized is linear, and hence both convex and concave, in its arguments;

- the operators $\tilde{\sigma} \in \mathcal{D}(\mathbb{C}^m)$ and $\tilde{\sigma}_+ \in \mathcal{F}$ are chosen so to satisfy:

$$\begin{aligned} D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+) &= \mathfrak{D}_{\max}^{\epsilon_2}(\sigma) && (6.22)\\ &:= \min_{\omega\in\mathcal{F}} \min_{\sigma'\in\mathsf{B}^{\epsilon}(\sigma)} D_{\max}(\sigma'\|\omega) && (6.23)\\ &= \min_{\sigma'\in\mathsf{B}^{\epsilon}(\sigma)} \min_{\omega\in\mathcal{F}} D_{\max}(\sigma'\|\omega) && (6.24)\\ &= \mathfrak{D}_{\max}(\tilde{\sigma}) \ , && (6.25) \end{aligned}$$

  that is, $\tilde{\sigma}_+$ achieves the generalized free fraction for $\tilde{\sigma}$ as in Eq. (6.12), namely:

$$\tilde{\sigma}_+ = \mathfrak{F}_g(\tilde{\sigma})\tilde{\sigma} + (1 - \mathfrak{F}_g(\tilde{\sigma}))\tilde{\sigma}_0 \ . \tag{6.26}$$

In Ref. [38], condition (6.15) alone is shown to be sufficient for the existence of a test-and-prepare CPTP linear map $\mathcal{E}$ such that $\|\mathcal{E}(\rho) - \sigma\|_1 \leqslant 2(\epsilon_1 + \epsilon_2)$ and $\mathcal{E}(\omega^*) = \tilde{\sigma}_+$. Such a map is explicitly given as follows:

$$\mathcal{E}(\cdot) = \mathrm{Tr}\{P^* \cdot\}\, \tilde{\sigma} + \mathrm{Tr}\{(\mathbb{1} - P^*) \cdot\}\, \frac{M\tilde{\sigma}_+ - \tilde{\sigma}}{M-1}\ , \tag{6.27}$$

where, for convenience of notation, we have put $M := 1/\,\mathrm{Tr}\{P^*\,\omega^*\} = 2^{\mathfrak{D}_h^{\epsilon_1}(\rho)}$. Without loss of generality, we can assume that $1 < M < +\infty$ for the following reasons. First of all, notice that the assumption $\mathfrak{D}_h^{\epsilon_1}(\rho) < +\infty$ implies $M < +\infty$. Moreover, we can also assume that $\mathrm{Tr}\{P^*\,\omega^*\} < 1$, that is $M > 1$, otherwise $\mathfrak{D}_h^{\epsilon_1}(\rho) = 0$ and, by (6.15), $\mathfrak{D}_{\max}^{\epsilon_2}(\sigma) = 0$, thus making the situation trivial.

As shown in [38], the above map is CPTP; in order to show that it is a resource morphism, we only need to show that $\mathcal{E}(\mathcal{F}) \subseteq \mathcal{F}$. To this end, let us assume that the input to $\mathcal{E}$ is an arbitrary $\varphi \in \mathcal{F}$. We need to show that $\mathcal{E}(\varphi) \in \mathcal{F}$. By arranging terms, we obtain,

$$\mathcal{E}(\varphi) = \left(1 - \frac{1 - \mathrm{Tr}\{P^*\,\varphi\}}{1 - \mathrm{Tr}\{P^*\,\omega^*\}}\right)\tilde{\sigma} + \left(\frac{1 - \mathrm{Tr}\{P^*\,\varphi\}}{1 - \mathrm{Tr}\{P^*\,\omega^*\}}\right)\tilde{\sigma}_+\ . \tag{6.28}$$

Again for convenience of notation, let us put $t := \frac{1-\mathrm{Tr}\{P^*\,\varphi\}}{1-\mathrm{Tr}\{P^*\,\omega^*\}}$ and $R := \mathfrak{F}_g(\tilde{\sigma})$. We now recall the optimal decomposition (6.26): by inserting it into (6.28) and rearranging terms once more, we arrive at

$$\mathcal{E}(\varphi) = (1 - t + tR)\tilde{\sigma} + (t - tR)\tilde{\sigma}_0\ . \tag{6.29}$$

The above relation tells us that $\mathcal{E}(\varphi)$ lies somewhere on the affine line passing through both $\tilde{\sigma}$ and $\tilde{\sigma}_0$. Therefore, in order to have $\mathcal{E}(\varphi) \in \mathcal{F}$, the coefficient $(1 - t + tR)$ weighing $\tilde{\sigma}$ must be carefully bounded both from above and from below, so that $\mathcal{E}(\varphi)$ is neither too close to $\tilde{\sigma}$ nor too close to $\tilde{\sigma}_0$, in which case it could end up lying outside $\mathcal{F}$ (see Fig. 6.2 for a schematic picture).

The upper bound is computed as follows. Since the free fraction is exactly defined as the *maximum* weight of $\tilde{\sigma}$ so that a convex mixture with $\tilde{\sigma}_0$ lies in $\mathcal{F}$, we want to show that the weight of $\tilde{\sigma}$ in (6.29) does not exceed R, that is,

$$1 - t + tR \leqslant R\ ,$$

or, equivalently,

$$1 - t \leqslant (1 - t)R\ . \tag{6.30}$$

Since, starting from Eq. (6.21),

$$\begin{aligned}\mathrm{Tr}\{P^*\,\omega^*\} &= \min_{P \in \mathsf{P}^{\epsilon_1}(\rho)} \max_{\omega \in \mathcal{F}} \mathrm{Tr}\{P\,\omega\} \\ &= \max_{\omega \in \mathcal{F}} \mathrm{Tr}\{P^*\,\omega\} \\ &\geqslant \mathrm{Tr}\{P^*\,\varphi\} \geqslant 0\ ,\end{aligned}$$

we see that $t \geqslant 1$, that is, $1 - t \leqslant 0$, and inequality (6.30) automatically holds for any $R \in [0, 1]$, without the need to invoke any extra condition.

Hence, condition (6.16) or condition (6.17) are only required to obtain the correct lower bound, that is, to prevent that $\mathcal{E}(\varphi)$ crosses the border of $\mathcal{F}$ when approaching $\tilde{\sigma}_0$. In order to derive the lower bound, we resort to the construction introduced in Eq. (6.13) and depicted in Fig. 6.2. Once a decomposition achieving the generalized free fraction of $\tilde{\sigma}$ is found, $\tilde{\sigma}_-$ is the state on the boundary of $\mathcal{F}$, which is "antipodal" with respect to $\tilde{\sigma}_+$. If we get past it, we end up outside $\mathcal{F}$: we need to make sure this does not happen.

We begin by assuming that $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) < +\infty$, that is, $\tilde{\sigma}_+ \neq \tilde{\sigma}_-$. (We recall that $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) > 0$ is a consequence of the assumption $\mathfrak{D}^{\epsilon_2}_{\max}(\sigma) > 0$.) In this case, we need to impose that

$$1 - t + tR \geqslant \frac{R\, 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1} \,. \tag{6.31}$$

Before proceeding, we notice that the above inequality, if satisfied, implies in particular $1 - t + tR \geqslant 0$, because $R\, 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} = 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) - D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+)} \geqslant 1$.

