UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP)
INSTITUTO DE FÍSICA GLEB WATAGHIN (IFGW)
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

RODRIGO ISAÚ RAMOS MOREIRA

Information protocols and nonlocality in probabilistic theories

Protocolos de informação e não-localidade em teorias probabilísticas

CAMPINAS
2026

RODRIGO ISAÚ RAMOS MOREIRA

Information protocols and nonlocality in probabilistic theories

Protocolos de informação e não-localidade em teorias probabilísticas

Dissertation presented to the Gleb Wataghin Institute of Physics of the University of Campinas (UNICAMP) as part of the requirements for the degree of Master of Physics, in the area of Physics.

Dissertação apresentada ao Instituto de Física Gleb Wataghin da Universidade Estadual de Campinas (UNICAMP) como parte dos requisitos exigidos para a obtenção do título de Mestre em Física, na área de Física.

Supervisor: Prof. Dr. Marcelo de Oliveira Terra Cunha

ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA.

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8-9174

M813i

Moreira, Rodrigo Isaú Ramos, 2000-
Information protocols and nonlocality in probabilistic theories / Rodrigo Isaú Ramos Moreira. – Campinas, SP : [s.n.], 2026.

Orientador: Marcelo de Oliveira Terra Cunha.
Dissertação (mestrado) – Universidade Estadual de Campinas (UNICAMP), Instituto de Física Gleb Wataghin.

1. Teoria quântica. 2. Teorias probabilísticas generalizadas. 3. Não-localidade de Bell. 4. Não-classicalidade. I. Terra-Cunha, Marcelo de Oliveira, 1973-. II. Universidade Estadual de Campinas (UNICAMP). Instituto de Física Gleb Wataghin. III. Título.

Informações complementares

Título em outro idioma: Protocolos de informação e não-localidade em teorias probabilísticas
Palavras-chave em inglês:
Quantum theory
Generalized probabilistic theories
Bell nonlocality
Nonclassicality
Área de concentração: Física
Titulação: Mestre em Física
Banca examinadora:
Marcelo de Oliveira Terra Cunha [Orientador]
Rafael Luiz da Silva Rabelo
Barbara Lopes Amaral
Data de defesa: 31-07-2026
Programa de Pós-Graduação: Física

Objetivos de Desenvolvimento Sustentável (ODS)
ODS: 4. Educação de qualidade
ODS: 9. Indústria, inovação e infraestrutura

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0009-0000-2146-2405
- Currículo Lattes do autor: http://lattes.cnpq.br/4037051914735535

FOLHA DE APROVAÇÃO

RODRIGO ISAÚ RAMOS MOREIRA

Dissertação apresentada ao Programa de Pós-Graduação em Física, defendida na data de 31/07/2026.

Composição da Comissão Examinadora:

Prof. Dr. Marcelo de Oliveira Terra Cunha

(IMECC/ UNICAMP)

Prof. Dr. Rafael Luiz da Silva Rabelo

(IFGW/ UNICAMP)

Profa. Dra. Barbara Lopes Amaral

(Universidade de São Paulo)

To my mother, Simonete, whom I always remember fondly

Acknowledgements

Although I wrote this thesis, it is certainly not the fruit of my efforts alone. First, because the knowledge herein is the result of centuries of human effort, both of those who directly contributed to the advance of scientific knowledge and of those who indirectly did so. Second, and more importantly to me, I myself am the product of many interactions I have had over the last years. In light of this, I would like to express my gratitute to several people.

First and foremost, I am deeply thankful to my father, Dário. Without your support, none of what is written here would have been possible, and I would not be who I am today. You have my everlasting gratitude.

Next, I would like to thank my supervisor, Marcelo, for giving me the freedom to explore subjects that interest me and for being receptive to the ideas the I brought to you over the last years. I am equally grateful to Roberto Baldijão for the very elucidating conversations that we have had. I always left them with a clearer picture in mind.

To the current and former members of our research group - Alexandre, Amanda, Arthur, Carlos, Denis, Gabriel, João, Rabelo, Sabrina, Trimarco, and Zé -, thanks. Despite seeing you once a week, you have made (and still make) the group meetings enjoyable occasions in which one can learn a lot.

I also thank my friends, whether or not distanced by time or space. In particular, Luís, Naim, Rafael and the Resolvay Conference crew: Alberto, Artur, Barretto, Boechat, Flores, Gabriel, Giovanni, both Rodrigos, Mateus, and Sasa. The undergraduate and master's years would not have been as enjoyable without you.

I would also like to thank the staff of the IFGW's graduate studies office, in special Cris, for always helping me navigate through the uncountably many bureaucratic procedures of graduate studies.

This study was partly financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and process number 88887.902226/2023-00; and partly financed by the Ministry of Science, Technology and Innovation and the National Council for Scientific and Technological Development (CNPq), process number 400587/2023-8.

Lastly, I would like to thank anyone else who feels like they played a part into making this work possible.

You, the people, have the power to make this life free and beautiful, to make this life a wonderful adventure. Then, in the name of democracy, let us use that power. Let us all unite! Let us fight for a new world. A decent world.

Charles Chaplin's speech in *The Great Dictator (1940)*

Modern methods of production have given us the possibility of ease and security for all; we have chosen, instead, to have overwork for some and starvation for the others. Hitherto we have continued to be as energetic as we were before there were machines; in this we have been foolish, but there is no reason to go on being foolish for ever.

Bertrand Russell, *In Praise of Idleness (1932)*

FUJINO: I... don't even like drawing manga. It is not fun, the entire process is a hassle. It's super unglamorous. You can draw all day long and still be nowhere near finished. Really, I'd be better off just sticking to read manga. I shouldn't draw it.

KYOMOTO: Then why do you draw, Fujino?

Fujino and Kyomoto's last dialogue in Fujimoto Tatsuki's *Look Back (2021)*

Abstract

Quantum theory exhibits several distinctive features that challenge intuitions shaped by our experience of the macroscopic world. Among these are entanglement and measurement incompatibility, which together can give rise to violations of Bell inequalities (constraints satisfied by classical models); as well as the no-cloning theorem, which establishes the impossibility of perfectly copying the information stored in an unknown quantum state. In this thesis, after introducing quantum theory and its (im)possibilities, I use the framework of generalized probabilistic theories (GPTs) to discuss how many of these distinctive features – such as entanglement and no-cloning – are not idiosyncrasies of quantum theory, but instead are ubiquitous across non-classical theories. Besides similarities, I also examine, within the contexts of Bell nonlocality and monogamy of entanglement, differences between quantum theory and other non-classical possibilities. This work is intended as an accessible introduction to the GPT framework, thus several specific results are not discussed in the text. I briefly mention some of these in the concluding remarks, where I also raise some open questions that have emerged during my years of research on this subject.

RESUMO

A teoria quântica apresenta diversas características marcantes que desafiam as intuições que construímos a partir da nossa experiência com o mundo macroscópico. Dentre essas, destacamos, por exemplo, o emaranhamento e as medições incompatíveis, que em conjunto possibilitam violações desigualdades de Bell (desigualdades satisfeitas por modelos clássicos); e o teorema da não-clonagem, que versa sobre a impossibilidade de clonar perfeitamente a informação armazenada num estado quântico desconhecido. Nesta dissertação, após introduzir a teoria quântica e suas (im)possibilidades, eu utilizo o formalismo de teorias probabilísticas generalizadas (GPTs) para discutir como tais características marcantes – como o emaranhamento e a não-clonagem – não são idiosincrasias desta teoria, mas sim são características ubíquas em teorias probabilísticas não-clássicas. Para além das semelhanças, também discuto, nos contextos da não-localidade de Bell e de monogamia de emaranhamento, diferenças entre a teoria quântica e outras possibilidades não-clássicas. Este trabalho foi escrito na intenção de servir como uma introdução acessível ao formalismo de GPTs, de modo que muitos resultos específicos não são discutidos no texto. Eu menciono alguns destes nas considerações finais, espaço onde também levanto algumas questões em aberto que se originaram ao longo dos anos em que estive fazendo pesquisa neste tópico.

List of Figures

List of Symbols

$\mathbb{N}$	Natural numbers.
$\mathbb{R}$	Real numbers.
$\mathbb{R}_+$	Non-negative real numbers.
$\mathbb{C}$	Complex numbers.
$\mathcal{H}$	Finite-dimensional, complex Hilbert space.
$\dim \mathcal{H}$	Dimension of $\mathcal{H}$.
$\mathrm{Lin}(\mathcal{H})$	$\mathbb{C}$-vector space of linear operators on $\mathcal{H}$.
$\mathrm{Her}(\mathcal{H})$	$\mathbb{R}$-vector space of Hermitian operators on $\mathcal{H}$.
$\mathcal{D}(\mathcal{H})$	Set of density operators on $\mathcal{H}$.
ρ	Quantum state.
$\mathbb{I}_d$	Identity operator on a d-dimensional complex Hilbert space.
V	Finite-dimensional $\mathbb{R}$-vector space.
V^*	Algebraic dual of V.
C_S^+	Convex cone generated by a set $S \subseteq V$.
$\mathcal{A}(S)$	$\mathbb{R}$-vector space of affine functionals on a set S.
$\mathcal{S}_\mathrm{A}$	State space of a GPT system A.
$\mathcal{E}_\mathrm{A}$	Effect space of a GPT system A.
u_A	Unit effect of a GPT system A.
$\square_{gbit}$	State space of a generalized bit.
Δ_d	$(d-1)$-dimensional probability simplex.
$\mathrm{vert}(\Delta_d)$	Vertices of Δ_d.
$\otimes$	Tensor product.
$\boxtimes$	Composition rule.
$\otimes_\mathrm{max}$	Maximal tensor product.
$\otimes_\mathrm{min}$	Minimal tensor product.

CONTENTS

INTRODUCTION

> Tanto que fazer!
> livros que não se lêem, cartas que não se escrevem,
> línguas que não se aprendem,
> amor que não se dá,
> tudo quanto se esquece.
>
> Cecília Meireles. Humildade

Without a doubt, quantum theory is one of the most well-tested scientific theories of the last century. At this point in time, shortly after the theory's hundredth anniversary, we can observe its impact in several segments of our society. To name a few: the development of transistors, which are indispensable for many everyday electronic devices; the advent of quantum chemistry and spectroscopy techniques such as MRI and IR; photonics and its applications in communications; and even in literature, most notably in science fiction. All of last century's technological advances that are due to quantum theory are grouped in what we now deem as the first Quantum Revolution.

The hallmark of the second Quantum Revolution will be our capacity to manipulate quantum systems for information-processing tasks (such as computation, cryptography, communications, and sensing) in a controlled fashion, allowing us to leverage what we call quantum resources, *e.g.* entanglement, nonlocality and contextuality, to outperform what is classically achievable.

Given this ever-growing utility of quantum systems for information-processing tasks, one might wonder where all of the theory's power comes from. The answers are diverse, but can ultimately be traced back to the its postulates. When we look at them, however, we are faced with a list of prescriptions for *describing physical systems in a general manner*, but no physical principles seem to underlie the mathematical content of these postulates. For example, in one of its versions, the first postulate of quantum theory states the following:

QT1) Associated with any isolated physical system is a Hilbert space $\mathcal{H}$, known as the state space of the system. This physical system is completely described by its state vector $|\psi\rangle$, which is a unit vector in the state space.

In addition, we also have postulates that define how a state transforms and how to assign probabilities from measurements:

QT2) The evolution of a closed quantum system is described by a unitary operator.

QT3) A measurement is described by linear operators $\{M_k : \mathcal{H} \to \mathcal{H}\}_{k\in\mathcal{O}}$, where $\mathcal{O}$ is the set of the measurement's outcomes. When the system is described by a state $|\psi\rangle$, the probability of observing an outcome k is given by

$$p(k) = \langle\psi|M_k^\dagger M_k|\psi\rangle. \qquad \text{(Born rule)}$$

Let us contrast this situation with another prominent theory from the 20th century: relativity. In its simplest presentation — *i.e.* when we neglect curvature effects — the theory is described by two postulates:

SR1) The laws of physics are the same for all inertial observers.

SR2) The speed of light in vacuum is the same for all inertial observers.

With these two simple-to-state postulates one can justify the need for a Lorentzian metric and derive the Lorentz transformations as the appropriate transformations that relate inertial reference frames. These imply interesting consequences such as the light-cone structure; length contraction and time dilation; the mass-energy equivalence; and corrections to the Doppler effect and aberration of light.

Now, when we look back at postulates **QT1-QT3**, none of them have the same character as **SR1** and **SR2**, and, although many have dedicated themselves to derive quantum theory from physical principles, we still do not have a definitive answer to what set of principles are the correct ones.

In their current formulation, however, the postulates have a rather operational flavour: they specify how we prepare a system, *i.e.* what state we assign to it; how we transform it via unitary operations; and how measurement statistics are obtained via the Born rule.

By abstracting away finer details, the postulates describe how a generic experiment proceeds: prepare (P) a system, possibly subject it to some transformation (T) and measure (M) the system, collecting some classical outcome k. If we frame a theory solely by what sort of statistics it can yield in this prepare-transform-measure (PTM) paradigm, then we could ask what sort of information-processing tasks are possible within such a theory. Remarkably, it happens that some of the most distinctive quantum features are also present in other hypothetical non-classical theories.

These general (probabilistic) theories (GPTs) are not intended to overthrow quantum theory; instead, they are better understood as part of a framework that can assist us physicists in understanding the idiosyncrasies of quantum theory, and, perhaps, help to pave a way for those who wish to derive quantum theory from more physically inclined postulates.

In this thesis, we present the GPT framework and highlight the similarities and differences between quantum theory and other non-classical possibilities. This

work is divided into two parts. In the first part, we review the basics of quantum theory (Chapter 1), introducing some necessary concepts that will be required in the following chapters. Next, in Chapter 2, we examine the capabilities of the theory from two points of view: the first one regards the possibility of teleporting information, despite the impossibility of cloning it. The second one regards Bell nonlocality and the failure of quantum theory to conform to statistical models that admit a classical explanation.

In the second part, we devote our attention to generic probabilistic theories. In Chapter 3, starting only from the operational primitives previously mentioned, we provide a thorough construction of the mathematical structure that characterizes any probabilistic theory. In Chapter 4, we tie the content of the previous three chapters together. We provide proofs of the no-cloning and the no-broadcasting theorems; extensions of the teleportation and BB84 protocols and their connections to remarkable aspects of quantum theory are discussed; we also discuss how monogamy of entanglement would differ if quantum theory were based on real numbers. Lastly, we present two original results, one regarding nonlocality in a modified version of quantum theory and another related to a notion dual to broadcasting.

We conclude the thesis by pointing out several questions related to what has been discussed throughout the text and mention some works within research in GPTs that were not be included in the main text, since the purpose of this thesis is to provide an accessible introduction to the GPT framework.

Part I

Quantum world

1 RUDIMENTS OF QUANTUM THEORY

> In a strict sense, quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow specified preparations.
>
> Asher Peres

Chapter Contents

In this chapter, we remind the reader of some basic notions of quantum theory by presenting its postulates and discussing examples that illustrate the formalism. For a more detailed introduction to quantum theory and its applications – especially from an information-theoretic point of view – please refer to [1], [2], [3], [4], [5], [6] or [7].

1.1 Quantum states, measurements and a minimal interpretation

In the late 19th and early 20th centuries, the inadequacy of classical physics in explaining a series of phenomena led, a few decades later, to the development of a new fundamental theory of Nature: quantum theory. But what *is* quantum theory, exactly? As hinted by the epigraph at the beginning of this chapter, a perhaps uncontroversial answer to that question is that, *at the very least* [2]:

> "Quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow specified preparations."

A physical system possessing any property which obeys these rules is called a *quantum system*.

Throughout this and the following sections, we gradually present these rules, which from now on are referred to as postulates, and discuss some elementary examples for clarity.

Remark 1.1. Henceforth, it is assumed that the reader is familiar with Dirac's bra-ket notation and with some basic principles of quantum mechanics (quantum states as kets, observables as Hermitian operators, and how projective measurements work, for instance), but not necessarily with concepts such as density operators, POVMs, and CPTP maps.

Remark 1.2. In this thesis, we will work solely with finite-dimensional quantum systems, since quantum information theory is typically formulated within this context. Typically, we denote $\dim \mathcal{H}_A = d_A$ and omit the subscript whenever labelling the system is irrelevant.

1.1.1 States

Postulate I. To every isolated physical system A, we ascribe a complex Hilbert space $\mathcal{H}_A$. Additionally, the set of possible (quantum) states that describe the system is defined as

$$\mathcal{D}(\mathcal{H}_A) \doteq \{\rho \in \mathrm{Lin}(\mathcal{H}_A) \mid \mathrm{Tr}\,\rho = 1, \rho \succcurlyeq 0\}. \tag{1.1}$$

Here, $\mathrm{Lin}(\mathcal{H}_A)$ denotes the vector space of linear operators $T : \mathcal{H}_A \to \mathcal{H}_A$. The last condition in equation (1.1) reads as "ρ is positive semidefinite" and it means that for any $|\psi\rangle \in \mathcal{H}_A$:

$$\langle\psi|\rho|\psi\rangle \geq 0. \tag{1.2}$$

At first sight, defining states as elements of $\mathcal{D}(\mathcal{H}_A)$ might seem too great a departure from the standard idea of quantum states simply as kets $|\psi\rangle$ in $\mathcal{H}_A$. However, we will show that in settings where there may be some uncertainty about how the system under analysis was prepared, equation (1.1) corresponds to the appropriate generalisation [1, 3]. Before doing so, at the risk of getting ahead of ourselves, we recall that a **projective measurement** is a collection $\{P_i \mid P_i \in \mathrm{Lin}(\mathcal{H})\}_{i=1}^{d'}, d' \leq d$, such that

$$\begin{cases} P_i = P_i^\dagger, \\ P_i P_j = \delta_{ij} P_i, \end{cases} \tag{1.3a}$$

and

$$\sum_{i=1}^{d'} P_i = \mathbb{I}_d, \tag{1.3b}$$

where P_i denotes the projector associated to outcome i and $\mathbb{I}_d$ is the identity operator on $\mathcal{H}$.

Now, suppose that someone, Alice, hands us a quantum system and the only information she tells us is that the system was prepared in some state $|\psi_x\rangle$, where x was chosen according to a probability distribution p_X over a finite alphabet[1] $\mathcal{X}$. This defines an *ensemble* of quantum states

$$\mathcal{E} = \{(p_X(x), |\psi_x\rangle)\}_{x\in\mathcal{X}}.$$

Let us denote by I the random variable that corresponds to the outcome of a projective measurement $\{P_i\}_{i=1}^{d'}$. For a particular $x \in \mathcal{X}$, we know that the conditional probability of obtaining outcome i conditioned on x is:

$$p_{I|X}(i|x) = \langle\psi_x|P_i|\psi_x\rangle. \tag{1.4}$$

Therefore, using equation (1.4), the probability of reading outcome i can be written as:

$$p_I(i) = \sum_{x\in\mathcal{X}} \langle\psi_x|P_i|\psi_x\rangle\, p_X(x). \tag{1.5}$$

Since $\langle\psi_x|P_i|\psi_x\rangle$ is equal to $\mathrm{Tr}\big(|\psi_x\rangle\langle\psi_x|\,P_i\big)$, we can rewrite equation (1.5) as

$$p_I(i) = \mathrm{Tr}\big(\rho_{\mathcal{E}} P_i\big), \tag{1.6}$$

where

$$\rho_{\mathcal{E}} = \sum_{x\in\mathcal{X}} p_X(x)\,|\psi_x\rangle\langle\psi_x| \tag{1.7}$$

is what we call a **density operator**. Since each $|\psi_x\rangle\langle\psi_x|$ has unit trace and the $p_X(x)$ sum to 1, $\rho_{\mathcal{E}}$ has unit trace too. Additionally, for any $|\varphi\rangle \in \mathcal{H}$:

$$\langle\varphi|\rho_{\mathcal{E}}|\varphi\rangle = \sum_{x\in\mathcal{X}} p_X(x)\big|\langle\varphi|\psi_x\rangle\big|^2,$$

which is nonnegative. Hence, $\rho_{\mathcal{E}}$ is positive semidefinite.

From the previous two calculations we see that $\rho_{\mathcal{E}}$ is a legitimate member of the set $\mathcal{D}(\mathcal{H})$ defined in equation (1.1). More interestingly, **$\mathcal{D}(\mathcal{H})$ consists only[2] of operators like $\rho_{\mathcal{E}}$**, *i.e.* operators for which one can find an ensemble of quantum states and express them as in equation (1.7). Thus, any quantum state (in the sense of **Postulate I**) is a density operator.

Notice that if there is no uncertainty regarding the preparation of the sys-

[1] By alphabet, we mean any collection of symbols.

[2] This comes from the fact that every positive semidefinite operator is also a Hermitian operator — meaning it is equal to its adjoint —, and this implies that it has nonnegative eigenvalues. The unit trace condition forces these eigenvalues to be probabilities.

tem, that is, if $p_X(x) = 1$ for a state $|\psi\rangle$ from the ensemble $\mathcal{E}$, then we can express

$$\rho_{\mathcal{E}} = |\psi\rangle\langle\psi|. \tag{1.8}$$

Whenever this is the case, we say that $\rho_{\mathcal{E}}$ is a **pure state**; whenever it is not, we say that $\rho_{\mathcal{E}}$ is a **mixed state**. Up to global phases, there is a one-to-one correspondence between kets and pure states, since the latter are precisely the projectors onto the former. We shall use this correspondence throughout the chapter whenever we refer to a ket as a pure state.

From equation (1.7), we see that the set of pure states of a quantum system is somehow special: any quantum state is a probabilistic mixture of a collection of pure states. This is by no means a coincidence, and it is actually rooted in the geometry of the state space $\mathcal{D}(\mathcal{H})$. This connection will be clarified later on in this chapter (remark 1.4).

Having established what state spaces are, we now turn to another essential aspect of quantum theory: measurements.

1.1.2 Measurements

In our derivation of the expression for $\rho_{\mathcal{E}}$, we had to invoke the notion of a projective measurement (equations (1.3a) and (1.3b)), and at the end of the process, we obtained an expression for computing probabilities using density operators (equation (1.6)). This notion, together with the derived expression, is formalized by the following:

Postulate II. *Let $\mathcal{H}_A$ be the Hilbert space associated with a physical system A. An n-outcome measurement on A is a collection of linear operators $\{M_i\}_{i=1}^n$ that satisfy a resolution to the identity*

$$\sum_{i=1}^{n} M_i^\dagger M_i = \mathbb{I}_{d_A}, \tag{1.9a}$$

where each operator is associated with a unique measurement outcome. If prior to measurement the system is described by a state ρ, the probability of obtaining outcome i is given by

$$p(i) = \mathrm{Tr}\left(M_i \rho M_i^\dagger\right). \tag{Born's rule}$$

The post-measurement state after a particular outcome is

$$\frac{M_i \rho M_i^\dagger}{\mathrm{Tr}\left(M_i \rho M_i^\dagger\right)}. \tag{1.9b}$$

In light of **Postulate II**, projective measurements are a particular type of measurement. They form a partition of $\mathcal{H}_A$ into orthogonal subspaces and, since $\dim \mathcal{H}_A = d_A$, such a partition has at most d_A elements. Orthogonality guarantees that there is no

ambiguity in the post-measurement state, as it could only have been obtained from one specific outcome. Additionally, for an observable X with spectral decomposition

$$X = \sum_{i=1}^{d_A} x_i P_i, \tag{1.10}$$

Born's rule implies that its expected value is

$$\langle X \rangle_\rho = \sum_{i=1}^{d_A} x_i \operatorname{Tr}\big(\rho P_i\big). \tag{1.11}$$

Another relevant type of measurement is one defined by a positive operator-valued measure (POVM for short); it is useful in settings where the post-measurement state is not relevant or even nonexistent (for instance if the system is destroyed after the measurement procedure). An **n-outcome POVM** is defined by a collection of operators E_i, named effects, that satisfy

$$E_i \succcurlyeq 0, \tag{1.12a}$$

and

$$\sum_{i=1}^{n} E_i = \mathbb{I}_d. \tag{1.12b}$$

Any n-outcome measurement yields a unique n-outcome POVM, since one can set the effects as $E_i = M_i^\dagger M_i$. The converse is not true, however: a single POVM is compatible with several n-outcome measurements. This happens because, although we can decompose each E_i in terms of its positive square-root $\sqrt{E_i}$, any choice of unitary U_i yields an $M_i = U_i\sqrt{E_i}$ that satisfies $M_i^\dagger M_i = E_i$. As a consequence of this unitary freedom, post-measurement states are not always well-defined for this class of measurements. Despite this drawback, POVMs are useful in operational tasks, such as quantum state discrimination and quantum state estimation [8, 9].

1.1.3 An operational view

Postulates **I** and **II** indicate that states and measurements are not on equal footing in quantum theory. From **postulate II**, a measurement device need not be described as a genuine quantum system; it suffices to treat it as a sort-of black-box which consumes quantum systems and outputs an outcome whilst potentially perturbing the state of the system. This is in contrast with classical physics: there, both the system under analysis and the measurement device can be modelled within the same theory [10] – for instance, the device must obey Newton's laws. Moreover, the post-measurement state update rule implies that measurement is an active process in quantum theory, as it can change a system's state rather drastically. Since the quantum state of a system determines the predictions of our experiments (cf. Born's rule), the interplay between

states and measurements leads to the following question [11, p. 71]:

"When a quantum state $|\psi\rangle$ is assigned to a physical system, does this mean that there is some independently existing property of the individual system that is in one-to-one correspondence with $|\psi\rangle$ (up to a global phase), or is $|\psi\rangle$ simply a mathematical tool for determining probabilities, existing only in the minds and calculations of quantum theorists?"

Although such a question may sound more appropriate to the realm of Philosophy rather than Physics, its answer has a deep implication to the latter: for if the quantum state corresponds to an element of reality, then there must be a physical mechanism that explains how the post-measurement assignment (equation (1.9b)) comes to be. To this day, however, this question remains a subject of debate and with no definitive resolution [10–12].

For our purposes, we adhere to a "minimal", also called operational, view on what the quantum state is [13, 14]. A quantum state is nothing more than a representation of a preparation procedure that one performs in the laboratory when running an experiment. We make no assumption regarding its ontological status. In Chapter 3 we will explore how this stance allows us to lay down a framework for designing foil theories, *i.e.*, theories that, when contrasted with quantum theory, highlight its properties.

Having explicitly made our stance and presented some of the essential postulates of the theory, we now examine a paradigmatic quantum system.

1.2 The qubit

The simplest, non-trivial quantum systems are qubits (short for quantum bits), which lie at the heart of quantum computation and quantum information theory; hence their importance. These are two-dimensional systems, in the sense that the Hilbert space describing a single qubit $\mathcal{H}$ is isomorphic to $\mathbb{C}^2$.

The positive semidefinite constraint imposes that qubit states belong to the (real) vector space of Hermitian operators. A suitable basis in this vector space is given by the identity and the Pauli matrices:

$$\mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.13}$$

Consequently, any qubit state can be expressed as follows:

$$\rho = \frac{1}{2}(\mathbb{I}_2 + \vec{r} \cdot \vec{\sigma}), \tag{1.14}$$

where $\vec{r}\cdot\sigma$ represents the inner product between $\vec{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ and $\vec{\sigma} = (\sigma_X, \sigma_Y, \sigma_Z)$. The overall $1/2$ factor accounts for the unit trace condition. The latter together with positive semidefiniteness ($\rho \succcurlyeq 0$), imply that

$$\|\vec{r}\| \leq 1. \tag{1.15}$$

The previous inequality suggests that the state space of qubits can be represented as a unit ball, named the Bloch ball (figure 1.1). To do that, it suffices to set

$$\vec{r} = (r\cos(\varphi)\sin(\theta), r\sin(\varphi)\sin(\theta), r\cos(\theta)) \tag{1.16}$$

for $r \in [0,1]$, $\varphi \in [0, 2\pi)$, and $\theta \in [0, \pi)$. Pure states are characterized by the condition $r = 1$, meaning that they lie on the boundary of the Bloch ball, which is called the Bloch sphere.

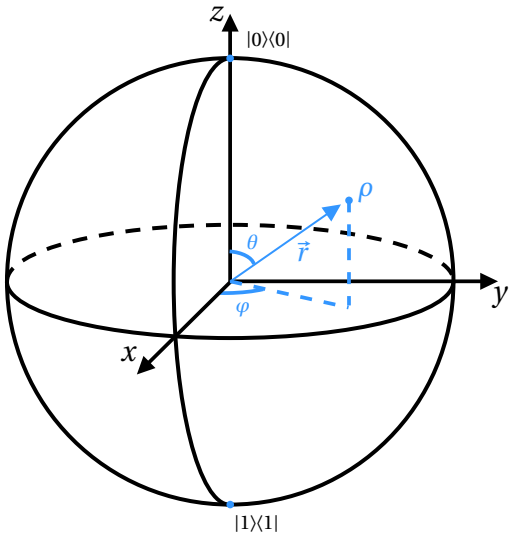


Figure 1.1: Illustration of the Bloch Ball.

Two important pure states are

$$|0\rangle\langle 0| = \frac{1}{2}(\mathbb{I}_2 + \sigma_Z) \tag{1.17a}$$

and

$$|1\rangle\langle 1| = \frac{1}{2}(\mathbb{I}_2 - \sigma_Z), \tag{1.17b}$$

where $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are the projectors associated with the canonical basis of $\mathbb{C}^2$:

$$|0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.18}$$

In the Bloch ball representation, these states form a pair of antipodal points,

i.e. they lie at polar opposites of each other. Their importance lies in the fact that they enable the embedding of a random/noisy bit inside a qubit. These are elementary systems in classical information theory, modelling 'yes-no' questions with a certain bias towards one of the two possibilities (0 or 1). That is, a random bit is a binary random variable with distribution

$$p(0) = q, \quad \text{and} \quad p(1) = 1 - q, \tag{1.19}$$

for $q \in [0, 1]$.

A possible embedding of this classical[3] system within the qubit is given by the state

$$\rho(q) = q\,|0\rangle\langle 0| + (1-q)\,|1\rangle\langle 1|\,, \tag{1.20}$$

which, from the previous discussion about density operators, arises from an ensemble that uses a random bit to prepare the system either in state $|0\rangle\langle 0|$ or in state $|1\rangle\langle 1|$.

What matters in equation (1.20) is that the states involved in its decomposition are **perfectly distinguishable**. Operationally, this condition translates to the existence of a measurement $\{M_0, M_1\}$ such that

$$\begin{aligned} \mathrm{Tr}\Big(M_0\,|0\rangle\langle 0|\,M_0^\dagger\Big) = 1, \quad \mathrm{Tr}\Big(M_0\,|1\rangle\langle 1|\,M_0^\dagger\Big) = 0, \\ \mathrm{Tr}\Big(M_1\,|0\rangle\langle 0|\,M_1^\dagger\Big) = 0, \quad \mathrm{Tr}\Big(M_1\,|1\rangle\langle 1|\,M_1^\dagger\Big) = 1. \end{aligned} \tag{1.21}$$

Equivalently, the probability distributions obtained from this measurement for each state unequivocally tell us whether the prepared state was $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$.

Since the states on the right-hand side of equation (1.20) are pure, perfect distinguishability amounts to the orthogonality of the kets $|0\rangle$ and $|1\rangle$, thence we can construct $\{M_0, M_1\}$ as the projective measurement $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$. For a general construction, see remark 1.3.

By varying q, we notice that states which describe a random bit belong to the line segment whose extremal points are $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. This expresses that $\rho(q)$ is a *convex combination* of those two points, and the parameter q sets $\rho(q)$ as the point that divides this segment into parts of length q and $1 - q$.

[3] Classical in the sense that it is a system from classical information theory.

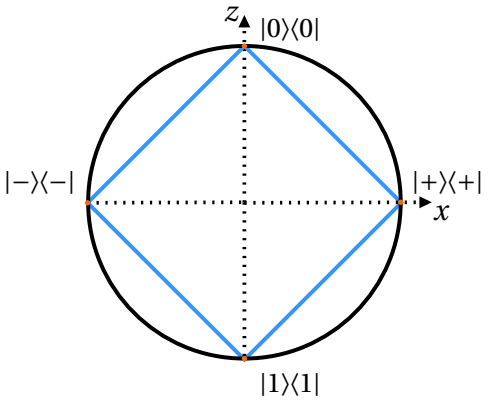


Figure 1.2: The square whose vertices are the BB84 states lies in the x-z plane of the Bloch ball. The dotted lines represent random bits.

