Bit Commitment in Operational Probabilistic Theories

Tesi per la Laurea Magistrale di
Lorenzo Giannelli

Relatore
Chiar.mo Prof. Giacomo Mauro D’Ariano
Correlatore
Dott. Alessandro Tosini

Anno Accademico 2019-2020
A Piero
Abstract

The aim of this thesis is to investigate the bit commitment protocol in the framework of operational probabilistic theories. In particular a careful study is carried on the feasibility of bit commitment in the non-local boxes theory and in order to do this new aspects of the theory are presented.

Lo scopo di questa tesi è di investigare il protocollo di bit commitment all’interno delle teorie probabilistiche operazionali. In particolare si è analizzato attentamente la fattibilità del protocollo all’interno della teoria dei non-local boxes e i nuovi aspetti della teoria emersi in questa analisi sono presentati.
# Contents

## Introduction

## 1 Operational Probabilistic Theory

1.1 The Framework .......................... 6
   1.1.1 Primitive notions and notation .......... 6
   1.1.2 States and effects ..................... 11
   1.1.3 Transformations ......................... 12
   1.1.4 Coarse-graining and refinement .......... 13
1.2 Quantum Theory as an OPT .................. 14
   1.2.1 Atomicity of composition ............... 15
   1.2.2 Perfect discriminability ............... 15
   1.2.3 Ideal compression ...................... 16
   1.2.4 Causality ................................ 17
   1.2.5 Local discriminability ................. 21
   1.2.6 Purification .......................... 22

## 2 Bit Commitment

2.1 From Coin Tossing to no-go Theorem ........ 26
2.2 A Formal Definition ....................... 28
   2.2.1 The protocol .......................... 28
   2.2.2 OP bit commitment ...................... 29
2.3 Lightening No Bit Commitment .............. 31
   2.3.1 Reversible dilation of channels .......... 33
   2.3.2 Casually ordered channels and channels with memory .... 37
   2.3.3 No bit commitment ...................... 39

## 3 PR-Boxes

3.1 Why PR-Boxes? .......................... 41
3.2 PR-Boxes as OPT ......................................................... 43
  3.2.1 States, effects and transformations ................................. 43
  3.2.2 Discriminability between PR-boxes ................................. 47
  3.2.3 Purification in the PR-box (bipartite restriction) model .... 48
  3.2.4 N-partite PR-boxes .................................................. 51

4 No Bit Commitment in PR-Boxes ........................................ 53
  4.1 No-Perfectly Secure Bit Commitment ................................. 54
    4.1.1 First Proof ...................................................... 55
    4.1.2 Alternative Proof ............................................... 57
  4.2 Unconditionally Secure Bit Commitment ............................. 58
  4.3 Bit Commitment in Tripartite Scenario ............................. 60

Conclusions ...................................................................... 64

Appendix A Tables ................................................................ 67

Bibliography ...................................................................... 70
Introduction

The study of quantum foundations is a discipline of science that seeks to understand the most characterizing aspects of quantum theory, to reformulate it and even propose new generalizations. An active area of research in quantum foundations is therefore to find alternative formulations of quantum theory which rely on physically compelling principles in attempt to find a re-derivation of the quantum formalism in terms of operational axioms. One of the most interesting effort in this direction is made investigating the relations between the current operational axioms and the main results of quantum information theory.

As a recent example it has been proved in Ref. [1] that the no information without disturbance (NIWD) theorem, i.e. the impossibility in quantum theory to extract information without disturbing the state of the system or its correlations with other systems, is independent of both local discriminability and purification, two of the defining axioms of quantum theory. Especially the latter, as we will see thoroughly in the first Chapter, is considered as a characteristic and distinctive quantum trait but now NIWD can be exhibited in absence of it and also of most of the principles of quantum theory.

The NIWD property spawns other no-go theorems, which represent some of the most famous and classic results in quantum information theory. Among these results one can certainly list the no-cloning theorem, the no-programming and the impossibility of perfectly secure bit commitment. In our thesis our efforts are focused on the latter.
A bit commitment (BC) protocol is meant to allow one party, Alice, to send a bit to a second party, Bob, in such a way that Bob cannot read the bit until Alice allows for its disclosure, while Alice cannot change the value of the bit after she encoded it. The bit commitment protocol is a very important primitive in cryptography, and perfectly secure protocols are known to be impossible in classical information theory. This is the case in quantum theory as well. The proof involves a very important characterization theorem for general theories with purification [2].

Numerous bit commitment protocols have been proposed in literature and this cryptography primitive has been deeply studied, especially the possibility of unconditionally secure bit commitment, both in quantum and classical information theory, due to its importance in practical applications. We would like to investigate the relations between the theorem of no-bit commitment and the operational axioms that characterize quantum theory.

A general strategy to apprehend the nature of these links is to test the validity of the theorem in a theory that lacks one or more principles; naturally the first try has to be made excluding the purification, the most quantum feature.

The answer of how to start in the analysis comes directly from the literature on bit commitment protocol itself. In fact, after it was proved to be impossible in quantum theory the protocol has been tested in more general scenarios, in particular in more non-local scenarios. To understand what it means we need to take a step back to the two parties of the protocol, Alice e Bob. Assume that they are not able to communicate but have access to physical states that they can use to generate joint correlations. In this experiment the outcome of the measurements on the state of their local systems are given by random variables. Obviously, causality constrains the correlations to be non-signalling, and on the other side quantum theory prevents the strength of the non-local correlation to violate Bell’s inequalities [3], where the maximal value is known as Cirel’son’s bound [4]. A well-known variant of a Bell inequality is the Clauser, Horne, Shimony & Holt (CHSH) inequality [5], which can be expressed as [6]

\[ \sum_{x,y \in \{0,1\}} \Pr(a_x \oplus b_y = x \cdot y) \leq 2\sqrt{2}. \]

Where \(x\) and \(y\) denote the choice of Alice’s and Bob’s measurement, respectively, \(a_x \in \{0,1\}, b_y \in \{0,1\}\) the respective binary outcomes, and \(\oplus\) addition modulo 2. However, if we only care about the causality constrain, Cirel’son’s bound can be violated up to the maximal value of 4. Popescu and Rohrlich, who first noticed [7], raised
the question of why nature is not more non-local and why does quantum mechanics not allow a stronger violation of the CHSH inequality. Following this lead, bit commitment has been studied in Popescu Rohrlich (PR) non-local boxes theory.

It has been claimed, as in Ref. [8], that bit commitment was admissible in PR-box theory but a counter-proof reached from Short, Gisin, and Popescu [9]. It was formulated on the behalf that, since non-locality is the main reason to prevent bit commitment in quantum theory, it would not be possible that in a theory more non-local than the quantum one BC would work. However, thereafter a new protocol was proposed by Buhrman et al. [10] and the argument of the counter-proof of Short, Gisin and Popescu is not able to deny it.