Condition (6.31), after writing t explicitly again, reads as follows:

$$1 - \frac{1 - \mathrm{Tr}\{P^*\,\varphi\}}{1 - \mathrm{Tr}\{P^*\,\omega^*\}} + R\left(\frac{1 - \mathrm{Tr}\{P^*\,\varphi\}}{1 - \mathrm{Tr}\{P^*\,\omega^*\}}\right) \geqslant \frac{R\, 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1} \,.$$

Since we are assuming that $\mathrm{Tr}\{P^*\,\omega^*\} < 1$, multiplying both sides by $1 - \mathrm{Tr}\{P^*\,\omega^*\}$ does not change the inequality, so we obtain the equivalent condition:

$$(1 - R)\,\mathrm{Tr}\{P^*\,\varphi\} \geqslant \frac{R\, 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}(1 - \mathrm{Tr}\{P^*\,\omega^*\}) + \mathrm{Tr}\{P^*\,\omega^*\} - R \,.$$

After rearranging the right-hand side, we arrive at

$$(1 - R)\,\mathrm{Tr}\{P^*\,\varphi\} \geqslant (1 - R)\frac{\mathrm{Tr}\{P^*\,\omega^*\}\, 2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}{2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1} \,.$$

Since $R < 1$ (because we assumed that $\tilde{\sigma} \notin \mathcal{F}$, that is, $\mathfrak{D}_{\max}(\tilde{\sigma}) > 0$), we can divide both sides by $(1 - R)$ and obtain

$$\mathrm{Tr}\{P^*\,\varphi\} \geqslant \frac{\frac{1}{M} 2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1}{2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)} - 1} \,. \tag{6.32}$$

The above condition must be satisfied for any $\varphi \in \mathcal{F}$. Hence, what we really want is a lower bound on $\min_{\varphi\in\mathcal{F}} \mathrm{Tr}\{P^*\,\varphi\}$. Noticing that

$$\min_{\varphi\in\mathcal{F}} \mathrm{Tr}\{P^*\,\varphi\} \geqslant \min_{P\in\mathsf{P}^{\epsilon_1}(\rho)} \min_{\varphi\in\mathcal{F}} \mathrm{Tr}\{P\,\varphi\} \tag{6.33}$$

$$= \min_{\omega\in\mathcal{F}} \min_{P\in\mathsf{P}^{\epsilon_1}(\rho)} \mathrm{Tr}\{P\,\omega\} \tag{6.34}$$

$$= 2^{-\max_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega)} \,, \tag{6.35}$$

condition (6.32) holds whenever the following, stricter condition holds, that is,

$$2^{-\max_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega)} \geqslant \frac{\frac{1}{M}2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)}-1}{2^{D_{\max,F}(\tilde{\sigma}\|\tilde{\sigma}_+)}-1} = \frac{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)-\min_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega)}-1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)}-1}\ .$$

Hence, condition (6.16) guarantees that $\mathcal{E}(\varphi)\in\mathcal{F}$ for any $\varphi\in\mathcal{F}$, that is, that the operation $\mathcal{E}$ defined in (6.27) is a valid resource morphism.

Let us finally consider the case in which $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)=+\infty$, that is, $\tilde{\sigma}_+=\tilde{\sigma}_-$. In this case, lower and upper bounds have to coincide, so that the map defined in (6.27) is a resource morphism if and only if $1-t+tR=R$. This can only happen if $R=1$ (but this is excluded because $\tilde{\sigma}\notin\mathcal{F}$) or if $1-t=0$, that is, if $t=1$ independently of the input $\varphi\in\mathcal{F}$. This is guaranteed if the operator P^* in (6.27) has the same trace on all free states, which is exactly the content of (6.17). ∎

A less general, but simpler, statement stemming from Theorem 8 is the following:

Corollary 4. *With the same notations of Theorem 8, then the following statement holds,*

$$\mathfrak{D}_h^{\epsilon_1}(\rho)\geqslant D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) \implies \rho\succ_{\epsilon_1+\epsilon_2}\sigma\ . \tag{6.36}$$

Proof. Assuming (6.36), if $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)=+\infty$, then also $\mathfrak{D}_h^{\epsilon_1}(\rho)=+\infty$. In such a case, we know from Theorem 8, case (i), that $\rho\succeq_{\epsilon_1}\sigma$, which of course implies also $\rho\succeq_{(\epsilon_1+\epsilon_2)}\sigma$.

On the other hand, if $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)=0$, we know that $\tilde{\sigma}\in\mathcal{F}$, so that, in fact, $\mathfrak{D}_{\max}^{\epsilon_2}(\sigma)=0$. In other words, we are in case (ii) of Theorem 8, and again $\rho\succeq_{(\epsilon_1+\epsilon_2)}\sigma$ holds.

We are hence left to consider the case

$$+\infty>\mathfrak{D}_h^{\epsilon_1}(\rho)\geqslant D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)>0. \tag{6.37}$$

We show that condition (6.37) alone implies both conditions (6.15) and (6.16) of case (iii.a) in Theorem 8.

Since by definition $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)\geqslant D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+)=\mathfrak{D}_{\max}^{\epsilon_2}(\sigma)$, we immediately see that condition (6.37) implies condition (6.15). Hence, we only need to show that also condition (6.16) is implied. In fact, we can show that (6.37) implies a condition that is even stronger than (6.16). Such a condition is the following:

$$0\geqslant\frac{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)-\mathfrak{D}_h^{\epsilon_1}(\rho)}-1}{2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)}-1}\ .$$

If the above is satisfied, also (6.16) is satisfied, and we can conclude that $\rho\succeq_{(\epsilon_1+\epsilon_2)}\sigma$. The above inequality is satisfied because, as a consequence of $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+)>$

0, the denominator in the right-hand side is strictly positive, so that the above inequality is equivalent to

$$1 \geqslant 2^{D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) - \mathfrak{D}_h^{\epsilon_1}(\rho)} \ ,$$

which is satisfied if and only if condition (6.37) is satisfied. ∎

A merit of Corollary 4 is to provide a simple compact sufficient condition, free of supplementary *caveat* like condition (6.16), which is difficult to interpret operationally. However, the right-hand side of (6.36) is not yet a valid resource monotone. The following result fills the gap.