The previous construction of a random bit is not unique. In fact, since antipodal points in the Bloch sphere represent pairs of orthogonal states, there are uncountably many ways to implement such a construction! Another possibility, for example, is to consider the states

$$|+\rangle\langle+| = \frac{1}{2}(\mathbb{I}_2 + \sigma_X) \tag{1.22a}$$

and

$$|-\rangle\langle-| = \frac{1}{2}(\mathbb{I}_2 - \sigma_X), \tag{1.22b}$$

which are associated with the rotated basis

$$|+\rangle \doteq \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle \doteq \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \tag{1.23}$$

Since the four states $|0\rangle\langle 0|$, $|1\rangle\langle 1|$, $|+\rangle\langle +|$ and $|-\rangle\langle -|$ are vertices of a square inscribed in an equator of the Bloch ball (figure 1.2), we can see that the random bits of these two pairs lie on orthogonal axes. We will refer to these four states as the BB84 states, due to their role in a seminal protocol in quantum cryptography [15].

We end this discussion by pointing out one of the fundamental differences between the qubit and the (random) classical bit. From our construction, each realization of the latter is determined by a fixed pair of pure states; this implies that any system described by a classical bit only has two pure states, and decompositions of mixed states are unique. Qubits, on the other hand, have uncountably many pure states (they are the points of the Bloch sphere), and mixed states do not have a unique decomposition in terms of them.

Remark 1.3. In general, we can construct a random bit from any two orthogonal states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, where orthogonality now is defined according to the Hilbert-Schmidt inner product $\langle \cdot, \cdot \rangle_{HS} : \mathrm{Lin}(\mathcal{H}) \times \mathrm{Lin}(\mathcal{H}) \to \mathbb{C}$

$$\langle X, Y \rangle_{HS} \doteq \mathrm{Tr}\left(X^\dagger Y\right). \tag{1.24}$$

The measurement $\{M_0, M_1\}$ is given by the projector onto the support of ρ (the subspace spanned by its eigenvectors with non-zero eigenvalues) and the projector onto its orthogonal complement. Orthogonality guarantees that the support of σ is contained in this complement.

Remark 1.4. This simple example shows that the Bloch ball illustration of quantum states perfectly captures the general geometry of quantum state spaces and the nature of pure states: any quantum state space $\mathcal{D}(\mathcal{H})$ is a **convex set**, meaning that for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and any $p \in [0,1]$ it is guaranteed that

$$p\rho + (1-p)\sigma \tag{1.25}$$

also belongs to $\mathcal{D}(\mathcal{H})$. This state lies on the line segment between ρ and σ, and it is the point that divides said segment into parts of length p and $1-p$. Moreover, the pure states are precisely the extremal points of this set, in the sense that they are not convex combination of other states.

1.2.1 An application of POVMs: minimum-error state discrimination

We end this section examining one central feature that makes the BB84 states interesting: the lack of perfect distinguishability.

Let $I = \{0, 1, +, -\}$ and suppose that we receive a quantum system from a source that generates a state $|i\rangle\langle i|$, $i \in I$, with uniform probability. In the case that we were not informed in which state the system was prepared, our task reads as follows:

Find the measurement that best distinguishes these four states.

Since projective measurements on a qubit can have at most two outcomes, they are insufficient for distinguishing four states. We therefore turn to POVMs, since they allow an arbitrary number of outcomes.

For a given four-outcome POVM $\{E_i\}_{i\in I}$, if we associate each outcome i to one of the states that we were asked to discriminate, then the *average success probability* of discriminating them is given by

$$p_{\text{succ}}(\{E_i\}_{i\in I}) = \frac{1}{4}\sum_{i\in I} \text{Tr}\big(\rho_i E_i\big), \tag{1.26}$$

where $\rho_i \equiv |i\rangle\langle i|$.

The task then becomes an optimization problem: we seek the measurement which, on average, minimizes the chances of a wrong guess, *i.e.*, it maximizes the average probability of success [8]. Such problem takes the form of a semidefinite program[4]

[4]In rough terms, an SDP is a constrained optimization problem in which the involved variables are

(SDP):

$$\begin{aligned} \max \quad & p_{\text{succ}}(\{E_i\}_{i\in I}) \\ \text{subject to} \quad & E_i \succcurlyeq 0, \quad \forall i \in I, \\ & \sum_{i\in I} E_i = \mathbb{I}_2. \end{aligned} \tag{1.27}$$

From the geometry of the states involved, a reasonable candidate POVM is the one that assigns an outcome to each vertex of the square in figure 1.2:

$$E_i = \frac{1}{2}|i\rangle\langle i| \tag{1.28}$$

for all $i \in I$. Such POVM yields

$$p_{\text{succ}}\left(\left\{\frac{1}{2}|i\rangle\langle i|\right\}_{i\in I}\right) = \frac{1}{2},$$

which means that on average, half of our guesses will be wrong. Indeed, this POVM is a solution of the SDP in equation (1.27), and it reveals we can do no better than just randomly guessing which state we were given.

1.3 Beyond single systems

So far in our discussions, it has not been necessary to describe systems composed of multiple quantum systems. We now turn to this matter, as it will play an important role in the remainder of this chapter and in subsequent ones.

1.3.1 Assigning Hilbert Spaces

Suppose that two parties, Alice (A) and Bob (B), have a quantum system each (which are labeled after the initials of each party). If the systems A and B are treated as a single quantum system AB, then by **postulate I** we must assign a Hilbert space $\mathcal{H}_{AB}$ to this joint system. This assignment is formalized by the

Postulate III. For a pair of isolated quantum systems A and B, the Hilbert space associated with the joint system AB is the tensor product of each party's Hilbert space:

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \tag{1.29}$$

The case for three or more parties follows immediately from the postulate: we just consider the tensor product of all Hilbert spaces involved. As an example, for three

positive semidefinite matrices. For precise definitions, see [16, see Chapters 2 and 4]. What is relevant here is that if a problem can be cast as an SDP (e.g. as in eq. (1.27)), then it is efficiently solvable and the solution is not approximate.

parties, A, B and C, we can first consider the joint system AB and then apply the postulate to systems AB and C; the other possibilities (A + BC, AC + B, A + CB, etc.) all result in Hilbert spaces isomorphic to $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$.

But why equation (1.29)? In our view, the reason for this choice is that it ties the whole to its parts in an economical manner. Mathematically, this is reflected by the fact that $\dim \mathcal{H}_{AB} = d_A d_B$, where $d_A = \dim \mathcal{H}_A$ and $d_B = \dim \mathcal{H}_B$. In particular, if $\{|\psi_i\rangle_A\}_{i=1}^{d_A}$ and $\{|\varphi_j\rangle_B\}_{j=1}^{d_B}$ are orthonormal basis of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, then

$$\mathcal{B}_{AB} = \{|\psi_i\rangle_A \otimes |\varphi_j\rangle_B \mid 1 \leq i \leq d_A,\ 1 \leq j \leq d_B\}$$

is an orthonormal basis of $\mathcal{H}_A \otimes \mathcal{H}_B$, *i.e.* $\mathcal{H}_A \otimes \mathcal{H}_B = \mathrm{span}(\mathcal{B}_{AB})$. Consequently, we can write a bipartite pure state $|\Psi\rangle_{AB}$ as

$$|\Psi\rangle_{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} c_{ij} |\psi_i\rangle_A \otimes |\varphi_j\rangle_B, \tag{1.30}$$

where $c_{ij} = \mathrm{Tr}\big(|\Psi\rangle\langle\Psi|_{AB} |\psi_i\rangle\langle\psi_i|_A \otimes |\varphi_j\rangle\langle\varphi_j|_B\big)$.

We can think that the local bases $\{|\psi_i\rangle_A\}_{i=1}^{d_A}$ and $\{|\varphi_j\rangle_B\}_{j=1}^{d_B}$ diagonalize some observables O_A and O_B of Alice's and Bob's systems, respectively. This viewpoint then shows that any pure state of AB, and thus all states in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, can always be completely characterized using only local[5] measurements. This is a property known as **local tomography** [17], which we will further explore in Chapters 3 and 4.

1.3.2 Entanglement

Beyond ensuring local tomography, the assignment in equation (1.29) is also compatible with both independent preparations and operations. Indeed, if each system is prepared in a pure state from an independent source, the joint state takes the form

$$\rho_{AB} = |\psi\rangle\langle\psi|_A \otimes |\varphi\rangle\langle\varphi|_B, \tag{1.31}$$

which is called a **product state**. If U and V are unitaries implemented respectively by Alice and Bob, then after performing these operations, the joint state is mapped to

$$\rho'_{AB} = (U|\psi\rangle\langle\psi|_A U^\dagger) \otimes (V|\varphi\rangle\langle\varphi|_B V^\dagger), \tag{1.32}$$

which demonstrates that independent local operations work as intended. Notice that the state still takes product form.

If Alice's and Bob's systems are prepared by random procedures from inde-

[5]The term "local" expresses that, in $\mathrm{Lin}(\mathcal{H}_{AB})$, these observables factorize with the identity, *i.e.*, they are written as $O_A \otimes \mathbb{I}_{d_B}$ and $\mathbb{I}_{d_A} \otimes O_B$.

pendent sources, the joint state can only be

$$\rho_{AB} = \rho_A \otimes \rho_B, \tag{1.33}$$

for arbitrary states $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and $\rho_B \in \mathcal{D}(\mathcal{H}_B)$, which remains a product state. But what if the sources are not independent? In this situation, the preparations are correlated by a discrete random variable taking values in some alphabet I. The generated state takes the form

$$\rho_{AB} = \sum_{i \in I} p(i) \rho_i^A \otimes \rho_i^B, \tag{1.34}$$

where $p(i)$ is the probability of outcome i, and both ρ_i^A and ρ_i^B are arbitrary states. These states, which exhibit classical correlations between systems, are called **separable states**.

Since separable states are convex combinations of product states, and product states themselves can be expressed as convex combinations of product pure states, it follows that separable states can always be decomposed in terms of product pure states.

Due to equation (1.29), *not every pure state is of product form*[6], thus not every bipartite quantum state is a separable state. When that is case, we say that the state is an **entangled state**.

More formally, let us denote the set of separable states as

$$\mathrm{Sep}(AB) \doteq \{\rho \in \mathcal{D}(\mathcal{H}_{AB}) \mid \ \rho \text{ satisfies equation (1.34)}\}. \tag{1.35}$$

Then we have the following definition:

Definition 1.1: Entangled state

A state $\rho \in \mathcal{D}(\mathcal{H}_{AB})$ is entangled if $\rho \notin \mathrm{Sep}(AB)$.

For bipartite qubit systems, a well-known collection of entangled pure states is:

$$\begin{aligned} |\Phi^+\rangle &\doteq \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |\Phi^-\rangle \doteq \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \\ |\Psi^+\rangle &\doteq \frac{|01\rangle + |10\rangle}{\sqrt{2}}, \quad |\Psi^-\rangle \doteq \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \end{aligned} \tag{1.36}$$

They are called **Bell states**. Their importance stems, for example, from their usefulness in quantum protocols such as quantum teleportation and superdense coding, as well as from their capacity to violate Bell inequalities. The latter provides the basis for security proofs in quantum key distribution (QKD) and a means for self-testing quantum

[6] Consider equation (1.30). Separability implies that $c_{ij} = a_i b_j$ for all i and j, which need not be true in general.

systems, both of which are fundamental for quantum cryptography.

We can use the Bell states to point out a geometric difference between Sep(AB) and the collection of entangled states. Notice that the state

$$\frac{1}{4}\mathbb{I} = \frac{1}{4}(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|)$$

is separable, since it is a convex combination of $|00\rangle, |10\rangle, |01\rangle$ and $|11\rangle$. However, we can also express it as

$$\frac{1}{4}\mathbb{I} = \frac{1}{4}(\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + |\Phi^-\rangle\langle\Phi^-| + \left|\Psi^+\right\rangle\left\langle\Psi^+\right| + |\Psi^-\rangle\langle\Psi^-|),$$

which is a convex combination of the Bell states. Hence, differently from Sep(AB), the set of entangled states is not a convex set (remark 1.4) as illustrated in figure 1.3.

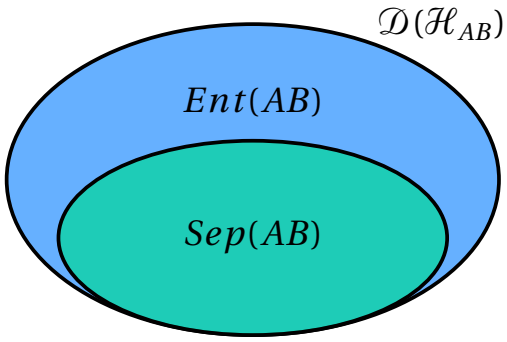


Figure 1.3: An illustration of the sets of separable (Sep(AB)) and entangled (Ent(AB)) states within $\mathcal{D}(\mathcal{H}_{AB})$ (outer ellipse).

For more than two parties, the classification between separable and entangled becomes more subtle. For instance, if a tripartite qubit system is described by a state $\left|\Phi^+\right\rangle\left\langle\Phi^+\right|_{AB} \otimes |0\rangle\langle 0|_C$, is it a separable state or an entangled one? Certainly, Alice's and Bob's systems are entangled, and both are unentangled from Charlie's (C) system. Hence, there is some ambiguity on whether this state is entangled or separable: it depends on which parties we look at. There exist, however, tripartite states that exhibit entanglement among all three parties [18, 19] and thus the answer to the previous question is unambiguous: the systems are genuinely entangled.

From this example, we see that there are many ways to classify states for more than two parties. They can be (a) fully separable, which are defined by the immediate generalization of equation (1.34); (b) partially separable, meaning that they exhibit some entanglement between a subset of parties, but not all of them; and (c) genuinely entangled.

Understanding multipartite entanglement is still an active area of research and we refer to [20, Chapter 14] for an introduction to the subject. Henceforth, we restrict ourselves to the more well-understood bipartite case.

1.4 Detecting and quantifying bipartite entanglement

Several methods have been devised to detect bipartite entanglement. Here, we mention only two of them and refer the reader to [19], [21, Chapter 1] and references therein for more in-depth discussions.

1.4.1 Reduced state criterion

The first method is a simple criterion regarding the reduced states of a pure bipartite quantum state. Here, by reduced states we mean the quantum states ρ^A and ρ^B that describe each party's system. To obtain these states, first we must define maps $\mathrm{Tr}_B : \mathrm{Lin}(\mathcal{H}_{AB}) \to \mathrm{Lin}(\mathcal{H}_A)$ and $\mathrm{Tr}_A : \mathrm{Lin}(\mathcal{H}_{AB}) \to \mathrm{Lin}(\mathcal{H}_B)$, named **partial trace** operations, by setting their action on product operators as

$$\begin{aligned} \mathrm{Tr}_B(X_A \otimes Y_B) &\doteq \mathrm{Tr}(Y_B) X_A, \\ \mathrm{Tr}_A(X_A \otimes Y_B) &\doteq \mathrm{Tr}(X_A) Y_B \end{aligned} \tag{1.37}$$

for any $X_A \in \mathrm{Lin}(\mathcal{H}_A)$ and $Y_B \in \mathrm{Lin}(\mathcal{H}_B)$ and linearly extending them to the entirety of $\mathrm{Lin}(\mathcal{H}_{AB})$[7]. Thence, for a bipartite state ρ_{AB}, we set

$$\begin{aligned} \rho_A &\doteq Tr_B(\rho_{AB}), \\ \rho_B &\doteq Tr_A(\rho_{AB}). \end{aligned} \tag{1.38}$$

Given these definitions, we can state the following proposition:

Proposition 1.1: Bipartite entangled pure states have mixed reduced states

Let $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$ be a pure state. If its reduced states ρ_A and ρ_B are mixed, then $|\psi\rangle_{AB}$ is entangled.

Proof. Suppose that $|\psi\rangle_{AB}$ is separable, *i.e.*, not entangled. Since it is a pure state, there exist states $|u\rangle_A \in \mathcal{H}_A$ and $|v\rangle_B \in \mathcal{H}_B$ such that

$$|\psi\rangle\langle\psi|_{AB} = |u\rangle\langle u|_A \otimes |v\rangle\langle v|_B .$$

Direct application of the partial trace operations results in ρ_A equal to $|u\rangle\langle u|_A$ and ρ_B equal to $|v\rangle\langle v|_B$, which are both pure states. The proposition follows by contrapositive. ∎

[7] This linear extension is possible because any operator $X_{AB} \in \mathrm{Lin}(\mathcal{H}_{AB})$ is a linear combination of product operators.

By noticing that the state $\frac{1}{2}\mathbb{I}$ is mixed and that it is also the reduced state of any state in equation (1.36), a direct application of the previous proposition shows that Bell states are indeed entangled.

The criterion in proposition 1.1 is actually sufficient and necessary to characterize entanglement of bipartite pure states. Necessity can be proven using the following theorem [1, Theorem 2.7]:

Theorem 1.1: Schmidt decomposition

Let $|\psi^{AB}\rangle \in \mathcal{H}_{AB}$ be a pure state and set $d = \min\{\dim\mathcal{H}_A, \mathcal{H}_B\}$. There exists an orthonormal set $\{|0\rangle, \ldots, |d-1\rangle\}$ and nonnegative coefficients such that $\lambda_0, \ldots, \lambda_{d-1}$

$$|\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i |i\rangle_A \otimes |i\rangle_B, \tag{1.39}$$

and $\sum_i \lambda_i^2 = 1$.

Since proposition 1.1 actually provides a necessary and sufficient condition for entanglement in bipartite pure states, it brings a new perspective on mixed states. In section 1.1.1, a mixed state reflected our ignorance about which pure state a system is prepared in. In light of the previous proposition, however, we can also understand the mixedness of ρ_A as arising from its entanglement with another system B, even when we have perfect knowledge of the composite system AB.

1.4.2 Entanglement witnesses

While the previous detection criterion only applies to pure states, a general way to detect entanglement is by means of entanglement witnesses [22, 23].

To define them, first, notice that $\mathrm{Sep}(AB)$ is a convex and closed[8] subset of the real vector space of Hermitian operators. Bearing this in mind, we can make use of the following result [24, Theorem 1.11, adapted]:

Theorem 1.2: Hyperplane separation

Let $\mathrm{Her}(\mathcal{H})$ be the real vector space of Hermitian operators on $\mathcal{H}$, let $C \subset \mathrm{Her}(\mathcal{H})$ be a convex and closed subset and consider a Hermitian operator $\rho \notin C$. There exists an operator $S \in \mathrm{Her}(\mathcal{H})$ and a scalar $\alpha \in \mathbb{R}$ such that

$$\mathrm{Tr}(\rho S) < \alpha \leq \mathrm{Tr}(\sigma S) \tag{1.40}$$

for all $\sigma \in C$.

[8] Meaning that the limit of any convergent sequence of separable states is also a separable state.

That is, the operator S defines a hyperplane $\mathrm{Tr}(SX) = \alpha$ such that every point inside C lies above said hyperplane, whereas the point ρ lies below it. Setting $\mathcal{H} = \mathcal{H}_{AB}$ and $C = \mathrm{Sep}(AB)$, we have the following definition:

Definition 1.2: Entanglement Witness

An operator $W \in \mathrm{Her}(\mathcal{H}_{AB})$ is an entanglement witness if there exists an entangled state ρ_{AB} such that

$$\mathrm{Tr}\left(\rho_{AB} W\right) < 0 \leq \mathrm{Tr}(\sigma W) \tag{1.41}$$

for all $\sigma \in \mathrm{Sep}(AB)$.

Hence, for an arbitrary bipartite state, the existence of an entanglement witness W for it is a necessary and sufficient condition to characterize the state as entangled.

1.4.3 Entanglement quantification and monogamy of entanglement

Besides being able to detect entanglement, quantifying it is also a relevant task as entanglement can be a useful resource in communication scenarios – we will examine such an example in the next chapter. Similarly to detection methods, there exist several entanglement quantifiers. Any reasonable quantifier E, which from now on we refer to as an **entanglement measure**, must satisfy at least the following three properties [25]:

EM1) E maps $\rho \in \mathcal{D}(\mathcal{H}_{AB})$ to a nonnegative number $E(\rho) \in \mathbb{R}_+$, and such that

$$E(|\Phi\rangle\langle\Phi|) = \log_2 d \tag{1.42}$$

for the state $|\Phi\rangle = \frac{1}{\sqrt{d}}(|0\rangle|0\rangle + \ldots + |d-1\rangle|d-1\rangle)$, with $d = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$.

EM2) For any $\rho \in \mathrm{Sep}(AB)$, $E(\rho) = 0$.

EM3) It is non-increasing under local operations aided by classical communications (LOCC). In other words, if we obtain σ from ρ via a LOCC operation, then

$$E(\sigma) \leq E(\rho). \tag{1.43}$$

Other desirable properties might also be required, such as continuity, convexity and a few more [25], which would result in what a "gold standard" entanglement measure should be, but no known E satisfies them all.

Entanglement of formation and concurrence

For bipartite pure states, a wildly adopted entanglement measure is the entanglement entropy. It is defined as

$$E_{ent}(|\psi\rangle\langle\psi|_{AB}) \doteq -\mathrm{Tr}(\rho_A \log \rho_A), \tag{1.44}$$

where ρ^A is obtained using equation (1.38). One can use the Schmidt decomposition (theorem 1.1) and verify that properties **EM1)** to **EM3)** hold[9].

An extension of E_{ent} to mixed states is the **entanglement of formation**. Recall from equation (1.7) and remark 1.4 that any mixed state has a pure state decomposition

$$\rho_{AB} = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|_{AB}. \tag{1.45}$$

From such a decomposition, we can compute the average entanglement entropy

$$\bar{E}_{ent}\left(\{(p_i, |\psi_i\rangle_{AB})\}_i\right) = \sum_{i=1}^{n} p_i E_{ent}\left(|\psi_i\rangle\langle\psi_i|_{AB}\right). \tag{1.46}$$

The entanglement of formation of ρ_{AB} is defined as the infimum over all such averages, *i.e.*

$$E_{for}(\rho_{AB}) \doteq \inf\left\{ \bar{E}_{ent}\left(\{(p_i, |\psi_i\rangle_{AB})\}_i\right) \mid \rho_{AB} = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|_{AB} \right\}. \tag{1.47}$$

Although reasonably motivated, $E_{for}(\rho_{AB})$ is not trivial to compute in general. For the applications which we will discuss in Chapter 4, it suffices to consider the two-qubit case. There, the entanglement of formation is given by

$$E_{for}(\rho_{AB}) = h\left(\frac{1+\sqrt{1-\tau_{AB}}}{2}\right), \tag{1.48a}$$

where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ and the quantity

$$\tau(\rho_{AB}) \doteq (\max\{\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4, 0\})^2, \tag{1.48b}$$

named **tangle**, is computed via the square roots of the eigenvalues (in decreasing order) $\lambda_1, \ldots, \lambda_4$ of the operator

$$R = \rho_{AB}[(\sigma_Y \otimes \sigma_Y)\rho_{AB}(\sigma_Y \otimes \sigma_Y)]. \tag{1.48c}$$

Although the tangle is obtained from a slightly cumbersome procedure, its importance is twofold. First, it can replace the two-qubit entanglement of formation

[9] At least when we consider local unitary operations.

altogether, since equation (1.48a) is monotonic in τ_{AB}. Second, it satisfies an inequality [26, 27] that expresses the fact that entanglement cannot be shared among an unbounded number of parties, a concept called **monogamy of entanglement**.

Suppose that $n+1$ parties, A, $B_1, \ldots, B_n$, share a multi-qubit state $\rho_{AB_1 \ldots B_n}$ (either pure or mixed). If ρ_{AB_i} denotes the two-qubit state between A and i-th party B_i, the Coffman-Kundu-Wootters inequality (CKW) [26, 27] says that

$$\tau(\rho_{AB_1}) + \ldots + \tau(\rho_{AB_n}) \leq 1. \tag{1.49}$$

Since the tangle between any two qubits can be at most 1, inequality (1.49) demonstrates that there is a trade-off between how much Alice's system is entangled with the system of any two distinct Bobs. In this vein, whenever the tangle of ρ_{AB_i} is close to 1, each ρ_{AB_j} is barely entangled.

As an extreme example of this observation, let $|\omega\rangle_{ABC}$ be a three-qubit state such that $\rho_{AB} = \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|_{AB}$, where $\left|\Phi^+\right\rangle$ is the top left Bell state from eq. (1.36). Using equation (1.48b), direct computation yields $\tau(\rho^{AB}) = 1$, meaning that it is maximally entangled according to equation (1.48a). Also, we have $\rho_{AC} \in \text{Sep}(AC)$ and $\rho_{BC} \in \text{Sep}(BC)$ – as consequence of the CKW inequality and of equation (1.48a). Now, direct application of the Schmidt decomposition (thm. 1.1), implies that $|\omega\rangle_{ABC} = \left|\Phi^+\right\rangle_{AB} \left|\varphi\right\rangle_C$ is the only state structure compatible with these reduced-state constraints. Hence, if two systems are in the $\left|\Phi^+\right\rangle$ state (or any Bell state), they cannot be entangled with any other system.

As we will explore in Chapter 4, a quantum theory based on real numbers allows the existence of multipartite states where each two-system reduced state is maximally entangled, something forbidden in standard quantum theory thanks to the previous inequality.

1.5 Quantum channels

After our discussions on composite systems and entanglement, the last element we need to introduce are the valid transformations between quantum systems.

Suppose that one party, Alice, sends a quantum system from her laboratory to another party's (Bob) laboratory via some physical medium. We refer to this medium as a quantum channel from A to B, as it allows Alice to send quantum information to Bob.

Now, since quantum states are positive operators (eq. (1.1)) Alice's action must be described by a map $T : \text{Lin}(\mathcal{H}_A) \to \text{Lin}(\mathcal{H}_B)$ that:

1. Preserves positivity, *i.e.*, for any $X^A \in \mathrm{Lin}(\mathcal{H}_A)$

$$X^A \succcurlyeq_A 0 \implies T(X^A) \succcurlyeq_B 0, \tag{1.50}$$

where $\succcurlyeq_A$ and $\succcurlyeq_B$ denote the order in the appropriate spaces.

2. It is trace-preserving, *i.e.*, $\mathrm{Tr}\left(X^A\right) = \mathrm{Tr}\left(T(X^A)\right)$ for any $X^A \in \mathrm{Lin}(\mathcal{H}_A)$.

3. It is convex-linear in $\mathcal{D}(\mathcal{H}_A)$, meaning that for any convex combination of states $\sum_{i=1}^n p_i \rho_i^A$

$$T\left(\sum_{i=1}^n p_i \rho_i^A\right) = \sum_{i=1}^n p_i T(\rho_i^A). \tag{1.51}$$

Any map that satisfies equation (1.50) is called a **positive map**. *If* we were only concerned with single systems, then trace-preserving, linear positive maps would be natural candidates for quantum channels. After all, it is immediate from the conditions above that any such T maps $\mathcal{D}(\mathcal{H}_A)$ into $\mathcal{D}(\mathcal{H}_B)$.

However, consider a setting as depicted in figure 1.4. Alice now has a composite system AA' described by a (possibly entangled) state $\rho_{AA'}$ and now she sends half of this joint system to Bob via this channel.

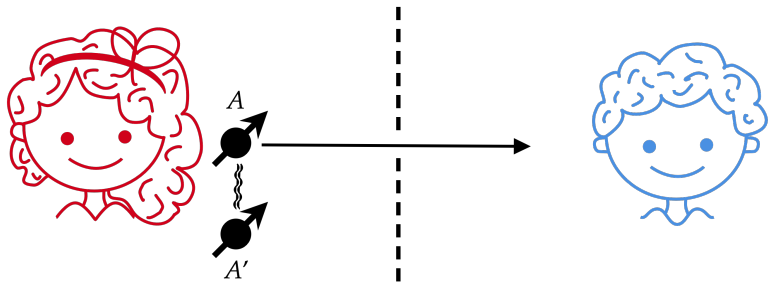


Figure 1.4: Alice possesses two quantum systems, labelled A and A' and sends one of them to Bob via a channel. The dashed line represents the separation between their laboratories.

Since she does not act on A', the output operator is

$$\sigma^{BA'} = (T \otimes I_{A'})\rho^{AA'}, \tag{1.52}$$

where $I_{A'} : \mathrm{Lin}(\mathcal{H}_{A'}) \to \mathrm{Lin}(\mathcal{H}_{A'})$ is the identity map. Even if T is a positive map, it is not true in general that $T \otimes I_{A'}$ will be a positive map.

To illustrate this previous observation, fix an orthonormal basis $\{|i\rangle\}_{i=0}^{d_A-1}$ and let t_A be the transposition map with respect to this basis, *i.e*,

$$t_A\left(|i\rangle\langle j|_A\right) \doteq |j\rangle\langle i|_A. \tag{1.53}$$

Consider the following pure bipartite state:

$$|\Phi\rangle = \frac{1}{\sqrt{d}_A} \sum_{i=0}^{d_A-1} |i\rangle_A \otimes |i\rangle_A . \tag{1.54}$$

The **partial transposition map** $t_a \otimes I_A$ takes $|\Phi\rangle$ to the operator

$$t_A \otimes I_A (|\Phi\rangle\langle\Phi|) = \frac{1}{d_A} \sum_{i,j=0}^{d_A-1} |j\rangle\langle i|_A \otimes |i\rangle\langle j|_A ,$$

which has eigenvalues $+1$ and -1, *i.e.*, it is not a positive operator. As a side note, $t_A \otimes I_A$ is intimately connected with entanglement detection and we refer the reader to refs. [22, 28] for more details.

To remedy cases such as the previous one, we restrict the valid transformations T to the class of completely positive maps:

Definition 1.3: Completely positive map

Let $T : \mathrm{Lin}(\mathcal{H}_A) \to \mathrm{Lin}(\mathcal{H}_B)$ be a positive linear map. It is said to be **k-positive** if $T \otimes I_k : \mathrm{Lin}(\mathcal{H}_A \otimes \mathcal{H}_k) \to \mathrm{Lin}(\mathcal{H}_B \otimes \mathcal{H}_k)$ is a positive map, where $\dim \mathcal{H}_k = k$. It is **completely positive** if it is k-positive for every $k \in \mathbb{N}$.

With this definition in mind, we can proceed and establish that:

Definition 1.4: Quantum channel

Given quantum systems A and B, a quantum channel is a completely positive, trace preserving (CPTP) linear map $T : \mathrm{Lin}(\mathcal{H}_A) \to \mathrm{Lin}(\mathcal{H}_B)$.

Definition 1.4 extends the notion of what is a valid transformation between systems beyond the unitary paradigm, since the input and output spaces need not have the same dimension. This definition, despite its operational motivation, is not so practical, since k-positiveness must be checked for every $k \in \mathbb{N}$. Fortunately, a fundamental result in the theory of positive maps [29] makes this useful (but impractical) condition easier to verify.

Theorem 1.3: Choi representation

A linear map $T : \mathrm{Lin}(\mathcal{H}_A) \to \mathrm{Lin}(\mathcal{H}_B)$ is completely positive if and only if the operator

$$J(T) \doteq (T \otimes I_A)\,|\Omega\rangle\langle\Omega| \tag{1.55}$$

is positive, where

$$|\Omega\rangle = \sum_{i=0}^{d_A-1} |i\rangle \otimes |i\rangle\,, \tag{1.56}$$

for an orthonormal basis $\{0, \ldots, d_A - 1\} \subset \mathcal{H}_A$.

Proof. See the original work [29] or Theorem 4.4.1 in reference [3] for a proof within the language of quantum information theory. ■

2 TELEPORTATION AND BELL NONLOCALITY

I travel in worlds you can't even imagine! You can't conceive of what I am capable of!

Jimmy McGill, *Better Call Saul*

Chapter Contents

In the this chapter, we explore some consequences of the formalism presented in Chapter 1 applied to two distinct problems involving two parties, namely: sending quantum information, and devising strategies to win an input-output game.

2.1 The no-cloning theorem

Before we examine the aforementioned applications, we must present a pivotal result from quantum information theory, which we will have to invoke in the next section.

As discussed in the previous chapter, a classical random bit can be expressed as a convex combination of any pair of perfectly distinguishable states. Let these states be ρ_0 and ρ_1 and denote the discriminating measurement operators by M_0 and M_1. For any $n \in \mathbb{N}$ we can define the following map:

$$T_n(\sigma) = \mathrm{Tr}(\sigma M_0)\rho_0^{\otimes n} + \mathrm{Tr}(\sigma M_1)\rho_1^{\otimes n}. \tag{2.1}$$

Clearly, this is a trace-preserving map for any n. Furthermore, this map is also completely positive, which can be verified using Choi's representation theorem (theorem 1.3). Hence, T_n is a quantum channel.