The first aim of our work is to bring some clarity, pointing out the limitations of the framework of validity about the results that have been claimed until now. In our path new aspects of PR-box theory emerged and have been studied. Even if the theory is still far to be considered complete, important characterizations have been made. We will show that all the BC protocols proposed so far respect the limitations of PR-box theory under which we are able to prove a no-bit commitment theorem. Furthermore we are able to create, with very similar arguments, a counter-proof of the protocol proposed in Ref. [10].

Finally, a surprising result is observed. Relaxing the constraint of PR-box theory and including some recent developments, a scheme of bit commitment that is perfectly secure seems possible. Even if there are actually no "operational" evidences within the theory to deny it, by the same fundamental reason expressed by Short, Gisin, and Popescu, we think that future advancements of the theory would reconsider the protocol as cheatable. However, as a matter of fact, now the answer to the existence of a theory with entanglement that admits bit commitment, seems to be “Yes!”.

A synopsis of the thesis is the following:

In the first Chapter the framework of operational probabilistic theories (OPTs) is presented by first introducing the operational language that expresses the possible connections between events, and then by dressing the elements of the language with a probabilistic structure. After that, the principles for the OPT of quantum mechanics are stated (given the framework, the rule of connectivity among events are given).

In the second Chapter, we will discuss the bit commitment protocol. We will start
with an historical perspective and then we will rigorously define the protocol in the language of OPT. We will mainly deal only with perfectly secure bit commitment since our analysis is carried in the OP framework and there are not yet the technical tools to analyse unconditionally secure BC in OPTs. In the third section of the Chapter, we will study the proof of the impossibility of perfectly secure bit commitment in quantum theory done in Ref. [2]. This is a very elegant and solid demonstration and our purpose is to adapt the proof in order to comprehend other theories than the quantum one. As we will see, non-locality and entanglement are the key reasons that plays in favor of the impossibility of BC and so it seems reasonable to try to extend the impossibility proof also to other non-local theories.

In the third Chapter, we will analyse the probabilistic theory corresponding to the popular PR-box model. A comprehensive study has never been made and an organic theory is not available. We propose to bring clarity to the actual model considering only bipartite correlated boxes, highlighting its limitations, and some progress by analysing new aspects such as perfect discriminability and states purification. We will also add some considerations and prevision on \( N \)-partite correlated boxes.

Finally, in the fourth Chapter we propose a proof of impossibility of perfectly secure bit commitment in PR-boxes. Even if the PR-box model does not constitute a proper theory it has been often used in numerous application in literature, such as in protocol of perfectly or unconditionally secure BC. However this led to neglect important elements of the theory and in return numerous results published can be proved false when contextualized in the PR-box OPT (also if it is still incomplete).

As the last remark, we conjecture that including tripartite correlated boxes in the theory, secure bit commitment would be possible. The reason is the following. We will see that what prevents BC in quantum theory is that every state can be purified and its purification is unique up to local reversible transformations. Exactly these local operations allow the cheating in the protocol. When we deal with PR-box theory limited to bipartite correlated boxes, only one internal state can be purified, but its purification is unique and this is enough to allow cheating. On the contrary, when we admit tripartite correlated boxes, the uniqueness of purification is lost and exactly this uncommon feature could open the door to the possibility of secure bit commitment.
The purpose of this Chapter is to introduce the framework of operational probabilistic theories (OPTs) and to express quantum theory as an OPT. In this Chapter we will follow Ref. [11].

The framework of operational probabilistic theories consists of two distinct conceptual ingredients: an operational structure, describing circuits that produce outcomes, and a probabilistic structure, which assigns probabilities to the outcomes in a consistent way. The operational structure summarizes all the possible circuits that can be constructed in a given physical theory, in this setting a rigorous formulation of the elements of the circuits: systems, transformations, and their composition is given, which constitutes the grammar for the probabilistic description of an experiment. However, it is only the probabilistic structure that promotes the operational language from a merely descriptive tool to a framework for predictions, the predictive power being the crucial requirement for any scientific theory and for its testability - the essence of science itself. Different OPTs will have different rules for assigning the joint probabilities of events. Working in this framework allows us to deal with a wide range of probabilistic theories, including not only quantum and classical theories, but also the theory of Popescu-Rohrlich (PR), or non-local, boxes.

In the first part of the Chapter the framework is provided by first introducing the operational language that expresses the possible connections between events, and then by dressing the elements of the language with a probabilistic structure. Then, in the second part of the Chapter we formulate the principles for the OPT
of quantum theory. In fact, once the framework is defined, the rule of connectivity among events are given.

1.1 The Framework

A theory for making predictions about joint events depending on their reciprocal connections is what we call an operational probabilistic theory. We see that OPT is a non-trivial extension of probability theory.

To the joint events we associate not only their joint probability, but also a circuit that connects them. When the events in the circuit have a well-defined order, the circuit is mathematically described by a directed acyclic graph (a graph with directed edges and without loops).

The basic element of an OPT - the notion of event - gets dressed with wires that allow us to connect it with other events. Such wires are the systems of the theory. In agreement with the directed nature of the graph, there are input and output systems. The events are the transformations, whereas the transformations with no input system are the states (corresponding to preparations of systems), and those with no output system are the effects (corresponding to observations of systems). Since the purpose of a single event is to describe a process connecting an input with an output, the full circuit associated to a probability is a closed one, namely a circuit with no input and no output.

The circuit framework is mathematically formalized in the language of category theory. In this language, an OPT is a category, whose systems and events are objects and arrows, respectively. Every arrow has an input and an output object, and arrows can be sequentially composed. The associativity, existence of a trivial system, and commutativity of the parallel composition of systems of quantum theory technically correspond to having a strict symmetric monoidal category.

1.1.1 Primitive notions and notation

The primitive notions of any operational theory are those of test, event, and system. A test \( \{\mathcal{A}_i\}_{i \in I} \) is the collection of events \( \mathcal{A}_i \), where \( i \) labels the element of the outcome space \( I \). In addition to comprising a collection of events, the notion of test carries also the event connectivity of the theory that is achieved by the systems. These can represent the input and the output of the test. The resulting representation of a test is the following diagram:

\[ A \xrightarrow{\{\mathcal{E}_x\}_{x \in X}} B. \]
1.1. The Framework

The wire on the left labeled as A represents the input system, whereas the wire on the right labeled as B represents the output system. The same diagrammatic representation is also used for any of the events, namely for \( x \in X \):

$$\begin{array}{ccc} \text{A} & \xrightarrow{E} & \text{B} \end{array}$$

In the following, the systems will be denoted by capital Roman letters A,B,...,Z, whereas the events by capital calligraphic letters \( \mathcal{A}, \mathcal{B}, \ldots, \mathcal{Z} \).