Theorem 9. *Given $\rho, \sigma \in \mathcal{D}(\mathbb{C}^m)$ and $\epsilon_1, \epsilon_2 \in [0,1]$, the following relation holds,*

$$\textit{if} \quad \mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma) \quad \textit{holds, then} \quad \rho \succeq_{(\epsilon_1+\epsilon_2)} \sigma \ . \tag{6.38}$$

Theorem 9, when $\epsilon_2 = 0$ and σ is rank-one, recovers Theorem 5 in Ref. [31] (see also Corollary 17 of [39]). Theorem 8 and Theorem 9 are independent of each other. This is because, on the one hand, it is possible that $\mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max}^{\epsilon_2}(\sigma)$ even though $\mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma)$, so that Theorem 9 would be inconclusive. On the other hand, it is possible that $\mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma)$ even though neither condition (6.16) nor (6.17) hold, so that Theorem 8 would be inconclusive. In other words, Theorem 8 and Theorem 9 in general apply to two different regimes and are logically independent of each other. Nonetheless, since $\tilde{\sigma} \in \mathcal{B}^{\epsilon_2}(\sigma)$ and $\tilde{\sigma}_+ \in \mathcal{F}$, we see that $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) \geqslant \mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma)$. This implies that Corollary 4 above can be as well derived as a consequence of Theorem 9.

Proof. We begin by noticing that, if $\mathfrak{D}_h^{\epsilon_1}(\rho) = +\infty$, we are back to case (i) of Theorem 8. Also, if $\mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma) = 0$, then also $\mathfrak{D}_{\max}^{\epsilon_2}(\sigma) = 0$, and we are back to case (ii) of Theorem 8. In what follows we will hence assume that $+\infty > \mathfrak{D}_h^{\epsilon_1}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma) > 0$.

Let us define $P^*, \omega^*, \tilde{\sigma}, \tilde{\sigma}_+$ as the optimizers achieving the quantities that appear in condition (6.38), that is,

$$\mathfrak{D}_h^{\epsilon_1}(\rho) := \min_{\omega \in \mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega) = D_h^{\epsilon_1}(\rho\|\omega^*) = -\log \mathrm{Tr}\{P^* \, \omega^*\} \tag{6.39}$$

$$\mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma) := \min_{\omega \in \mathcal{F}} D_{\max,\mathcal{F}}^{\epsilon_2}(\sigma\|\omega) = D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) \ . \tag{6.40}$$

Notice that while $\tilde{\sigma}, \tilde{\sigma}_+$ were used in Theorem 8 to denote the optimizers achieving $\mathfrak{D}_{\max}^{\epsilon_2}(\sigma)$, for the sake of this proof the same symbols are used to denote the optimizers achieving $\mathfrak{D}_{\max,\mathcal{F}}^{\epsilon_2}(\sigma)$.

Writing $M := 1/\,\mathrm{Tr}\{P^* \, \omega^*\}$, that is,

$$\frac{1}{M} = \mathrm{Tr}\{P^* \, \omega^*\} = \max_{\omega \in \mathcal{F}} \min_{P \in \mathsf{P}^{\epsilon_1}(\rho)} \mathrm{Tr}\{P \, \omega\} \ ,$$

we define the map

$$\mathcal{E}(\cdot) = \mathrm{Tr}\{P^* \cdot\}\, \tilde{\sigma} + (1 - \mathrm{Tr}\{P^* \cdot\}) \frac{M\tilde{\sigma}_+ - \tilde{\sigma}}{M-1} \ . \tag{6.41}$$

Notice that, with respect to the map constructed in (6.27), the above map uses the same operator P^*, but prepares different states depending on the outcome. As before, moreover, it is possible to assume without loss of generality that $1 < M < +\infty$.

Since $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) \geqslant D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+)$, condition (6.38) implies that,

$$D_h^{\epsilon_1}(\rho\|\omega^*) \geqslant D_{\max}(\tilde{\sigma}\|\tilde{\sigma}_+) \ , \tag{6.42}$$

A direct consequence of [38] is that condition (6.42), together with the fact that $\tilde{\sigma} \in \mathsf{B}^{\epsilon_2}(\sigma)$, imply that the map $\mathcal{E}$ defined in (6.41) is a valid CPTP map such that $\frac{1}{2}\left\|\mathcal{E}(\rho) - \sigma\right\|_1 \leqslant \epsilon_1 + \epsilon_2$. In what follows we show that $\mathcal{E}$ is, in particular, a resource morphism.

Because $\tilde{\sigma}$ and $\tilde{\sigma}_+$ have been chosen as the states that optimize $\mathfrak{D}^{\epsilon_2}_{\max,\mathcal{F}}(\sigma)$, we have $D_{\max,\mathcal{F}}(\tilde{\sigma}\|\tilde{\sigma}_+) = \mathfrak{D}_{\max,\mathcal{F}}(\tilde{\sigma}) = -\log \mathfrak{F}(\tilde{\sigma})$. Therefore, we obtain the following decomposition of σ_+,

$$\tilde{\sigma}_+ = \mathfrak{F}(\tilde{\sigma})\tilde{\sigma} + (1 - \mathfrak{F}(\tilde{\sigma}))\tilde{\sigma}_0 \ , \tag{6.43}$$

with $\tilde{\sigma}_0 \in \mathcal{F}$. By plugging (6.43) in (6.41), and considering as input to the map an arbitrary free state $\varphi \in \mathcal{F}$, we reach the following

$$\mathcal{E}(\varphi) = (1 - t + tR)\tilde{\sigma} + (t - tR)\tilde{\sigma}_0 \ , \tag{6.44}$$

where, for the sake of notation, we put $t := \frac{1-\mathrm{Tr}\{P^*\ \varphi\}}{1-\mathrm{Tr}\{P^*\ \omega^*\}}$ and $R := \mathfrak{F}(\sigma)$. Notice that while the proof of Theorem 8 is obtained by working with the *generalized* free fraction, in this proof we are mostly working with the free fraction.

We need to show that $\mathcal{E}(\varphi) \in \mathcal{F}$, for all $\varphi \in \mathcal{F}$. To that end, we only need to show that the weight in front of $\tilde{\sigma}$ in (6.44) is non-negative and upper bounded by R.

In order to show that it does not exceed R, we proceed as follows. In the proof of Theorem 8, we have shown that $t \geqslant 1$, so that $1 - t + tR \leqslant R$, that is, $R(t-1) \leqslant t - 1$, holds automatically for any $R \in [0, 1]$.

In order to show the weight of $\tilde{\sigma}$ is non-negative, it suffices to show that

$$R \geqslant 1 - \frac{1}{t} \ .$$

Since $t \leqslant \frac{1}{1-\mathrm{Tr}\{P^*\ \omega^*\}} = \frac{M}{M-1}$, we have that $1 - t^{-1} \leqslant 1 - \frac{M-1}{M}$, so that the above is satisfied whenever

$$R \geqslant \frac{1}{M} = \mathrm{Tr}\{P^*\ \omega^*\} \ ,$$

that is to say

$$\mathfrak{F}(\tilde{\sigma}) = 2^{-\mathfrak{D}^{\epsilon_2}_{\max,\mathcal{F}}(\sigma)} \geqslant \mathrm{Tr}\{P^*\ \omega^*\} = 2^{-\mathfrak{D}^{\epsilon_1}_h(\rho)} \ .$$

∎

6.5 Applications and examples

In this section we apply Theorems 8 and 9 to some situations of physical interest, and show how we can not only rederive, but sometimes also strengthen, previous results.