It follows from perfect distinguishability that $T_n(\rho_i) = \rho_i^{\otimes n}$ for $i \in \{0, 1\}$. From this, we observe that such channels can be used to create as many copies of ρ_0 and ρ_1 as one wants. In other words, the information stored in these states, *i.e.* the bit values 0 and 1, can be arbitrarily cloned.

One might wonder whether the previous observation generalizes to any quantum state, that is, if there is a quantum channel that acts as a universal cloning machine. As first shown by Wootters and Zurek [30] and later generalized by Werner [31], no such channel exists. This is the content of the following theorem:

Theorem 2.1: No-cloning (quantum version)

There is no quantum channel $T : \mathrm{Lin}(\mathcal{H}_A) \to \mathrm{Lin}(\mathcal{H}_A^{\otimes 2})$ such that

$$T(\rho) = \rho \otimes \rho \tag{2.2}$$

for all $\rho \in \mathcal{D}(\mathcal{H}_A)$.

We postpone its proof to Chapter 4, where we will demonstrate that absence of universal cloning is not exclusive to quantum theory, but it is actually a characteristic of any non-classical probabilistic theory. In this vein, classical theory stands as the odd one out as it allows arbitrary cloning of information.

2.2 The teleportation protocol

As our first application of the formalism, let us consider the following setting:

A party named Alice receives a qubit system prepared in an arbitrary state σ^A. She is instructed to transmit this system to a distant party, whom we name Bob, without perturbing its state.

If Alice had access to an arbitrary number of copies of the system, she could find out which state was sent to her by measuring the Pauli observables, and then (classically) communicate to him the angles of the state's Bloch vector (see eq. (1.14)). Due to the no-cloning theorem, there is no procedure that lets her create an arbitrary number of copies of the state σ^A, which hinders her from adopting the previous strategy.

At first, let us assume that Alice and Bob can communicate through a quantum channel (figure 2.1a). If the channel is noiseless – *i.e.*, a map $T(\sigma^A) = \sigma^A$ for any input state – the task is trivialized. Conversely, if it is noisy, she would not always succeed in her task[1].

If we now assume that Alice and Bob only have access to a perfect classical communication channel[2], she can always succeed in her task with the aid of preestablished entanglement between Bob and herself (figure 2.1b). The protocol that ac-

[1] For a given channel T, she only succeeds if σ^A is a fixed point of T.

[2] By a "classical communication channel" we mean any way in which Alice and Bob can communicate bits. For instance, they may call each other's laboratories.

(a) Quantum communication.

(b) Classical comm. + entanglement.

Figure 2.1: Two distinct communication scenarios between Alice and Bob.

complishes this is named the **teleportation protocol** and it was originally proposed in reference [32]. It consists of the following steps:

TP0) Alice and Bob share an entangled system $A'B$ prepared in the Bell state (eq. (1.36))

$$\left|\Phi^+\right\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

TP1) Alice performs a measurement on systems AA' in the Bell basis, whose projectors are

$$\begin{aligned} \Pi_{00} &= \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|, \quad \Pi_{01} = |\Phi^-\rangle\langle\Phi^-|, \\ \Pi_{10} &= \left|\Psi^+\right\rangle\!\left\langle\Psi^+\right|, \quad \Pi_{11} = |\Psi^-\rangle\langle\Psi^-|, \end{aligned} \tag{2.3}$$

and obtains 2 bits $a, a' \in \{0,1\}$.

The (unnormalized) global post-measurement state is

$$\rho^{AA'B}_{aa'} = \Pi^{AA'}_{aa'} \otimes \mathbb{I}^B \left(\sigma^A \otimes \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|^{A'B}\right) \Pi^{AA'}_{aa'} \otimes \mathbb{I}^B. \tag{2.4}$$

TP2) She communicates the pair (a, a') to Bob via a classical channel.

TP3) Based on the received bits, Bob applies a unitary $U_{aa'}$ on his system and obtains the state Alice wanted to send, that is:

$$U_{aa'} \rho^B_{aa'} U^\dagger_{aa'} = \sigma^B, \tag{2.5}$$

where $\rho^B_{aa'} = \mathrm{Tr}_{AA'} \rho^{AA'B}_{aa'}$ and σ^B is the same as σ^A.

The state $\rho^B_{aa'}$ can be computed using the following facts:

Fact 1) In an arbitrary bipartite system XY, given any projective measurement $\{\Pi^X_i\}_i$ on system X:

$$\mathrm{Tr}_Y \left[\left(\Pi^X_i \otimes \mathbb{I}^Y\right) \rho^{XY} \left(\Pi^X_i \otimes \mathbb{I}^Y\right)\right] = \mathrm{Tr}_Y \left[\rho^{XY} \left(\Pi^X_i \otimes \mathbb{I}^Y\right)\right]$$

for any state ρ^{XY}.

Fact 2) For any $a, a' \in \{0,1\}$, the projectors in equation (2.3) satisfy:

$$\Pi_{aa'} = (\mathbb{I} \otimes V_{aa'}) \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right| (\mathbb{I} \otimes V_{aa'}^\dagger),$$

where $V_{aa'} = \sigma_X^a \sigma_Z^{a'}$.

Fact 3) For any linear operators P, Q, R, S:

$$(P \otimes Q)(R \otimes S) = PR \otimes QS.$$

Fact 4) For any operator P:

$$(\mathbb{I} \otimes P) \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right| = \left(\mathbb{I} \otimes P^T\right) \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|,$$

where T represents the transposition operation.

A short calculation results in:

$$\rho_{aa'}^B = \frac{1}{4} V_{aa'} \sigma^B V_{aa'}^\dagger. \tag{2.6}$$

Hence, up to normalization of the state, the unitaries Bob implements are $U_{aa'} = V_{aa'}^\dagger$.

Notice that if Alice does not communicate the bits (a, a') to Bob, then his system is described by

$$\frac{1}{4} \sum_{a,a'} V_{aa'} \sigma^B V_{aa'}^\dagger = \frac{1}{4} \left(\sigma^B + \sigma_X \sigma^B \sigma_X^\dagger + \sigma_Y \sigma^B \sigma_Y^\dagger + \sigma_Z \sigma^B \sigma_Z^\dagger\right) = \frac{1}{2} \mathbb{I}^B, \tag{2.7}$$

since $\sigma_X \sigma_Z = -i\sigma_Y$. This means that he cannot learn anything about the system[3] unless the parties follow the protocol.

Despite its apparent simplicity, the teleportation protocol can be used as a subroutine to establish entanglement at long distances (this is called entanglement swapping), which is highly desirable for the development of a quantum internet.

2.3 Bell nonlocality

The final application of the quantum formalism that we explore in this chapter is a phenomenon known as Bell nonlocality. A seminal work in the study of nonlocality was John Bell's 1964 paper "On the Einstein Podolsky Rosen paradox" [33]. In it, Bell showed that, under mild assumptions, any model that attempts to render quantum

[3]The state in the right-hand side (RHS) of equation (2.7) is called the maximally mixed state and it is a quantum analogue of the discrete uniform distribution.

theory deterministic by supplementing it with additional variables cannot fully reproduce the theory's predictions. This debate traces back at least to the work of the EPR trio [34]. For further details regarding historical aspects, we refer the reader to the entries *The Einstein-Podolsky-Rosen Argument in Quantum Theory* [35] and *Bell's theorem* [36] in the Stanford Encyclopedia of Philosophy.

Beyond its foundational importance, Bell nonlocality has also found practical applications. It forms the basis of device-independent quantum cryptography [37], where security is guaranteed solely by the observed violation of what is called a Bell inequality; and it also enables the self-testing of quantum devices [38].

The remainder of this section is structured as follows. First, we introduce the basics of Bell nonlocality via a simple two-party game. We then present an important result, proved by Boris Tsirelson [39], that relates to the optimal quantum strategy for this game. We end the section by formally introducing the notion of Bell inequalities. For more detailed discussions on nonlocality, we refer the reader to [40, 41] (which form the basis of the following subsections), [42] (only available in Brazilian Portuguese), or [43].

2.3.1 The CHSH game and pre-established agreement

In a modern approach to Bell nonlocality, it is customary to present the subject in terms of multiparty input-output games, which are suitably called *nonlocal games*. Such presentation brings conceptual clarity to the phenomenon, hence we will make use of it. In what follows, we provide one rendition (in the form of a thought experiment) of the CHSH[4] game, which is the simplest nonlocal game there is.

The thought experiment goes as follows: suppose that two siblings, Alice ($\mathcal{A}$) and Bob ($\mathcal{B}$), accept a challenge from a maze's master ($\mathcal{M}$) to reach a prize that lies at the end of his maze. Right before $\mathcal{A}$ and $\mathcal{B}$ enter the maze, the master warns them that, once they are close to the prize room, a sign will indicate that they must separate and each go down a dead-end hall (figure 2.2).

Once each sibling reaches the end of their hall, the maze's master, who is monitoring them in real time, randomly opens a passage either to their left or to their right, closing it immediately after the sibling crosses it. Alice's passage is denoted by $x \in \{0, 1\}$, and similarly Bob's passage is denoted by $y \in \{0, 1\}$.

Behind each passage is a two-door room and each sibling indicates to $\mathcal{M}$ which door they want opened; we label the doors using $a, b \in \{-1, 1\}$. Once the master has registered both players' choices, he only opens the chosen doors if the following

[4]Named after John Clauser, Michael Horne, Abner Shimony and Richard Holt, who together proposed an experiment that could test a generalization of Bell's 1964 inequality [44].

condition is met:

$$ab = (-1)^{xy}. \tag{2.8}$$

In other words, if Alice and Bob both walk past a right passage ($x = y = 1$), they must choose distinct doors, whereas they must choose the same doors in all remaining cases. In case the siblings' choices satisfy equation (2.8), they walk past the doors and reach the prize. Otherwise, $\mathcal{M}$ expels them from the maze, rearranges all of its corridors, and challenges the siblings once more.

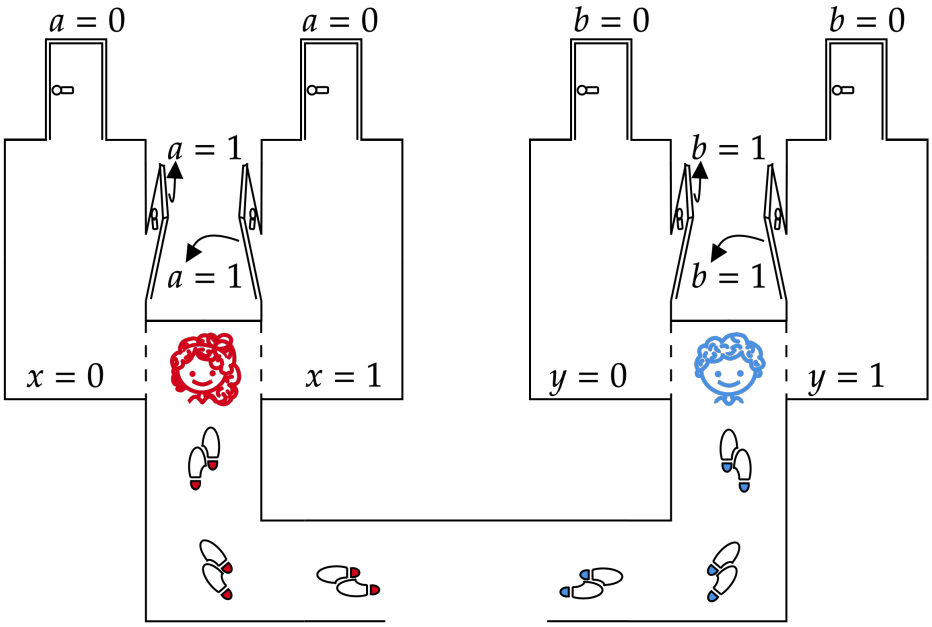


Figure 2.2: The CHSH game.

With this setup in mind, we are interested in how well Alice and Bob may perform in this game. More formally, we assume that they play the CHSH game for many rounds and, at the start of each round, the siblings agree on a strategy[5] λ prior to entering the maze. The number of rounds is sufficient for $\mathcal{M}$ to correctly estimate the probabilities

$$p(a, b|x, y) = \int_{\lambda\in\Lambda} q(\lambda)\, p(a, b|x, y, \lambda)\mathrm{d}\lambda, \tag{2.9}$$

for any pair x and y. The probabilities $p(a, b|x, y, \lambda)$ dictate how Alice and Bob behave according to a strategy λ and the previous expression captures the idea that the siblings are free to choose whatever strategy they wish according to a probability distribution $q(\lambda)$ defined on the set of all strategies Λ.

[5]Here, by strategy we mean any process that Alice and Bob can use to help them choose their doors. For instance, a particular λ can represent a strategy where Alice and Bob choose their doors in a deterministic fashion, *i.e.*, as functions $a = f(x)$ and $b = g(y)$. Another possibility would be one in which λ represents a strategy where they choose their doors by means of some random process (like the tossing of a coin).

From the game's setup, for each λ we have

$$p(a,b|x,y,\lambda) = p(a|x,\lambda)\,p(b|y,\lambda), \tag{2.10}$$

since each sibling's choice does not depend on which passage the other one went through, but only on their respective passage and on whatever strategy they had agreed before going on separate ways. Consequently:

$$p(a,b|x,y) = \int_{\lambda\in\Lambda} q(\lambda)\,p(a|x,\lambda)\,p(b|y,\lambda). \tag{2.11}$$

Any distribution that admits a decomposition as in equation (2.11) is called **local**, due to the factorization in equation (2.10). We name the collection $\{p(a,b|x,y)\}_{x,y,a,b}$ as the **behaviour** of Alice and Bob.

For each x, y, let A_x and B_y represent the binary random variables associated to how Alice and Bob choose a door. We can evaluate how well the siblings perform using the following quantity as a score:

$$\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle\,,$$

where expectations are calculated using the probabilities $p(a,b|x,y)$:

$$\left\langle A_xB_y\right\rangle = \sum_{a,b\in\{-1,1\}} abp(a,b|x,y).$$

From the interpretation of equation (2.8), we see that the maximal value of this quantity is 4, and it is achieved if they always win the game, i.e. if their choices always satisfy eq. (2.8).

Whenever Alice and Bob's behaviour is a local one, it can be shown (see Section 1.3 and theorem 2.1 from [40]) that:

$$\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leq 2. \tag{CHSH}$$

This inequality, from now on referred as the CHSH inequality, shows that, as long as Alice and Bob produce their answers according to some pre-established shared randomness (encoded in λ and $q(\lambda)$), there exists an upper bound to how well they can perform in the CHSH game.

Informally, a *Bell inequality* is an inequality at the level of the behaviours obtained in an experiment and it attests whether such behaviours admit a local decomposition. From the many known Bell inequalities [43], CHSH ranks amongst the most well-known ones.

In light of the last paragraph, the fact that quantum theory can violate Bell inequalities (example 2.1 below) shows that certain quantum correlations cannot be

generated from the same processes that generate local correlations. Thus the name **nonlocality**.

2.3.2 Sufficient conditions for locality in quantum systems & Tsirelson's bound

Before we show what is the maximal violation of the CHSH inequality allowed by quantum theory, it is important to assess *when* quantum systems produce local correlations.

First of all, we need the following definition:

Definition 2.1: Quantum behaviour

Given finite sets $\mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}$ and a behaviour $\{p(a,b|x,y)\}_{(a,b,x,y)\in\mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}}$, we say that this behaviour is a quantum behaviour if and only if there exist Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$, a quantum state $\rho \in \mathcal{D}(\mathcal{H}_{AB})$, and POVMs $\{E_{a|x}\}_{a,x} \subset \mathrm{Lin}(\mathcal{H}_A)$, $\{F_{b|y}\} \subset \mathrm{Lin}(\mathcal{H}_B)$ such that

$$p(a,b|x,y) = \mathrm{Tr}\big(\rho E_{a|x} \otimes F_{b|y}\big). \tag{2.12}$$

Since any quantum behaviour depends both on a state and on measurements, the fact that it is local might stem from at least one of the previous two objects. The next propositions make this observation more precise.

For states, the following is sufficient:

Proposition 2.1: Separability implies local behaviour

Let ρ^{AB} be a bipartite quantum state and consider any choice of local POVMs $\{E_{a|x}\}_{x,a}$ and $\{F_{b|y}\}_{b,y}$ for each x and y. If ρ^{AB} is separable, then the behaviour with probabilities as in equation (2.12) is local.

Proof. Let $\{E_{a|x}\}_{x,a}$ and $\{F_{b|y}\}_{b,y}$ be as stated and assume that ρ^{AB} is separable, i.e.

$$\rho^{AB} = \sum_{k=1}^{n} q_k \rho_k^A \otimes \rho_k^B,$$

for states $\rho_k^A \in \mathcal{D}(\mathcal{H}_\mathrm{A})$, $\rho_k^B \in \mathcal{D}(\mathcal{H}_\mathrm{B})$ and a probability distribution $\{q_k\}_{k=1}^n$. A quantum behaviour using these measurements and state has the following form:

$$p(a,b|x,y) = \sum_{k=1}^{n} q_k \,\mathrm{Tr}\big[(\rho_k^A \otimes \rho_k^B)(E_{a|x} \otimes F_{b|y})\big].$$

Using that $(P\otimes Q)(R\otimes S) = (PR)\otimes(QS)$ for any operators P, Q, R and S, and $\mathrm{Tr}(P\otimes Q) =$

$\mathrm{Tr}(P)\,\mathrm{Tr}(Q)$ we can rewrite the equality above as

$$p(a,b|x,y) = \sum_{k=1}^{n} q_k \,\mathrm{Tr}\big(\rho_k^A E_{a|x}\big)\,\mathrm{Tr}\big(\rho_k^B F_{b|y}\big).$$

Set $\Lambda = \{1,\dots,n\}$, $\lambda \equiv k$, $p(a|x,\lambda) = \mathrm{Tr}\big(\rho_k^A E_{a|x}\big)$ and $p(b|y,\lambda) = \mathrm{Tr}\big(\rho_k^B F_{b|y}\big)$. Then $p(a,b|x,y)$ is local. ■

As for measurements, first recall that two observables are compatible if and only if they commute. In terms of POVMs (equations (1.12a) and (1.12b)), this condition is reformulated as follows:

Definition 2.2: Compatible measurements

Two POVMs $\{E_i\}_{i\in I}, \{E'_j\}_{j\in J} \subset \mathrm{Lin}(\mathcal{H})$ are called compatible if there exists a third POVM $\{G_{ij}\}_{(i,j)\in I\times J}$ such that

$$\sum_{j\in J} G_{ij} = E_i \quad \text{and} \quad \sum_{i\in I} G_{ij} = E'_j. \tag{2.13}$$

We then have another sufficient condition for locality in quantum theory:

Proposition 2.2: Compatibility implies local behaviour

Let ρ^{AB} be a bipartite quantum state, let $\{E_{a|x}\}_{x,a}$ and $\{F_{b|y}\}_{b,y}$ be POVMs for each x and y. If the measurements in one of these families are all compatible, then the behaviour with probabilities as in equation (2.12) is local.

We prove the theorem for the case with only two measurements, *i.e.* x or y belong to $\{0,1\}$. The arguments we provide are easily adaptable to an arbitrary number of measurements.

Proof. Let ρ^{AB}, $\{E_{a|x}\}_{x,a}$ and $\{F_{b|y}\}_{b,y}$ be as stated. Without loss of generality, suppose that all of Alice's measurements are compatible. Denote by a_x the outcome a obtained when measurement x is performed. Compatibility between $\{E_{a|0}\}_a$ and $\{E_{a|1}\}_a$ means that

$$E_{a|0} = \sum_{a_1} G_{aa_1} \quad \text{and} \quad E_{a|1} = \sum_{a_0} G_{a_0 a},$$

where $\{G_{a_0a_1}\}_{(a_0,a_1)\in\mathcal{A}\times\mathcal{A}}$ is a POVM. This is equivalent to the existence of a POVM $\{G_\lambda\}_{\lambda\in\Lambda}$ and probabilities $\{p(a|x=0,\lambda)\}_{a\in\mathcal{A}}$, $\{p(a|x=1,\lambda)\}_{a\in\mathcal{A}}\}$ such that

$$E_{a|0} = \sum_{\lambda} p(a|x=0,\lambda) G_\lambda \quad \text{and} \quad E_{a|1} = \sum_{\lambda} p(a|x=1,\lambda) G_\lambda.$$

To see this, set $\Lambda = \mathcal{A}\times\mathcal{A}$, $\lambda = (a_0,a_1)$, $p(a|x,(a_0,a_1)) = \delta_{aa_x}$ and $G_\lambda = G_{a_0a_1}$. If $p(a,b|x,y)$

is a quantum behaviour obtained from the state and measurements under consideration, direct computation yields

$$p(a,b|x,y) = \sum_{\lambda\in\Lambda} p(a|x,\lambda)\,\mathrm{Tr}\big(\rho^{AB} G_\lambda \otimes F_{b|y}\big).$$

Now, set $q(\lambda) = \mathrm{Tr}\big(\rho^{AB} G_\lambda \otimes \mathbb{I}_B\big)$. The normalized conditional state in system B is:

$$\rho^B_\lambda = \frac{\mathrm{Tr}_A(\rho^{AB} G_\lambda \otimes \mathbb{I}_B)}{q(\lambda)}.$$

Consequently:

$$\mathrm{Tr}\big(\rho^{AB} G_\lambda \otimes F_{b|y}\big) = \mathrm{Tr}\big(\rho^{AB}(G_\lambda \otimes \mathbb{I}_B)(\mathbb{I}_A \otimes F_{b|y})\big) = \mathrm{Tr}\big(\mathrm{Tr}_A(\rho^{AB} G_\lambda \otimes \mathbb{I}_B) F_{b|y}\big) = q(\lambda)\,\mathrm{Tr}\big(\rho^B_\lambda F_{b|y}\big).$$

Define $p(b|y,\lambda) = \mathrm{Tr}\big(\rho^B_\lambda F_{b|y}\big)$. Then

$$p(a,b|x,y) = \sum_{\lambda\in\Lambda} q(\lambda)\,p(a|x,\lambda)\,p(b|y,\lambda),$$

which is a local behaviour. ■

As consequence of propositions 2.1 and 2.2, any statistics that does not admit a local description can only be obtained using entangled states and incompatible measurements. It is a well-known fact that quantum theory allows violations of the CHSH inequality. We exemplify this below.

Example 2.1: Quantum theory violates the CHSH inequality

Consider that Alice and Bob use quantum systems to determine their choices in the CHSH game. Using the state

$$\rho^{AB} = |\Psi^+\rangle\langle\Psi^+| \tag{2.14a}$$

and the observables

$$\begin{aligned} A_0 &= \sigma_Z, \quad A_1 = \sigma_X, \\ B_0 &= \frac{\sigma_X + \sigma_Z}{\sqrt{2}}, \quad B_1 = \frac{\sigma_X - \sigma_Z}{\sqrt{2}}, \end{aligned} \tag{2.14b}$$

they obtain

$$\langle A_0 \otimes B_0\rangle + \langle A_0 \otimes B_1\rangle + \langle A_1 \otimes B_0\rangle - \langle A_1 \otimes B_1\rangle = 2\sqrt{2}.$$

The next theorem shows that the violation of the CHSH inequality obtained in the previous example is rather special: it is the maximal value of the CHSH game allowed by quantum theory itself. This was first noticed by Tsirelson in [39] and

nowadays one typically refers to the maximal quantum violation of a Bell inequality as its **Tsirelson's bound**.

Theorem 2.2: Tsirelson's bound of the CHSH inequality

Let $\rho^{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a bipartite quantum state. For any choice of local observables $A_0, A_1 \in \mathrm{Her}(\mathcal{H}_A)$ and $B_0, B_1 \in \mathrm{Her}(\mathcal{H}_B)$ with eigenvalues in $[-1,1]$, define the CHSH operator as

$$S_{CHSH} \doteq A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \tag{2.15}$$

We have that

$$\left|\mathrm{Tr}\left(\rho^{AB} S_{CHSH}\right)\right| \leq 2\sqrt{2}, \tag{2.16}$$

and there is a choice of state and observables that attains the value $2\sqrt{2}$.

Proof. Choosing ρ^{AB} and A_0, A_1, B_0, B_1 as in example 2.1 attains the right-hand side (RHS) of inequality (2.16). We follow [40, Section 3.2.2] for the remainder of the proof.

The CHSH operator can be written as

$$S_{CHSH} = P - Q,$$

where P and Q are positive operators, and, in particular:

$$P = \frac{1}{\sqrt{2}}(A_0^2 \otimes \mathbb{I}_B + A_1^2 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_0^2 + \mathbb{I}_A \otimes B_1^2).$$

Now, regarding a Hermitian operator X, its operator norm is given by

$$\|X\|_\infty = \max_{\psi \in \mathcal{H}_{AB} : \langle\psi|\psi\rangle = 1} \left|\langle\psi|X|\psi\rangle\right|.$$

Additionally, if $X \succcurlyeq Y$ (*i.e. their difference is a positive operator)*, then

$$\|Y\|_\infty \leq \|X\|_\infty.$$

Noting that $P \succcurlyeq S_{CHSH}$, the previous observation implies that

$$\|S_{CHSH}\|_\infty \leq \left\| \frac{1}{\sqrt{2}}(A_0^2 \otimes \mathbb{I}_B + A_1^2 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_0^2 + \mathbb{I}_A \otimes B_1^2) \right\|_\infty.$$

Using the triangle inequality and $\|\alpha X\|_\infty = |\alpha| \|X\|_\infty$ for any $\alpha \in \mathbb{C}$, we obtain

$$\|S_{CHSH}\|_\infty \leq \frac{1}{\sqrt{2}} \left(\left\| A_0^2 \otimes \mathbb{I}_B \right\|_\infty + \left\| A_0^2 \otimes \mathbb{I}_B \right\|_\infty + \left\| \mathbb{I}_A \otimes B_0^2 \right\|_\infty + \left\| \mathbb{I}_A \otimes B_1^2 \right\|_\infty \right).$$

Now, the norm of each operator in the RHS is the product of the norms of each factor. Also, it is immediate from the definition of $\|\cdot\|_\infty$ that the identity operator has norm 1. Thus the RHS is equal to:

$$\frac{1}{\sqrt{2}}\left(\left\|A_0^2\right\|_\infty + \left\|A_1^2\right\|_\infty + \left\|B_0^2\right\|_\infty + \left\|B_1^2\right\|_\infty\right).$$

For any Hermitian operator X, one can check that $\left\|X^2\right\|_\infty = \|X\|_\infty^2$. Additionally, since each of the operators above has eigenvalues in $[-1,+1]$, their expected values also belong to this same interval, irrespective of the pure state $|\psi\rangle$. Thus, all of the four operators have norm equal or less than 1. Consequently, the sum above is equal or less than $2\sqrt{2}$. Thus:

$$\|S_{CHSH}\|_\infty \leq 2\sqrt{2}.$$

In other words, for any pure state $|\psi\rangle \in \mathcal{H}_{\mathrm{AB}}$:

$$\left|\mathrm{Tr}\left(|\psi\rangle\langle\psi| S_{CHSH}\right)\right| \leq 2\sqrt{2}.$$

Since any quantum state is at most a convex combination of pure states, the inequality also holds for arbitrary $\rho^{\mathrm{AB}} \in \mathcal{D}(\mathcal{H}_{\mathrm{AB}})$. ■

Remarkably, since the maximal attainable value is less than 4, the previous theorem demonstrates that not even quantum theory allows Alice and Bob to devise strategies that always win the CHSH game. In the next chapter we will discuss how a hypothetical system, called a *generalized bit*, can be used to always win this game.

2.3.3 Bell inequalities

We end our discussion of Bell nonlocality by briefly presenting the broader framework within which one can discuss settings that extend the one from Section 2.3.1.

Abstracting away from siblings, mazes, and prizes, the CHSH game is specified by the following: a number of parties, each of whom possess a physical system; a collection of possible measurements each party may perform on their respective system; and the set of observable outcomes of each measurement. The conjunction of these three concepts specifies a **Bell scenario**.

Definition 2.3: n-party Bell scenario

For a given $n \in \mathbb{N}$, $n \geq 2$, an n-party Bell scenario is a $2n$-uple $(\mathcal{X}_1,\ldots,\mathcal{X}_n;\mathcal{O}_1,\ldots,\mathcal{O}_n)$ where:

1. Each $\mathcal{X}_i$ is a finite set whose elements are labels of possible measurement choices.

2. Each $\mathcal{O}_i$ is a collection of finite sets, and to each $x \in \mathcal{X}_i$ corresponds an unique $O_x \in \mathcal{O}_i$. The set O_x represents the outcomes of measurement x.

Alternatively, an n-party Bell scenario is specified by a triple (n, m, r), where n denotes the number of parties, m denotes the number of measurements of each party, and r denotes the number of outcomes of each measurement.

In light of the previous definition, the CHSH game determines an inequality of the $(2,2,2)$ Bell scenario. A 4-uple that realizes this scenario is given $\mathcal{X}_1 = \mathcal{X}_2 = \{0,1\}$ and $\mathcal{O}_1 = \mathcal{O}_2 = \{\{-1,1\}\}$. The fact that the outcomes sets of both parties are equal is represent by $\mathcal{O}_i$ consisting of one and the same set. In figure 2.3, we depict generic examples of 2-party and 3-party Bell scenarios.

We emphasize that in a standard Bell scenario, there is a single source that produces and distributes the physical systems to each party. This assumption could be relaxed, and one could also include more than one source, as depicted in figure 2.4. The signature of nonlocality in these more elaborate setups characterizes a phenomenon named Bell nonlocality in networks. For an introduction to this matter, we refer the reader to the review in ref. [45]. The Bell scenarios discussed in this thesis will always consist of a single source.

The **behaviour** associated to an n-party Bell scenario is the collection of all probabilities $p(a_1, \ldots, a_n | x_1, \ldots, x_n)$. The main point of a Bell scenario is to investigate the correlations between the parties' measurement outcomes, under the assumption that they perform their experiments under conditions that prevent them from communicating with each other.

The assumption introduced in the last paragraph is called the **no-signalling condition**. Note that if two parties can communicate with each other while performing their experiments, then their systems could potentially influence one another, because they are causally connected. This connection might create additional correlations between the systems, correlations which did not originate from the source that prepared the systems.

In light of the previous observation, the no-signalling condition is an imposition that must be made to assure that the only causal influences between the systems arise from the source that distributed them.

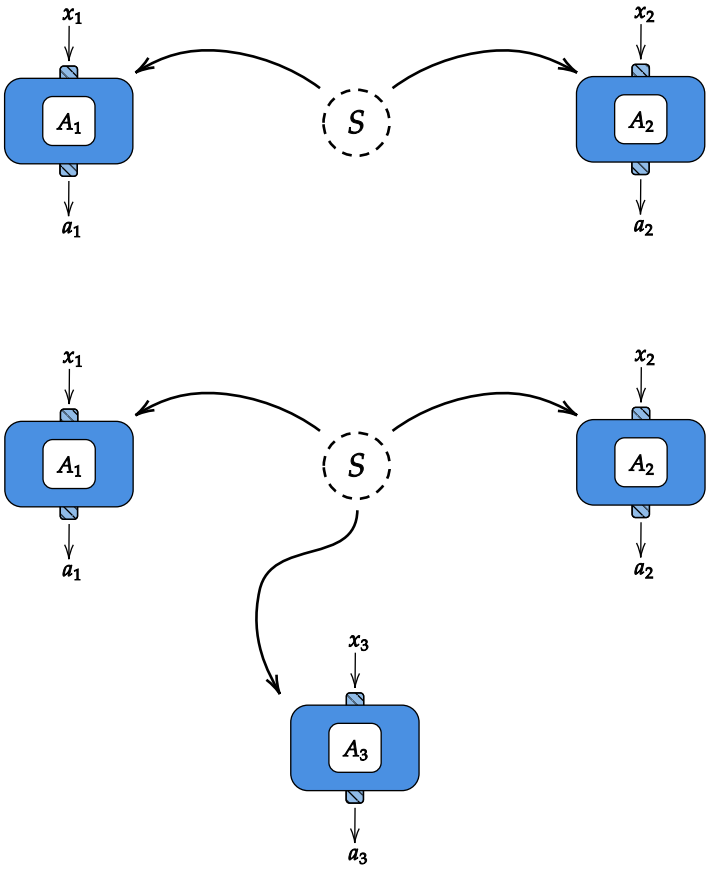


Figure 2.3: Examples of standard Bell scenarios. A source S distributes physical systems to parties A_i. Each party chooses a measurement, denoted by x_i, from a set of possibilities. After their measurements are performed, each party reads an outcome a_i. Details about how measurements are performed and how outcomes are obtained are cast aside and the whole process is represented by a black-box.