Different tests can be combined in a *circuit*, which is a directed acyclic graph where the links are the systems (oriented from left to right, namely from input to output) and the nodes are the boxes of the tests. The same graph can be built up for a single test instance, namely with the network nodes being events instead of tests, corresponding to a joint outcome for all tests.

The circuit graph is obtained precisely by using the following rules.

**Sequential Composition of Test** When the output system of test \( \{\mathcal{C}_x\}_{x \in X} \) and the input system of test \( \{\mathcal{D}_y\}_{y \in Y} \) coincide, the two tests can be composed in sequence as follows:

$$\begin{array}{ccc} \text{A} & \xrightarrow{\{\mathcal{C}_x\}_{x \in X}} & \text{B} \xrightarrow{\{\mathcal{D}_y\}_{y \in Y}} \text{C} \end{array} = \begin{array}{ccc} \text{A} & \xrightarrow{\mathcal{E}_{(x,y)}} & \text{C} \end{array},$$

resulting in the test \( \{\mathcal{E}_{(x,y)}\}_{(x,y) \in X \times Y} \) called *sequential composition* of \( \{\mathcal{C}_x\}_{x \in X} \) and \( \{\mathcal{D}_y\}_{y \in Y} \). In formulas we will also write \( \mathcal{E}_{(x,y)} := \mathcal{D}_y \mathcal{C}_x \).

**Identity Test** For every system A, one can perform the *identity test* (shortly identity) that "leaves the system alone". Formally, this is the deterministic test \( \mathcal{I}_A \) with the property

$$\begin{array}{ccc} \text{A} & \xrightarrow{\mathcal{I}_A} & \text{A} \xrightarrow{\mathcal{C}} \text{B} = \begin{array}{ccc} \text{A} & \xrightarrow{\mathcal{C}} & \text{B} \end{array}, \end{array}$$

$$\begin{array}{ccc} \text{B} & \xrightarrow{\mathcal{D}} & \text{A} \xrightarrow{\mathcal{I}_A} \text{A} = \begin{array}{ccc} \text{B} & \xrightarrow{\mathcal{D}} & \text{A} \end{array}, \end{array}$$

where the above identities must hold for any event \( \text{A} \xrightarrow{\mathcal{C}} \text{B} \) and \( \text{B} \xrightarrow{\mathcal{D}} \text{A} \), respectively. The sub-index A will be dropped from \( \mathcal{I}_A \) where there is no ambiguity.

**Operationally Equivalent Systems** We say that two systems A and A’ are *operationally equivalent* - denoted as \( A' \simeq A \) or just \( A' = A \) - if there exist two deterministic events \( \text{A} \xrightarrow{U} \text{A'} \) and \( \text{A'} \xrightarrow{V} \text{A} \) such that

$$\begin{array}{ccc} \text{A} & \xrightarrow{U} & \text{A'} \xrightarrow{V} \text{A} = \begin{array}{ccc} \text{A} & \xrightarrow{\mathcal{I}} & \text{A} \end{array}, \end{array}$$

$$\begin{array}{ccc} \text{A'} & \xrightarrow{V} & \text{A} \xrightarrow{U} \text{A'} = \begin{array}{ccc} \text{A'} & \xrightarrow{\mathcal{I}} & \text{A'} \end{array}. \end{array}$$
Accordingly, if \( \{C_x\}_{x \in X} \) is any test for system A, performing an equivalent test on system \( A' \) means performing the test \( \{C'_x\}_{x \in X} \) defined as

\[
A' C'_x A' = A' \bigtriangledown A C_x A \bigtriangledown A'.
\]

**Composite System** Given two systems A and B, one can join them into the single composite system AB. As a rule, the system AB is operationally equivalent to the system BA, and we will identify them in the following. This means that the system composition is commutative,

\[
AB = BA.
\]

We will call a system *trivial system*, reserving for it the letter I, if it corresponds to the identity in the system composition, namely

\[
AI = IA = A.
\]

The trivial system corresponds to having no system, namely I carries no information.

Finally we require the composition of systems to be associative, namely

\[
A(BC) = (AB)C.
\]

In other words, if we iterate composition on many systems we always end up with a composite system that only depends on the components, and not on the particular composition sequence according to which they have been composed. Systems then make an Abelian monoid. A test with input system AB and output system CD represents an interaction process (see the parallel composition of tests in the following).

**Parallel Composition of Tests** Any two tests \( A \{C_x\}_{x \in X} B \) \( C \{D_y\}_{y \in Y} D \) can be composed in parallel as follows:

\[
\begin{array}{c}
A \\
\bigtriangledown \\
C \\
\{C_x\}_{x \in X} \\
\{D_y\}_{y \in Y} \\
\bigtriangledown \\
D
\end{array} = \begin{array}{c}
AC \\
\bigtriangledown \\
BD \\
\{F_{(x,y)}\}_{(x,y) \in X \times Y}
\end{array}.
\]

The test \( AC \{F_{(x,y)}\}_{(x,y) \in X \times Y} BD \) is the *parallel composition* of tests \( A \{C_x\}_{x \in X} B \) and \( C \{D_y\}_{y \in Y} D \), where \( \{F_{(x,y)}\}_{(x,y) \in X \times Y} \equiv \{C_x \otimes D_y\}_{(x,y) \in X \times Y} \). Parallel and
1.1. The Framework

sequential composition of tests commute, namely one has

\[
\begin{align*}
A \begin{array}{c}
\mathcal{C}_z \\
\mathcal{B}_y \\
\mathcal{E}_x
\end{array}
B \\
\mathcal{A}_x \\
\mathcal{D}_w \\
\mathcal{F}_w
\end{array} C &= \begin{array}{c}
\mathcal{C}_z \\
\mathcal{B}_y \\
\mathcal{E}_x
\end{array}
A \\
\mathcal{A}_x \\
\mathcal{D}_w \\
\mathcal{F}_w
\end{array} C.
\end{align*}
\]

(1.1)

When one of the two operation is the identity, we will omit the identity box and drawn only a straight line:

\[
\begin{array}{c}
A \mathcal{C}_x \\
\mathcal{D}_y
\end{array} B.
\]

Therefore, as a consequence of commutation between sequential and parallel composition, we have the following identity:

\[
\begin{array}{c}
A \mathcal{C}_x \\
\mathcal{D}_y
\end{array} B = \begin{array}{c}
A \mathcal{C}_x \\
\mathcal{D}_y
\end{array} B.
\]

Preparation Tests and Observation Tests Tests with a trivial input system are called preparation tests, and tests with a trivial output system are called observation tests. They will be represented as follows:

\[
\begin{align*}
A \{\rho_x\}_{x \in X} B &:= I \{\rho_x\}_{x \in X} B, \\
A \{a_y\}_{y \in Y} B &= A \{a_y\}_{y \in Y} I.
\end{align*}
\]

The corresponding events will be called preparation events and observation events. In formulas we will also write \(|\rho_i|_A\) to denote a preparation event and \((a_j|_A\) to denote an observation event.