6.5.1 Singleton Resource Theories

We begin this section by considering the special case of singleton resource theories, in which the set of free states $\mathcal{F}$ comprises only one element. This scenario includes the resource theory of athermality, namely, the case in which free operations are those that preserve the thermal state of the system [25, 40, 41, 42], which in turns provide the backbone of the resource theory of quantum thermodynamics [43]. More generally, when the output singleton is allowed to differ from the input one, this is referred to as the resource theory of asymmetric distinguishability, whose optimal rates have been studied in [38, 44, 45].

In the singleton case, the log-robustness typically is infinite, and the applicability of Theorem 9 is quite limited. On the contrary, Theorem 8 can still be useful, even in the case of a singleton $\mathcal{F}$. Indeed, Theorem 8 reduces in the singleton case to Lemma 3.3 of [38], which is good enough to serve as the starting point to study optimal asymptotic interconversion rates.

Proposition 5. *Consider an input system, with initial state $\rho \in \mathcal{D}(\mathbb{C}^m)$ and free singleton $\mathcal{F} = \{\gamma\}$, and an output system, with target state $\sigma \in \mathcal{D}(\mathbb{C}^n)$ and free singleton $\mathcal{F}' = \{\gamma'\}$. we have,*

$$D_h^{\epsilon_1}(\rho\|\gamma) \geqslant D_{\max}^{\epsilon_2}(\sigma\|\gamma') \qquad \Longrightarrow \qquad \rho \succeq_{(\epsilon_1+\epsilon_2)} \sigma \ . \tag{6.45}$$

Proof. We can restrict ourselves to consider only case (iii.b) of Theorem 8, because for a singleton $\mathcal{F}' = \{\gamma'\}$, whenever $\sigma \neq \gamma'$, one has $D_{\max,\mathcal{F}'}(\sigma\|\gamma') = +\infty$. But since also the input free set $\mathcal{F}$ is a singleton, we have

$$\min_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega) = \max_{\omega\in\mathcal{F}} D_h^{\epsilon_1}(\rho\|\omega) \ ,$$

and condition (6.17) is automatically satisfied. ∎

6.5.2 Resource Theory of Bipartite Entanglement

Next, we specialize our results to the resource theory of entanglement. We begin by considering bipartite entanglement, namely, the case in which $\mathcal{F}$ is the set of all separable states of a given bipartite system. Resource morphisms are given by separability-preserving (or non-entangling) CPTP maps, usually denoted as SEPP. One-shot entanglement distillation and dilution under SEPP have been studied in [35]. In what follows we show how our Corollary 4 is able to guarantee the existence of a SEPP transition directly mapping ρ to σ, even in situations in which the

results of Ref. [35] cannot guarantee the existence of a "distill-and-dilute" transition.

In order to illustrate the point, it is enough to consider the exact case, that is, $\epsilon_1 = \epsilon_2 = 0$. The same conclusions hold also in the approximate case, however, some care must be taken in that while here we use the trace-distance to measure approximations, Ref. [35] uses the fidelity. Trace-distance and fidelity are well-known to be equivalent [26, 27] , but approximation parameters must be changed: we leave it to the interested reader to work out the exact factors.

By rewriting the main results of [35] using our notation, the zero-error one-shot SEPP-distillable entanglement $E^{(1)}_{D,\mathrm{SEPP}}(\rho)$ and the zero-error one-shot SEPP-entanglement cost $E^{(1)}_{C,\mathrm{SEPP}}(\sigma)$ satisfy

$$E^{(1)}_{D,\mathrm{SEPP}}(\rho) \geqslant \lfloor \mathfrak{D}_{\min}(\rho) \rfloor \tag{6.46}$$

and

$$E^{(1)}_{C,\mathrm{SEPP}}(\sigma) \leqslant \mathfrak{D}_{\max,\mathcal{F}}(\sigma) + 1 \ , \tag{6.47}$$

respectively. These two relations together guarantee that it is possible to exactly go from ρ to σ via SEPP (passing through the maximally entangled state) if

$$\lfloor \mathfrak{D}_{\min}(\rho) \rfloor \geqslant \mathfrak{D}_{\max,\mathcal{F}}(\sigma) + 1 \ ,$$

which is more restrictive than what Theorem 9 says, that is,

$$\mathfrak{D}_{\min}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}(\sigma) \ .$$

This is possible because we do not require the transformation to pass through the maximally entangled state, but we allow it to go directly from ρ to σ.

Remark. *When working within the resource theory of entanglement, especially in the one-shot regime, it is customary to allow the output system to differ from the input one. Consequently, also the set of free states changes from $\mathcal{F}$ to $\mathcal{F}'$. As already noticed, our bounds can be straightforwardly extended to cover this situation as well: in such a case, all quantities related to the input state ρ will be computed with respect to the input free set $\mathcal{F}$, while all quantities related to the output state σ will be computed with respect to the output free set $\mathcal{F}'$.*

6.5.3 Existence of a Maximally Resourceful State and Weak-Converse Bounds for Distillation and Dilution

In this section we show how Corollary 3 and Theorem 9 can be used to formulate sufficient conditions that guarantee that an element α is maximally resourceful, in the sense of Definition 6.3.2. We also address the related problem of deciding when Corollary 3 and Theorem 9 are optimal, i.e., when the sufficient conditions they formulate become also necessary. For the sake of the presentation, we focus

here on the case of exact transitions, that is, $\epsilon_1 = \epsilon_2 = 0$, keeping in mind, however, that the main results allow us to go beyond the exact case and to speak of, e.g., almost-maximally resourceful elements.

We begin with the following fact (see also Corollary 4 in [31]):

Proposition 6. *The following statements hold.*

1. *Let $\alpha \in \mathcal{D}(\mathbb{C}^d)$ be such that $\mathfrak{D}_{\min}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max}(\rho)$, and $\mathrm{Tr}\{\omega\,\Pi_\alpha\} = \mathrm{constant}$, for any $\omega \in \mathcal{F}$. Then α is maximally resourceful in $\mathcal{D}(\mathbb{C}^d)$.*

2. *Let $\alpha \in \mathcal{D}(\mathbb{C}^d)$ be such that $\mathfrak{D}_{\min}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max,\mathcal{F}}(\rho)$. Then α is maximally resourceful in $\mathcal{D}(\mathbb{C}^d)$.*

Proof. Case (1): being $\mathrm{Tr}\{\omega\,\Pi_\alpha\}$ constant for any $\omega \in \mathcal{F}$, the assumptions in Corollary 3 are satisfied with $\epsilon_1 = \epsilon_2 = 0$. The proof then follows trivially, from the assumption that $\mathfrak{D}_{\min}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max}(\rho) \geqslant \mathfrak{D}_{\max}(\sigma)$ for any $\sigma \in \mathcal{D}(\mathbb{C}^d)$.