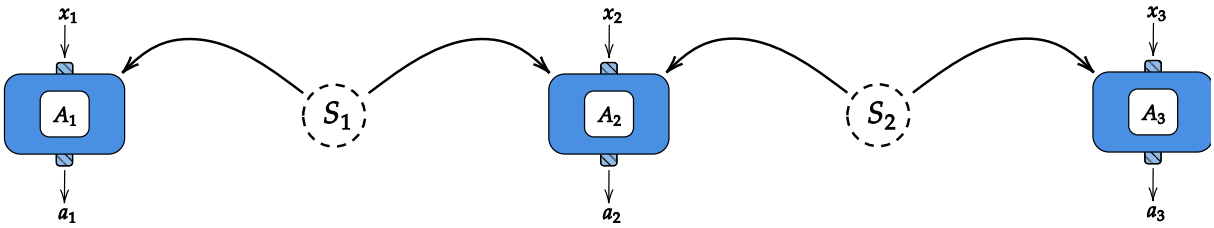


Figure 2.4: An example of a Bell scenario with three parties and more than one source. In this setting, source S_1 prepares and distributes systems to parties A_1 and A_2, whereas S_2 prepares and distributes systems to parties A_2 and A_3.

At the level of the probabilities $p(a_1,\ldots a_n|x_1,\ldots,x_n)$, the no-signalling condition means that for any subset of parties $\{i_1,\ldots,i_k\} \subset \{1,\ldots,n\}$, we have:

$$p(a_{i_1},\ldots,a_{i_k}|x_1,\ldots,x_n) = p(a_{i_1},\ldots,a_{i_k}|x_{i_1},\ldots,x_{i_k}), \tag{2.17}$$

for all $x_{i_j} \in \mathcal{X}_{i_j}$ and for all $a_{i_j} \in O_{x_{i_j}}$. In other words, the outcomes obtained by this subset of parties only rely on their measurement choices.

If a behaviour has a quantum realization (in the sense of eq. (2.12), but with more parties), direct computation shows that the no-signalling condition holds. This

prompts the following question: *do all non-signalling behaviours have a quantum realization?* In the following chapter we will provide an example that answers this question in the negative. Consequently, despite the no-signalling condition being in accordance with relativistic causality, it is not a strong-enough of an imposition to completely characterize the correlations produced by quantum systems in Bell scenarios.

Although the probabilities of a non-signalling behaviour might not be quantum realizable, the incapacity stems only from the state side of the story. As shown in ref. [46], one can always realize a no-signalling behaviour using quantum measurements.

Proposition 2.3: Non-signalling behaviours and POVMs

In an n-party Bell scenario $(\mathcal{X}_1,\ldots,\mathcal{X}_n;\mathcal{O}_1,\ldots,\mathcal{O}_n)$, a behaviour satisfies the no-signalling condition if and only if there exists a Hermitian operator ω and local POVMs $\{E_{a_i|x_i}\}_{a_i\in O_{x_i}}$, $x_i \in \mathcal{X}_i$, such that

$$p(a_1,\ldots,a_n|x_1,\ldots,x_n) = \mathrm{Tr}\big(\omega E_{a_1|x_1}\otimes\ldots\otimes E_{a_n|x_n}\big). \tag{2.18}$$

Proof. See Theorem 1 in [46]. ■

Proceeding, in order to define what a Bell inequality is, we have to restrict ourselves to a particular class of non-signalling behaviours: those that admit a causal explanation that is local. Since no-signalling forbids any other correlations between the involved systems, other than those generated by the source, it sounds reasonable that, besides their measurement choices, the only other factor that influences each party's outcome pertains to information about the configuration in which the source produced the systems. Setting λ as a variable that embodies this information, we assume that:

$$p(a_1,\ldots.a_n|x_1,\ldots,x_n,\lambda) = p(a_1|x_1,\lambda)\ldots p(a_n|x_n,\lambda). \tag{2.19}$$

In ordinary language, λ is a latent variable that makes $p(a_1,\ldots.a_n|x_1,\ldots,x_n)$ conditionally independent. If Λ represents the set of all values λ can take, then in general, λ is distributed according to some probability distribution q over Λ. The law of total probability in conjunction with equation (2.19) imply the decomposition

$$p(a_1,\ldots,a_n|x_1,\ldots,x_n) = \int_\Lambda q(\lambda)p(a_1|x_1,\lambda)\ldots p(a_n|x_n,\lambda)\mathrm{d}\lambda. \tag{2.20}$$

Whenever it is possible to write the probabilities of a behaviour as in equation (2.20), we say that this behaviour if **fully local**.

We point out that in the 2-party case, a fully local behaviour is the the same as a local behaviour. For three or more parties the situation is more delicate, since λ

might make $p(a_1,\ldots,a_n|x_1,\ldots,x_n)$ only partially factorable. In any case, we have the following definition:

Definition 2.4: Fully local set and Bell inequality

Let $(\mathcal{X}_1,\ldots,\mathcal{X}_n;\mathcal{O}_1,\ldots,\mathcal{O}_n)$ be an n-party Bell scenario. The **fully local set** of this scenario, denoted by $\mathcal{L}$, is the collection of all of its behaviours whose probabilities are fully local. A **Bell inequality** is any inequality of the form

$$\vec{c}\cdot\vec{p}(\vec{a}|\vec{x}) \leq I_{\mathcal{L}}, \tag{2.21}$$

where $\vec{p}(\vec{a}|\vec{x})$ is the vector whose components are $p(a_1,\ldots,a_n|x_1,\ldots,x_n)$, $\vec{c}$ is a real vector and $I_{\mathcal{L}}$ is the largest value of $\vec{c}\cdot\vec{p}(\vec{a}|\vec{x})$ attainable using fully local behaviours.

If a behaviour violates inequality (2.21), all we can say is that it is not fully local. For example, in the 3-party case, this signifies that $p(a_1,a_2,a_3|x_1,x_2,x_3)$ could still admit one of the the following factorizations

$$p(a_1,a_2,a_3|x_1,x_2,x_3) = \int_{\Lambda_1} q_{A_1|A_2A_3}(\lambda_1)p(a_1|x_1,\lambda_1)p(a_2,a_3|x_2,x_3,\lambda_1)\mathrm{d}\lambda_1 \tag{2.22a}$$

or

$$p(a_1,a_2,a_3|x_1,x_2,x_3) = \int_{\Lambda_2} q_{A_2|A_1A_3}(\lambda_2)p(a_2|x_2,\lambda_1)p(a_1,a_3|x_1,x_3,\lambda_2)\mathrm{d}\lambda_2 \tag{2.22b}$$

or

$$p(a_1,a_2,a_3|x_1,x_2,x_3) = \int_{\Lambda_3} q_{A_3|A_1A_2}(\lambda_3)p(a_3|x_3,\lambda_1)p(a_1,a_2|x_1,x_2,\lambda_3)\mathrm{d}\lambda_3. \tag{2.22c}$$

It could even be a convex combination of the decompositions in equations (2.22a) to (2.22c). In any of these instances, we say that it is a **biseparable behaviour**. This condition expresses that $p(a_1,a_2,a_3|x_1,x_2,x_3)$ is local only relative to two-element partitions of the set $\{A_1,A_2,A_3\}$. We highlight that, since $p(a_1,a_2,a_3|x_1,x_2,x_3)$ is non-signalling, if it is also biseparable, then the probability distributions of parties that have been grouped together should be non-signalling with respect to the remaining party. This however, does not prevent signalling between parties that have been grouped together.

Any inequality that also rules out the possibility of biseparable decompositions certifies (in the 3-party case) that the correlations amongst the parties' systems are genuinely tripartite. An example of such an inequality was proposed by Svetlichny in ref. [47], which was later generalized to settings with more parties [48, 49].

The setting is that of a $(3,2,2)$ Bell scenario. Let us relabel the parties as A, B

and C, and their respective outcome sets as $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ – all of which are taken to be $\{0,1\}$. If A_x, B_y, C_z are quantities assuming values in $\{-1,1\}$ for each $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$, then the inequality reads as:

$$|\langle A_0B_0C_1\rangle + \langle A_0B_1C_1\rangle + \langle A_1B_0C_1\rangle - \langle A_1B_1C_1\rangle + \langle A_1B_1C_0\rangle + \langle A_1B_0C_0\rangle + \langle A_0B_1C_0\rangle - \langle A_0B_0C_0\rangle| \leq 4, \tag{2.23}$$

where for each triple (x, y, z), the expected value is given by

$$\langle A_xB_yC_Z\rangle = \sum_{a,b,c\in\{-1,1\}} abcp(a,b,c|x,y,z).$$

By grouping parties A and B together, we can rewrite it as

$$|\langle (A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1)C_1\rangle + \langle (A_1B_1 + A_1B_0 + A_0B_1 - A_0B_0)C_0\rangle| \leq 4. \tag{2.24}$$

Thus, we can associate Svetlichny's inequality to a game in which parties A and B either play the original CHSH game or a modified version, denoted by $CHSH'$, depending on whether party C chooses to measure C_0 or C_1. In turn, this implies that Tsirelson's bound for Svetlichny's inequality is $4\sqrt{2}$. More concretely:

Example 2.2: Quantum theory violates Svetlichny's inequality

Consider the following tripartite state

$$\rho_{GHZ} = |GHZ\rangle\langle GHZ|, \tag{2.25a}$$

where

$$|GHZ\rangle \doteq \frac{|000\rangle + |111\rangle}{\sqrt{2}},$$

and the following observables:

$$\begin{aligned} \hat{A}_0 &= \sigma_Y, \quad \hat{A}_1 = -\sigma_X, \\ \hat{B}_0 &= \frac{\sigma_X + \sigma_Z}{\sqrt{2}}, \quad \hat{B}_1 = \frac{\sigma_X - \sigma_Z}{\sqrt{2}}, \\ \hat{C}_0 &= -\sigma_Y, \quad \hat{C}_1 = \sigma_X. \end{aligned} \tag{2.25b}$$

Inspired by the left-hand side of equation (2.24), define

$$S_3 = S_{CHSH} \otimes \hat{C}_1 + S_{CHSH'} \otimes \hat{C}_0. \tag{2.26}$$

For our particular choice of state and observables, we have:

$$\mathrm{Tr}\big(\rho_{GHZ} S_3\big) = 4\sqrt{2}.$$

As presented in reference [49], one can recursively extend Svetlichny's setting to an arbitrary number $n \in \mathbb{N}$ of parties.

1. For $n = 2$, and binary variables $A_0^{(1)}, A_1^{(1)}, A_0^{(2)}, A_1^{(1)}$, set

$$s_2 \doteq A_0^{(1)} A_0^{(2)} + A_0^{(1)} A_1^{(2)} + A_1^{(1)} A_0^{(2)} - A_1^{(1)} A_1^{(2)}, \tag{2.27a}$$

   which is just the polynomial associated to the CHSH inequality.

2. For $n \geq 3$, set

$$s_n \doteq s_{n-1} A_1^{(n)} + s'_{n-1} A_0^{(n)}, \tag{2.27b}$$

   where s'_{n-1} is obtained by the mapping $A_0^{(1)} \longrightarrow -A_1^{(1)}$ and $A_1^{(1)} \longrightarrow A_0^{(1)}$.

In other words, for $n \geq 3$, we construct s_n by grouping the first $n-1$ parties together and letting the n-th party's choice of observable determine which $(n-1)$-party Svetlichny game the remaining parties will play.

We observe that there is nothing special in separating the last party from the previous ones, the resulting polynomial can be factored as in equation (2.27b) using any other party. What is being tested by an n-party Svetlichny setting is whether correlations admit "1 vs $n-1$" factorizations, which we still call biseparable, or whether they are genuinely n-partite.

As mentioned in reference [40], for each $n \geq 2$, using biseparable correlations we have:

$$|\langle s_n \rangle| \leq_{Bisep} 2^{n-1}. \tag{2.28}$$

The maximal value achievable using quantum correlations in the n-party Svetlichny's setting is $2^{n-\frac{1}{2}}$. This represents a $\sqrt{2}$ advantage with respect to biseparable correlations.

In chapter 4, we will examine whether a modified version of quantum theory allows any form of genuine multipartite nonlocality that surpasses Tsirelson's bound for an n-party Svetlichny setting.

Part II

Foil theories

3 Generalized Probabilistic Theories

> I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space anymore.
>
> John von Neumann, in a letter to Garrett Birkhoff (1935)

Chapter Contents

3.1 The story so far and the next step

In the previous chapters we discussed several aspects about quantum theory: how quantum states differ from classical ones; the impossibility of universally cloning quantum information; the capacity to assist classical communication using entanglement; and the existence of quantum correlations that have no classical explanation in terms of pre-established agreement.

All of these aspects highlight distinct ways in which the theory departs from our classical experience. If one were asked why quantum theory admits such deviations, the plain answer would be "Because of the postulates" – after all, they are the basis of the formalism. To a certain extent, this is unsatisfying, for the postulates are devoid of any physical content. Hence, it would be far more compelling if an answer to questions related to the power and limitations of quantum theory could be traced back to physical grounds. That would require one to find a set of physical principles that allows us to derive (or at least justify) the mathematical structure of the theory, in a fashion similar to how the postulates of special relativity lead to the Lorentz group as the symmetry group of flat spacetime. The line of research in this direction characterizes the Reconstruction Programme of quantum theory [50, 51].

A sensible starting point that sets us in the direction of achieving the programme's ultimate goal is to develop mathematical frameworks that let us "look at quantum theory from the outside" [13]. One such framework, largely adopted in the quantum foundations community over the past few years, is the Generalized Probabilistic Theories (GPTs) framework [13, 52–54]. It allows us to undress quantum theory from its Hilbert space garments and contrast it with different classes of probabilistic theories, each of which reproduces some of its features, so that we can gain a better understanding of quantum theory itself.

Over the course of this and the following chapter, we will present the GPT framework and explore how some of the features of quantum theory discussed in the previous chapters might be similar or different in generic probabilistic theories.

3.2 The GPT framework

3.2.1 Operational Primitives

The mathematical framework of generalized probabilistic theories [13, 53] is built upon a minimal understanding of what the role of a physical theory is. More precisely, it is desirable that any and all candidate physical theories must be capable of (a) describing physical systems, and (b) making successful predictions about these systems [10]. Based on this understanding, the GPT framework adopts operational principles to erect the arena in which theories are formulated.

Under operational lenses, the objects of a theory are defined only if they can be characterized by a series of well-defined processes (also referred to as operations) [55]. This stance deliberately avoids committing to any particular ontology regarding the entities described by a theory [56]. Bearing in mind that predictive power is a fundamental aspect of any physical theory, and that this power is only verified experimentally, the language of generalized probabilistic theories breaks down an experiment into three primitive operational procedures: **preparations, transformations, and measurements**. This is illustrated in figure 3.1.

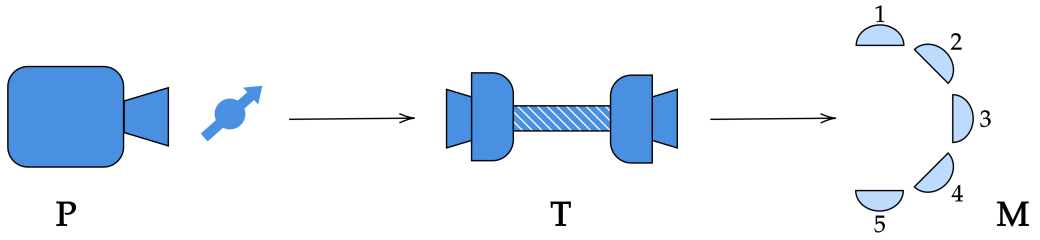


Figure 3.1: Generic steps involved in an experiment. A system is prepared according to **P**, then subjected to a transformation **T** and finally some quantity is measured according to a procedure **M**, producing an outcome (in this case, one of five possibilities).

In the generic experiment from figure 3.1, a preparation (**P**) consists of a sequence of steps that takes an input system to a desired initial configuration. Following any preparation, a transformation (**T**) is any operation which takes the initial configuration **P** to a configuration $\mathbf{P_T}$ of an output system. Finally, a measurement **M** is a procedure performed on the (output) system that generates a definite classical outcome, which marks the end of the experiment.

For any triple $(\mathbf{P}, \mathbf{T}, \mathbf{M})$ present in an experiment, we assume that, after sufficiently many iterations, one can construct a probability table

$$p(k|\mathbf{P}, \mathbf{T}, \mathbf{M}), \tag{3.1}$$

where k denotes one of the possible outcomes of measurement **M**. In doing this, we assume that k belongs to a finite list of possibilities, that is, we consider measurement procedures with discrete outcomes.

Despite its restrictiveness, as we will see next, the previous assumption suffices to construct classes of probabilistic theories that already reproduce features of quantum theory, without invoking any of the Hilbert space machinery. Thus, we can justify its usefulness *a posteriori*.

Before proceeding, we introduce the following notation: $\mathbf{Prep}_\mathrm{A}$ represents the collection of all possible preparations of a system A. Similarly, $\mathbf{Mes}_\mathrm{A}$ and $\mathbf{Transf}_\mathrm{A}$ correspond to the collection of all measurements and transformations (from A to itself), respectively.

We point out that the following comprise the most mathematically dense parts of this chapter.

3.2.2 States, effects and transformations

State spaces

The probabilities in equation (3.1) are key to define the basic objects we will need for the remainder of this thesis, the first of which are *states*.

For all pairs consisting of a measurement **M** and a transformation **T**, suppose that two distinct preparation procedures, **P** and **P**′, yield the same probability tables, *i.e.*

$$\forall\ \mathbf{T} \in \mathbf{Transf}_{\mathrm{A}}, \forall\ \mathbf{M} \in \mathbf{Mes}_{\mathrm{A}}:\ \vec{p}(\mathbf{P},\mathbf{T},\mathbf{M}) = \vec{p}(\mathbf{P}',\mathbf{T},\mathbf{M}), \tag{3.2}$$

where $\vec{p}(\mathbf{P},\mathbf{T},\mathbf{M})$ denotes the vector with entries given by equation (3.1), and similarly for $\vec{p}(\mathbf{P}',\mathbf{T},\mathbf{M})$.

This means that, regardless of which of them we choose at the start of each iteration of an experiment, they always result in the same predictions about the system. Consequently, **P** and **P**′ are statistically indistinguishable, and, as long as we are only interested in the probabilities of an experiment, we can treat both preparations as equivalent. This leads to the following definition

Definition 3.1: Equivalent class of preparations

Two preparations $\mathbf{P},\mathbf{P}'$ are **statistically equivalent** (symbolically, $\mathbf{P} \simeq_{\mathbf{Prep}_{\mathrm{A}}} \mathbf{P}'$) if equation (3.2) holds. For a given **P**, we define its equivalence class as

$$[\mathbf{P}] \doteq \{\mathbf{P}' \in \mathbf{Prep}_{\mathrm{A}} \mid \mathbf{P} \simeq_{\mathbf{Prep}_{\mathrm{A}}} \mathbf{P}'\}. \tag{3.3}$$

Since any equivalent preparations determine the same equivalence class, which we denote by ρ, distinct classes must have an empty intersection. Besides that, since any preparation is trivially contained in its own equivalence class, we conclude that $\mathbf{Prep}_{\mathrm{A}}$ is the disjoint union of all distinct classes ρ. Thus, for a given system A, it is useful to define the following.

Definition 3.2: State space (provisory)

The **state space** of a system A is the collection of all equivalence classes of $\mathbf{Prep}_{\mathrm{A}}$ under the relation $\simeq_{\mathbf{Prep}_{\mathrm{A}}}$. It is denoted by $\mathcal{S}_{\mathrm{A}}$.

Naturally, the elements of $\mathcal{S}_{\mathrm{A}}$ are called **states**, and from the construction above a state is merely a mathematical representation of a list of probabilities. From now on, we denote states using lowercase Greek letters (ρ,σ,ω, etc).

What structure should we add to $\mathcal{S}_{\mathrm{A}}$? We argue that it is desirable that it should be a "convex" set (see remark 1.4 for definition of convexity). To see why, first, choose any $n \in \mathbb{N}$. Note that for all $i \in \{0,\ldots,n\}$, any procedure that randomly

implements a preparation $\mathbf{P}_i$ with probability p_i is also a valid preparation procedure. Denote it by $\mathbf{P}_{\Sigma p_i \mathbf{P_i}}$. For any $\mathbf{M} \in \mathbf{Mes}_\mathrm{A}$ and $\mathbf{T} \in \mathbf{Transf}_\mathrm{A}$:

$$p(k|\mathbf{P}_{\Sigma p_i \mathbf{P_i}}, \mathbf{M}, \mathbf{T}) = \sum_{i=1}^{n} p_i \, p(k|\mathbf{P_i}, \mathbf{M}, \mathbf{T}). \tag{3.4}$$

Without loss of generality, we can assume that the $\mathbf{P_i}'s$ are all statistically inequivalent. Hence, each $\mathbf{P_i}$ is associated with a distinct state ρ_i. Denote the state associated with $\mathbf{P}_{\Sigma p_i \mathbf{P_i}}$ by ρ. Then the previous expression is the same as

$$p(k|\rho, \mathbf{M}, \mathbf{T}) = \sum_{i=1}^{n} p_i \, p(k|\rho_i, \mathbf{M}, \mathbf{T}). \tag{3.5}$$

We use the expression above to define ρ as

$$\rho = \sum_{i=1}^{n} p_i \rho_i. \tag{3.6}$$

More explicitly: a state ρ can be written as in equation (3.6) if and only if we can find inequivalent preparations $\mathbf{P}_i$ such that equation (3.4) is true. Strictly speaking, this does not make $\mathcal{S}_\mathrm{A}$ truly convex, since there is no addition operation defined on it, and while $\mathbf{P}_{\Sigma p_i \mathbf{P_i}}$ has a straightforward operational interpretation, it is not clear what should be meant by the "sum" of arbitrary preparation procedures. Hence, *a priori*, equation (3.6) should be taken more as a placeholder: it represents the list of probabilities of the form in equation (3.4).

We now show that the state space can be identified with a *de facto* convex subset of a real vector space. For that, we need the following definition:

Definition 3.3: Affine functional and affine map

Let S be a set such that for any finite probability distribution $(p_1, \ldots, p_n)$ and for any collection of elements $\{s_1, \ldots, s_n\} \subset S$, there is an element in S denoted by $\sum_{i=1}^{n} p_i s_i$. A map $f : S \to \mathbb{R}$ is an **affine functional on S** if:

$$f\left(\sum_{i=1}^{n} p_i s_i\right) = \sum_{i=1}^{n} p_i f(s_i). \tag{3.7}$$

More generally, when we replace $\mathbb{R}$ by any real vector space V, we say that f is an **affine map**.

For a fixed transformation $\mathbf{T}$ and a fixed d-outcome measurement $\mathbf{M}$, pick a particular outcome k. The mapping $\rho \in \mathcal{S}_\mathrm{A} \mapsto p(k|\rho, \mathbf{T}, \mathbf{M})$ is an example of an affine functional due to equations (3.5) and (3.6). We emphasize that **if states ρ and σ are different, there is an affine functional f such that $f(\rho) \neq f(\sigma)$.**

Now, note that the collection of all affine functions on S, *i.e.*, the set

$$\mathcal{A}(S) \doteq \{f : S \to \mathbb{R} |\ f \text{ satisfies equation (3.7)}\}, \tag{3.8}$$

is a real vector space. We represent its algebraic dual as

$$\mathcal{A}^*(S) \doteq \{\varphi : \mathcal{A}(S) \to \mathbb{R} |\ \varphi \text{ is linear}\}. \tag{3.9}$$

The following theorem guarantees that $\mathcal{S}_\mathrm{A}$ is isomorphic to a convex subset of $\mathcal{A}^*(\mathcal{S}_\mathrm{A})$.

Theorem 3.1: Embedding of $\mathcal{S}_\mathrm{A}$ [57, Proposition 1.2.1]

Let $\mathcal{S}_\mathrm{A}$ be a state space. There exists an injective affine map $E : \mathcal{S}_\mathrm{A} \to \mathcal{A}^*(\mathcal{S}_\mathrm{A})$.

Proof. The proof follows ref. [57]. For better readability, we write $\mathcal{S}_\mathrm{A} \equiv \mathrm{S}$. Let $\rho \in S$ be a state. We define the embedding map $\rho \mapsto E(\rho) \in \mathcal{A}^*(S)$ as:

$$E(\rho)(f) \doteq f(\rho)$$

for all $f \in \mathcal{A}(S)$. Now, let us consider any "convex" combination as in equation (3.6). In this case, due to $f \in \mathcal{A}(S)$ being affine-linear (equation (3.7)), we have

$$E\left(\sum_{i=1}^{n} p_i \rho_i\right)(f) = \sum_{i=1}^{n} p_i f(\rho_i) = \sum_{i=1}^{n} p_i [E(\rho_i)(f)] = \left(\sum_{i=1}^{n} p_i E(\rho_i)\right)(f).$$

Hence:

$$E\left(\sum_{i=1}^{n} p_i \rho_i\right) = \sum_{i=1}^{n} p_i E(\rho_i),$$

which proves that E is an affine map.

To prove injectivity, let $\rho_1, \rho_2 \in S$ be states such that $E(\rho_1) = E(\rho_2)$. By definition of the embedding map, this means that $f(\rho_1) = f(\rho_2)$ for all $f \in \mathcal{A}(S)$. Now, as we stressed earlier: if $\rho_1 \neq \rho_2$ then $f(\rho_1) \neq f(\rho_2)$ for some functional. Using the contrapositive of this implication, we can prove that $E(\rho_1) = E(\rho_2)$ implies that $\rho_1 = \rho_2$. Thus, E is injective. ■

For all intents and purposes, from now on, we can treat $\mathcal{S}_\mathrm{A}$ as a convex set, since there is a one-to-one correspondence between it and its image by the embedding map E.

Before we carry on to discuss about measurements, we have to make an important consideration regarding the dimension of $\mathcal{A}(\mathcal{S}_\mathrm{A})$. A basic fact from linear algebra is that any affine map $f : V \to W$ between vector spaces can be written as a linear map $L : V \to W$ plus a fixed translation $f(0_V) \in W$, where 0_V denotes the zero vector in

V. Now, if both V and W are finite-dimensional, then $\mathcal{A}(V;W)$ is a $(\dim V \dim W + 1)$-dimensional vector space. In the case of state spaces, $W = \mathbb{R}$, but $V = \mathrm{span}(\mathcal{S}_A)$, which need not be finite-dimensional. Henceforth, we make the following

Dimensional assumption: The state space $\mathcal{S}_A$ generates a finite-dimensional real vector space V.

There are two main reasons for why we assume this:

Reason 1) The assumption situates us within the familiar realm of finite-dimensional linear algebra, so we do not have to work with the more sophisticated machinery of functional analysis, which is the arena of infinite-dimensional vector spaces. We refer the interested reader to Chapter 1 from [58] for an introduction to the matter in the context of GPTs.

Reason 2) In quantum information theory. it is typical to work with finite-dimensional systems. As we want to include quantum systems as a particular case of our construction, it suffices to work under the assumption above.

Additionally, we assume that $\mathcal{S}_A$ is a compact[1] set. As we will see shortly, this will turn out to be an important feature.

Definition 3.4: Extremal points

Let S be a convex subset of a real vector space. An element $s \in S$ is an **extremal point** if for every $s_1, s_2 \in S$ such that

$$s = p s_1 + (1-p) s_2 \tag{3.10}$$

for $p \in (0,1)$, we have $s_1 = s_2 = s$. We represent by $\mathrm{Ext}(S)$ the collection of all extremal points of S.

In terms of preparations, extremal points represent preparations that cannot be generated from any randomized procedure. Adopting the nomenclature used in quantum theory, states associated with these preparations are called **pure states**. If a state is not pure, then it is called a **mixed state**.

Due to convexity, finite-dimensionality and compactness, we have the following adapted version of the Minkowski-Carathéodory theorem:

Theorem 3.2: Minkowski-Carathéodory

Any state in $\mathcal{S}_A$ is a convex combination of pure states. Additionally, if $\dim(V) = d$, with $V = \mathrm{span}(\mathcal{S}_A)$, no more than $d+1$ pure states are needed.

[1] Assuming an underlying norm, this means that the state space is closed and bounded.

Proof. See theorem 8.11 in ref. [59]. ■

In light of the embedding (thm. 3.1) and Minkowski-Carathéodory theorems, from now on, we will adopt the following definition of a state space [53].

Definition 3.5: State space

Given a system A, its state space is a compact and convex subset of a finite-dimensional real vector space V_A, and $V_A = \mathrm{span}(\mathcal{S}_A)$.

Recall from Chapter 1 that the state space of a finite-dimensional quantum system is the collection of positive operators with unit trace. As it was demonstrated in that chapter, this set is a convex subset of the real vector space of Hermitian operators $\mathrm{Her}(\mathcal{H}_A)$.

To verify compactness, first define the Hilbert-Schmidt norm of an operator as $\|X\|_{HS} = \sqrt{\mathrm{Tr}(X^\dagger X)}$. With this norm, $\mathrm{Her}(\mathcal{H}_A)$ is a finite-dimensional normed vector space, where compactness of a subset is equivalent to it being closed and bounded. A simple calculation shows that both conditions hold, hence $\mathcal{D}(\mathcal{H}_A)$ is compact. Therefore, we see that the previous definition at least encompasses finite-dimensional quantum state spaces.

Effects

To define the next essential object needed, we turn our attention to equation (3.1) once more. Just as it was the case for preparations, we can also define a notion of equivalent measurement outcomes.

In detail, any n-outcome measurement $\mathbf{M}$ can be seen as a collection of n 'yes-no' questions about the occurrence of its outcomes. Upon observation of outcome $k \in \{1,\dots,n\}$, the question associated with it is positively answered, otherwise it is answered in the negative. Let us represent by $k|\mathbf{M}$ the effect of obtaining a 'yes' answer for a particular outcome k.

Now, let $\mathbf{M}$ and $\mathbf{M}'$ be measurement procedures with finitely many outcomes[2] $k \in \{1,\dots,n\}$ and $k' \in \{1,\dots,n'\}$, respectively. If for a particular pair (k,k') we have

$$\forall\ \mathbf{P} \in \mathbf{Prep}_A, \forall\ \mathbf{T} \in \mathbf{Transf}_A :\ p(k|\mathbf{P},\mathbf{T},\mathbf{M}) = p(k'|\mathbf{P},\mathbf{T},\mathbf{M}'), \tag{3.11}$$

then we say that the effects $k|\mathbf{M}$ and $k'|\mathbf{M}'$ are statistically equivalent. We indicate this equivalence by $k|\mathbf{M} \simeq_{\mathbf{Eff}_A} k'|\mathbf{M}'$, where $\mathbf{Eff}_A$ is the collection of all effects $k|\mathbf{M}$, *i.e.*, of all 'yes' answers for every measurement in $\mathbf{Mes}_A$. Similarly to definition 3.2, we have:

[2] Their outcome sets need not be of the same size, because what matters is if there are any two outcomes that cannot be distinguished from the probability tables alone.

Definition 3.6: Effect space (provisory)

The **effect space** of a system A is the collection of all equivalence classes of $\mathbf{Eff}_{\mathrm{A}}$ under the relation $\simeq_{\mathbf{Eff}_{\mathrm{A}}}$, and it is denoted by $\mathcal{E}_{\mathrm{A}}$:

Similarly as it was for state spaces, $\mathcal{E}_{\mathrm{A}}$ does not have any *a priori* structure. In spite of that, it must have elements that represent outcomes obtained from randomized measurement procedures (symbolically denoted by "convex" combinations). Additionally, it must have elements that represent the grouping of measurement outcomes (symbolically denoted by "sums" of the outcomes), also known as coarse-grainings.