Closed Circuits Using the above rules we can build up closed circuits, i.e. circuits with no input and no output system. An example is given by the following circuit:

\[
\begin{array}{c}
A \{\Psi_i\} \\
B \{A_j\} \\
C \{C_i\} \\
D \{G_q\}
\end{array}
\begin{array}{c}
B \{\mathcal{E}_m\} \\
G \{D_m\} \\
M \{\mathcal{F}_p\}
\end{array}
\begin{array}{c}
L \{B_k\} \\
H \{\mathcal{C}_l\} \\
E \{F_q\}
\end{array}
\begin{array}{c}
E \{\mathcal{C}_z\} \\
\mathcal{B}_y \\
\mathcal{E}_x
\end{array} C
\]

(1.2)

where we omitted the probability spaces of each test.
**Independent Systems** For any (generally open) circuit constructed according to the above rules we call a set of systems *independent* if for each couple of systems in the set the two are not connected by a unidirected path (i.e. following the arrow from the input to the output). For example, in Eq. (1.2) the sets \{A,E\}, \{H,O\}, \{A,E,H,O\}, \{A,L\}, \{A,E,L,P\} are independent, whereas e.g. the sets \{A,M\}, \{A,B\}, \{A,E,N\} are not. A maximal set of independent systems is called a *slice*.

We are now in position to move towards the general purpose of an operational probabilistic theory: predicting and accounting for the joint probability of events corresponding to a particular circuit of connections.

Given a closed circuit, as in Eq. (1.2), we are left with just a joint probability distribution. Therefore, to a closed circuit of event as the following:

\[
\Psi_i A_j B C D E F G H I J K L M N O P
\]

we will associate a joint probability \(p(i, j, k, l, m, n, p, q)\) which we will consider as *parametrically dependent* on the circuit, namely, for a different choice of events and/or different connections we will have a different joint probability.

Since we are interested only in the joint probabilities and their corresponding circuits, we will build up probabilistic equivalence classes, and define:

*Two events from system A to system B are equivalent if they occur with the same joint probability with the other events within any circuit.*

We will call *transformation from A to B* - denoted as \(\mathcal{A} \in \text{Transf}(A \to B)\) - the equivalence class of events from A to B that are equivalent in the above sense. Likewise we will call *instrument* an equivalence class of tests, *state* an equivalence class of preparation events, and *effect* an equivalence class of observation events. We will denote the set of states of system A as \(\mathcal{St}(A)\), and the set of its effects as \(\mathcal{Eff}(A)\). Clearly, the input systems belonging to two different elements of an equivalence class will be operationally equivalent, and likewise for output systems.

We now can define an operational probabilistic theory as follows:

*An operational probabilistic theory (OPT) is a collection of systems and transformations, along with rules for composition of systems and parallel and*
sequential composition of transformations. The OPT assigns a joint probability to each closed circuit.

Therefore, in an OPT every test from the trivial system I to itself is a probability distribution \( \{ p_i \}_{i \in X} \) for the set of joint outcomes \( X \), with \( p(i) := p_i \in [0, 1] \) and \( \sum_{i \in X} p(i) = 1 \). Compound events from the trivial system to itself are independent, namely their joint probability is given by the product of the respective probabilities for both the parallel and the sequential composition, namely

\[
\rho_{i_1} A_{i_2} = \rho_{i_1} A_{i_2} \sigma_{j_1} B_{j_2} = p(i_1, i_2)q(j_1, j_2).
\]

A special case of OPT is the deterministic OPT, where all probabilities are 0 or 1.

### 1.1.2 States and effects

Using the parallel and sequential composition of transformation it follows that any closed circuit can be regarded as the composition of a preparation event and an observation event, for example the circuit in Eq. (1.3) can be cut along a slice as follows:

\[
(\Psi_i, A_j, B_k, D_m, F_p) \xrightarrow{BGN} (C_l, E_n, G_q)
\]

and thus is equivalent to the following state-effect circuit:

\[
(\Psi_i, A_j, B_k, D_m, F_p) \xrightarrow{BGN} (C_l, E_n, G_q)
\]

Therefore, a state \( \rho \in St(A) \) is a functional over effects \( Eff(A) \), the functional being denoted with the pairing \( (a|\rho) \) with \( a \in Eff(A) \) and analogously an effect \( a \in Eff(A) \) is a functional over states \( St(A) \).

By taking linear combinations of functionals we see that \( St_\mathbb{R}(A) := \text{Span}_\mathbb{R} [St(A)] \) and \( Eff_\mathbb{R}(A) := \text{Span}_\mathbb{R} [Eff(A)] \) are dual spaces, and states are positive linear functionals over effects, and effects are positive linear functional over states \( (St_\mathbb{R}(A) \text{ and } Eff_\mathbb{R}(A) \text{ assume finite dimensional and we denote as } D_A := \dim St_\mathbb{R}(A) \equiv \dim Eff_\mathbb{R}(A) \text{ also called size of system } A). \) In the following we also denote by \( St_1(A) \) and \( Eff_1(A) \) the sets of deterministic states and effects, respectively.
1.1. The Framework

According to the above definition, two states are different if and only if there exists an effect which occurs on them with different joint probabilities. We also have that two effects are different if and only if there exists a state on which they have different probabilities.

In particular, given two states $\rho_0 \neq \rho_1 \in \text{St}(A)$ we will say that an effect $a \in \text{Eff}(A)$ separates the state $\rho_0$ and $\rho_1$ when $(\rho_1|a) \neq (\rho_0|a)$, namely when the effect occurs with different joint probabilities over the two states (the analogous relation holds for separable states respect to effects). Therefore we conclude that:

*States are separating for effects and effects are separating for states.*

It is possible to demonstrate that in any convex OPT if two states (effects) $\rho_0, \rho_1 \in \text{St}(A)$ ($a_0, a_1 \in \text{Eff}(A)$) are distinct, then one can discriminate them with error probability strictly smaller than $\frac{1}{2}$.