Case (2): in this case we apply Theorem 9, and again the proof follows trivially, from the assumption that $\mathfrak{D}_{\min}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max,\mathcal{F}}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}}(\sigma)$. ∎

Remark. *Since, for any ρ, σ, $D_{\min}(\rho\|\sigma) \leqslant D_{\max}(\rho\|\sigma) \leqslant D_{\max,\mathcal{F}}(\rho\|\sigma)$, condition (1) of Case (1) in Proposition 6 above implies that $\mathfrak{D}_{\min}(\alpha) = \mathfrak{D}_{\max}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\min}(\rho) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max}(\rho)$; analogously, condition (2) of Case (2) implies $\mathfrak{D}_{\min}(\alpha) = \mathfrak{D}_{\max,\mathcal{F}}(\alpha) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\min}(\rho) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max,\mathcal{F}}(\rho)$.*

The two sufficient conditions considered in Proposition 6 are independent. For example, both in the resource theory of bipartite entanglement and in the resource theory of coherence a golden state exists, namely, the maximally entangled state and the maximally coherent state, respectively. It is also known that these are both in fact maximally resourceful in their respective theories. However, while the maximally coherent state satisfies condition (1) but not condition (2), the maximally entangled state satisfies condition (2) but not (1): see [39] for the explicit calculation.

Another example is provided by the resource theory of genuine multipartite entanglement, in which the free set is taken to be the set of all biseparable states and resource morphisms correspondingly are defined as biseparability-preserving maps. In this case, it is possible to show by explicit calculation [39, 46] that the generalized GHZ state, that is,

$$|\Psi_{\mathsf{GHZ}}^{(N,d)}\rangle := \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle^{\otimes N} \ ,$$

satisfies condition (2) of Proposition 6. We conclude, therefore, that $|\Psi_{\mathsf{GHZ}}^{(N,d)}\rangle$ is maximally resourceful.

The following propositions provide sufficient conditions so that the bounds in Corollary 3 and Theorem 9 are optimal. In the following proposition, we make it explicit that the input system (with state space $\mathcal{D}(\mathbb{C}^m)$ and free set $\mathcal{F}$) in general may differ from the output system (with state space $\mathcal{D}(\mathbb{C}^n)$ and free set $\mathcal{F}'$). A related result is Theorem 2 of Ref. [31].

Proposition 7 (Weak-converse bounds for dilution)**.** *When dealing with transitions from an input system $(\mathbb{C}^m, \mathcal{F})$ to an output system $(\mathbb{C}^n, \mathcal{F}')$, the following statements hold.*

1. *Suppose that $\alpha \in \mathcal{D}(\mathbb{C}^m)$ satisfies $\mathfrak{D}_{\min}(\alpha) = \mathfrak{D}_{\max}(\alpha)$; then, for any $\sigma \in \mathcal{D}(\mathbb{C}^n)$*

$$\alpha \succeq \sigma \qquad \Longrightarrow \qquad \mathfrak{D}_{\min}(\alpha) \geqslant \mathfrak{D}_{\max}(\sigma) \; . \tag{6.48}$$

2. *Suppose that $\alpha \in \mathcal{D}(\mathbb{C}^m)$ satisfies $\mathfrak{D}_{\min}(\alpha) = \mathfrak{D}_{\max,\mathcal{F}}(\alpha)$; then, for any $\sigma \in \mathcal{D}(\mathbb{C}^n)$*

$$\alpha \succeq \sigma \qquad \Longrightarrow \qquad \mathfrak{D}_{\min}(\alpha) \geqslant \mathfrak{D}_{\max,\mathcal{F}'}(\sigma) \; . \tag{6.49}$$

Proof. Case (1): suppose that $\alpha \succeq \sigma$, so that there exists a resource morphism $\mathcal{E} : \mathcal{D}(\mathbb{C}^m) \to \mathcal{D}(\mathbb{C}^n)$ such that $\mathcal{E}(\alpha) = \sigma$; then,

$$\begin{aligned} \mathfrak{D}_{\min}(\alpha) &= \mathfrak{D}_{\max}(\alpha) \\ &\geqslant \mathfrak{D}_{\max}(\mathcal{E}(\alpha)) \\ &= \mathfrak{D}_{\max}(\sigma) \; , \end{aligned}$$

where the inequality in the second line comes from the fact that $\mathfrak{D}_{\max}$ is a resource monotone.

Case (2): suppose that $\alpha \succeq \sigma$, then

$$\begin{aligned} \mathfrak{D}_{\min}(\alpha) &= \mathfrak{D}_{\max,\mathcal{F}}(\alpha) \\ &\geqslant \mathfrak{D}_{\max,\mathcal{F}'}(\mathcal{E}(\alpha)) \\ &= \mathfrak{D}_{\max,\mathcal{F}'}(\sigma) \; , \end{aligned}$$

where the inequality in the second line comes from the fact that $\mathfrak{D}_{\max,\mathcal{F}}$ is a resource monotone. ∎

An analogous weak converse for distillation is the following (see also Theorem 5 of [31] for a related result).

Proposition 8 (Weak-converse bound for distillation)**.** *Consider an input system $(\mathbb{C}^m, \mathcal{F})$ and an output system $(\mathbb{C}^n, \mathcal{F}')$, and let $\alpha \in \mathcal{D}(\mathbb{C}^n)$ be a target state such that $\mathfrak{D}_{\max,\mathcal{F}'}(\alpha) = \mathfrak{D}_{\min}(\alpha)$. Then, for any $\rho \in \mathcal{D}(\mathbb{C}^m)$,*

$$\rho \succeq \alpha \qquad \Longrightarrow \qquad \mathfrak{D}_{\min}(\rho) \geqslant \mathfrak{D}_{\max,\mathcal{F}'}(\alpha) \; . \tag{6.50}$$

Proof. If $\rho \succeq \alpha$ then $\mathfrak{D}_{\min}(\rho) \geqslant \mathfrak{D}_{\min}(\mathcal{E}(\rho)) = \mathfrak{D}_{\min}(\alpha) = \mathfrak{D}_{\max,\mathcal{F}'}(\alpha)$. ∎

Remark. *By looking at the proofs of Theorem 8 and Theorem 9, we see that the resource morphisms used there have been constructed as test-and-prepare quantum channels. As a consequence, Propositions 7 and 8 above can be interpreted as giving sufficient conditions for which test-and-prepare channels are provably optimal in resource manipulation, despite constituting a very special class among all CPTP maps.*

A natural question to ask, at this point, is whether density matrices always exist, for which Propositions 7 and 8 hold, namely, for which test-and-prepare channels provide the optimal resource morphisms. As it turns out, perhaps surprisingly, in any resource theory with non-empty closed and convex $\mathcal{F}$, even if a maximally resourceful element may not exist, a *golden state*, namely, a rank-one density matrix $\Psi_+ \in \mathcal{D}(\mathbb{C}^d)$ such that $\max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\min}(\rho) = \mathfrak{D}_{\min}(\Psi_+) = \mathfrak{D}_{\max}(\Psi_+) = \max_{\rho\in\mathcal{D}(\mathbb{C}^d)} \mathfrak{D}_{\max}(\rho)$, can always be found [31, 39]. However, before concluding that test-and-prepare morphisms are optimal for golden states, one still needs to verify that, either Ψ_+ satisfies $\mathrm{Tr}\{\omega\ \Psi_+\} =$ constant for all free ω, or $\mathfrak{D}_{\max}(\Psi_+) = \mathfrak{D}_{\max,\mathcal{F}}(\Psi_+)$ also holds, and both such extra conditions depend on the actual resource theory at hand. The resource theories of coherence and bipartite entanglement again provide two paradigmatic examples in this sense.