Just as states can be embedded into $\mathcal{A}^*(\mathcal{S}_{\mathrm{A}})$, we can identify any effect $e \in \mathcal{E}_{\mathrm{A}}$ with an affine functional on $\mathcal{S}_{\mathrm{A}}$. To verify this claim, first notice that any preparation $\mathbf{P} \in \mathbf{Prep}_{\mathrm{A}}$ followed by a transformation $\mathbf{T} \in \mathbf{Transf}_{\mathrm{A}}$ can be regarded as a new preparation of system A. Next, without loss of generality, we see that the mapping

$$\rho \longmapsto p(k|\rho, \mathbf{M})$$

from $\mathcal{S}_{\mathrm{A}}$ to $\mathbb{R}$ depends exclusively on $k|\mathbf{M}$. Since this effect belongs to one, and just one, equivalence class $e_k \in \mathcal{E}_{\mathrm{A}}$, we set that

$$e_k(\rho) \doteq p(k|\rho, \mathbf{M}), \tag{3.12}$$

where $k|\mathbf{M}$ is any representative of the class e_k. Thus, by virtue of definition 3.3, the effect e_k is an affine functional on $\mathcal{S}_{\mathrm{A}}$. From this, we conclude that $\mathcal{E}_{\mathrm{A}} \subset \mathcal{A}(\mathcal{S}_{\mathrm{A}})$. In the same spirit of what we did when going from definition 3.2 to definition 3.5, next, we will derive a concrete characterization of the effect space.

Given an effect e, notice that the inequality

$$e(\rho) \geq 0 \tag{3.13}$$

is true for any state $\rho \in \mathcal{S}_{\mathrm{A}}$, since this number is interpreted as a probability. This condition of positivity on $\mathcal{S}_{\mathrm{A}}$ induces an order in the effect space. To see this, we conveniently introduce the following set:

$$\mathcal{A}(\mathcal{S}_{\mathrm{A}})^+ \doteq \{f \in \mathcal{A}(\mathcal{S}_{\mathrm{A}}) \mid f(\rho) \geq 0,\ \forall\ \rho \in \mathcal{S}_{\mathrm{A}}\}. \tag{3.14}$$

Simply put, $\mathcal{A}(\mathcal{S}_{\mathrm{A}})^+$ is the collection of all affine functionals that are positive on the state space $\mathcal{S}_{\mathrm{A}}$. We call it the **positive cone of effects** for it is a (convex) cone[3] that contains the effect space $\mathcal{E}_{\mathrm{A}}$. This cone introduces a partial order in $\mathcal{A}(\mathcal{S}_{\mathrm{A}})$, which is

[3] In a vector space V, a cone is any subset C that is stable under nonnegative scaling. That is, if $\alpha \geq 0$ and $v \in C$, then $\alpha v \in C$. It is convex if $v_1 + v_2 \in C$ for any $v_1, v_2 \in C$.

represented as $\preccurlyeq_A$ and defined by

$$g \preccurlyeq_A f \iff f - g \in \mathcal{A}(\mathcal{S}_A)^+, \tag{3.15}$$

where f and g are affine functionals. From its definition, we see that the positive cone is precisely the set of positive elements according to the order $\preccurlyeq_A$, *i.e.*, it is the set of elements satisfying $o_A \preccurlyeq_A f$, where o_A is the zero functional $o_A(\rho) = 0$ for all $\rho \in \mathcal{S}_A$.

Additionally, let u_A be the effect such that $u_A(\rho) = 1$ for all states ρ. Any effect e satisfies $o_A \preccurlyeq_A e \preccurlyeq_A u_A$. In light of this, define the unit interval in $\mathcal{A}(\mathcal{S}_A)$ as

$$[o_A, u_A] \doteq \{f \in \mathcal{A}(\mathcal{S}_A) \mid o_A \preccurlyeq_A f \preccurlyeq_A u_A\}. \tag{3.16}$$

We observe that:

1. The unit interval is a convex set.

2. For any $f \in [o_A, u_A]$, there is $f^\perp \in [o_A, u_A]$ such that

$$f + f^\perp = u_A. \tag{3.17}$$

The first observation is trivial. The second one can be verified by setting $f^\perp = u_A - f$ and evaluating it on an arbitrary state. If f is an effect (def. 3.6), the interpretation of $f^\perp(\rho)$ is that it represents the probability of not observing the outcome associated with f in an n-outcome measurement.

From these two observations, we can redefine the effect space as:

Definition 3.7: Effect space

The **effect space** $\mathcal{E}_A$ of system A is a closed convex subset of $[o_A, u_A]$ containing both o_A and u_A, such that equation (3.17) holds. Additionally, we assume that $\mathrm{span}(\mathcal{E}_A) = V_A^*$, where V_A^* is the vector space dual to $\mathrm{span}(\mathcal{S}_A)$.

In other words, the effect space of a system represents the collection of all possible 'yes-no' questions that are experimentally accessible via measurement procedures in $\mathbf{Mes}_A$. These questions are all the two-outcome measurements $\{e, e^\perp\}$ and, by construction, **an n-outcome measurement** is a collection of effects $\{e_1, \ldots, e_n\}$ such that

$$\sum_{i=1}^{n} e_i = u_A. \tag{3.18}$$

In this sense, the effect u_A plays the same role as the identity operator does in a projective measurement or in a POVM (equations (1.3b) and (1.12b)). It represents that, in any given measurement, an outcome has been observed. Typically it is called

the **unit effect** of the system under consideration.

The fact that the space generated by $\mathcal{E}_A$ is dual to V_A translates the idea that effects are separating for states, meaning that any pair of distinct states ρ and σ can be distinguished by an effect $e \in \mathcal{E}_A$. Since any effect belongs to a two-outcome measurement, this condition also means that there exists a measurement that distinguishes any two different states. From this point of view, the separating condition is also referred to as **tomographic completeness** [60, 61].

We bring to attention that, in principle, the unit interval $[o_A, u_A]$ qualifies as an effect space, since the action of any of its elements on an arbitrary state returns a number between 0 and 1. Whenever this interval is taken as the effect space, we say that the system satisfies the **no-restriction hypothesis**: any mathematically well-defined effect represents the effect of an actual measurement from $\mathbf{Mes}_A$.

Even though the no-restriction hypothesis may look like just a mathematical convenience, we note that standard quantum theory already satisfies it. There, the unit interval is given by

$$[0, \mathbb{I}] = \{E \in \mathrm{Her}(\mathcal{H}) \mid 0 \preccurlyeq E \preccurlyeq \mathbb{I}\},$$

that is, it consists of all positive operators with eigenvalues between 0 and 1. Since the measurement operators of a given POVM (or a projective measurement) are positive operators which satisfy a resolution to the identity, they necessarily lie within the interval above.

Unless otherwise stated, we will implicitly consider that systems satisfy the no-restriction hypothesis. For relaxations thereof, see [62].

Remark 3.1 (**Self-duality**)**.** At this stage it is worth pointing out that states and effects are represented by distinct mathematical entities in general probabilistic theories. From definitions 3.5 and 3.7 we see that states are vectors, whereas effects are functionals. From the previous example, however, we see that in quantum theory both states and effects are treated as elements from the same vector space. From the formalism, a quantum effect should be a linear functional e that maps ρ to $e(\rho) \in [0, 1]$. The fact that this is not the case, and that we can still use Born's rule to compute probabilities, comes from a feature known as self-duality, which is deeply tied to the fact that all classical systems in quantum theory are equivalent [63].

Using the unit effect, one can write down the state space in a way that resembles $\mathcal{D}(\mathcal{H})$, as the following theorem shows.

Theorem 3.3: State space from effect space

Let $\mathcal{E}_A$ be an effect space. Define the cone generated by it as

$$C^+_{\mathcal{E}_A} \doteq \{\alpha e : \alpha \in \mathbb{R}_+, e \in \mathcal{E}_A\}.$$

Under the no-restriction hypothesis, we have that

$$\mathcal{S}_A = \{\rho \in V_A \mid \rho \succcurlyeq_A 0_A,\ u_A(\rho) = 1\}, \tag{3.19}$$

where the positivity condition $\rho \succcurlyeq_A 0_A$ means that $f(\rho) \geq 0$ for all $f \in C^+_{\mathcal{E}_A}$.

Proof. The proof is not difficult, but we omit it so that we do not extend ourselves too much. It is essentially an application of the hyperplane separation theorem (thm. 1.2). For the proof, see theorem 3.19 from ref. [53]. ■

The state space in equation (3.19) is, by all accounts, the generalization of the state space from quantum theory (equation (1.1)) to any probabilistic theory. Even when the no-restriction hypothesis is no longer assumed, it is often taken that the state space can be written as in equation (3.19). We will frequently use it to construct the state spaces in most of our examples.

Transformations

The last elements we need before discussing examples are the transformations between systems. This time, in the probabilities from equation (3.1), we must consider that the input and output systems need not be the same, that is, **T** maps an input system A into an output system B. We construct statistically equivalent transformations between A and B just as we did with preparations and measurements. These make up the collection of the physically allowed transformations of a theory, where now we view a transformation as a linear map $T : V_A \to V_B$.

A minimal requirement that any physically allowed transformation must satisfy is that it maps states to (sub-)normalized states[4]. By sub-normalized state, we mean any vector of the form $p\omega$, where $p \in [0, 1]$ and ω belongs to a state space as in equation (3.19). This requirement implies the following definition:

Definition 3.8: Valid transformation

Given systems $\mathcal{S}_A \subset V_A$ and $\mathcal{S}_B \subset V_B$, a **valid transformation** is a linear transform-

[4]This requirement is inspired, for example, by the idea that quantum operations can be trace non-increasing, instead of just trace-preserving.

ation $T : V_A \to V_B$ such that:

$$\rho \in \mathcal{S}_A \implies T(\rho) \in \text{ConvHull}(\{0_B, \mathcal{S}_B\}), \tag{3.20}$$

where ConvHull(X) denotes the convex hull of a set X and 0_B is the zero vector in V_B. The collection of all valid transformations from A to B is indicated by $\mathcal{T}(A;B)$.

Definition 3.8 should be seen as a working definition of transformation. The collection $\mathcal{T}(A;B)$ plays a role similar to the unit interval $[o_A, u_A]$, in the sense that if we take it as the collection of physically allowed transformations, then we are assuming a no-restriction hypothesis for transformations. For example, if we were working with quantum systems, $\mathcal{T}(A;B)$ would be the collection of trace non-increasing positive maps. However, we know from Chapter 1 that quantum operations should be *completely positive maps*, which demonstrates that even in quantum theory we do not adhere to this sort of no-restriction.

Despite this apparent drawback, definition 3.8 is general enough so that whenever one proves that there is no valid transformation that has some property P, then we are certain that no physically allowed transformation could ever have P. This will prove to be useful, since we are yet to prove the no-cloning theorem (thm. 2.1).

3.2.3 GPT system

After thoroughly discussing all the necessary notions, the previous subsection can be summarized in the following definition:

Definition 3.9: GPT system

A GPT system A is a tuple $(\mathcal{S}_A, \mathcal{E}_A, u_A, \mathcal{T}_{allowed}(A))$, where

1. $\mathcal{S}_A$ and $\mathcal{E}_A$ are, respectively, a state space and an effect space such that $V_A = \text{span}(\mathcal{S}_A)$ is a finite-dimensional vector space and $\text{span}(\mathcal{E}_A) = V_A^*$;
2. u_A is the unique element in $\mathcal{E}_A$ such that $u_A(\rho) = 1$ for all states ρ and equation (3.19) both hold;
3. $\mathcal{T}_{allowed}(A)$ is a subset of $\mathcal{T}(A;A)$ representing the physically allowed transformations from A to itself.

Let us mention two paradigmatic examples of GPT systems.

Example 3.1: Quantum systems

The canonical example of a GPT system is a finite-dimensional quantum system. In this example, we consider:

$$\mathcal{S}_A = \mathcal{D}(\mathcal{H}_A), \quad \mathcal{E}_A = [0, \mathbb{I}_A], \quad u_A = \mathbb{I}_A,$$

and $\mathcal{T}_{allowed}(A)$ is the collection of all trace non-increasing completely positive maps.

Example 3.2: Classical systems

A finite classical system is characterized by a finite set of perfectly distinguishable preparations, which we denote by Ω. These preparations can be identified with the canonical basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$, where $n = |\Omega|$. The state space is

$$\mathcal{S}_\Omega = \text{ConvHull}(\{v_1, \ldots, v_n\}),$$

which we write as Δ_Ω. In other words, classical states are probability distributions over the configurations Ω, and each vertex v_i is a pure state. Since the vertices are affinely independent[a], Δ_Ω is a $(n-1)$-dimensional simplex. Consequently, any state $\omega \in \mathcal{S}_\Omega$ has a **unique pure state decomposition**.
If $\{e_1, \ldots, e_n\} \subset (\mathbb{R}^n)^*$ is the basis dual to the canonical basis, the effects of a classical system are precisely the functionals

$$e = \sum_{i=1}^{n} c_i e_i,$$

with $c_i \in [0,1]$ for $i = 1, \ldots, n$. Additionally, in terms of the dual basis, the unit effect is $u_\Omega = \sum_i e_i$, which just sums all components of a state vector.
The effect space $\mathcal{E}_\Omega \equiv \Delta_\Omega^*$ is the convex hull of the zero functional o_Ω, the unit effect u_Ω, plus all possible groupings of the outcomes $1, \ldots, n$. This makes Δ_Ω^* a hypercube with 2^n vertices (illustrated in figure 3.2 for $n = 2$ and $n = 3$).
The set of allowed transformations consists of all substochastic matrices, *i.e.*, $n \times n$ matrices with nonnegative entries, and with the sum of elements in each row being less than or equal to 1.

[a] This means that $\{v_2 - v_1, \ldots, v_n - v_1\}$ is linearly independent.

An important feature of classical systems is that any finite collection of measurements can be performed together, a property called **joint measurability**. Formally, for any measurements $\{e_{i_1}\}_{i_1 \in I_1}$, ..., $\{e_{i_N}\}_{i_N \in I_N}$, joint measurability means that there exists a measurement $\{E_{\vec{i}}\}_{\vec{i} \in I_1 \times \ldots \times I_N}$ whose outcomes $\vec{i} = (i_1, \ldots, i_N)$ represent simul-

taneously measuring each individual measurement.

Additionally, each outcome of the j-th measurement is obtained by grouping all possibilities in which it could occur, *i.e*:

$$e_{i_j} = \sum_{\vec{i}:\ i_j \in \vec{i}} E_{\vec{i}}. \tag{3.21}$$

As demonstrated in reference [64], this property is actually sufficient and necessary to characterize a system as classical. Consequently, non-classical systems always have at least a pair of incompatible measurements.

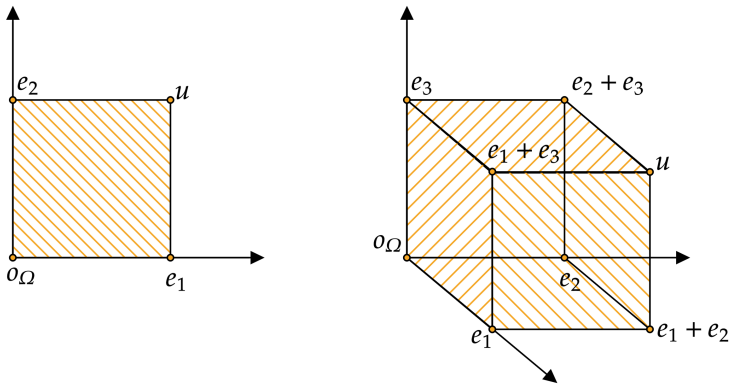


Figure 3.2: The effect spaces of classical systems Δ_2 (left) and Δ_3 (right). The systems are named bit and trit, respectively.

3.3 First proper example: the generalized bit

The first non-classical and non-quantum system that we will construct is the *gbit*, which is short for generalized bit. This system was introduced in ref. [54] as part of a probabilistic theory originally called generalized no-signalling theory, but which gradually became referred to as *Boxworld*.

To define the state spaces of this theory, one begins by choosing a pair (n, k), where n fix a number of measurements, each of which has k outcomes. That is, each $x \in \{0, \ldots, n-1\}$ defines a measurement $\{e_{a|x}\}_{a=0}^{k-1}$ whose effects belong to an effect space $\mathcal{E}_\mathrm{A} \subset V_\mathrm{A}^*$. In practice, this effect space is the convex hull of $o_\mathrm{A}, u_\mathrm{A}$ and all the $e_{a|x}$'s.

Using the state space construction from theorem 3.3, $\mathcal{S}_\mathrm{A}$ consists of all $\rho \in V_\mathrm{A}$ such that

$$0 \leq e_{a|x}(\rho) \leq 1 \tag{3.22}$$

for all effects $e_{a|x}$ and such that $o_\mathrm{A}(\rho) = 0$ and $u_\mathrm{A}(\rho) = 1$.

The gbit consists of the case $(n, k) = (2, 2)$. For concreteness, we can imagine that a party named Alice measures the Pauli observables σ_Z $(x = 0)$ and σ_X $(x = 1)$.

Using the identification $e_{+1|0} = |0\rangle\langle 0|$, $e_{-1|0} = |1\rangle\langle 1|$, $e_{+1|1} = |+\rangle\langle +|$ and $e_{-1|1} = |-\rangle\langle -|$, the effect space is given by

$$\mathcal{E}_{\text{gbit}} = \text{ConvHull}(\{0, \mathbb{I}_2, |0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -|\}). \tag{3.23}$$

As illustrated in figure 3.3, this effect space is an octahedron. Furthermore, $V^*_{\text{gbit}} = \text{span}(\mathcal{E}_{\text{gbit}}) = \text{span}(\{\mathbb{I}_2, \sigma_X, \sigma_Z\})$ is the three-dimensional space in which $\mathcal{E}_{\text{gbit}}$ lives.

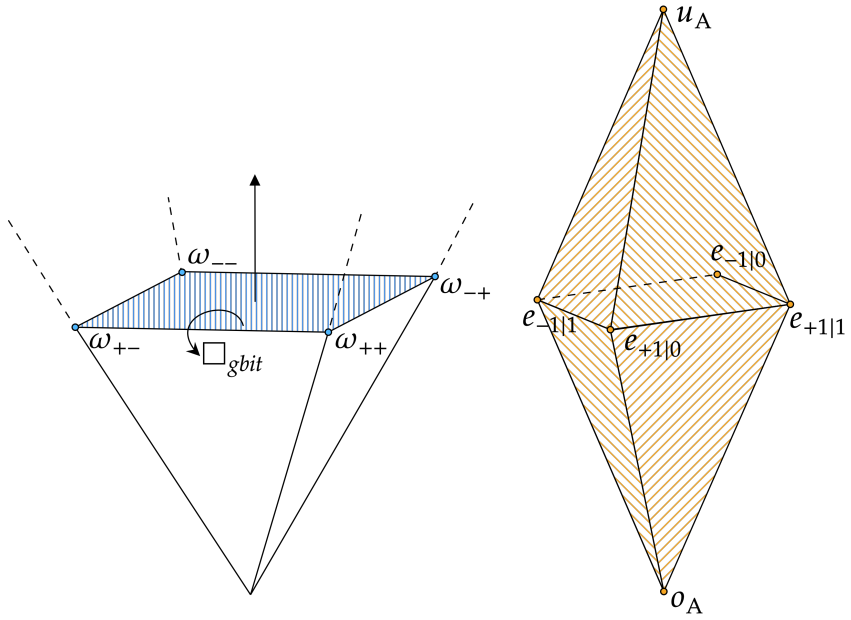


Figure 3.3: The gbit's state space (left); sub-normalized states lie below the blue square. The effect space (right) generated by the convex hull of the effects in equation (3.23).

The state space has exactly four pure states, denoted by ω_{ij}, where $i, j \in \{-1, +1\}$. They represent the four possible preparations in which the system would behave deterministically upon measuring the given observables. For example, the state ω_{++} represents a preparation in which we always observe the values +1 and +1 of σ_Z and σ_X, respectively. Each of them is constructed by imposing that $\text{Tr}(\omega_{ij} e_{a|x}) \in \{0, 1\}$ accordingly with the previous reasoning.

Using this representation of gbits as operators in the space of Hermitian operators, we have:

$$\omega_{++} = \frac{1}{2}(\mathbb{I}_2 + \sigma_Z + \sigma_X), \quad \omega_{+-} = \frac{1}{2}(\mathbb{I}_2 + \sigma_Z - \sigma_X), \tag{3.24a}$$

and

$$\omega_{-+} = \frac{1}{2}(\mathbb{I}_2 - \sigma_Z + \sigma_X), \quad \omega_{--} = \frac{1}{2}(\mathbb{I}_2 - \sigma_Z - \sigma_X). \tag{3.24b}$$

As we can see in figure 3.3, these states are the vertices of a square. The sub-normalized states are precisely the operators in $\text{ConvHull}(\{0, \omega_i j\}_{i,j})\}$ that lie below this square *i.e.*

they have trace less than one.

Notice that all qubit states which lie in the x-z plane of the Bloch ball are valid gbit states, since they belong to a disk inscribed in $\mathcal{S}_{\text{gbit}} \equiv \square_{\text{gbit}}$. All other states, *i.e.* those outside this disk, are non-quantum states in which a gbit can be prepared.

From their construction, we observe that any pair of diagonally opposing pure states can be perfectly distinguished by either measuring σ_Z or σ_X. This is in stark contrast with qubits, where antipodal points (pure states) in the Bloch sphere are distinguished by a single observable.

In order to discuss an interesting application of gbits in Bell nonlocality, we must first explain how to compose GPT systems.

3.4 Composition rules

3.4.1 Requirements and locally tomographic systems

Given two GPT systems A and B, a composition rule is an assignment that maps both systems to a larger GPT system AB, which consists of a tuple $(\mathcal{S}_{\text{AB}}, \mathcal{E}_{\text{AB}}, u_{\text{AB}}, \mathcal{T}_{allowed}(\text{AB}))$. Let us represent this assignment by $\boxtimes$.

Any composition rule must allow descriptions of independent experiments. That is, for any state-effect pairs ρ_{A} and e_{A} from system A, and ρ_{B} and e_{B} from system B:

CR1) $\rho_{\text{A}} \boxtimes \rho_{\text{B}}$ is a valid bipartite state and $e_{\text{A}} \boxtimes e_{\text{B}}$ is a valid bipartite effect. If both states (effects) are pure, then the resulting bipartite state (effect) is also pure. Furthermore:

$$e_{\text{A}} \boxtimes e_{\text{B}}(\rho_{\text{A}} \boxtimes \rho_{\text{B}}) = e_{\text{A}}(\rho_{\text{A}}) e_{\text{B}}(\rho_{\text{B}}). \tag{3.25}$$

Additionally:

CR2) The composition of unit effects is the bipartite unit effect, *i.e.* $u_{\text{A}} \boxtimes u_{\text{B}} = u_{\text{AB}}$.

Another aspect of any composition rule is that conditional operations should result in valid states and effects. More precisely:

CR3) Fix effects $e_{\text{A}} \in \mathcal{E}_{\text{A}}$ and $e_{\text{B}} \in \mathcal{E}_{\text{B}}$. For any $\rho_{\text{AB}} \in \mathcal{S}_{\text{AB}}$, define $\rho_{\text{A}}^{e_{\text{B}}} \in V_{\text{A}}$ and $\rho_{\text{B}}^{e_{\text{A}}} \in V_{\text{B}}$ as the vectors that satisfy:

$$\begin{aligned} f_{\text{A}}(\rho_{\text{A}}^{e_{\text{B}}}) &\doteq (f_{\text{A}} \boxtimes e_{\text{B}})\rho_{\text{AB}}, \ \forall f_{\text{A}} \in \mathcal{E}_{\text{A}}, \\ f_{\text{B}}(\rho_{\text{A}}^{e_{\text{B}}}) &\doteq (e_{\text{A}} \boxtimes f_{\text{B}})\rho_{\text{AB}}, \ \forall f_{\text{B}} \in \mathcal{E}_{\text{B}}. \end{aligned} \tag{3.26}$$

Both $\rho_{\text{A}}^{e_{\text{B}}}$ and $\rho_{\text{B}}^{e_{\text{A}}}$ should be (sub)-normalized states of their respective systems.

In particular, when $e_B = u_B$, we simply write $\rho_A^{u_B} = \rho_A$ and refer to this state as the **reduced state** of system A. The reduced state of system B is defined similarly.

CR4) Fix states $\rho_A \in \mathcal{S}_A$ and $\rho_B \in \mathcal{S}_B$. For any $e_{AB} \in \mathcal{E}_{AB}$, define $e_A^{\rho_B} \in V_A^*$ and $e_B^{\rho_A} \in V_B^*$ as the functionals that satisfy:

$$\begin{aligned} e_A^{\rho_B}(\omega_A) &\doteq e_{AB}(\omega_A \boxtimes \rho_B), \ \forall \omega_A \in \mathcal{S}_A, \\ e_B^{\rho_A}(\omega_B) &\doteq e_{AB}(\rho_A \boxtimes \omega_B), \ \forall \omega_B \in \mathcal{S}_B. \end{aligned} \tag{3.27}$$

Both $e_A^{\rho_B}$ and $e_B^{\rho_A}$ should be valid effects of their respective systems.

The assignment $\boxtimes$ should also be bilinear in its entries, meaning that when any statistical mixture of states on any subsystem is composed with a state of the remaining subsystem, the result should also be a valid bipartite state.

For any pair of GPT systems, a composition rule is called a **minimal tensor product for states** [53], if any state $\rho_{AB} \in \mathcal{S}_{AB}$ is a convex combination of states $\rho_A \boxtimes \rho_B$. In other words, the bipartite state space is defined as

$$\mathcal{S}_A \otimes_{\min} \mathcal{S}_B \doteq \text{ConvHull}(\{\rho_A \boxtimes \rho_B \mid \rho_A \in \mathcal{S}_A \text{ and } \rho_B \in \mathcal{S}_B\}). \tag{3.28}$$

Similarly, it is a **minimal tensor product for effects** if the bipartite effect space is

$$\mathcal{E}_A \otimes_{\min} \mathcal{E}_B \doteq \text{ConvHull}(\{e_A \boxtimes e_B \mid e_A \in \mathcal{E}_A \text{ and } e_B \in \mathcal{E}_B\}). \tag{3.29}$$

Example 3.3: Minimal tensor product of quantum systems

As a concrete example, let us consider that each GPT system is a quantum system. A very natural minimal tensor product rule is given by $\boxtimes = \otimes$, where $\otimes$ is the usual tensor product operation. In this case, for state spaces $\mathcal{D}(\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_B)$, their minimal tensor product is

$$\mathcal{D}(\mathcal{H}_A) \otimes_{\min} \mathcal{D}(\mathcal{H}_B) = \{\rho_{AB} \mid \rho_{AB} \text{ is separable}\}.$$

Hence, there are no entangled states in a "minimal" quantum theory. Similarly, the minimal tensor product of quantum effects has no entangled effects. Clearly, the composition rule of quantum systems, $\otimes_Q$, strictly contains the states and effects of both minimal tensor products, and we represent this by $\otimes_{\min} \subset \otimes_Q$.

In light of the previous example, we say that the minimal tensor product consists only of separable states or effects, where separability now must be understood in terms of the composition rule $\boxtimes$. In order to have entangled states and entangled effects, the composition rule must satisfy $\otimes_{\min} \subset \boxtimes$. Now, notice the following:

Lemma 3.1: Minimal tensor product generates $V_A \otimes V_B$

Let $V_{AB}^{\min}$ be the smallest vector space containing $\mathcal{S}_A \otimes_{\min} \mathcal{S}_B$, *i.e.*

$$V_{AB}^{\min} = \operatorname{span}(\{\rho_A \boxtimes \rho_B \mid \rho_A \in \mathcal{S}_A \text{ and } \rho_B \in \mathcal{S}_B\}).$$

Then $V_{AB}^{\min} \simeq V_A \otimes V_B$. Similarly, $(V_{AB}^{\min})^* \simeq V_A^* \otimes V_B^*$.

Proof. See Lemma 1 from [65]. ■

This means that the bipartite vector space V_{AB} should at least accommodate $V_A \otimes V_B$, thence $\dim V_A \dim V_B \leq \dim V_{AB}$. As we discussed in Section 1.3, the quantum composition rule $\otimes_Q$ is such that dimensions match. This comes from the **local tomography** property, which we now formulate:

Definition 3.10: Local tomography

A bipartite GPT system AB has the local tomography property if, for any pair of states ρ_{AB} and ω_{AB}, the following conditions are equivalent:

1. $\rho_{AB} = \omega_{AB}$;
2. For any effects $e_A \in \mathcal{E}_A$ and $e_B \in \mathcal{E}_B$:

$$e_A \boxtimes e_B(\rho_{AB}) = e_A \boxtimes e_B(\omega_{AB}).$$

If a bipartite system satisfy this property, we say it is a **locally tomographic** system.

Not all probabilistic theories must be locally tomographic. For instance, although standard quantum theory satisfies local tomography, under superselection rules (which often appear when a system has a conserved quantity), this property no longer holds, as exemplified in ref. [66]. Additionally, it is even possible to formulate a non-locally tomographic quantum theory that yields the same statistical predictions of standard quantum theory, as recently explored in ref. [67]. For a more detailed discussion about non-locally tomographic GPTs, we refer the reader to ref. [65].

In the next chapter, we will examine how a non-locally tomographic modification of quantum theory allows violations of monogamy of entanglement (as discussed in Section 1.4). With this sole exception, unless otherwise stated, from now on we assume that all bipartite systems are locally tomographic, that is, we adhere to the following principle:

Principle 3.1: Local tomography

Every multipartite GPT system is locally tomographic.

Regarding locally tomographic systems, we have the following property:

Proposition 3.1: Dimension matching

Let AB be a locally tomographic system. Then $\dim V_{\mathrm{AB}} = \dim V_{\mathrm{A}} \dim V_{\mathrm{B}}$.

Proof. By lemma 3.1, we know that $(V_{\mathrm{AB}}^{\min})^* \simeq V_{\mathrm{A}}^* \otimes V_{\mathrm{B}}^*$. This lets us identify each $e_{\mathrm{A}} \boxtimes e_{\mathrm{B}}$ with $e_A \otimes e_B$. Now, recall that we can write

$$V_{\mathrm{A}}^* \otimes V_{\mathrm{B}}^* = \mathrm{span}(\{e_{\mathrm{A}} \otimes e_{\mathrm{B}} \mid e_{\mathrm{A}} \in \mathcal{E}_{\mathrm{A}} \text{ and } e_{\mathrm{B}} \in \mathcal{E}_{\mathrm{B}}\}).$$

Using the previous lemma, we can always express $V_{AB} = (V_{\mathrm{A}} \otimes V_{\mathrm{B}}) \oplus W$, where

$$W = \{v_{\mathrm{AB}} \in V_{\mathrm{AB}} \mid f(v_{\mathrm{AB}}) = 0,\ \forall f \in V_{\mathrm{A}}^* \otimes V_{\mathrm{B}}^*\}.$$

Since product effects span $V_{\mathrm{A}}^* \otimes V_{\mathrm{B}}^*$, we can rewrite this set as:

$$W = \{v_{\mathrm{AB}} \in V_{\mathrm{AB}} \mid e_{\mathrm{A}} \otimes e_{\mathrm{B}}(v_{\mathrm{AB}}) = 0,\ \forall e_{\mathrm{A}} \in \mathcal{E}_{\mathrm{A}}\ \forall e_{\mathrm{B}} \in \mathcal{E}_{\mathrm{B}}\}.$$

Now, we can always write $v_{\mathrm{AB}} = a\rho - b\omega$ for nonnegative scalars a, b and for states $\rho, \omega \in \mathcal{S}_{\mathrm{AB}}$ (lemma 3.35 from [53]). Hence, the condition $e_{\mathrm{A}} \otimes e_{\mathrm{B}}(v_{\mathrm{AB}}) = 0$ for all local effects, means, in particular, that $u_{\mathrm{A}} \otimes u_{\mathrm{B}}(v_{\mathrm{AB}}) = 0$. This implies that $a = b$. By local tomography, it follows that $\rho = \omega$, which means that $v_{AB} = 0$. From this, we conclude that $W = \{0\}$. Hence, $V_{\mathrm{AB}} = V_{\mathrm{A}} \otimes V_{\mathrm{B}}$, from which the proposition follows. ■

For locally tomographic GPT systems, there is also another important composition rule, one that allows the largest number of states consistent with local measurements. This rule is called the **maximal tensor product** (for states) and the state space is constructed using equation (3.19) in the particular case where $\mathcal{E}_{\mathrm{AB}} = \mathcal{E}_{\mathrm{A}} \otimes_{\min} \mathcal{E}_{\mathrm{B}}$:

$$\mathcal{S}_{\mathrm{A}} \otimes_{\max} \mathcal{S}_{\mathrm{B}} \doteq \{\rho \in V_{\mathrm{AB}} \mid e_{\mathrm{A}} \boxtimes e_{\mathrm{B}}(\rho) \in [0,1],\ \forall e_{\mathrm{A}} \in \mathcal{E}_{\mathrm{A}}, \forall e_{\mathrm{B}} \in \mathcal{E}_{\mathrm{B}},\ u_{\mathrm{AB}}(\rho) = 1\}. \tag{3.30}$$

The definition of $\mathcal{E}_{\mathrm{A}} \otimes_{\max} \mathcal{E}_{\mathrm{B}}$ is similar.