1.1.3 Transformations

From what we said before, the following circuit is a state of system BFHO:

```
  ψ
  
  A B
  E F
  H O
```

This means that any transformation connected to some output systems of a state maps the state into another state of generally different systems. Thus, while states and effects are linear functionals over each other, we can always regard a transformation as a map between states. In particular, a transformation $\mathcal{T} \in \text{Transf}(A \rightarrow B)$ is always associated to a map $\hat{\mathcal{T}}$ from $\text{St}(A)$ to $\text{St}(B)$, uniquely defined as

$$\hat{T}: |\rho\rangle \in \text{St}(A) \mapsto \hat{T}|\rho\rangle = |T\rho\rangle \in \text{St}(B).$$

Similarly the transformation can be associated to a map from $\text{Eff}(A)$ to $\text{Eff}(B)$. The map $\hat{T}$ can be linearly extended to a map from $\text{St}_R(A)$ to $\text{St}_R(B)$. Notice that the linear extension of $\mathcal{T}$ (which we will denote by the same symbol) is well defined. In fact, a linear combination of states of $A$ is null - in formula $\sum_i c_i |\rho_{oi}\rangle = 0$ - if and only if $\sum_i c_i (\rho_i) = 0$ for every $a \in \text{Eff}(A)$, and since for every $b \in \text{Eff}(B)$ we have $(b|\mathcal{T} \in \text{Eff}(A)$, then $(b|\mathcal{T}(\sum_i c_i |\rho_i\rangle) = \sum_i c_i (b|\mathcal{T}|\rho_i\rangle = (b|\sum_i c_i \mathcal{T}|\rho_i\rangle = 0$, and finally $\sum_i c_i \mathcal{T}|\rho_i\rangle = 0$.

We want to stress that if two transformations $\mathcal{T}, \mathcal{T}' \in \text{Transf}(A \rightarrow B)$ correspond to the same map $\hat{T}$ from $\text{St}(A)$ to $\text{St}(B)$, this does not mean that the two transformations

12
are the same, since as an equivalence class, they must occur with the same joint probability in all possible circuits. In terms of state mappings, the same definition of the transformation as equivalence class corresponds to say that $\mathcal{T}, \mathcal{T}' \in \text{Trans}f(A \rightarrow B)$ as maps from states of $AR$ to states of $BR$ are the same for all possible systems $R$ of the theory, namely $\mathcal{T} = \mathcal{T}' \in \text{Trans}f(A \rightarrow B)$ if and only if

$$\forall R, \forall \Psi \in \text{St}(AR) \quad \Psi \begin{array}{c} A \\ R \end{array} \mathcal{T} \begin{array}{c} B \\ R \end{array} = \Psi \begin{array}{c} A \\ R \end{array} \mathcal{T}' \begin{array}{c} B \\ R \end{array}. \quad (1.4)$$

Indeed, there exist cases of OPT where there are transformations $\mathcal{T} = \mathcal{T}' \in \text{Trans}f(A \rightarrow B)$ corresponding to the same map when applied to $\text{St}(A)$ and not when applied to $\text{St}(AR)$ for some system $R$, a relevant example is fermionic theory [12, 13]. Since we can take linear combinations of linear transformations, $\text{Trans}f(A \rightarrow B)$ can be embedded in the vector space $\text{Trans}f(A \rightarrow B)$. The deterministic transformations, whose set will be denoted as $\text{Trans}f_{1}(A \rightarrow B)$, will be also called channels. Finally, a transformation $\mathcal{U} \in \text{Trans}f(A \rightarrow B)$ is reversible if there exists another transformation $\mathcal{U}^{-1} \in \text{Trans}f(B \rightarrow A)$ such that $\mathcal{U}^{-1} \mathcal{U} = \mathcal{I}_A$ and $\mathcal{U} \mathcal{U}^{-1} = \mathcal{I}_B$. The set of reversible transformations from $A$ to $B$ will be denoted by $\text{Rev} \text{Trans}f(A \rightarrow B)$.

### 1.1.4 Coarse-graining and refinement

When dealing with probabilistic events, a natural notion is that of *coarse-graining*, corresponding to merging events into a single event. According to probability theory, the probability of a coarse-grained event $S \subseteq X$ subset of the outcome space $X$ is the sum of probabilities of the elements of $S$, namely $p(S) = \sum_{i \in S} p(i)$. We then correspondingly have that the coarse-grained event $\mathcal{T}_S$ of a test $\{\mathcal{T}_i\}_{i \in X}$ will be given by

$$\mathcal{T}_S = \sum_{i \in S} \mathcal{T}_i. \quad (1.5)$$

We stress that the equal sign in Eq. (1.5) is to be meant in the sense of Eq. (1.4). In addition to the notion of coarse-grained event we have also that of coarse-grained test, corresponding to the collection of a coarse-grained events $\{\mathcal{T}_{X_i}\}_{i \in Z}$ from a partition $X = \cup_{i \in Z} X_i$ of the outcome space $X$, with $X_i \cap X_j = \emptyset$ for $i \neq j$. The converse procedure of coarse-graining is what we call *refinement*. If $\mathcal{T}_S$ the coarse-graining in Eq. (1.5), we call any sum $\sum_{i \in S'} \mathcal{T}_i$ with $S' \subseteq S$ a refinement of $\mathcal{T}_S$. The same notion can be analogously considered for a test. Intuitively, a test that refines another is a test that extracts more detailed information, namely it is a test with better "resolving power".

The notion of refinement is translated to transformations (hence also to states, and
1.2 Quantum Theory as an OPT

effects), as equivalence classes of events. Refinement and coarse-graining define a
partial ordering in the set of transformations $\text{Transf}(A \rightarrow B)$, writing $\mathcal{D} \prec \mathcal{C}$ if $\mathcal{D}$ is a
refinement of $\mathcal{C}$. A transformation $\mathcal{C}$ is atomic if it has only trivial refinement, namely
$\mathcal{C}_i$ refines $\mathcal{C}$ implies that $\mathcal{C}_i = p\mathcal{C}$ for some probability $p \geq 0$. A test that consists of
atomic transformations is a test whose "resolving power" cannot be further improved.
It is often useful to refer to the set of all possible refinements of a given event $\mathcal{C}$. This
set is called refinement set of the event $\mathcal{C} \in \text{Transf}(A \rightarrow B)$, and is denoted by $\text{RefSet}(\mathcal{C})$. In formula,
$\text{RefSet}(\mathcal{C}) := \{ \mathcal{D} \in \text{Transf}(A \rightarrow B) \mid \mathcal{D} \prec \mathcal{C} \}$.
In the special case of states, we will use the word pure as a synonym of atomic.
A pure state describes an event providing maximal knowledge about the system’s
preparation, namely a knowledge that cannot be further refined (we will denote with $\text{PurSt}(A)$ the set of pure states of system $A$).
As usual, a state that is not pure will be called mixed. An important notion is that of
internal state. A state is called internal when any other state can refine it: precisely,
$\omega \in \text{St}(A)$ is internal if for every $\rho \in \text{St}(A)$ there is a non-zero probability $p > 0$
such that $pp$ is a refinement of $\omega$, i.e. $pp \in \text{RefSet}(\omega)$. The adjective "internal" has a
precise geometric connotation, since the state cannot belong to the border of $\text{St}(A)$.
An internal state describes a situation in which there is no definite knowledge about
the system preparation, namely a priori we cannot in principle exclude any possible
preparation.