6.6 Summary

In this chapter, we formulated a general resource theory with F-morphisms as the restricted transformation and complex matrices as the resource element. The resource element is taken from a closed and convex set with a subset as the null resource set. The core idea from a resource-theoretic viewpoint is that it is possible to quantify resources without referring to the maximal one. Moreover, it is enough to compare the entropic quantity of resources to ensure the existence of F-morphisms between them. With our resource-theoretic frame, it is also possible to derive a sufficient and necessary condition of the transformation between an arbitrary resource and the maximal resource or a typical reference resource, called the golden states in terms of quantum states.

Chapter 7

RWDC in the quantum world

Real-world data circulation (RWDC) is the core idea behind the graduate program for Real World Data Circulation Leaders, one of the leading graduate school programs[1] designed for some of the top Japanese universities. In this chapter, I will demonstrate the concept of RWDC in the quantum world.

7.1 Introduction to RWDC

There are three ingredients of RWDC: Data acquisition, data analysis, and real-world implementation. The concept can be illustrated as in Figure 7.1.

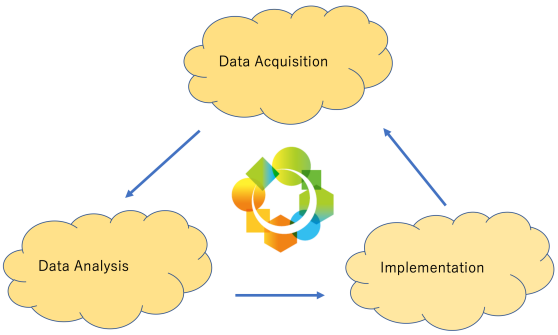

Figure 7.1: RWDC concept

The most important feature of these ingredients is the circulation among them. One of the reasons that behind this circulation is the ability from recent technologies to acquire and store large amount of data. We are living in a world where the amounts of data is growing exponentially rapidly. Many types of technological products, such as smartphones, IoT devices, personal computers, etc., make it easier than before to generate a large amount of data. All these data are potential information to help to speed up the development of those artificial intelligence

[1] Refer to Japan Society for the Promotion of Science: https://www.jsps.go.jp/j-hakasekatei/

technologies, which in return generates more data. Moreover, the growing technology for analyzing various data set also help us find additional value from the newly generated data. This process can become an iteration process if the data analysis remains yielding positive feedbacks. This virtuous circle is crucial to artificial intelligence, product development and other industries. In addition to the applications of the RWDC concept in the field of artificial intelligence, an example of a novel application explains well this concept is in [47], in which the authors apply data analysis techniques to manufacturing metal cutting process. This new method will in turn help to better develop the cutting process.

This concept of RWDC seems to have no relations with the micro-world, i.e., quantum world, being only suitable to address real-world data in some sense. However, all data that we can collect are fundamentally from the quantum world. The following section explains data circulation in the view of the quantum world.

7.2 Quantum measurement as a data acquisition process

In our daily life in which we can feel and see our surroundings, we use measurement extensively. For example, we measure the length of a table through a ruler, the weight of a human body through a weighing-machine, the life span of human beings through time. These measured quantities are what we call the data. In the field of artificial intelligence, the "big data" comes from any device that can produce data in the measuring process. It includes images, audio signals, videos, GPS data, and Lidar data, among others.

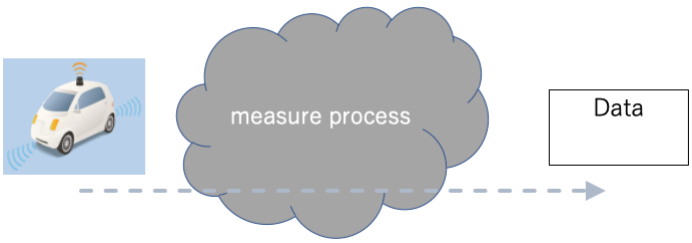

Figure 7.2: Measuring process

Figure 7.2 gives an example to illustrate the process of obtaining data from a measuring process. Image data comes from a camera that requires a sensor to sense the light, while audio signal data comes from a sensor device that can sense the sound. The same process applies to all possible types of data. The only difference is how we process the source that can generate data and what process to use to obtain the data.

In general, we understand that these measurable quantities are fixed values even before the measurement, and we need a device to conduct the measurement

to obtain that quantity. In the quantum world, which is sometimes counterintuitive compared to the macro-world when it comes to measurement, we measure the quantities of atomic-scale particles. For example, we measure the polarization of photons, we measure the momentum of electrons, etc. Research has shown that instead of the deterministic outcomes from measurement in the macro-world, the outcome of measurement in the quantum world is stochastic. Although the interpretation of the whole process of measurement is still under debate, for the moment researchers have come to the common understanding that quantum measurement is indeterministic. Being indeterministic means that the measurement does not output a fixed value; instead, it outputs a probabilistic distribution of outcomes. Counter-intuitively, the measurement outcome might not exist even before the measurement; it only exists when there is a measure device to probe the quantum system. The quantum process is illustrated in Figure 7.3.

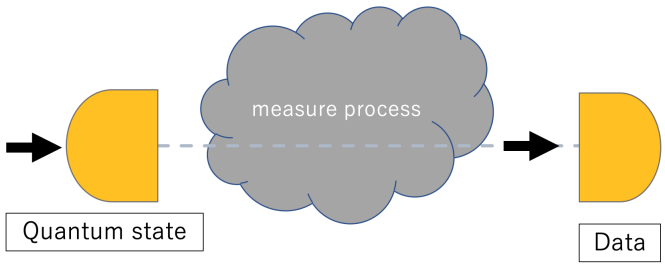

Figure 7.3: Measuring process

In above Figure 7.3, the quantum state represents the quantum system being measured. The data is the outcomes of quantum measurement. Normally, we read out the outcomes of quantum measurements as classical bits that can in return represent any real world data. Since any data that is used for artificial intelligence is just classical bits fundamentally, in this sense, we see the quantum measurement process the same as the process of obtaining data. A good example of using quantum measurement is in the construction of quantum computer's hardware, in which the measurement is the process to read out the bits for programming.

Hence, in the most general viewpoint that combines both classical mechanics and quantum mechanics, we can generalize the measurement process as the process to acquire the data. Hence, we have made it clear that quantum measurement is a particular process to acquire data.

7.3 Statistical analysis as data analysis

Due to the large number of various kinds of data available and the super computing power nowadays to deal with them, we have benefited largely from revealing the hidden value in those big amounts of data through data analysis techniques, such as machine learning. In the previous section, we have discussed how to col-

lect data from the "data acquisition" process from the viewpoint of quantum measurement. In this section, we discuss the process of data analysis and its importance in the quantum world.