For any bipartite systems, under any composition rule $\boxtimes$, all of its states satisfy the conditions in the RHS of equation (3.30). Consequently, at the level of states, we have $\boxtimes \subseteq \otimes_{\max}$. We thus have that multipartite state spaces of any composition rule satisfy $\otimes_{\min} \subseteq \boxtimes \subseteq \otimes_{\max}$.

We bring to attention a crucial feature of the maximal tensor product of state spaces: the probabilities generated by its states satisfy the no-signalling conditions (eq.

(2.17)).

To verify the previous claim, for each x in a finite collection $\mathcal{X}$, let $\{e_{a|x}\}_{a=1}^{n}$ be an n-outcome measurement in system A. Fix any m-outcome measurement $\{e_b\}_{b=1}^{m}$ in system B. For an arbitrary state $\rho_{\mathrm{AB}} \in \mathcal{S}_{\mathrm{A}} \otimes_{\max} \mathcal{S}_{\mathrm{B}}$, we write:

$$p(a,b|x) = (e_{a|x} \boxtimes e_b)\rho_{\mathrm{AB}}.$$

Summing over all outcomes a, we get:

$$\sum_{a=1}^{n} p(a,b|x) = \sum_{a=1}^{n} (e_{a|x} \boxtimes e_b)\rho_{\mathrm{AB}} = (\sum_{a=1}^{n} e_{a|x} \boxtimes e_b)\rho = (u_{\mathrm{A}} \boxtimes e_b)\rho_{\mathrm{AB}} = e_b(\rho_{\mathrm{B}}),$$

where we used the bilinearity of $\boxtimes$ to go from the third to the fourth equality, and ρ_{B} represents the reduced state of system B (**CR3**). Clearly, the RHS only depends on e_b and it is precisely the probability of observing outcome b. From this, we conclude that:

$$p(b|x) = p(b), \ \forall x \in \mathcal{X},$$

which is the no-signalling condition from Alice to Bob. By similar calculations, it is also the case Bob cannot signal to Alice. Since state spaces satisfy $\boxtimes \subseteq \otimes_{\max}$, we see that the maximal tensor product is the largest state space compatible with the no-signalling condition.

In what follows, we will apply the maximal tensor product to compose gbits, which will allow violations of the CHSH inequality.

3.4.2 Two gbits maximally violate the CHSH inequality

Let us revisit the CHSH game from Section 2.3. Recall that it is characterized by two parties, Alice (A) and Bob (B), each of whom can choose between two binary-outcome measurements (these represent which passage they walk through in figure 2.2) to perform on their respective systems. Let us denote these measurements by A_x and B_y, $x, y \in \{0,1\}$.

As demonstrated in theorem 2.2, the maximal quantum violation of the CHSH inequality is $2\sqrt{2}$, where the inequality reads as

$$|\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle| \leq 2.$$

Now, let us suppose that instead of classical or quantum systems, Alice and Bob have gbits at their disposal. From Section 3.3, the gbit state space $\square_{\mathrm{gbit}}$ was constructed from the effect space obtained from the observables σ_X and σ_Z. Since these are incompatible, by adapting proposition 2.2, we already have one of the necessary

requirements to produce nonlocal correlations. We set these as the observables that Alice and Bob measure in the CHSH game, *i.e.*

$$A_0 = \sigma_Z,\ A_1 = \sigma_X,\ B_0 = \sigma_Z,\ B_1 = \sigma_X.$$

Due to the structure of the CHSH game, we know that Alice and Bob's correlations must be no-signalling. As discussed previously, the maximal tensor product of states yields the largest state space compatible with this condition. This leads us to consider $\square_{\text{gbit}} \otimes_{\max} \square_{\text{gbit}}$ as the state space of two gbits.

From the definition of the maximal tensor product, we see that the states (see equations (3.24a) and (3.24b)) $\omega_{ij} \otimes \omega_{kl}$, $i, j, k, l \in \{+, -\}$ are pure states in $\square_{\text{gbit}} \otimes_{\max} \square_{\text{gbit}}$ and their convex hull generates the separable states of two gbits.

Besides these 16 product states, there exist 8 additional pure states in $\square_{\text{gbit}} \otimes_{\max} \square_{\text{gbit}}$, which are entangled. They can be obtained by working through all the constraints imposed by equation (3.30) or via numerical means[5]. In any case, these additional pure states are:

$$\begin{aligned}
\omega_1^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} - \sigma_X\otimes\sigma_X + \sigma_X\otimes\sigma_Z + \sigma_Z\otimes\sigma_X + \sigma_Z\otimes\sigma_Z)\\
\omega_2^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sigma_X\otimes\sigma_X - \sigma_X\otimes\sigma_Z + \sigma_Z\otimes\sigma_X + \sigma_Z\otimes\sigma_Z)\\
\omega_3^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sigma_X\otimes\sigma_X + \sigma_X\otimes\sigma_Z - \sigma_Z\otimes\sigma_X + \sigma_Z\otimes\sigma_Z)\\
\omega_4^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sigma_X\otimes\sigma_X + \sigma_X\otimes\sigma_Z + \sigma_Z\otimes\sigma_X - \sigma_Z\otimes\sigma_Z)\\
\omega_5^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sigma_X\otimes\sigma_X - \sigma_X\otimes\sigma_Z - \sigma_Z\otimes\sigma_X - \sigma_Z\otimes\sigma_Z)\\
\omega_6^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} - \sigma_X\otimes\sigma_X + \sigma_X\otimes\sigma_Z - \sigma_Z\otimes\sigma_X - \sigma_Z\otimes\sigma_Z)\\
\omega_7^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} - \sigma_X\otimes\sigma_X - \sigma_X\otimes\sigma_Z + \sigma_Z\otimes\sigma_X - \sigma_Z\otimes\sigma_Z)\\
\omega_8^{PR} &= \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} - \sigma_X\otimes\sigma_X - \sigma_X\otimes\sigma_Z - \sigma_Z\otimes\sigma_X + \sigma_Z\otimes\sigma_Z).
\end{aligned} \tag{3.31}$$

[5]For two gbits, the conditions in equation (3.30) imply that the bipartite state space is a convex set obtained from the intersection of a finite number of half-spaces. This makes it a convex polytope, which has a finite number of vertices (ext. points). Enumerating these vertices can be done numerically, for instance, with specialized python libraries (pypoman or pycddlib). That is how we obtained the states in equation (3.31) [68].

Let $a, b \in \{-1, +1\}$. Using state ω_1^{PR} we have

$$p_1^{PR}(a,b|x,y) = \begin{cases} \frac{1}{2}, \text{ if } ab = (-1)^{xy}, \\ 0, \text{ otherwise.} \end{cases} \tag{3.32}$$

From these, we compute the expectation value of the CHSH operator (equation (2.15)) and achieve the value 4, the maximal value possible in the CHSH game. From this, we conclude that Alice and Bob always win the CHSH game if they share two gbits.

The probabilities in equation (3.32) are very special: they comprise the most nonlocal behaviour of the CHSH scenario that is compatible with the no-signalling condition. They were first introduced by Popescu and Rohrlich in ref. [69], where the authors investigated whether nonlocal correlations and relativistic causality where sufficient conditions to recover all quantum correlations. From our construction, we immediately see that this is not the case: first, because the gbit has non-quantum states; second, because two gbits violate Tsirelson's bound.

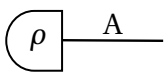
We note that all the remaining states in equation (3.31) also generate PR correlations, and consequently, they too can be used to maximally violate different versions of CHSH inequality[6].

An analysis of Bell nonlocality using polygonal systems (generalizations of the gbit which include more vertices), was done in ref. [70]. There, among other results, the authors find that Tsirelson's bound is always satisfied for any polygonal system with an odd number of vertices, whilst the same is not true for state spaces with an even number of vertices. For generalizations of the gbit which include entangled states *and* entangled effects, we refer the reader to ref. [71].

3.5 Diagrammatic representation

We end this chapter introducing a diagrammatic/graphical representation of states, transformations and effects that is often adopted in works involving GPTs [61, 65, 72–74]. We will make short use of it in the next chapters in places where we judge that a visual representation of an equation or an object makes it easier to understand what is going on.

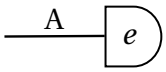
States are represented by half-oval shapes with a wire stemming from their flat side, which always faces rightwards. The label over the wire represents the system the state belongs to. For instance, $\rho \in \mathcal{S}_{\mathrm{A}}$ is visually represented as



(3.33)

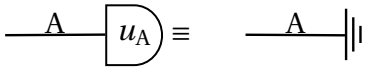
[6]These different versions are obtained by relabelling either the inputs x, y or the outputs a, b.

Effects are represent in similar fashion, the only difference being that they are horizontally flipped. That is, an element $e \in \mathcal{E}_{\mathrm{A}}$ is represented as



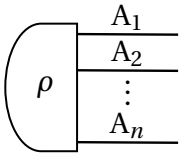
(3.34a)

The unit effect u_{A} is specially represented as



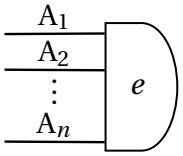
(3.34b)

The idea naturally extends to multipartite settings. If now we have $\rho \in \mathcal{S}_{\mathrm{A}_1 \ldots \mathrm{A}_n}$ and $e \in \mathcal{E}_{\mathrm{A}_1 \ldots \mathrm{A}_n}$, they are respectively represented as:



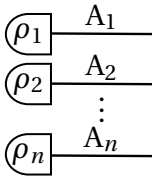
(3.35a)

and



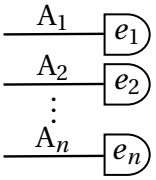
(3.35b)

When $\rho = \rho_1 \otimes \rho_2 \ldots \otimes \rho_n$ and $e = e_1 \otimes e_2 \ldots \otimes e_n$, we simply stack the diagrams of each state (effect) involved in the product:



(3.35c)

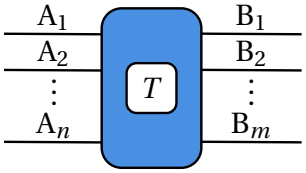
and



(3.35d)

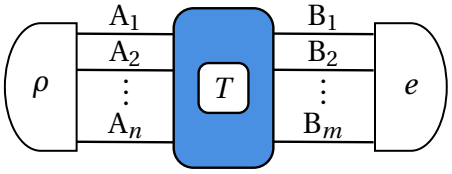
A transformation T from systems $\mathrm{A}_1 \ldots \mathrm{A}_n$ to systems $\mathrm{B}_1 \ldots \mathrm{B}_m$ is represented by a blue box with left wires representing the input systems and right wires represent-

ing the output systems:



(3.36)

The probability given by $e(T(\rho))$ is the closed circuit obtained by concatenating the diagrams of each object:



(3.37)

4 THE FRAMEWORK IN ACTION

From the side, a whole range;
from the end, a single peak:
far, near, high, low, no two parts alike.
Why can't I tell the true shape of Lushan?
Because I myself am in the mountain.

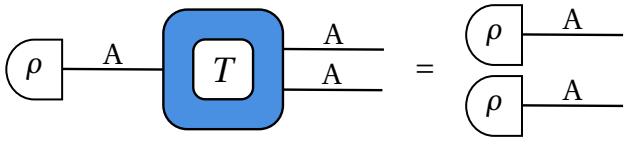
Su Shi, Written on the Wall at West Forest Temple

Chapter Contents

In this chapter we present a few applications of the GPT formalism. We begin the chapter by presenting the no-cloning theorem, a hallmark of any non-classical GPT system, and its generalization, known as the no-broadcasting theorem. Next, we discuss two protocols that rely on the no-cloning theorem: the teleportation protocol, whose quantum version was introduced in Chapter 2, and the BB84 cryptographic protocol. Then, we turn our attention to two modified versions of quantum theory and discuss how they contrast with standard quantum theory. We finish the chapter presenting a notion that plays a role dual to broadcasting.

4.1 No-cloning and no-broadcasting

At the beginning of Chapter 2, we mentioned how an arbitrary quantum state cannot be cloned. This means there is no CPTP map T such that, for any quantum state $\rho \in \mathcal{D}(\mathcal{H}_A)$, the following diagram equality holds:



That no CPTP map can clone an arbitrary mixed state is a true statement follows directly from observing that such a map would not be convex. It could still be the case, though, that a CPTP map clones arbitrary pure states. That also would not be possible, once again due to violation of convexity.

Since arbitrary cloning is prohibited in quantum theory, we are inclined to ask the following:

When is it possible to clone a finite collection of states?

Notice that, the way it was framed, this question does not require one to assume that the states are quantum states. In fact, this a type of question that does not depend on any particularities of whatever underlying theory is considered true. Mostly, applications of the GPT framework deal with these sort of "theory-independent" questions: one typically restricts oneself to a class of GPTs systems and then prove statements about these systems. In our case, this class is that of locally tomographic GPTs that satisfy the no-restriction hypothesis.

To answer the previous question, we need to introduce a few definitions, the first one being:

Definition 4.1: Informationally complete measurement

Let $E = \{e_i\}_{i\in I}$ be a measurement in a GPT system A. We say that E is informationally complete if for any two states $\rho, \sigma \in \mathcal{S}_\mathrm{A}$,

$$\forall\ i \in I:\ e_i(\rho) = e_i(\sigma) \tag{4.1}$$

implies $\rho = \sigma$.

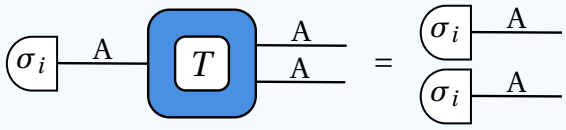
Thus, an informationally complete (IC) measurement is one which we can use to uniquely reconstruct any state solely from the probability distribution it determines. It is a fact that an IC measurement always exists in any GPT system (see Lemma 1 from [75]).

The next definitions we need are grouped together:

Definition 4.2: Clonability and joint distinguishability

Let $\sigma_1, \ldots, \sigma_n$ be states of a GPT system A. We say that these states are:

(a) Cloneable if and only if there exists a valid transformation $T:\ V_A \to V_A \otimes V_A$ such that the following diagram equality holds for all σ_i:



(4.2)

(b) Jointly distinguishable if and only if there exists an n-outcome measurement $\mathrm{M} = \{e_i\}_{i=1}^n$ such that for all $i, j \in \{1, \ldots, n\}$:

$$e_i(\sigma_j) = \delta_{ij}. \tag{4.3}$$

We are now ready to state and prove a theorem that answers the previous question. Additionally, it implies the no-cloning theorem from Chapter 2.

Theorem 4.1: Cloning is equivalent to perfect discrimination

Let A be a system from any GPT whose composition rule for states satisfies $\otimes_{\min} \subseteq \boxtimes \subseteq \otimes_{max}$. A finite collection of states $\{\sigma_i\}_{i\in I}$ is cloneable if and only if it is jointly distinguishable.

Proof. We essentially follow the steps of the original proof (Thm. 1 in [75]), but flesh out some of its passages to improve clarity.

We begin by proving necessity ($\Leftarrow$): assume that the states in the collection $\{\sigma_i\}_{i\in I}$ are jointly distinguishable and let $\{a_i\}_{i\in I}$ be the measurement that satisfies equation (4.3). Define $T : V_\mathrm{A} \to V_\mathrm{A} \otimes V_\mathrm{A}$ as the map whose action on any $v \in V_\mathrm{A}$ is

$$T(v) \doteq \sum_{i\in I} a_i(v)\sigma_i \boxtimes \sigma_i.$$

Direct computation shows that its action on each σ_i satisfies equation (4.2). Additionally, any $\rho \in \mathcal{S}_\mathrm{A}$ is mapped to a (normalized) bipartite state, which makes T a valid transformation. Consequently, the states in $\{\sigma_i\}_{i\in I}$ are cloneable.

Now, we turn to sufficiency ($\Rightarrow$): assume that $\{\sigma_i\}_{i\in I}$ is cloneable, and let T be as in equation (4.2).

The proof is the formalization of the following idea: first, we fix an informationally complete n-outcome measurement E. Next, we generate an arbitrary number of copies of each state using the transformation T. We then apply the measurement E to each copy. Since each state σ_i determines a probability distribution of E, this last step is akin to many independent rounds of an experiment that samples E, where the distribution of outcomes is identical throughout the rounds. Finally, since E is informationally complete, each state σ_i can be estimated with increasing accuracy using the probability distribution obtained in this experiment. As different states yield distinct distributions, we can use those probabilities to differentiate them.

It is a long proof, we subdivide it in small steps for better readability.

Step 1) For a given $m \in \mathbb{N}$, the map T induces a transformation $T_m : V_{\mathrm{A}} \to V_{\mathrm{A}}^{\otimes 2^m}$ which produces 2^m copies of each state σ_i. This transformation is defined recursively by setting $T_1 = T$ and

$$T_m = \underbrace{(T \otimes \ldots \otimes T)}_{2^{m-1}\ times} T_{m-1}.$$

Step 2) Let $E = \{e_1, \ldots, e_n\}$ be an informationally complete (IC) measurement in A, and denote its outcome set by O_E. Since E is IC, each state σ_i is uniquely identified with a probability vector

$$\vec{p}_i \equiv (e_1(\sigma_i), \ldots, e_n(\sigma_i)).$$

Now, for a given $m \in \mathbb{N}$, let E_{2^m} be the measurement in which each effect is a product of 2^m effects from E. That is, for every $\vec{J} = (j_1, \ldots, j_{2^m}) \in O_E^{2^m}$, an element of E_{2^m} has the form

$$e_{\vec{J}} = e_{j_1} \boxtimes \ldots \boxtimes e_{j_{2^m}}.$$

Each $\vec{J}$ induces an empirical probability distribution $q_{\vec{J}}$ over O_E, which is given by the relative frequency of each $j \in O_E$ in the entries of $\vec{J}$, *i.e.*:

$$q_{\vec{J}}(j) \doteq \frac{\left|\{x \mid j_x = j\}\right|}{2^m}.$$

In what follows, for a given $\varepsilon > 0$, we will have to consider the instances $\vec{J}$ in which the empirical distribution is ε-close to the distribution $\vec{p}_i$. For that, we define the set

$$A^{\varepsilon}_{i,m} \doteq \left\{\vec{J} \mid \left\| \vec{q}_{\vec{J}} - \vec{p}_i \right\| < \varepsilon \right\},$$

where the norm of a vector is taken to be the maximal absolute value of its entries.

The probability of observing any sequence of outcomes in $A^{\varepsilon}_{i,m}$ is given by

$$Pr(A^{\varepsilon}_{i,m}) = e_{A^{\varepsilon}_{i,m}}(T_m(\sigma_i)),$$

where the effect $e_{A^{\varepsilon}_{i,m}}$ is the sum of all $e_{\vec{J}}$ such that $\vec{J} \in A^{\varepsilon}_{i,m}$. By virtue of the weak law of large numbers[1], we have

$$Pr(A^{\varepsilon}_{i,m}) > 1 - \varepsilon$$

[1] In plain language, this version of the law says that with sufficiently many rounds, the expected value of the sample average of i.i.d random variables converges to the true expected value. For a precise formulation, check here or here.

for sufficiently large m.

Step 3) Let $a^{\varepsilon}_{i,m} : V_{\mathrm{A}} \to \mathbb{R}$ be the functional whose action on an arbitrary $\rho \in \mathcal{S}_{\mathrm{A}}$ is set as

$$a^{\varepsilon}_{i,m}(\rho) \doteq e_{A^{\varepsilon}_{i,m}}(T_m(\omega)).$$

It is an effect in system A, and, by construction, $a^{\varepsilon}_{i,m}(\sigma_i) > 1 - \varepsilon$, for sufficiently large m. We take m large enough so that the previous inequality is true for all $i \in I$.

The sum $\sum_{i\in I} a^{\varepsilon}_{i,m}$ is still a valid effect, because its action on any state ρ is the same as the action the sum of finitely many $e'_j s$ acting on $T_m(\rho)$, which cannot exceed 1.

We set $a^{\varepsilon}_{0,m}$ as $u_{\mathrm{A}} - \sum_{i\in I} a^{\varepsilon}_{i,m}$. Consequently, $\{a^{\varepsilon}_{i,m}\}_{i\in I\cup\{0\}}$ is an observable in A.

Step 4) Since the construction of $a^{\varepsilon}_{i,m}$ works for any ε, we take a sequence $\varepsilon_m = 1/2^m$. Then, $(a^{\varepsilon_m}_{i,m})_{m\in\mathbb{N}}$ is a sequence in the effect space $\mathcal{E}_{\mathrm{A}}$. Since $\mathcal{E}_{\mathrm{A}}$ is compact, the previous sequence admits a convergent subsequence $(a^{\varepsilon_{m_k}}_{i,m_k})_{m_k\in\mathbb{N}}$. For each $i \in I \cup \{0\}$, let a_i be the limit of such subsequence. Then, if $i \neq 0$:

$$a_i(\sigma_i) = \lim_{m_k\to\infty} 1 - \frac{1}{2^{m_k}} = 1.$$

As consequence, $a_i(\sigma_l)$ is equal to 0 whenever $l \neq i$.

When $i = 0$, we have

$$a_0(\sigma_l) = 0$$

for all $l \in I$.

Step 5) Finally, without loss of generality, fix an $i_0 \in I$. The measurement $\{b_i\}_{i\in I}$ that perfectly distinguishes the states in $\{\sigma_i\}_{i\in I}$ is given by grouping a_0 with a_{i_0}. More precisely:

$$b_i = \begin{cases} a_0 + a_{i_0}, \text{ if } i = i_0; \\ a_i, \text{ otherwise} \end{cases}.$$

■

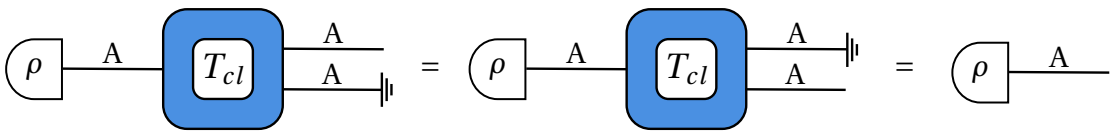
Theorem 4.1 is a quintessential example of how the GPT framework can be employed to prove results that do not rely on particularities of a theory. In the current case, whether the theory is classical or non-classical, the theorem illustrates that the states that can be cloned are precisely those which can be perfectly distinguished. This immediately implies the following characterization of non-classical state spaces:

Theorem 4.2: No-cloning

The pure states of a GPT system are cloneable if and only if the system is classical. Equivalently, a state space is non-classical if and only if there is no map that clones an arbitrary pure state.

By the result above, the ability to create copies of an arbitrary pure state in a theory is so strong of an imposition that it immediately restricts a system to be classical. In view of that, one could investigate if conditions that resemble perfect cloning could lift this restriction. We now discuss one such condition.

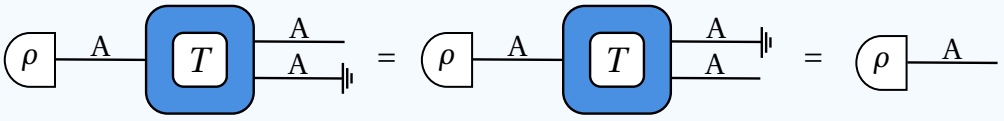
Notice that if a transformation T maps a state ρ to $\rho \boxtimes \rho$, then the reduced states of each system are precisely ρ. Diagrammatically, this is represented by the following equalities:



These diagrams indicate that the cloning map T_{cl} broadcasts the input state to each output system by simply creating two copies of the input state. Assuming that systems are non-classical, the no-cloning theorem forbids the existence of the cloning map T_{cl}, but it does not rule out the existence of other maps T that have the property above. This motivates the following definition:

Definition 4.3: Broadcasting

Let $T : V_A \to V_A \otimes V_A$ be a valid transformation. A state $\rho \in \mathcal{S}_A$ is **broadcast** by T if the following equality holds:



(4.4)

The set of states broadcast by T is denoted by Δ_T and we call it the broadcastable set of T.

Broadcasting is a relaxation of cloning, since it only demands the marginal states to be equal to the input state. In particular, T could map a state ρ to an entangled state whose marginals are ρ, whereas cloning would require the output to be precisely $\rho \boxtimes \rho$.

Example 4.1: Broadcasting in quantum theory

As a simple example, we consider a qubit system. Let T be the map

$$T(\rho) = (CNOT)\rho \otimes |0\rangle\langle 0| (CNOT),$$

where $CNOT$ is the operator that implements $|10\rangle \mapsto |11\rangle$ and vice-versa, but keeps $|00\rangle$ and $|01\rangle$ fixed. Direct computation shows that the reduced states are equal:

$$\rho_A = \rho_B = \mathrm{Tr}\big(\rho\,|0\rangle\langle 0|\big)\,|0\rangle\langle 0| + \mathrm{Tr}\big(\rho\,|1\rangle\langle 1|\big)\,|1\rangle\langle 1|\,.$$

Hence, if ρ is broadcast by T, this means that

$$\rho = p_0\,|0\rangle\langle 0| + p_1\,|1\rangle\langle 1|\,,$$

that is, ρ represents the state of a bit. Hence, Δ_T is a subset of $\mathcal{D}(\mathbb{C}^2)$ which can be identified with a classical system.

The fact that the broadcastable states in the last example are essentially classical states is an instance of the following theorem (also proved in ref. [75]):

Theorem 4.3: Broadcastable set is classical

For any valid transformation $T : V_A \to V_A \otimes V_A$, Δ_T is a simplex whose vertices are jointly distinguishable states.

Proof. See Thm. 2 in [75]. ■

Consequently, one can also characterize state spaces in terms of broadcasting. In the same vein of the no-cloning theorem, we have that a state space is non-classical if and only if there is no map that broadcasts an arbitrary state. Hence, we conclude that not even broadcasting of arbitrary states is allowed in non-classical systems, a fact known as the **no-broadcasting** theorem.

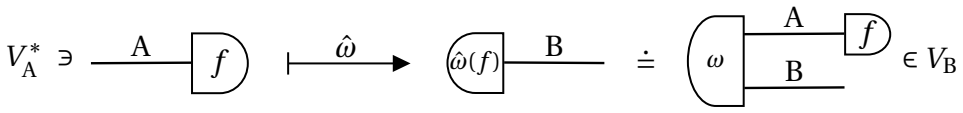
Later on, we will present a notion that is dual to broadcasting, in the sense that it is a diagram equality similar to equation (4.4), but which involves effects rather than states. In this new scenario, we will investigate if there is an analogue of theorem 4.3, that is, if the effects that satisfy the proposed diagram equality are also classical. That would potentially imply in a "no-broadcasting theorem", but for effects.

4.2 Two protocols

4.2.1 Teleportation, purification and Hilbert spaces

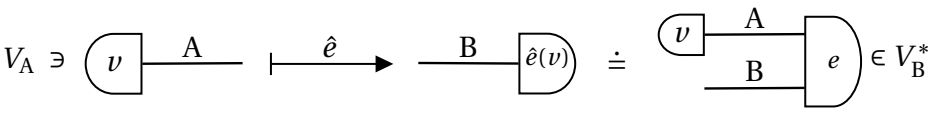
In Section 2.2, we presented a protocol which allowed a party (Alice) to transmit an unknown qubit state to a (potentially) distant party (Bob), only by requiring classical communication and shared entanglement between the parties. Below, we present a looser version of this protocol, called **conclusive teleportation**, where the success of teleportation is based on the occurrence of an event [76]. After that, we will mention how this protocol is tied to Hilbert space representations of GPT systems that satisfy a principle known as **purification** [72].

To state the protocol, we observe that any bipartite state $\omega \in \mathcal{S}_A \boxtimes \mathcal{S}_B$ induces a linear transformation $\hat{\omega}: V_A^* \to V_B$ that is given as follows:



(4.5)

Similarly, any bipartite effect $e \in \mathcal{E}_A \boxtimes \mathcal{E}_B$ induces a linear transformation $\hat{e}: V_A \to V_B^*$ given by

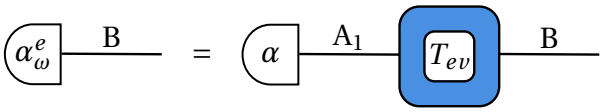


(4.6)

As the diagrams indicate, these maps take effects to (sub)normalized states and states to effects, respectively. From now on, we suppose that the composition rule $\boxtimes$ allows both entangled states and entangled effects.

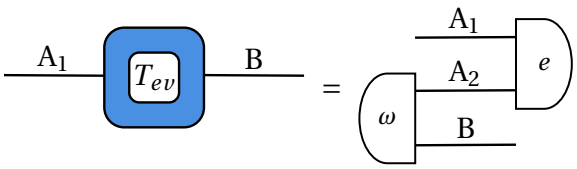
Let A_1A_2B be a tripartite GPT system, where A_1 is the system to be teleported. The protocol setting is the same as in Section 2.2, *i.e*, we suppose that Alice and Bob only have access to classical communication and share a bipartite entangled state $\omega \in \mathcal{S}_{A_2} \boxtimes \mathcal{S}_B$.

Now, suppose that Alice performs a binary measurement $\{e, u_{A_1A_2} - e\}$ on A_1A_2, where e is an entangled effect, and let A_1 be in an arbitrary state $\alpha \in \mathcal{S}_{A_1}$. Upon occurrence of outcome e, system B is described the (sub-normalized) conditional state α_ω^e, which is given by



(4.7)

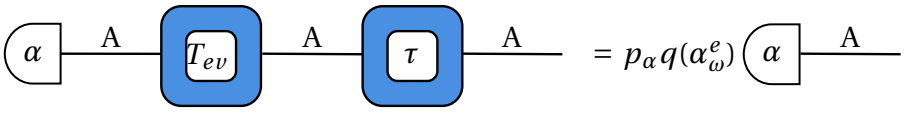
where T_{ev} is the valid transformation[2] obtained by appropriately combining the diagrams in equations (4.5) and (4.6):



(4.8)

This transformation maps $\mathcal{S}_{\mathrm{A}_1}$ into all possible conditional states in B given that Alice read outcome e.

Without loss of generality, we can assume that A_1 and B are the same system[3] and we relabel both as A. Teleportation is only possible if there exists a valid transformation $\tau : V_{\mathrm{A}} \to V_{\mathrm{A}}$ (dependent on the pair (ω, e)), such that, for any $\alpha \in \mathcal{S}_{\mathrm{A}}$:



(4.9)

where $p_\alpha \in (0, 1]$ is the probability that τ correctly implements the correction of the conditional state α^e_ω, and $q(\alpha^e_\omega)$ is the normalization of said state. Equivalently, τ corrects the normalized state $\alpha^e_\omega / q(\alpha^e_\omega)$ with probability p_α.

In ref. [76], Barnum, Barrett, Leifer and Wilce proved that if a given pair (ω, e) can be used for conclusive teleportation, then the map T_{ev} is invertible, and vice versa. However, these conditions do not answer the following question:

In a composite GPT system $\mathrm{AA}_2\mathrm{A}$*, does a pair* (ω, e) *that supports conclusive teleportation exist?*

In general, there is no guarantee that such a pair exists, which means that, differently from the no-cloning theorem, the teleportation protocol is not ubiquitous across all non-classical theories that have entangled states and entangled effects. In light of this, is there any principle which guarantees the possibility of teleportation?

In ref. [72], Chiribella, D'Ariano and Perinotti showed that conclusive teleportation is one, among many, of the consequences[4] of a principle called purification. We state it below, after defining what a purification is.

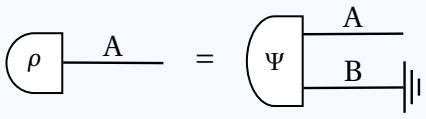
[2] Here, "ev" is short for "evaluation".

[3] Otherwise, we would have to require the systems to be isomorphic. In the original work where the GPT formulation of the protocol was presented [76], the authors considered this general case.

[4] To cite just few of them, purification implies that: there is an analogue of the Choi representation theorem; any two pure states are connected by a reversible transformation; entanglement-breaking channels are measure-and-prepare channels.