1.2 Quantum Theory as an OPT

In this section we provide an overview of the six principles used for constructing quan-
tum theory as an OPT. All features of quantum theory - ranging from the superpo-
sition principle, entanglement, no cloning, teleportation, Bell’s inequalities violation,
quantum cryptography - can be understood and proved using only the principles,
without using Hilbert spaces. However, our aim is only to introduce the principle and
analyse them from an operative point of view.
All the six principles are operational, in that they stipulate whether or not certain
tasks can be accomplished: they set the rules of the game for all the experiments and
all the protocols that can be carried out in the theory. They also provide a great
insight into the worldview at which quantum theory hints.
We review the list of the principles:

1. Atomicity of composition
2. Perfect discriminability
3. Ideal compression
4. Causality
5. Local discriminability
6. Purification

All six principles, with the exception of purification, express standard features that are shared by both classical and quantum theory. The principle of purification picks up uniquely quantum theory among the theories allowed by the first five, partly explaining the magic of quantum information.

1.2.1 Atomicity of composition

In the general framework we encountered the notions of coarse-grained and atomic operation. A coarse-grained operation is obtained by joining together outcomes of a test, corresponding to neglect some information. The inverse process of coarse-graining is that of refining. An atomic operation is one where no information has been neglected, namely an operation that cannot be refined. When the operation is atomic, the experimenter has maximal knowledge of what’s happening in the lab. A test consisting of atomic operations represents the highest level of control achievable according to our theory.

The principle of atomicity of composition states that it possible to maintain such a level of control throughout a sequence of experiments, stating precisely what follows:

**Axiom 1** (Atomicity of composition). The sequence of two atomic operations is an atomic operation.

One of the immediate consequences granted by atomicity of composition is the following:

**Corollary 1** (Parallel composition of pure states). Given two pure states $\alpha \in \text{St}(A)$ and $\beta \in \text{St}(B)$, the parallel composition of $\alpha$ and $\beta$ is a pure state of $\text{St}(AB)$.

1.2.2 Perfect discriminability

Two deterministic states $\rho_0$ and $\rho_1$ are perfectly discriminable if there exists a measurement $\{m_y\}_{y \in \{0,1\}}$ such that

$$(m_y|\rho_x) = \delta_{xy} \; \forall x, y \in \{0,1\}.$$
The existence of perfectly discriminable states is important, because these states can be used to communicate classical information without errors. In a communication protocol, the sender can encode the value of a bit $x$ into the state $\rho_x$ and then transmit the system to the receiver, who can decode the value of the bit using the measurement $\{m_y\}_{y \in \{0,1\}}$.

The perfect discriminability axiom ensures that our ability to discriminate states is as sharp as it could possibly be: except for trivial cases, every state can be perfectly discriminated from some other state. The "trivial cases" are those states that cannot be discriminated from anything else because they contain every other state in their convex decomposition. We can call them internal, or completely mixed.

**Axiom 2 (Perfect discriminability).** *Every deterministic state that is not completely mixed is perfectly discriminable from some other state.*

As anticipated, the perfect discriminability axiom guarantees that every non-trivial system has at least two perfectly discriminable states:

**Proposition 1.** *In a theory satisfying perfect discriminability, every physical system has at least two perfectly discriminable states, unless the system is trivial (i.e. it has only one deterministic state).*

**Proof.** Pick a pure state $\alpha \in \text{St}(A)$. If $\alpha$ is not internal, then perfect discriminability guarantees that $\alpha$ is perfectly discriminable from some other state $\alpha'$, hence $A$ has two perfectly discriminable states. If $\alpha$ is internal every pure state belongs to its refinement set. Moreover, since it is also pure, i.e. extremal, one has that every other deterministic state $\rho_1 \in \text{St}_1(A)$ must be equal to $\alpha$, i.e. $A$ has only one deterministic state.

An easy consequence of this result is that the theory can describe noiseless classical communication.

### 1.2.3 Ideal compression

Ideal compression guarantees that information can be transferred faithfully from one system to another. Namely, suppose that Alice has a preparation device, which prepares system $A$ in some state $\alpha$. Alice does not know the state $\alpha$, but she knows that on average the device prepares the deterministic state $\rho \in \text{St}_1(A)$. Now, suppose Alice wants to transfer the state of her system to Bob's laboratory, but unfortunately
she cannot send system A directly. Instead, she has to encode the state \( \alpha \) into the state of another system B, by applying a suitable deterministic operation \( \mathcal{E} \) (the \textit{encoding}), which transforms the state \( \alpha \) into the state
\[
\beta := \mathcal{E}\alpha.
\]
We say that the encoding is \textit{lossless} for the state \( \rho \) iff there exists another deterministic operation \( \mathcal{D} \) (the \textit{decoding}) such that
\[
\mathcal{D}\mathcal{E}\alpha = \alpha \quad \forall \alpha \in F_\rho
\]
where \( F_\rho \) is the refinement set of \( \rho \), which is made of the set of all states \( \alpha \) that are compatible with \( \rho \) (on the convex set of states this would be the face to which \( \rho \) belongs).
This third axiom establishes the possibility of a particular type of lossless encoding, called \textit{ideal compression}. The ultimate limit to the lossless compression of a given state \( \rho \) is reached when every state of the encoding system B is a codeword for some state in \( F_\rho \), namely when every state \( \beta \in \text{St}(B) \) is of the form \( \mathcal{E}\alpha \) for some \( \alpha \in F_\rho \). When this is the case, we say that the compression is \textit{efficient}, and we call the triple \((B, \mathcal{E}, \mathcal{D})\) an \textit{ideal compression protocol}.

\textbf{Axiom 3} (Ideal Compression). \textit{Every state can be compressed in a lossless and efficient way.}

\subsection*{1.2.4 Causality}

The causality axiom identifies the input–output ordering of a circuit with the direction along which information flows, identifying such ordering with a proper-time arrow, corresponding to the request that future choices cannot influence the present.