Data analysis is closely related to data collection and further collection and analysis will be motivated to form iterations if feedbacks from previous data analysis can always show some greatness. An early example that explains the greatness of data analysis can date back to World War II, during which statistician Abraham Wald introduced the idea of survivorship bias from the statistical analysis of the data that the damaged areas in those survived aircraft that were not shot down during the war. His analysis showed the exact opposite opinion from the aircraft engineers who claimed that those places in the aircraft where the fuselage and wings got more shots needed to be heavily armored for protection. Wald oppositely concluded that the areas with fewer recorded shots needed the most armor due to the idea of survivorship bias. His conclusion helped the army to build stronger aircraft. The most important data is those that answered the question "where do planes that don't come back get shot?". The planes that returned safely had more shots on the areas that can handle more shots. His well-known quote "The most important data is the data you don't have." has been popular.

As the above example showed the statistical analysis of data plays a very important role in unrevealing value from available data. In what follows we show that the statistical analysis of the outcome of quantum measurement also makes a difference in determining if a quantum resource is truly useful or not. Let us recall the resource theory of quantum measurement. The core idea in quantum resource theory of measurement is to answer the following question: if we treat quantum measurement as resources (as discussed in Chapter 4) that are necessarily useful for building quantum technologies, how do we know whether a given resource is genuinely useful or not? This is the main barrier in what is a so-called device-independent test of quantum resources. We do not want to get a resource that we are not sure if it is truly useful or not. The game-theoretic approach has played an important role in determining it. From the viewpoint of RWDC, the application of our resource theory (Theorem 4) of measurement plays the role in data analysis, since the game-theoretic approach is a statistical analysis of the results from quantum measurement. We will re-demonstrate the game-theoretic approach in Chapter 4 to recall its connection to data analysis.

Suppose we are given an index set $m \in \mathcal{M}$, the quantum data, and we encode them into quantum states. We then input these quantum states into the measurement, and we are allowed to choose whatever encoder and decoder to get a better result. We are interested in finding the best strategy that gives us the maximum guessing probability for the index m, i.e. the statistical analysis of quantum data. The process will repeat itself enough amount of times until we collect sufficient statistical data. We illustrate the guessing game in the following Figure 7.4.

In the Figure 7.4, we see that the boxes before and after the measuring devices are seen as encoders and decoders of the measuring device up to some shared randomness. This game setting allows us to analyze the statistical outcomes from the measuring device and based on the statistical results from which we will make a

Figure 7.4: Guess the index set through quantum measurement device

decision on the performance of the measuring device. In this game-theoretic approach to quantum measurement, we were able to testify a genuine quantum measurement device by benchmarking the statistical outcome from the above guessing game as stated in the Corollary 1 in Chapter 4: the idea is that if a device is perfect genuine, it has to satisfy the statistical benchmark. We rewrite the details of main idea in a loose way.

Corollary 5. *If a quantum measurement $\mathbb{M}$ is genuine, it is equivalent to say we can always find a statistical guessing game such that the payoff of that game has to reach beyond a known threshold—the benchmark.*

To sum up, statistical analysis of the data from the quantum device has demonstrated its power in determining a genuine quantum device.

7.4 Implementation of genuine quantum device

In the above circulation in Figure 7.1,the most important part after the data acquisition and data analysis is the real-world implementation as the engineering part. This idea could also be applied to detecting the genuine quantum measurement device. One application from the genuineness detection is to use the statistical benchmark to build a better quantum device through trials and errors from the statistical analysis of data obtained from those engineered devices. These trials and errors come from the circulation of RWDC in the quantum world. Since a genuine quantum device can lead closer to the statistical benchmark that serves as the significance level of the test, we can always engineer a new device if until the device can go beyond the significance level. In a real-world case for engineering a quantum device, the test would take many times for certifying a genuine one. In those circulations, we will deny those devices that do not go beyond the benchmark. With this circulation, we will be able to create a real-world device that is useful and genuine for building any quantum technologies that are based on quantum measurement devices.

The demonstration of the idea of RWDC in quantum world seems to be only theoretical, but it gives a novel potential tool to benchmark the quantum devices that need to be tested, which is almost the same as the development of big data analysis. The theory for machine learning was early developed many years before the real world has the power to implement (or test) the theory and it has been

tested from the large amount of data since recent years. We can safely say that we are only benefiting from the tested results of data analysis for machine learning, i.e., we have been doing data analysis "experiment" in the macro-world. This coincides exactly what has been happening in the quantum world. In the early stage of quantum technology development, to build a reliable theory will be always as important as implementing the technologies in the near future when the engineering part is ready.

At last, we conclude this chapter by providing the following circulation Figure 7.5 in terms of the concept of RWDC that has been demonstrated in details through this chapter.

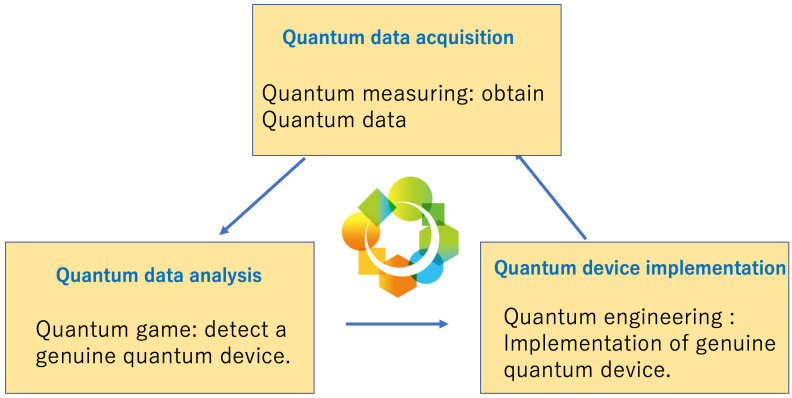

Figure 7.5: The concept of RWDC in quantum world

7.5 Social value: for the next generation technologies

Quantum measurement devices are the core of most quantum technology devices. Developing quantum technologies has been mainstream in the technology industry. In what follows, I introduce two main areas that can benefit directly from advanced quantum technologies. At first, in the key distribution industry, mature quantum key distribution technologies have been widely used to prepare for the disadvantage of the current classical key distribution technologies when quantum computers are strong enough to break the RSA cryptosystem. Secondly, the quantum computer has an absolute advantage in solving complex optimization problems when compared to current classical computing power. For example, if solving complex optimization problems can be much easier than before, many new medicines that depend on simulating complex molecules can be made easily.

Chapter 8

Summary and prospect

In this Ph.D thesis, we developed a generalized resource theory from the viewpoint of statistical decisions theory, and we applied it to both the resource theory of quantum state and the resource theory of quantum measurement. We found the sufficient and necessary conditions for the transformation between POVMs and PMDs through a game-theoretic approach. We also derived sufficient (sometimes necessary as well) for the existence of restricted transformations between resources. The game-theoretic approach is the core of most resource theories, in which we are comparing resources with pre-or partial- ordering by game payoffs. This is the same idea in statistical comparing theories as well, in which the statistical comparison aims to express the possibility of transforming an initial statistical model into another one, in terms of the utility that the two statistical models provide an operationally motivating scenario. The utility could be formed in terms of game scenarios or as simply as entropic quantities. A possible future direction of the presented researches is to look for potential application in the real world in quantum technologies. Our resource theory of PMD provides a possible method to certify genuine quantum memory and the certification of it could be applied to the certification true quantum key distribution as well.