Definition 4.4: Purification

Let A be a GPT system. A purification of a state $\rho \in \mathcal{S}_A$ is a bipartite pure state $\Psi \in \mathcal{S}_{AB}$, where B is some other GPT system, such that



(4.10)

In this case, the system B is called a purifying system.

Principle 4.1: Purification

Every GPT system admits a purifying system.

The purification principle captures the idea that any source of randomness in the preparation of a system originates from incomplete knowledge about how said system is correlated to its surroundings.

Now, an interesting consequence of the existence of teleportation in theories that additionally satisfy the purification principle is the following:

Theorem 4.4: Upper bound on teleportation probability

Suppose that a theory is locally tomographic. Given $e \in \mathcal{E}_{AA_2}$ and $\omega \in \mathcal{S}_{A_2A}$, if the pair (ω, e) supports conclusive teleportation with constant p_α, then

$$p_\alpha \leq \frac{1}{\dim V_A}. \tag{4.11}$$

Additionally, if this theory also satisfies the purification principle, then

$$p_\alpha \leq \frac{1}{d_A^2}, \tag{4.12}$$

where d_A is the maximal number of jointly distinguishable states in A. In both cases, there exists such a pair that saturates the given bounds.

Proof. See Theorems 13.6, 13.1 and 13.2 in [77]. ■

From the previous theorem, it follows that

$$\dim V_A = d_A^2. \tag{4.13}$$

In particular, this means that in a locally tomographic theory with purification, any state can be represented as a $d_A \times d_A$ real matrix M. One can equivalently represent

a state as a Hermitian matrix in a finite-dimensional complex Hilbert space, it only requires mapping $M \mapsto \frac{1}{2}(M + M^t) + \frac{i}{2}(M - M^t)$, where t denotes transposition.

In conjunction with local tomography, the purification principle is one of the principles required to derive the mathematical structure of quantum theory, as it was also shown by Chiribella, D'Ariano and Perinotti in ref. [78]. We point out that this is one among several proposed derivations of quantum theory (see ref. [51]), so the principles the authors chose are not the only possible ones consistent with the theory.

4.2.2 BB84 and existence of entanglement

Besides the teleportation protocol, another important protocol which we extend from quantum information theory is the BB84 cryptographic protocol [15], named after its proposers, Charles Bennett and Gilles Brassard, and its year of publication. Its motivation is as follows.

Suppose that two parties, which by now need no introduction, want to exchange messages (say bits) between each other, but they cannot assure that the communication channel they use is secure, *i.e.*, some malicious third party (Eve) might intercept and read the content of their messages. The essence of any cryptographic protocol is to devise a scheme that makes their messages look unintelligible to anyone but themselves.

So, suppose that $m = m_1 \dots m_n$, $m_i \in \{0, 1\}$, is the message that Alice wishes to send to Bob. A simple scheme they can use to protect their communication can be formulated by requiring that they share a common sequence of uncorrelated bits $k = k_1 \dots k_n$, which is kept secret from anyone else. We call this sequence a (secret) key.

If the parties have established a key, Alice can encrypt her message by summing the entries of m and k, producing a new string of bits $c = k \oplus m$, where $\oplus$ denotes sum modulo 2. As long as c looks sufficiently random, m is secured [79, 80]. Upon receiving c, Bob simply applies the map $k \oplus c$ and recovers m. The question is:

Can Alice and Bob devise a secure protocol that generates the key k?

This kind of task, which involves a mechanism to create and share a key, is a key distribution task. As the original BB84 is a key distribution protocol that uses quantum systems, it is called a quantum key distribution (QKD) protocol. The main advantage of this protocol is that its keys are generated by exploiting non-classical features of quantum theory, namely: incompatible measurements, the no-cloning theorem, and entanglement. Thus, its security is build upon physical constraints, rather than on typical computational hardness assumptions.

By now, we know that none of the mentioned features are exclusive to quantum theory, hence we have opted to present the protocol in a "theory-independent"

manner, following reference [81]. To the reader interested in QKD or more generally in quantum cryptography, references [6] and [80] provide an excellent introduction to the matter.

In order to formulate a version of the BB84 protocol in GPTs, we need the following theorem from [81]:

Theorem 4.5: Non-classicality of cones

A proper cone C is non-classical if and only if there are non-zero vectors $x_0, x_1, x_+, x_- \in C$ and functionals $f_0, f_1, f_+, f_- \in C^*$ such that:

1. $x_0 + x_1 = x_+ + x_-$ and $f_0 + f_1 = f_+ + f_-$;
2. $f_0(x_1) = f_1(x_0) = f_+(x_-) = f_-(x_+) = 0$;
3. $f_i + f_j$ is strictly positive for all $i \in \{0,1\}$, $j \in \{+,-\}$.

Assume that Alice's and Bob's systems are nonclassical. The preceding theorem guarantees the existence of states ρ_i and ρ_j, $i \in \{0,1\}$, $j \in \{+,-\}$, such that all three conditions hold: we just have to set $x_i = p_i\rho_i$ and $x_j = q_j\sigma_j$, $p_i, q_j > 0$. We can assume that $p_0 + p_1 = q_+ + q_- = 1$ and that $f_0 + f_1 = f_+ + f_- = l$, with $l \preccurlyeq u_A$. The protocol goes as follows:

1. Alice tosses a fair coin. If her outcome is heads, she prepares one of the states ρ_0,ρ_1 with prior probabilities p_0, p_1. If her outcome is tails, she prepares one of the states σ_+, σ_- with prior probabilities q_+, q_-. An eavesdropper cannot distinguish both scenarios, due to condition 1 from theorem 4.5.
2. She sends her system to Bob.
3. Bob tosses a fair coin and implements one of the following unambiguous discrimination[5] measurements: $\{f_0, f_1, u_A - l\}$ (if heads) or $\{f_+, f_-, u_A - l\}$ (if tails).
4. They publicly announce the results of their coin tosses and remove the rounds in which their results were different or when Bob obtained an inconclusive measurement outcome.
5. Due to the second condition in theorem 4.5, they have established a perfectly correlated sequence of bits.

In the quantum version of the protocol, the states ρ_0, ρ_1, ρ_+ and ρ_- are the eigenstates of σ_Z and σ_X, *i.e.* $|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|$ and $|-\rangle\langle -|$. The effects are also given by

[5]This is a measurement in which although Bob can be sure of what state he received, if the outcome is $u_A - l$, then he cannot conclude which state he was given.

these operators, and in this case they actually sum to the unit effect, which is $\mathbb{I}$. Hence, in the quantum case, Bob's measurements correctly tell him which state was prepared (provided that he measures in the same basis that Alice prepared her state).

Since the no-cloning theorem holds in any non-classical GPT system, if a malicious party, Eve, tried to intercept the system sent by Alice, she would not have the means to figure out in which state the system was prepared, since cannot create perfect clones to estimate the state via an IC measurement. A sensible strategy for Eve[6] is to try to mimic Bob by randomly choosing to measure either $\{f_0, f_1, u_A - l\}$ or $\{f_+, f_-, u_A - l\}$. This however, would reduce the number of rounds in which Alice and Bob should agree (~50% of the cases in the quantum version). Hence, if Eve intervenes enough times, Alice and Bob could actually detect her presence and abort the protocol.

To actually generate a key, Alice and Bob still need to perform some classical post-processing of the correlated bits they obtained (assuming Eve was undetected), which will also comprise of a step that minimizes Eve's knowledge of the key. For details, see section 4.2 in ref. [80].

Theorem 4.5 has an interesting consequence for bipartite systems. First, we bring to attention that classical subsystems (of locally tomographic GPTs) cannot be entangled with any other system[7]:

Proposition 4.1: Classical systems cannot be entangled with other systems

Let Ω be a classical system. If A is a non-classical GPT system, then

$$\mathcal{S}_A \otimes_{\min} \Delta_\Omega = \mathcal{S}_A \otimes_{\max} \Delta_\Omega.$$

Proof. See proposition 5.20 from ref. [53]. ■

Recall that any composition rule satisfies $\otimes_{\min} \subseteq \boxtimes \subseteq \otimes_{\max}$. Then, the previous proposition shows that if at least one of the subsystems is classical, there is no entanglement both at the level of states and of effects, regardless of the adopted composition rule. Using theorem 4.5, Aubrun, Lami, Palazuelos and Plávala showed in ref. [81] that this absence also implies that at least one of the subsystems is classical. The possibility of executing the BB84 protocol is thus a necessary and sufficient condition for the existence of entangled states or entangled effects in a theory.

[6] Here we have implicitly supposed that the malicious party only acts after Alice sent the system and before Bob measured it.

[7] For a construction of theories where classical systems can exhibit entanglement, see [82].

4.3 No monogamy of entanglement in $\mathrm{QT}_{\mathbb{R}}$

In this section, we will examine how quantum theory based on real numbers violates the monogamy of entanglement relations presented in Chapter 1. In order to distinguish this GPT from traditional quantum theory, we refer to it as $\mathrm{QT}_{\mathbb{R}}$. Most of our discussion is based on the in refs. [65, 83–85].

Systems in $\mathrm{QT}_{\mathbb{R}}$ are defined in a similar fashion as to how they are defined in quantum theory. That is, the state space is still given by:

$$\mathcal{D}(\mathcal{H}_{\mathbb{R}}) = \{\rho \in Lin(\mathcal{H}_{\mathbb{R}}) \mid \mathrm{Tr}\,\rho = 1,\ \rho \succcurlyeq 0\}, \tag{4.14}$$

but now $\mathcal{H}_{\mathbb{R}}$ is a *real* vector space. Consequently, states are not positive Hermitian operators, but actually positive *symmetric* operators.

The fact that states are symmetric operators implies that they belong to a vector space of dimension $\frac{d(d+1)}{2}$, where $d = \dim \mathcal{H}_{\mathbb{R}}$. Let us analyze the implications of this when we have two systems.

We still assume that the Hilbert space of bipartite systems is $\mathcal{H}_{\mathrm{A}}^{\mathbb{R}} \otimes \mathcal{H}_{\mathrm{B}}^{\mathbb{R}}$. Hence, bipartite states in $\mathrm{QT}_{\mathbb{R}}$ span a vector space of dimension

$$\dim V_{\mathrm{AB}} = \dim(\mathrm{span}(\mathcal{D}(\mathcal{H}_{\mathrm{AB}}^{\mathbb{R}}))) = \frac{d_{\mathrm{A}} d_{\mathrm{B}}(d_{\mathrm{A}} d_{\mathrm{B}} + 1)}{2}. \tag{4.15}$$

However, the product of the dimensions of the space generated by each system is:

$$\dim V_{\mathrm{A}} \dim V_{\mathrm{B}} = \frac{d_{\mathrm{A}}(d_{\mathrm{A}}+1)}{2}\frac{d_{\mathrm{B}}(d_{\mathrm{B}}+1)}{2} = \frac{d_{\mathrm{A}} d_{\mathrm{B}}(d_{\mathrm{A}} d_{\mathrm{B}}+1)}{4} + \frac{d_{\mathrm{A}} d_{\mathrm{B}}(d_{\mathrm{A}}+d_{\mathrm{B}})}{4}. \tag{4.16}$$

A short calculation shows that $d_{\mathrm{A}} d_{\mathrm{B}} + 1 \geq d_{\mathrm{A}} + d_{\mathrm{B}}$, with equality holding only if one of the systems is one-dimensional. Hence, for local dimensions $d_{\mathrm{A}}, d_{\mathrm{B}} \geq 2$, we have $\dim V_{\mathrm{A}} \dim V_{\mathrm{B}} < \dim V_{\mathrm{AB}}$. This shows that $\mathrm{QT}_{\mathbb{R}}$ is not a locally tomographic theory.

We now restrict ourselves to the simplest non-trivial case, which is when both systems are 2-dimensional. In analogy to qubits, these systems are called *rebits*. In this case, we have $\dim V_{\mathrm{A}} = \dim V_{\mathrm{B}} = 3$ and $\dim V_{\mathrm{AB}} = 10$.

Since rebits are $2x2$ symmetric operators, they can be written as

$$\rho^{\mathbb{R}} = \frac{1}{2}(\mathrm{I} + r_X \sigma_X + r_Z \sigma_Z), \tag{4.17}$$

where $\vec{r} = (r_X, r_Z)$ is the Bloch vector (equation (1.15)) for rebits. From this decomposition, we observe that rebits are trivially embedded in the Bloch ball: their state space is the disk obtained from the intersection of the ball with the x-z plane.

An arbitrary two-rebit state can be written as

$$\rho_{\mathrm{AB}}^{\mathbb{R}} = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \sum_{i\in\{X,Z\}} r_i\sigma_i\otimes\mathbb{I} + \sum_{j\in\{X,Z\}} s_j\mathbb{I}\otimes\sigma_j + \sum_{i,j\in\{X,Z\}} t_{ij}\sigma_i\otimes\sigma_j\right) + t_{YY}\sigma_Y\otimes\sigma_Y, \quad (4.18)$$

where $\sigma_Y\otimes\sigma_Y$ is the additional component not present in any of the local systems.

Notice that $\sigma_Y\otimes\sigma_Y$ is an entangled observable in real quantum theory, for product observables are linear combinations of products involving $\mathbb{I}, \sigma_X$ and σ_Z. Consequently, the states

$$\omega_\pm \doteq \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} \pm \sigma_Y\otimes\sigma_Y) \quad (4.19)$$

are entangled states in $\mathrm{QT}_{\mathbb{R}}$, even though they are separable states in ordinary quantum theory.

For any pair of rebits, their (real) tangle [85] is given by

$$\tau_{\mathrm{AB}}^{\mathbb{R}} = \left|\mathrm{Tr}\left(\rho_{\mathrm{AB}}^{\mathbb{R}}\sigma_Y\otimes\sigma_Y\right)\right|^2. \quad (4.20)$$

Hence, in $\mathrm{QT}_{\mathbb{R}}$ the states $\omega_\pm$ are maximally entangled states, since their real tangle is 1. In ordinary quantum theory, we know that maximally entangled systems cannot be entangled with any other system (as discussed in Section 1.4). This is not the case in $\mathrm{QT}_{\mathbb{R}}$, the maximally entangled state ω_+ is actually the reduced state of the following entangled tripartite state:

$$\frac{1}{2}(|000\rangle - |011\rangle - |101\rangle - |110\rangle) \quad (4.21)$$

for any choice of two subsystems. In fact, the state above can be generalized to n rebits and still have any two-reduced states being maximally entangled, as proved in ref. [83]. Thus, in real quantum theory there exist maximally entangled states that can share their entanglement with an arbitrary number of parties.

In a recent work [65], it was investigated how locally-inaccessible entanglement could be useful in Bell nonlocality or teleportation. It turned out that this type of entanglement is useless for both of them.

4.4 Original results

In this last section we present two original results. The first one is a negative result related to supra-quantum tripartite nonlocality in a GPT whose single systems are quantum. The second one is related to compatibility of measurements and a concept that is dual to broadcasting.

4.4.1 Nonlocality in $\mathrm{QT}_{\otimes_{\max}}$

In our discussion on Bell inequalities in Chapter 2, we pointed out how non-signalling correlations are always realizable using local quantum measurements and some Hermitian operator (proposition 2.3). In what follows, we will analyze the particular case where this Hermitian operator is a state in a composite GPT system whose elementary systems are quantum, but whose composition rule is given by the maximal tensor product. Let us denote by $\mathrm{QT}_{\otimes_{\max}}$ the GPT whose states are obtained precisely in this manner. We shall refer to this theory as "maximal quantum theory", where "maximal" alludes to the composition rule adopted.

Our main interest in $\mathrm{QT}_{\otimes_{\max}}$ comes from its apparent simplicity: it is obtained by only requiring that states of composite quantum systems should be consistent with the no-signalling condition and nothing else[8]. In light of what we have discussed about composition rules and about nonlocality in this thesis, this small modification raises two questions:

1. *What are the additional states in* $\mathrm{QT}_{\otimes_{\max}}$*?*
2. *Do the additional states allow supra-quantum nonlocality?*

Regarding the first question, the answer is surprisingly simple:

Proposition 4.2: Additional states in $\mathrm{QT}_{\otimes_{\max}}$

The non-quantum states in $\mathrm{QT}_{\otimes_{\max}}$ are precisely the unit-trace entanglement witnesses.

Proof. Let $\mathcal{D}(\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_B)$ be quantum systems. A state in $\mathcal{D}(\mathcal{H}_A) \otimes_{\max} \mathcal{D}(\mathcal{H}_B)$, is any unit-trace Hermitian operator ω such that

$$0 \leq \mathrm{Tr}(\omega E_A \otimes F_B) \leq 1,$$

where E_A and F_B are any effects in their respective systems. In particular, the inequalities above hold whenever E_A and F_B are rank-1 operators. Thus, for any normalized vectors $|\psi\rangle \in \mathcal{H}_A$ and $|\varphi\rangle \in \mathcal{H}_B$:

$$\langle \psi \otimes \varphi | \omega | \psi \otimes \varphi \rangle \geq 0.$$

This implies that $\mathrm{Tr}(\omega\rho)$ is always positive whenever ρ is a separable quantum state. Assuming that ω is not a quantum state itself, then by definition 1.2, it is an entanglement witness.

[8] That is a partial truth, because we are also implicitly adhering to the no-restriction hypothesis and local tomography.

The argument still holds when there are more than two systems. Consequently, non-quantum states in maximal quantum theory are precisely unit-trace entanglement witnesses. ■

Now, when it comes to the second question, the answer is less simple, but still surprising, at least in the bipartite Bell scenarios. The following lemma is not an original result, rather, it was first proved in ref. [86].

Lemma 4.1: Bipartite correlations in $\mathrm{QT}_{\otimes_{\max}}$

Let ω^{AB} be a bipartite state in $\mathrm{QT}_{\otimes_{\max}}$ and let $\{E_a\}_a$, $\{F_b\}_b$ be arbitrary POVMs in system A and B, respectively. The probabilities

$$p(a,b) = \mathrm{Tr}\big(\omega^{AB} E_a \otimes F_b\big)$$

have a quantum realization.

Proof. The Choi representation theorem (thm. 1.3) implies that that any unit-trace entanglement witness ω^{AB} can be written as

$$\omega^{AB} = (\mathbb{I}_A \otimes \Lambda)(|\Psi\rangle\langle\Psi|),$$

where $\Lambda : Lin(\mathcal{H}_A) \to Lin(\mathcal{H}_B)$ is a positive, trace-preserving map, and $|\Psi\rangle\langle\Psi|$ is a pure bipartite quantum state. We then have:

$$\begin{aligned}\mathrm{Tr}\big(\omega^{AB} E_a \otimes F_b\big) &= \mathrm{Tr}[(E_a \otimes F_b)(\mathbb{I}_A \otimes \Lambda)\,|\Psi\rangle\langle\Psi|] \\ &= \mathrm{Tr}\big(E_a \otimes \tilde{F}_b\,|\Psi\rangle\langle\Psi|\big),\end{aligned}$$

where $\tilde{F}_b$ is an effect in system A defined via the action of the adjoint map $\Lambda^* : Lin(\mathcal{H}_B) \to Lin(\mathcal{H}_A)$ defined via $\mathrm{Tr}\big(\Lambda(\rho^A)F^B\big) = \mathrm{Tr}\big(\rho^A \Lambda^*(F^B)\big)$.

Since Λ is positive, so is Λ^*, thus $\Lambda^*(F_b)$ is a positive operator for all b. The fact that it is trace-preserving implies that Λ^* preserves the identity. Thus $\{\Lambda^*(F_b)\}_b$ is a valid POVM in A. Consequently:

$$p(a,b) = \mathrm{Tr}\big(|\Psi\rangle\langle\Psi|\,E_a \otimes \tilde{F}_b\big)$$

is a quantum realization of the probabilities obtained from ω^{AB}, $\{E_a\}_a$ and $\{F_b\}_b$. ■

The immediate consequence of the previous lemma is that the behaviours in every bipartite Bell scenario in $\mathrm{QT}_{\otimes_{\max}}$ can be mapped to quantum behaviours. Hence, **bipartite systems in maximal quantum theory always obey Tsirelson's bound**. Con-

sequently, $\mathrm{QT}_{\otimes_{\max}}$ and standard quantum theory cannot be told apart if we only examine the statistics produced in bipartite Bell scenarios.

In the tripartite case, the situation starts to deviate. As shown in ref. [46], there is a Bell inequality for which $\mathrm{QT}_{\otimes_{\max}}$ surpasses the maximal quantum violation. The scenario is a $(3,2,2)$ Bell scenario and the inequality reads as

$$p(000|000) + p(110|011) + p(011|101) + p(101|110) \leq 1, \tag{4.22}$$

which is called the *guess your neighbour's input* (GYNI) inequality. It was introduced in ref. [87], and in the same work, the authors show that the maximal value of β_{GNYI} using quantum theory is 1, hence this is a Bell inequality that's actually never violated by quantum systems.

The strategy the authors of ref. [46] found to prove that $\mathrm{QT}_{\otimes_{\max}}$ violates the GYNI inequality relied in a construction of entanglement witnesses that witness bound entanglement, which, in a rough manner, is the kind of entanglement that cannot be used to generate Bell states. We do not delve into the explicit construction and refer the reader to ref. [46] for further details. What is relevant here is that $\mathrm{QT}_{\otimes_{\max}}$ can produce stronger-than-quantum correlations in the tripartite case. In light of this, we aimed at answering the following question:

Does $\mathrm{QT}_{\otimes_{\max}}$ *also violate Tsirelson's bound of Svetlichny's inequality (eq. (2.24))?*

A positive answer to this question would demonstrate that $\mathrm{QT}_{\otimes_{\max}}$ could be more nonlocal than standard quantum theory in a Bell scenario where there exists a gap between Tsirelson's bound $(4\sqrt{2})$ and optimal local strategies (4). It would also demonstrate a difference between genuine tripartite nonlocality in both theories.

We have managed to prove the following:

Theorem 4.6: $\mathrm{QT}_{\otimes_{\max}}$ **tripartite correlations in Svetlichny's setting**

Let ω^{ABC} be a tripartite state in maximal quantum theory. For any observables A_0, A_1, B_0, B_1, C_0 and C_1 with eigenvalues in $\{-1,+1\}$, we have

$$\left|\mathrm{Tr}\left(\omega^{\mathrm{ABC}} S_3\right)\right| \leq 4\sqrt{2},$$

where

$$S_3 = S_{CHSH} \otimes C_1 + S_{CHSH'} \otimes C_0,$$

is the observable associated to Svetlichny's inequality.

Proof. Let ω^{ABC} be a tripartite state in maximal quantum theory, and let $\langle S_3 \rangle$ be given

by $\mathrm{Tr}\big(\omega^{\mathrm{ABC}} S_3\big)$. By virtue of the triangle inequality, we have the following:

$$|\langle S_3\rangle| = |\langle S_{CHSH}\otimes C_1\rangle + \langle S_{CHSH'}\otimes C_0\rangle| \leq |\langle S_{CHSH}\otimes C_1\rangle| + |\langle S_{CHSH'}\otimes C_0\rangle|,$$

where the expected values $\langle S_{CHSH}\otimes C_1\rangle$ and $\langle S_{CHSH'}\otimes C_0\rangle$ are computed similarly to $\langle S_3\rangle$. Thus, if each term on the right-hand side is bounded by $2\sqrt{2}$, then $|\langle S_3\rangle|$ is necessarily bounded by $4\sqrt{2}$, which establishes the theorem. In what follows, we show that $|\langle S_{CHSH}\otimes C_1\rangle| \leq 2\sqrt{2}$. We divide the proof in small steps for better readability.

Step 1) Since C_1 is Hermitian and has eigenvalues $+1$ and -1, we can use the spectral theorem to write

$$C_1 = P_{+|1} - P_{-|1},$$

where $P_{+|1}$ and $P_{-|1}$ are orthogonal projectors whose sum is $\mathbb{I}_{\mathrm{C}}$. The (unnormalized) conditional states on system AB upon a measurement of C_1 are

$$\tilde{\omega}^{\mathrm{AB}}_{+|1} = \mathrm{Tr}_C\left[(\mathbb{I}_{\mathrm{AB}}\otimes P_{+|1})\omega^{\mathrm{ABC}}\right],$$

and

$$\tilde{\omega}^{\mathrm{AB}}_{-|1} = \mathrm{Tr}_C\left[(\mathbb{I}_{\mathrm{AB}}\otimes P_{-|1})\omega^{\mathrm{ABC}}\right].$$

The probabilities of obtaining outcomes ± 1, $p(\pm 1|1)$, are just the trace of these conditional states.

Step 2) The normalized states $\omega^{\mathrm{AB}}_{\pm|1}$ are bipartite states in $\mathrm{QT}_{\otimes_{\max}}$: let E_{A} and E_{B} be arbitrary effects in the appropriate systems. Then:

$$\begin{aligned}\mathrm{Tr}\big(\omega^{\mathrm{AB}}_{\pm 1|1} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\big) &= \frac{1}{p(\pm 1|1)}\mathrm{Tr}\big(\tilde{\omega}^{\mathrm{AB}}_{\pm 1|1} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\big)\\ &= \frac{1}{p(\pm 1|1)}\mathrm{Tr}\left\{\mathrm{Tr}_{\mathrm{C}}\left[(\mathbb{I}_{\mathrm{AB}}\otimes P_{\pm|1})\omega^{\mathrm{ABC}}\right] E_{\mathrm{A}}\otimes E_{\mathrm{B}}\right\}\\ &= \frac{1}{p(\pm 1|1)}\mathrm{Tr}\big(\omega^{\mathrm{ABC}} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\otimes P_{\pm|1}\big).\end{aligned}$$

Thus, $\mathrm{Tr}\Big(\omega^{\mathrm{AB}}_{\pm 1|1} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\Big)$ is non-negative, since it is a ratio of non-negative numbers. We claim that it is also bounded by 1. To see that, note that $E_{\mathrm{A}}\otimes E_{\mathrm{B}}\otimes P_{\pm|1} \preccurlyeq \mathbb{I}_{\mathrm{AB}}\otimes P_{\pm|1}$ (their difference is a positive operator). This implies that:

$$\mathrm{Tr}\big(\omega^{\mathrm{ABC}} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\otimes P_{\pm|1}\big) \leq \mathrm{Tr}\big[\omega^{\mathrm{ABC}}(\mathbb{I}_{\mathrm{AB}}\otimes P_{\pm|1})\big] = p(\pm|1).$$

Consequently, $\mathrm{Tr}\Big(\omega^{\mathrm{AB}}_{\pm 1|1} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\Big)$ is upper bounded by 1. Thus, we have shown that $\mathrm{Tr}\Big(\omega^{\mathrm{AB}}_{\pm 1|1} E_{\mathrm{A}}\otimes E_{\mathrm{B}}\Big) \in [0,1]$ for arbitrary product effects, which makes $\omega^{\mathrm{AB}}_{+1|1}$ and

$\omega^{AB}_{-1|1}$ valid bipartite states.

Step 3) Due to lemma 4.1, bipartite systems in $QT_{\otimes_{max}}$ satisfy the corresponding Tsirelson's bound of any 2-party Bell scenario. In particular, for the CHSH game, this implies that

$$-2\sqrt{2} \leq \mathrm{Tr}\big(S_{CHSH}\omega^{AB}_{\pm 1|1}\big) \leq 2\sqrt{2}.$$

Noting that

$$\langle S_{CHSH} \otimes C_1 \rangle = p(+1|1)\,\mathrm{Tr}\big(S_{CHSH}\omega^{AB}_{+1|1}\big) - p(-1|1)\,\mathrm{Tr}\big(S_{CHSH}\omega^{AB}_{-1|1}\big),$$

the previous inequalities imply that

$$-2\sqrt{2} \leq \langle S_{CHSH} \otimes C_1 \rangle \leq 2\sqrt{2},$$

as desired.

By similar steps, it also holds that

$$-2\sqrt{2} \leq \langle S_{CHSH'} \otimes C_0 \rangle \leq 2\sqrt{2}.$$

■

In view of the previous theorem, we see that $QT_{\otimes_{max}}$ is as as genuinely nonlocal as standard quantum theory when it comes to Svetlichny's inequality. Hence, similarly to bipartite Bell scenarios, the statistics produced in any test of this inequality cannot differentiate the two theories.

The previous theorem extends to an arbitrary number of parties, which establishes that genuine multipartite nonlocality, in the sense of Svetlichny, is the same in both theories.

Theorem 4.7: $QT_{\otimes_{max}}$ ***n*-partite correlations in Svetlichny's setting**

Let $\hat{\omega}^{A_1\ldots A_n}$ be a multipartite state in maximal quantum theory. For any choice of observables $\hat{A}^{(k)}_0, \hat{A}^{(k)}_1$ with eigenvalues $\{-1,+1\}$, $k = 1,\ldots,n$, we have

$$\left|\mathrm{Tr}\big(\hat{\omega}^{A_1\ldots A_n} S_n\big)\right| \leq 2^{n-\frac{1}{2}},$$

where

$$S_n = S_{n-1} \otimes \hat{A}^{(n)}_1 + S'_{n-1} \otimes \hat{A}^{(n)}_0$$

is recursively defined similarly as in equation (2.27b).

Proof. The proof is by induction on n. The base case $n = 2$ has already been proved in lemma 4.1. Now, let $n \geq 3$, suppose that the theorem holds for $n-1$ parties, and let $\hat{\omega}^{\mathrm{A_1 \ldots A_n}}$ and $\hat{A}_0^{(k)}, \hat{A}_1^{(k)}$, $k = 1, \ldots, n$ be as in the statement. By virtue of the triangle inequality, we have

$$\underbrace{\left|\mathrm{Tr}\left(\hat{\omega}^{\mathrm{A_1 \ldots A_n}} S_n\right)\right|}_{\equiv \left|\langle S_n \rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right|} \leq \underbrace{\left|\mathrm{Tr}\left(\hat{\omega}^{\mathrm{A_1 \ldots A_n}} S_{n-1} \otimes \hat{A}_1^{(n)}\right)\right|}_{\equiv \left|\left\langle S_{n-1} \otimes \hat{A}_1^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right|} + \underbrace{\left|\mathrm{Tr}\left(\hat{\omega}^{\mathrm{A_1 \ldots A_n}} S'_{n-1} \otimes \hat{A}_0^{(n)}\right)\right|}_{\equiv \left|\left\langle S'_{n-1} \otimes \hat{A}_0^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right|}.$$

Just as in the tripartite case, after we apply the spectral theorem to $\hat{A}_1^{(n)}$, we can write

$$\left\langle S_{n-1} \otimes \hat{A}_1^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}} = p(+1|1)\,\mathrm{Tr}\left(S_{n-1} \hat{\omega}_{+1|1}^{\mathrm{A_1 \ldots A_{n-1}}}\right) - p(-1|1)\,\mathrm{Tr}\left(S_{n-1} \hat{\omega}_{-1|1}^{\mathrm{A_1 \ldots A_{n-1}}}\right),$$

where

$$\hat{\omega}_{\pm 1|1}^{\mathrm{A_1 \ldots A_{n-1}}} \doteq \left(\frac{1}{p(\pm 1|1)}\right) \mathrm{Tr}_{A_n}\left[\left(\mathbb{I}_{\mathrm{A_1 \ldots A_{n-1}}} \otimes P_{\pm 1}^{(n)}\right) \hat{\omega}^{\mathrm{A_1 \ldots A_n}}\right]$$

is the conditional state of the first $n-1$ systems based on the measurement $\left\{P_{-1}^{(n)}, P_{+1}^{(n)}\right\}$ obtained from the spectral decomposition of $\hat{A}_1^{(n)}$ and

$$p(\pm 1|1) = \mathrm{Tr}\left[\left(\mathbb{I}_{\mathrm{A_1 \ldots A_{n-1}}} \otimes P_{\pm 1}^{(n)}\right) \hat{\omega}^{\mathrm{A_1 \ldots A_n}}\right].$$

Due to the induction hypothesis, the theorem holds for S_{n-1}, thus

$$\left|\left\langle S_{n-1} \otimes \hat{A}_1^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right| \leq 2^{(n-1)-\frac{1}{2}}.$$

By similar steps, we also have that

$$\left|\left\langle S'_{n-1} \otimes \hat{A}_0^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right| \leq 2^{(n-1)-\frac{1}{2}}.$$

As consequence of the previous two inequalities, it holds that

$$\left|\left\langle S_{n-1} \otimes \hat{A}_1^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right| + \left|\left\langle S'_{n-1} \otimes \hat{A}_0^{(n)} \right\rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right| \leq 2^{n-\frac{1}{2}}.$$

Hence, the theorem follows:

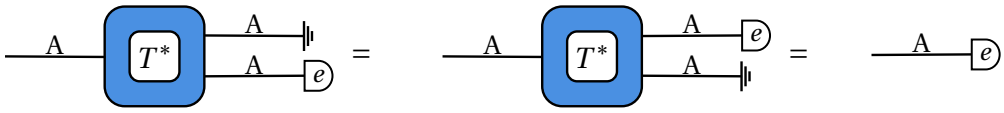
$$\left|\langle S_n \rangle_{\hat{\omega}^{\mathrm{A_1 \ldots A_n}}}\right| \leq 2^{n-\frac{1}{2}}.$$

■

4.4.2 A notion dual to broadcasting

Inspired by the broadcasting condition in equation (4.4), we investigated a notion of non-disturbance for effects. Consider the collection of single-system effects $\mathcal{E}_{T^*} \subset \mathcal{E}_{\mathrm{A}}$

such that



(4.23)

holds for all $e \in \mathcal{E}_{T^*}$. Here, $T^* : V_A^* \otimes V_A^* \to V_A^*$ is the transformation dual to a transformation T. It maps a bipartite effect e_{bip} into a single-system effect via the action

$$T^*(e_{bip})(v) \doteq e_{bip}(T(v)) \tag{4.24}$$

for all $v \in V_A$.