\textbf{Axiom 4} (Causality). \textit{The probability of the outcome of a preparation test is independent of the choice of observation tests connected at its output.}

To better understand the statement it is useful to consider the joint test consisting of a preparation test \( \mathcal{X} = \{\rho_i\}_{i \in X} \subset \text{St}(A) \) followed by the observation test \( \mathcal{Y} = \{a_j\}_{j \in Y} \subset \text{Eff}(A) \) performed on system A:
\[
\begin{tikzpicture}
  \node[coordinate] (X) at (0,0) {};
  \node[coordinate] (Y) at (2,0) {};
  \node at (1,.5) {$\mathcal{X}$};
  \node at (1,-.5) {$\mathcal{Y}$};
  \draw[->] (X) -- node[above] {$\Lambda$} (Y);
  \foreach \i in {1,2,...,5} {
    \fill (\i,0) circle (1pt);
  }
\end{tikzpicture}
\]
The joint probability of preparation \( \rho_i \) and observation \( a_j \) is given by
\[
p(i, j|\mathcal{X}, \mathcal{Y}) := (a_j|\rho_i) = \begin{tikzpicture}
  \node[coordinate] (X) at (0,0) {};
  \node[coordinate] (Y) at (1,0) {};
  \node at (0.5,.5) {$\rho_i$};
  \node at (0.5,-.5) {$a_j$};
  \draw[->] (X) -- node[above] {$\Lambda$} (Y);
\end{tikzpicture}.
\]
The marginal probability of the preparation alone does not depend on the outcome \( j \). Yet, it generally depends on which observation test \( \mathcal{Y} \) is performed, namely

\[
\sum_{a_j \in \mathcal{Y}} (a_j | \rho_i) =: p(i | \mathcal{X}, \mathcal{Y}).
\]

The marginal probability of preparation \( \rho_i \) is then generally conditioned on the choice of the observation test \( \mathcal{Y} \). What the causality axiom states is that \( p(i | \mathcal{X}, \mathcal{Y}) \) is actually independent of \( \mathcal{Y} \), namely for any two different observation tests \( \mathcal{Y} = \{a_j\}_{j \in Y} \) and \( \mathcal{Z} = \{b_k\}_{k \in Z} \) one has

\[
p(i | \mathcal{X}, \mathcal{Y}) = p(i | \mathcal{X}, \mathcal{Z}) = p(i | \mathcal{X}).
\]

In a causal OPT the choice of a test on a system can be conditioned on the outcomes of a preceding test, since causality guarantees that the probability distribution of the preceding test is independent of the choice of the following test. This leads us to introduce the notion of **conditioned test**.

**Definition 1** (Conditioned test). If \( \{\mathcal{A}_i\}_{i \in \mathcal{X}} \) is a test from \( A \) to \( B \), and \( \{\mathcal{B}^{(i)}_j\}_{j \in Y_i} \) is a test from \( B \) to \( C \) for every \( i \in \mathcal{X} \), then the conditioned test is a test from \( A \) to \( C \), with outcomes \( (i, j) \in \mathcal{Z} := \cup_i \{\{i\} \times Y_i\} \), and events \( \{\mathcal{B}^{(i)}_j \circ \mathcal{A}_i\}_{(i,j) \in \mathcal{Z}} \). Diagrammatically, the events \( \mathcal{B}^{(i)}_j \circ \mathcal{A}_i \) are represented as follows:

\[
\begin{array}{c}
A \quad \mathcal{A}_i \quad B \quad \mathcal{B}^{(i)}_j \quad C.
\end{array}
\]

Among conditioned test, a special role is played by the **observe-and-prepare test**, where the "connecting" system is the null system I. They are thus made of a preparation test conditioned by an observation test, as follows:

\[
\begin{array}{c}
A \quad \mathcal{A}_i \quad l_i \quad \mathcal{B}^{(i)}_j \quad C.
\end{array}
\]

which can be also represented as \( \{\omega^{(i)}(l_i)\}_{i \in \mathcal{X}} \).

Another remarkable way to characterize causal theories is to require the unicity of the deterministic effect, the equivalence of this formulation with Axiom 4 is given by the following lemma.

**Lemma 1.** An OPT is causal if and only if for every system \( A \) there is a unique deterministic effect.
Proof. We will prove the two directions separately, namely: (1) if the probability of preparation of states is independent of the observation test, then the deterministic effect is unique; (2) vice versa. (1) The probability of the preparation $\rho$ is given by the marginal of the joint probability with the observation, namely $p(\rho) = \sum_{i \in X} (a_i | \rho)$. Upon denoting the deterministic effects of two different tests as $a = \sum_{i \in X} a_i$ and $b = \sum_{j \in Y} b_j$, the statement that the preparation probability is independent of the observation tests translates to $(a | \rho) = (b | \rho)$ for every preparation $\rho \in \text{St}(A)$, which implies that $a = b$, since the set of states is separating for events. (2) Uniqueness of the deterministic effect implies that the preparation probability of each state is independent of the test, since the effect $a = \sum_{i \in X} a_i$ for any test $\{a_i\}_{i \in X}$ is deterministic, and $(a | \rho)$ for any deterministic effect $a \in \text{Eff}(a)$ is the probability of preparation $\rho$. \qed

We will denote the unique deterministic effect for system $A$ as $e_A$, and the subindex will be dropped when no confusion can arise.

In the following we will use the notation $\leq$ to denote the partial ordering between effects, defined as follows:

$$a, b \in \text{Eff}(A), a \leq b \iff (a | \rho) \leq (b | \rho), \forall \rho \in \text{St}(A).$$

It is immediate to show that the causality condition of Lemma 1 spawns the following lemmas.

**Lemma 2.** Causality is equivalent to the following statements regarding tests:

1. **Completeness of observation tests:** For any system $A$ and for every observation test $\{a_i\}_{i \in X}$ one has
   $$\sum_{i \in X} a_i = e_A.$$

2. **Completeness of tests:** For any systems $A, B$ and for every test $\{C_i\}_{i \in X}$ from $A$ to $B$ one has
   $$\sum_{i \in X} (e_B | C_i) = (e_A).$$

3. **Domination of transformations:** For any systems $A, B$ a transformation $C \in \text{Transf}(A \to B)$ satisfies the condition
   $$(e_B | C \leq (e_A),$$


with the equality if and only if \( \mathcal{C} \) is a channel, i.e. a deterministic transformation corresponding to a single-outcome test.

4. **Domination of effects**: For any system A all effects are dominated by a unique effect \( e_A \) which is deterministic

\[
\forall a \in \text{Eff}(A), \quad 0 \leq a \leq e_A.
\]

An immediate consequence of uniqueness of the deterministic effect is the identification of all transformations of the form

\[
\forall \mathcal{B} \in \text{Transf}(B \to C), \quad \sum_{i \in X} a_i \otimes \mathcal{B} = e_A \otimes \mathcal{B},
\]

for any observation test \( \{a_i\}_{i \in X} \) of system A. In particular, we have the factorization of the deterministic effect of composite systems

\[
e_{AB} = e_A \otimes e_B.
\]

The uniqueness of the deterministic effect naturally leads to the relevant notion of *marginal state* or also called *local state*.

**Definition 2** (Marginal state). The marginal state of \( |\sigma\rangle_{AB} \) on system A is the state

\[
|\rho\rangle_A := (e_B|\sigma\rangle_{AB}
\]

represented by the diagram

\[
\begin{array}{c}
\sigma_{AB} \\
\downarrow \\
B
\end{array}^{A} \quad =: \quad \begin{array}{c}
\rho_{A} \\
\downarrow \\
A
\end{array}^{B}.
\]

Finally, the last implication of causality we would outline is the impossibility of signaling without interaction, i.e. by just performing local tests.