Bibliography

[1] C. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379–423, 1948.

[2] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 1991.

[3] John von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, 1932.

[4] Albert Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780, may 1935.

[5] John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[6] Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: A new violation of bell's inequalities. *Phys. Rev. Lett.*, 49:91–94, Jul 1982.

[7] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal of Computing*, 26(5):1484–1509, oct 1997.

[8] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, Apr 2019.

[9] Ryuji Takagi and Bartosz Regula. General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks. *Phys. Rev. X*, 9:031053, Sep 2019.

[10] Bartosz Regula and Ryuji Takagi. Fundamental limitations on quantum channel manipulation. *arXiv e-prints*, page arXiv:2010.11942, October 2020.

[11] Ryuji Takagi and Hiroyasu Tajima. Universal limitations on implementing resourceful unitary evolutions. *Phys. Rev. A*, 101:022315, Feb 2020.

[12] Francesco Buscemi. Comparison of Quantum Statistical Models: Equivalent Conditions for Sufficiency. *Communications in Mathematical Physics*, 310(3):625–647, jan 2012.

[13] Akinori Kawachi Gen Kimura Tomohiro Ogawa Masahito Hayashi, Satoshi Ishizaka. *Introduction to Quantum Information Science*. Number 1 in 1868-4513. Springer-Verlag Berlin Heidelberg, DOI 10.1007/978-3-662-43502-1, 1 edition, 2015.

[14] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[15] D. Blackwell. Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102. University of California Press, 1951.

[16] S. Sherman. On a theorem of hardy, littlewood, polya, and blackwell. *Proceedings of the National Academy of Sciences of the United States of America*, 37(12):826–831, 1951.

[17] C. Stein. Notes on a seminar on theoretical statistics. i. comparison of experiments. *Report,University of Chicago*, 1951.

[18] C. E. Shannon. A note on a partial ordering for communication channels. *Information and control*, 1(4):390–397, 1958.

[19] Francesco Buscemi, Michael Keyl, Giacomo Mauro DAriano, Paolo Perinotti, and Reinhard F. Werner. Clean positive operator valued measures. *Journal of Mathematical Physics*, 46(8):082109, 2005.

[20] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.

[21] D Cavalcanti and P Skrzypczyk. Quantum steering: a review with focus on semidefinite programming. *Reports on Progress in Physics*, 80(2):024001, dec 2016.

[22] Paul Skrzypczyk and Noah Linden. Robustness of measurement, discrimination games, and accessible information. *Phys. Rev. Lett.*, 122:140403, Apr 2019.

[23] I. Devetak and A. Winter. Distilling common randomness from bipartite quantum states. *IEEE Transactions on Information Theory*, 50(12):3183–3196, 2004.

[24] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, Apr 1996.

[25] Fernando G. S. L. Brandao, Michał Horodecki, Jonathan Oppenheim, Joseph M. Renes, and Robert W. Spekkens. Resource Theory of Quantum States Out of Thermal Equilibrium. *Phys. Rev. Lett.*, 111:250404, Dec 2013.

[26] Michael A. Nielsen and I. Chuang. Quantum Computation and Quantum Information, 2000.

[27] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, jun 2013.

[28] H. Umegaki. Conditional Expectation in an Operator Algebra. *Kodai Math. Sem. Rep.*, 14:59–85, 1962.

[29] Francesco Buscemi and Nilanjana Datta. The Quantum Capacity of Channels With Arbitrarily Correlated Noise. *IEEE Transactions on Information Theory*, 56(3):1447–1460, mar 2010.

[30] Nilanjana Datta. Min- and Max- Relative Entropies and a New Entanglement Monotone. *IEEE Transactions on Information Theory*, 55(6):2816–2826, 2009.

[31] Zi-Wen Liu, Kaifeng Bu, and Ryuji Takagi. One-Shot Operational Quantum Resource Theory. *Phys. Rev. Lett.*, 123(2):020401, Jul 2019.

[32] Guifré Vidal and Rolf Tarrach. Robustness of entanglement. *Phys. Rev. A*, 59:141–155, Jan 1999.

[33] Michael Steiner. Generalized robustness of entanglement. *Phys. Rev. A*, 67:054305, May 2003.

[34] Nilanjana Datta. Max-relative entropy of entanglement, alias log robustness. *International Journal of Quantum Information*, 7(02):475–491, 2009.

[35] Fernando GSL Brandao and Nilanjana Datta. One-shot rates for entanglement manipulation under non-entangling maps. *IEEE Transactions on Information Theory*, 57(3):1754–1760, 2011.

[36] Shizuo Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Math. J.*, 8(3):457–459, 09 1941.

[37] J.B.G. Frenk, G. Kassay, and J. Kolumbán. On equivalent results in minimax theory. *European Journal of Operational Research*, 157(1):46–58, 2004. Smooth and Nonsmooth Optimization.

[38] Francesco Buscemi, David Sutter, and Marco Tomamichel. An information-theoretic treatment of quantum dichotomies. *Quantum*, 3:209, December 2019.

[39] Bartosz Regula, Kaifeng Bu, Ryuji Takagi, and Zi-Wen Liu. Benchmarking one-shot distillation in general quantum resource theories. *arXiv e-prints*, page arXiv:1909.11677, September 2019.

[40] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4(1), jun 2013.

[41] Fernando Brandão, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences*, 112(11):3275–3279, 2015.

[42] Francesco Buscemi. Fully quantum second-law–like statements from the theory of statistical comparisons. *arXiv e-prints*, page arXiv:1010.1030, 2015.

[43] Gilad Gour, David Jennings, Francesco Buscemi, Runyao Duan, and Iman Marvian. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics. *Nature Communications*, 9(1), dec 2018.

[44] Keiji Matsumoto. Reverse Test and Characterization of Quantum Relative Entropy. *arXiv e-prints*, page arXiv:1010.1030, October 2010.

[45] Xin Wang and Mark M. Wilde. Resource theory of asymmetric distinguishability. *Phys. Rev. Research*, 1:033170, Dec 2019.

[46] Patricia Contreras-Tejada, Carlos Palazuelos, and Julio I. de Vicente. Resource Theory of Entanglement with a Unique Multipartite Maximally Entangled State. *Phys. Rev. Lett.*, 122:120503, Mar 2019.

[47] Hongjin Zheng, Eiji Shamoto, Hyuntai Chin, and Anh Nguyen Viet. Monitoring of elliptical vibration cutting process by utilizing internal data in ultrasonic elliptical vibration device. *Proceedings of JSPE Semestrial Meeting*, 2016S:153–154, 2016.