The collection $\mathcal{E}_{T^*}$ is non-empty, since it at least contains the unit effect. Moreover, it is stable under coarse-grainings[9] and under convex combinations. Any measurement whose effects belong to $\mathcal{E}_{T^*}$ will be called an undisturbed measurement. Regarding this type of measurements, we have the following:

Proposition 4.3: Undisturbed measurements are pairwise compatible

Let $T : V_A \to V_A \otimes V_A$ be a valid transformation. Any two single-system measurements $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^m$ whose effects are undisturbed by T^* are compatible.

Proof. The measurement defined by effects $g_{ij} = \frac{1}{2} T^*(e_i \otimes f_j + f_j \otimes e_i)$ satisfies

$$\sum_{j=1}^{m} g_{ij} = e_i \tag{4.25a}$$

and

$$\sum_{i=1}^{n} g_{ij} = f_j \tag{4.25b}$$

for all i and j. Hence, it qualifies as a joint measurement of $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^m$, meaning that these are compatible observables. ■

As mentioned in Chapter 3, a remarkable feature of classical systems is the fact that all measurements are compatible. Bearing this in mind, the previous proposition demonstrates that the non-disturbance condition resembles some sort of classicality, due to the compatibility of any two undisturbed measurements. In what follows, we investigate under which conditions all the undisturbed measurements can be explained in terms of a single observable, that is: when is $\mathcal{E}_{T^*}$ the effect space of a classical subsystem of a non-classical GPT.

Recall that, by the no-broadcasting theorem (thm. 4.3), the set of states broadcast by a map T, Δ_T, is a simplex whose vertices are distinguishable states. Let

[9] Meaning we group effects by simply summing them

$\mathrm{vert}(\Delta_T) = \{\sigma_i\}_{i=1}^n$ denote the collection of such vertices, and let $\mathrm{M} = \{m_i\}_{i=1}^n$ be the measurement that perfectly discriminates them. In the following, let Δ_T^* be the hypercube dual to this simplex (see example 3.2 if necessary).

If each $m_i \in \mathrm{M}$ is not disturbed by T^*, that is, if it satisfies equation (4.23), then Δ_T^* is contained in $\mathcal{E}_{T^*}$. This follows from the stability of $\mathcal{E}_{T^*}$ under coarse-grainings and convex combinations. However, it is not clear whether

$$\mathcal{E}_{T^*} \subseteq \Delta_T^* \tag{4.26}$$

is also true under the hypothesis that M satisfies equation (4.23). Under an additional hypothesis, we managed to show the following:

Theorem 4.8: Sufficient conditions for $\Delta_T^* = \mathcal{E}_{T^*}$

Let $\mathrm{M} = \{m_i\}_{i=1}^n$ be the measurement that perfectly discriminates the vertices of the simplex Δ_T. If M only consists of pure effects and if each m_i satisfies equation (4.23), then the hypercube Δ_T^* is equal to the collection of effects undisturbed by T^*.

In other words, under its hypotheses, the theorem establishes that any measurement with undisturbed effects is explainable in terms of a unique measurement M.

Proof. If each effect in M satisfies the non-disturbance condition, we have already argued that this implies $\Delta_T^* \subset \mathcal{E}_{T^*}$.

Now, let $e \in \mathcal{E}_{T^*}$ be arbitrary. Consider the binary measurement $\{e, u - e\}$. Since both this measurement and M are not disturbed by T^*, by proposition 4.3 it follows that they are compatible. Hence there is a measurement $\mathrm{G} = \{g_{i0}, g_{i1}\}_{i=1}^n$ such that

$$g_{i0} + g_{i1} = m_i, \tag{4.27a}$$

and, for $j \in \{0, 1\}$,

$$\sum_{i=1}^{n} g_{ij} = e_j \tag{4.27b}$$

where we identify $e_0 = e$ and $e_1 = u - e$.

Since m_i is pure, the decomposition in equation (4.27a) is only possible if there exist nonnegative coefficients α_{i0} and α_{i1} such that

$$\begin{aligned} g_{i0} &= \alpha_{i0} m_i, \\ g_{i1} &= \alpha_{i1} m_i, \\ \alpha_{i0} + \alpha_{i1} &= 1. \end{aligned} \tag{4.28}$$

Consequently, due to equations (4.27b) and (4.28), we can write

$$e_j = \sum_{i=1}^{n} \alpha_{ij} m_i, \tag{4.29}$$

which implies that $\{e_0, e_1\}$ is a post-processing[10] of M, and thus it lies inside Δ_T^*. ■

The reason why equation (4.29) implies that e_j belongs to Δ_T^* is as follows. Notice that the vector $\vec{\alpha}_j = (\alpha_{1j}, \ldots, \alpha_{nj})$ belongs to the unit hypercube $[0,1]^n$. Since $[0,1]^n$ is compact and convex, we can express $\vec{\alpha}_j$ as a convex combination of points in $\mathrm{vert}([0,1]^n)$. Now, we identify $\mathrm{vert}([0,1]^n)$ with $\mathrm{vert}(\Delta_T^*)$ and write down e_j using the coefficients of the aforementioned convex combination.

[10] That is, we obtain it via a row stochastic matrix.

5 Conclusions and open questions

> Tanto que fazer!
> E fizemos apenas isto.
> E nunca soubemos quem éramos,
> nem pra quê.
>
> *Cecília Meireles, Humildade*

In this thesis, I have provided an introduction to the GPT framework by illustrating how features taken to be genuinely quantum also manifest themselves in generic non-classical theories. Since, as far as current scientific knowledge indicates, Nature is quantum mechanical, the previous observation raises an immediate question:

What makes quantum theory special among all other non-classical theories?

Although not part of the Millennium Prize problems, I think that this is also a one-million-dollar question. Light-hearted comments aside, several attempts to reconstruct quantum theory have been proposed within the operational approach (see for instance refs. [50, 78, 88]). However, after reading Chris Fuchs' "*Negative Remarks on Operational Approaches to Quantum Theory*" [51], I seem to agree with him (to an extent), in that formulating everything in operational terms feels a bit "lacking". Especially because this approach does not put too much emphasis on the dynamical laws of its systems. For instance, starting strictly from an operational approach, what guarantees that closed quantum systems derived from such an approach must satisfy Schrödinger's equation? In other words,

could one justify why the Hamiltonian is the generator of time translations of a closed system strictly from operational principles?

Although probabilistic theories were first and foremost inspired by an information-theoretic view on physical theories, trying to bridge the gap with the more familiar dynamical view appears to be an interesting long-term research programme.

Turning now to short-term or mid-term questions, mostly related to what we explored in this thesis, we list a few problems that look promising:

1. Despite perfect cloning being impossible, in quantum theory it is possible to devise optimal cloning machines [31, 89], where their quality is based on quantities

such as fidelity or trace distance. Could one devise optimal cloning machines in GPTs too? Analyzing the gbit case could be a sensible start[1].

2. In a similar vein to the previous question, we also know that asymptotic cloning via optimal quantum cloning machines is equal to a task known as state estimation [90]. Does the same hold in the class of GPTs we have considered throughout the text? The proof mostly relies on a monogamy of entanglement argument, so perhaps understanding the very technical work [91] might shed light on this question.

3. $\mathrm{QT}_{\otimes_{\max}}$ is a theory whose effects are just separable quantum effects. What happens if we now construct a theory in which the effects of any two parties could also be entangled? What about any three parties? How is nonlocality affected as we increase the number of parties that can have entangled effects? Surely, if we let this number be arbitrary, then we just recover standard quantum theory. Is there a point at which standard Bell scenarios cannot tell these theories apart?

4. Can a "no-broadcasting" theorem for effects be derived from the non-disturbance condition on effects of diagram (4.23)?

5. Computation has been explored in the framework of probabilistic theories in the works [92–96]. These focused mostly on extending known quantum algorithms to the GPT framework or analyzing the computational power of probabilistic theories. Now, given that the simplest model of computation is the finite automaton [97], what can one say about finite automata constructed using non-classical theories?

6. Can we use computational principles to constraint probabilistic theories? For instance, in quantum theory, by bounding the computational power of parties, a different notion of entanglement is possible [98, 99]. The references in item 5 are a reasonable starting point to investigate this and similar questions.

Lastly, a few subjects that were not discussed in this thesis, but that are worth mentioning.

The first one is generalized (non)contextuality [100, 101]. Roughly speaking, it represents a way of classifying the statistics of experiments into those that are classically explainable (a.k.a. noncontextual) and those that are not. Being able to distinguish these scenarios is relevant whenever a non-classical regime is required for a given task. In the GPT framework, noncontextuality is equivalent to a geometric condition on the state and effect spaces, which is called *simplex embeddability*.

[1] I tried this myself, but it does not seem too straightforward what figure of merit to use to assess the machine's quality.

The second one is paradoxes [102]. The Wigner's friend paradox and its many variations [103] pose intriguing questions on the consistency of quantum theory. In reference [102], Vilasini, Nurgalieva and del Rio investigated which theories allowed the existence of one these variations, known as the Frauchiger-Renner paradox. Perhaps, casting Wigner's original paradox or other variants under an operational light might elucidate what aspects of quantum theory allow the existence of such inconsistencies, or even what is necessary to lift them.

BIBLIOGRAPHY

[1] Nielsen, M. A. & Chuang, I. L. *Quantum computation and quantum information* (Cambridge University Press, Cambridge ; New York, 2010), 10th anniversary ed edn.

[2] Peres, A. (ed.) *Quantum Theory: Concepts and Methods*, vol. 57 of *Fundamental Theories of Physics* (Springer Netherlands, Dordrecht, 2002). URL http://link.springer.com/10.1007/0-306-47120-5.

[3] Wilde, M. M. *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013). URL https://www.cambridge.org/core/books/quantum-information-theory/9DC2CA59F45636D4F0F30D971B677623.

[4] Bengtsson, I. & Zyczkowski, K. *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2006), 1 edn. URL https://www.cambridge.org/core/product/identifier/9780511535048/type/book.

[5] Amaral, B., Baraviera, A. T. & Cunha, M. O. T. *Mecânica quântica para matemáticos em formação*. Colóquio Brasileiro de Matemática; Apostilas (IMPA, Rio de Janeiro, 2011). Meeting Name: Colóquio Brasileiro de Matemática(.

[6] Vidick, T. & Wehner, S. *Introduction to quantum cryptography* (Cambridge University Press, Cambridge, United Kingdom ; New York, NY, USA, 2024).

[7] Holevo, A. S. *Quantum Systems, Channels, Information: A Mathematical Introduction* (De Gruyter, 2019). URL https://www.degruyter.com/document/doi/10.1515/9783110642490/html.

[8] Bae, J. & Kwek, L.-C. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical* **48**, 083001 (2015). URL https://doi.org/10.1088/1751-8113/48/8/083001.

[9] Chiribella, G. On Quantum Estimation, Quantum Cloning and Finite Quantum de Finetti Theorems. In van Dam, W., Kendon, V. M. & Severini, S. (eds.) *Theory of Quantum Computation, Communication, and Cryptography*, 9–25 (Springer, Berlin, Heidelberg, 2011).

[10] Adlam, E. Foundations of Quantum Mechanics. *Elements in the Philosophy of Physics* (2021). URL https://www.cambridge.org/core/elements/foundations-of-quantum-mechanics/7D2F34BA2F54B51FBB33D557B2058D8E. ISBN: 9781108885515 9781108794442.

[11] Leifer, M. S. Is the Quantum State Real? An Extended Review of ψ-Ontology Theorems. *Quanta* **3**, 67–155 (2014). URL https://dankogeorgiev.com/ojs/index.php/quanta/article/view/16.

[12] Cabello, A. Interpretations of Quantum Theory: A Map of Madness. In Lombardi, O., Fortin, S., Holik, F. & López, C. (eds.) *What is Quantum Information?*, 138–144 (Cambridge University Press, 2017), 1 edn. URL https://www.cambridge.org/core/product/identifier/9781316494233%23CN-bp-7/type/book_part.

[13] Müller, M. Probabilistic theories and reconstructions of quantum theory. *SciPost Physics Lecture Notes* 028 (2021). URL https://scipost.org/SciPostPhysLectNotes.28.

[14] Adlam, E. Operational theories as structural realism. *Studies in History and Philosophy of Science* **94**, 99–111 (2022). URL https://www.sciencedirect.com/science/article/pii/S0039368122000814.

[15] Bennett, C. H. & Brassard, G. Quantum Cryptography: Public Key Distribution and Coin Tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (Bangalore, India, 1984). URL https://www.sciencedirect.com/science/article/pii/S0304397514004241.

[16] Skrzypczyk, P. *Semidefinite Programming in Quantum Information Science*. IOP Series in Quantum Technology Series (Institute of Physics Publishing, Bristol, 2023), 1st ed edn.

[17] Hardy, L. & Wootters, W. K. Limited Holism and Real-Vector-Space Quantum Theory. *Foundations of Physics* **42**, 454–473 (2012). URL https://doi.org/10.1007/s10701-011-9616-6.

[18] Dür, W., Vidal, G. & Cirac, J. I. Three qubits can be entangled in two inequivalent ways. *Physical Review A* **62**, 062314 (2000). URL https://link.aps.org/doi/10.1103/PhysRevA.62.062314.

[19] Gühne, O. & Tóth, G. Entanglement detection. *Physics Reports* **474**, 1–75 (2009). URL https://linkinghub.elsevier.com/retrieve/pii/S0370157309000623.

[20] Bruß, D. & Leuchs, G. (eds.) *Quantum Information: From Foundations to Quantum Technology Applications* (Wiley, 2016), 1 edn. URL https://onlinelibrary.wiley.com/doi/book/10.1002/9783527805785.

[21] Cunha, M. d. O. T. *Emaranhamento: caracterização, manipulação e consequências*. Ph.D. thesis, Universidade Federal de Minas Gerais (2005). URL

https://repositorio.ufmg.br/handle/1843/ESCZ-6L6GNL. Accepted: 2019-08-10T14:58:32Z.

[22] Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A* **223**, 1–8 (1996). URL https://linkinghub.elsevier.com/retrieve/pii/S0375960196007062.

[23] Terhal, B. M. Bell inequalities and the separability criterion. *Physics Letters A* **271**, 319–326 (2000). URL https://linkinghub.elsevier.com/retrieve/pii/S0375960100004011.

[24] Watrous, J. *The Theory of Quantum Information* (Cambridge University Press, 2018), 1 edn. URL https://www.cambridge.org/core/product/identifier/9781316848142/type/book.

[25] Bruß, D. Characterizing entanglement. *Journal of Mathematical Physics* **43**, 4237–4251 (2002). URL https://pubs.aip.org/jmp/article/43/9/4237/230925/Characterizing-entanglement.

[26] Coffman, V., Kundu, J. & Wootters, W. K. Distributed entanglement. *Physical Review A* **61**, 052306 (2000). URL https://link.aps.org/doi/10.1103/PhysRevA.61.052306.

[27] Osborne, T. J. & Verstraete, F. General Monogamy Inequality for Bipartite Qubit Entanglement. *Physical Review Letters* **96**, 220503 (2006). URL https://link.aps.org/doi/10.1103/PhysRevLett.96.220503.

[28] Peres, A. Separability Criterion for Density Matrices. *Physical Review Letters* **77**, 1413–1415 (1996).

[29] Choi, M.-D. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications* **10**, 285–290 (1975). URL https://linkinghub.elsevier.com/retrieve/pii/0024379575900750.

[30] Wootters, W. K. & Zurek, W. H. A single quantum cannot be cloned. *Nature* **299**, 802–803 (1982). URL https://doi.org/10.1038/299802a0.

[31] Werner, R. F. Optimal cloning of pure states. *Physical Review A* **58**, 1827–1832 (1998). URL https://link.aps.org/doi/10.1103/PhysRevA.58.1827.

[32] Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters* **70**, 1895–1899 (1993). URL https://link.aps.org/doi/10.1103/PhysRevLett.70.1895.

[33] Bell, J. S. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika* **1**, 195–200 (1964). URL https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195.

[34] Einstein, A., Podolsky, B. & Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review* **47**, 777–780 (1935). URL https://link.aps.org/doi/10.1103/PhysRev.47.777.

[35] Fine, A. The Einstein-Podolsky-Rosen Argument in Quantum Theory. In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2020), summer 2020 edn. URL https://plato.stanford.edu/archives/sum2020/entries/qt-epr/.

[36] Myrvold, W., Genovese, M. & Shimony, A. Bell's Theorem. In Zalta, E. N. & Nodelman, U. (eds.) *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2024), spring 2024 edn. URL https://plato.stanford.edu/archives/spr2024/entries/bell-theorem/.

[37] Ekert, A. K. Quantum cryptography based on Bell's theorem. *Physical Review Letters* **67**, 661–663 (1991). URL https://link.aps.org/doi/10.1103/PhysRevLett.67.661.

[38] Šupić, I. & Bowles, J. Self-testing of quantum systems: a review. *Quantum* **4**, 337 (2020). URL https://quantum-journal.org/papers/q-2020-09-30-337/.

[39] Cirel'son, B. S. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics* **4**, 93–100 (1980). URL https://doi.org/10.1007/BF00417500.

[40] Scarani, V. *Bell Nonlocality* (Oxford University PressOxford, 2019), 1 edn. URL https://academic.oup.com/book/36528.

[41] Scarani, V. The device-independent outlook on quantum physics (lecture notes on the power of Bell's theorem) (2015). URL http://arxiv.org/abs/1303.3081. ArXiv:1303.3081 [quant-ph].

[42] Rabelo, R. *Não-localidade quântica: matemática e fundamentos*. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais (2010).

[43] Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Reviews of Modern Physics* **86**, 419–478 (2014). URL https://link.aps.org/doi/10.1103/RevModPhys.86.419.

[44] Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters* **23**, 880–884 (1969). URL https://link.aps.org/doi/10.1103/PhysRevLett.23.880.

[45] Tavakoli, A., Pozas-Kerstjens, A., Luo, M.-X. & Renou, M.-O. Bell nonlocality in networks. *Reports on Progress in Physics* **85**, 056001 (2022). URL https://doi.org/10.1088/1361-6633/ac41bb.

[46] Acín, A. *et al.* Unified Framework for Correlations in Terms of Local Quantum Observables. *Physical Review Letters* **104**, 140404 (2010). URL https://link.aps.org/doi/10.1103/PhysRevLett.104.140404.

[47] Svetlichny, G. Distinguishing three-body from two-body nonseparability by a Bell-type inequality. *Physical Review D* **35**, 3066–3069 (1987). URL https://link.aps.org/doi/10.1103/PhysRevD.35.3066.

[48] Collins, D., Gisin, N., Popescu, S., Roberts, D. & Scarani, V. Bell-Type Inequalities to Detect True n-Body Nonseparability. *Physical Review Letters* **88**, 170405 (2002). URL https://link.aps.org/doi/10.1103/PhysRevLett.88.170405.

[49] Bancal, J.-D., Brunner, N., Gisin, N. & Liang, Y.-C. Detecting Genuine Multipartite Quantum Nonlocality: A Simple Approach and Generalization to Arbitrary Dimensions. *Physical Review Letters* **106**, 020405 (2011). URL https://link.aps.org/doi/10.1103/PhysRevLett.106.020405.

[50] Hardy, L. Quantum Theory From Five Reasonable Axioms (2001). URL http://arxiv.org/abs/quant-ph/0101012. ArXiv:quant-ph/0101012.

[51] Chiribella, G. & Spekkens, R. W. (eds.) *Quantum Theory: Informational Foundations and Foils*, vol. 181 of *Fundamental Theories of Physics* (Springer Netherlands, Dordrecht, 2016). URL https://link.springer.com/10.1007/978-94-017-7303-4.

[52] Janotta, P. & Hinrichsen, H. Generalized probability theories: what determines the structure of quantum theory? *Journal of Physics A: Mathematical and Theoretical* **47**, 323001 (2014). URL https://dx.doi.org/10.1088/1751-8113/47/32/323001.

[53] Plávala, M. General probabilistic theories: An introduction. *Physics Reports* **1033**, 1–64 (2023). URL http://arxiv.org/abs/2103.07469. ArXiv:2103.07469 [quant-ph].

[54] Barrett, J. Information processing in generalized probabilistic theories. *Physical Review A* **75**, 032304 (2007). URL https://link.aps.org/doi/10.1103/PhysRevA.75.032304.

[55] Chang, H. Operationalism. In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2021), fall 2021 edn. URL https://plato.stanford.edu/archives/fall2021/entries/operationalism.

[56] Adlam, E. The Operational Choi–Jamiołkowski Isomorphism. *Entropy* **22**, 1063 (2020). URL https://www.mdpi.com/1099-4300/22/9/1063.

[57] Holevo, A. *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Pisa, 2011). URL http://link.springer.com/10.1007/978-88-7642-378-9.

[58] Lami, L. *Non-classical correlations in quantum mechanics and beyond*. Ph.D. thesis, Universitat Autònoma de Barcelona, Spain (2017). URL http://arxiv.org/abs/1803.02902. ArXiv:1803.02902 [quant-ph].

[59] Simon, B. *Convexity: An Analytic Viewpoint* (Cambridge University Press, 2011), 1 edn. URL https://www.cambridge.org/core/product/identifier/9780511910135/type/book.

[60] Filippov, S. N., Gudder, S., Heinosaari, T. & Leppäjärvi, L. Operational Restrictions in General Probabilistic Theories. *Foundations of Physics* **50**, 850–876 (2020). URL https://doi.org/10.1007/s10701-020-00352-6.

[61] Schmid, D., Selby, J. H., Rossi, V. P., Baldijão, R. D. & Sainz, A. B. Shadows and subsystems of generalized probabilistic theories: when tomographic incompleteness is not a loophole for contextuality proofs. *Quantum* **9**, 1880 (2025). URL https://quantum-journal.org/papers/q-2025-10-13-1880/.

[62] Janotta, P. & Lal, R. Generalized probabilistic theories without the no-restriction hypothesis. *Physical Review A* **87**, 052131 (2013). URL https://link.aps.org/doi/10.1103/PhysRevA.87.052131.

[63] Müller, M. P. & Ududec, C. Structure of Reversible Computation Determines the Self-Duality of Quantum Theory. *Physical Review Letters* **108**, 130401 (2012). URL https://link.aps.org/doi/10.1103/PhysRevLett.108.130401.

[64] Plávala, M. All measurements in a probabilistic theory are compatible if and only if the state space is a simplex. *Physical Review A* **94**, 042108 (2016). URL https://link.aps.org/doi/10.1103/PhysRevA.94.042108.

[65] Baldijão, R. D., Erba, M., Schmid, D., Selby, J. H. & Sainz, A. B. Tomographically-nonlocal entanglement (2026). URL https://arxiv.org/abs/2602.16280. Version Number: 1.

[66] Centeno, D. *et al.* Symmetry-induced failures of tomographic locality: Constructing foil theories by twirling. *Physical Review A* **112**, L030202 (2025). URL https://link.aps.org/doi/10.1103/hpmv-15sf.

[67] Erba, M. & Perinotti, P. The composition rule for quantum systems is not the only possible one (2025). URL http://arxiv.org/abs/2411.15964. ArXiv:2411.15964 [quant-ph].

[68] Ramos Moreira, R. I. & Terra Cunha, M. Vértices do espaço de estados de 2 gbits (2026). URL https://doi.org/10.25824/redu/KSQKIU.

[69] Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Foundations of Physics* **24**, 379–385 (1994). URL https://doi.org/10.1007/BF02058098.

[70] Janotta, P., Gogolin, C., Barrett, J. & Brunner, N. Limits on nonlocal correlations from the structure of the local state space. *New Journal of Physics* **13**, 063024 (2011). URL https://dx.doi.org/10.1088/1367-2630/13/6/063024.

[71] Janotta, P. Generalizations of Boxworld. *Electronic Proceedings in Theoretical Computer Science* **95**, 183–192 (2012). URL http://arxiv.org/abs/1210.0618. ArXiv:1210.0618 [quant-ph].

[72] Chiribella, G., D'Ariano, G. M. & Perinotti, P. Probabilistic theories with purification. *Physical Review A* **81**, 062348 (2010). URL https://link.aps.org/doi/10.1103/PhysRevA.81.062348.

[73] Selby, J. H. *et al.* Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality. *Physical Review A* **107**, 062203 (2023). URL https://link.aps.org/doi/10.1103/PhysRevA.107.062203.

[74] Schmid, D., Selby, J. H. & Spekkens, R. W. Unscrambling the omelette of causation and inference: The framework of causal-inferential theories (2020). URL https://arxiv.org/abs/2009.03297. Version Number: 3.

[75] Barnum, H., Barrett, J., Leifer, M. & Wilce, A. Cloning and Broadcasting in Generic Probabilistic Theories (2006). URL http://arxiv.org/abs/quant-ph/0611295. ArXiv:quant-ph/0611295.

[76] Barnum, H., Barrett, J., Leifer, M. & Wilce, A. Teleportation in General Probabilistic Theories (2008). URL http://arxiv.org/abs/0805.3553. ArXiv:0805.3553 [quant-ph].

[77] D'Ariano, G. M., Chiribella, G. & Perinotti, P. *Quantum Theory from First Principles: An Informational Approach* (Cambridge University Press, 2016), 1 edn. URL https://www.cambridge.org/core/product/identifier/9781107338340/type/book.

[78] Chiribella, G., D'Ariano, G. M. & Perinotti, P. Informational derivation of quantum theory. *Physical Review A* **84**, 012311 (2011). URL https://link.aps.org/doi/10.1103/PhysRevA.84.012311.

[79] Katz, J. & Lindell, Y. *Introduction to Modern Cryptography* (Chapman and Hall/CRC, 2020), 3 edn. URL https://www.taylorfrancis.com/books/9781351133029.

[80] Wolf, R. *Quantum Key Distribution: An Introduction with Exercises*, vol. 988 of *Lecture Notes in Physics* (Springer International Publishing, Cham, 2021). URL https://link.springer.com/10.1007/978-3-030-73991-1.

[81] Aubrun, G., Lami, L., Palazuelos, C. & Plávala, M. Entanglement and Superposition Are Equivalent Concepts in Any Physical Theory. *Physical Review Letters* **128**, 160402 (2022). URL https://link.aps.org/doi/10.1103/PhysRevLett.128.160402.

[82] D'Ariano, G. M., Erba, M. & Perinotti, P. Classical theories with entanglement. *Physical Review A* **101**, 042118 (2020). URL https://link.aps.org/doi/10.1103/PhysRevA.101.042118.

[83] Wootters, W. K. Entanglement Sharing in Real-Vector-Space Quantum Theory. *Foundations of Physics* **42**, 19–28 (2012). URL https://doi.org/10.1007/s10701-010-9488-1.

[84] Wootters, W. K. The rebit three-tangle and its relation to two-qubit entanglement. *Journal of Physics A: Mathematical and Theoretical* **47**, 424037 (2014). URL https://iopscience.iop.org/article/10.1088/1751-8113/47/42/424037.

[85] Caves, C. M., Fuchs, C. A. & Rungta, P. Entanglement of Formation of an Arbitrary State of Two Rebits. *Foundations of Physics Letters* **14**, 199–212 (2001). URL https://doi.org/10.1023/A:1012215309321.

[86] Barnum, H., Beigi, S., Boixo, S., Elliott, M. B. & Wehner, S. Local Quantum Measurement and No-Signaling Imply Quantum Correlations. *Physical Review Letters* **104**, 140401 (2010). URL https://link.aps.org/doi/10.1103/PhysRevLett.104.140401.

[87] Almeida, M. L. *et al.* Guess Your Neighbor's Input: A Multipartite Nonlocal Game with No Quantum Advantage. *Physical Review Letters* **104**, 230404 (2010). URL https://link.aps.org/doi/10.1103/PhysRevLett.104.230404.

[88] Masanes, L. & Müller, M. P. A derivation of quantum theory from physical requirements. *New Journal of Physics* **13**, 063001 (2011). URL https://iopscience.iop.org/article/10.1088/1367-2630/13/6/063001.

[89] Scarani, V., Iblisdir, S., Gisin, N. & Acín, A. Quantum cloning. *Reviews of Modern Physics* **77**, 1225–1256 (2005). URL https://link.aps.org/doi/10.1103/RevModPhys.77.1225.

[90] Bae, J. & Acín, A. Asymptotic Quantum Cloning Is State Estimation. *Physical Review Letters* **97**, 030402 (2006). URL https://link.aps.org/doi/10.1103/PhysRevLett.97.030402.

[91] Aubrun, G., Müller-Hermes, A. & Plávala, M. Monogamy of entanglement between cones. *Mathematische Annalen* **391**, 1591–1609 (2025). URL https://link.springer.com/10.1007/s00208-024-02935-4.

[92] Lee, C. M. & Barrett, J. Computation in generalised probabilisitic theories. *New Journal of Physics* **17**, 083001 (2015). URL https://doi.org/10.1088/1367-2630/17/8/083001.

[93] Lee, C. M. & Hoban, M. J. Bounds on the power of proofs and advice in general physical theories. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **472**, 20160076 (2016). URL http://royalsocietypublishing.org/rspa/article/57211.

[94] Lee, C. M. & Selby, J. H. Generalised phase kick-back: the structure of computational algorithms from physical principles. *New Journal of Physics* **18**, 033023 (2016). URL https://doi.org/10.1088/1367-2630/18/3/033023.

[95] Barrett, J., de Beaudrap, N., Hoban, M. J. & Lee, C. M. The computational landscape of general physical theories. *npj Quantum Information* **5**, 41 (2019). URL https://www.nature.com/articles/s41534-019-0156-9.

[96] Gilligan-Lee, C. M. Computation in a general physical setting. *Journal of Physics A: Mathematical and Theoretical* **54**, 394001 (2021). URL https://doi.org/10.1088/1751-8121/ac2007.

[97] Sipser, M. *Introduction to the Theory of Computation* (Cengage Learning, USA, 2013), 3 edn.

[98] Aaronson, S. *et al.* Quantum Pseudoentanglement (2023). URL http://arxiv.org/abs/2211.00747. ArXiv:2211.00747 [quant-ph].

[99] Arnon, R., Brakerski, Z. & Vidick, T. Computational Entanglement Theory (2023). URL http://arxiv.org/abs/2310.02783. ArXiv:2310.02783 [quant-ph].

[100] Spekkens, R. W. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A* **71**, 052108 (2005). URL https://link.aps.org/doi/10.1103/PhysRevA.71.052108.

[101] Schmid, D., Selby, J. H., Wolfe, E., Kunjwal, R. & Spekkens, R. W. Characterization of Noncontextuality in the Framework of Generalized Probabilistic Theories. *PRX Quantum* **2**, 010331 (2021). URL https://link.aps.org/doi/10.1103/PRXQuantum.2.010331.

[102] Vilasini, V., Nurgalieva, N. & Rio, L. d. Multi-agent paradoxes beyond quantum theory. *New Journal of Physics* **21**, 113028 (2019). URL https://dx.doi.org/10.1088/1367-2630/ab4fc4.

[103] Schmid, D., Yīng, Y. & Leifer, M. A review and analysis of six extended Wigner's friend arguments (2024). URL http://arxiv.org/abs/2308.16220. ArXiv:2308.16220 [quant-ph].