**Theorem 1** (No signaling without interaction). In a causal OPT it is impossible to send signals by performing only local tests.

**Proof.** Suppose the general situation in which two "distant" parties Alice and Bob share a bipartite state \( |\Psi\rangle_{AB} \) of systems A and B. Alice performs her local test \( \{A_i\}_{i \in X} \)
on system A and similarly Bob performs his local test \( \{ B_j \}_{j \in Y} \) on system B. The joint probability of their outcomes is

\[
p_{ij} = \langle e |_{AB} (A_i \otimes B_j) |\Psi\rangle_{AB}.
\]

The marginal probabilities \( p_i^A \) at Alice and \( p_j^B \) at Bob are given by

\[
p_i^A := \sum_j p_{ij}, \quad p_j^B := \sum_i p_{ij}.
\]

Alice’s marginal does not depend on the choice of test \( \{ B_j \} \) of Bob, since

\[
p_i^A = \sum_j \langle e |_A (e |_B (A_i \otimes B_j) |\Psi\rangle_{AB} = \langle e |_A (A_i \otimes \sum_j \langle e |_B B_j \rangle) |\Psi\rangle_{AB}
\]

\[
= \langle e |_A A_i |\rho\rangle_A, \quad |\rho\rangle_A := \langle e |_B |\Psi\rangle_{AB},
\]

where we used Eq. (1.1) and the normalization condition \( \sum_j \langle e |_B B_j = \langle e |_B \). The same argument holds for Bob’s marginal.

\[
1.2.5 \quad \text{Local discriminability}
\]

Now we introduce the principle of local discriminability, which stipulates the possibility of discriminating states of composite systems via local measurements on the component systems.

**Axiom 5** (Local discriminability). *It is possible to discriminate any pair of states of composite systems using only local measurements.*

Mathematically the axiom asserts that for every two joint states \( \rho, \sigma \in \text{St}(AB) \), with \( \rho \neq \sigma \), there exist effects \( a \in \text{Eff}(A) \) and \( b \in \text{Eff}(B) \) such that the joint probabilities for the two states are different, namely, in circuits

\[
\begin{aligned}
\begin{array}{c}
\rho \quad \text{A} \\
\text{B}
\end{array}
\neq
\begin{array}{c}
\sigma \quad \text{A} \\
\text{B}
\end{array}
\Rightarrow
\begin{array}{c}
\rho \quad \text{A} \\
\text{B}
\end{array}
\begin{array}{c}
a \\
b
\end{array}
\neq
\begin{array}{c}
\sigma \quad \text{A} \\
\text{B}
\end{array}
\begin{array}{c}
a \\
b
\end{array}.
\end{aligned}
\]

We can now prove one of the main theorem following from the principle of local discriminability.
Theorem 2 (Product rule for composite system). A theory satisfies local discriminability if and only if, for every composite system AB, one has

\[ D_{AB} = D_A D_B. \]  

Proof. By Eq. (1.6), a theory satisfies local discriminability if and only if local effects \( a \otimes b \in \text{Eff}(AB) \), with \( a \in \text{Eff}(A) \) and \( b \in \text{Eff}(B) \), are separating for joint states \( \text{St}(AB) \). Equivalently, the set \( T := \{ a \otimes b | a \in \text{Eff}(A), b \in \text{Eff}(B) \} \) is a spanning set for \( \text{Eff}_{\mathbb{R}}(AB) \). Since the dimension of \( \text{Span}_{\mathbb{R}}(T) \) is \( D_A D_B \) and the spaces of states and effects have the same dimension, we have \( D_{AB} = D_A D_B \). Conversely, if Eq. (1.7) holds, then the product effects are a spanning set for the vector space \( \text{Eff}_{\mathbb{R}}(AB) \), hence they are separating, and local discriminability holds.

Along with the axiom of local discriminability we introduce the notion of entangled and separable states, where entangled states are defined, by negation, as those states that are not separable.

Definition 3 (Separable states). Given \( n \) systems \( A_1, A_2, \ldots, A_n \), the separable states of the composite system \( A_1 A_2 \ldots A_n \) are those of the form

\[
\sum_{i \in X} p_i \alpha_{i_1}^{A_1} \alpha_{i_2}^{A_2} \ldots \alpha_{i_n}^{A_n} = \sum_{i \in X} p_i \alpha_{i_1}^{A_1} \alpha_{i_2}^{A_2} \ldots \alpha_{i_n}^{A_n},
\]

where for \( j = 1, 2, \ldots, n \), \( \alpha_{i_j} \in \text{St}(A_j) \forall i \in X \).

1.2.6 Purification

Purification is the really distinctive and fundamental trait of quantum theory, in the sense that purification allows to distinguish it between all the other possible theories (all the ones we can think of). The statement of the axiom is the following.

\[ \]
Axiom 6 (Purification). For every system $A$ and for every state $\rho \in \text{St}(A)$, there exists a system $B$ and a pure state $\Psi \in \text{PurSt}(AB)$ such that

$$\rho_A = \Psi_{ABE}.$$ 

If two pure states $\Psi$ and $\Psi'$ satisfy

$$\Psi'_{ABE} = \Psi_{ABE},$$

then there exists a reversible transformation $\mathcal{U}$, acting only on system $B$, such that

$$\Psi'_{AB} = \Psi_{AB} \mathcal{U}_{BB'}.$$ (1.8)

Here we say that $\Psi$ is a purification of $\rho$ and that $B$ is the purifying system. The property stated in Eq. 1.8 is called uniqueness of purification and refers to the case in which the two purifications have the same purifying system. It can be easily generalized (the purifying systems are different):

Proposition 2. If two pure states $\Psi \in \text{PurSt}(AB)$ and $\Psi' \in \text{PurSt}(AB')$ are purifications of the same mixed state, then

$$\Psi'_{AB'} = \mathcal{C}_{BC'B'} \Psi_{AB}$$

for some deterministic transformation $\mathcal{C}$ transforming system $B$ into system $B'$.

Proof. Pick two pure states $\beta \in \text{PurSt}(B)$ and $\beta' \in \text{PurSt}(B')$. Since $\Psi \otimes \beta$ and $\Psi' \otimes \beta$ are purifications of the same state on $A$, the uniqueness of purification implies

$$\Psi'_{AB} = \Psi_{AB} \mathcal{U}_{BB'}.$$ 

$$\beta_{BB'} = \beta'_{BB'} \mathcal{U}_{BB'}.$$
for some reversible transformation $\mathcal{U}$ (we have also assumed local discriminability). Discarding system $B$ on both sides we then obtain $\Psi' = (\mathcal{I}_A \otimes C)\Psi$, where $C$ is the deterministic transformation defined by

\[
\begin{align*}
    B & \quad \square \quad B' \\
    \beta & \quad \circ \quad \beta'
\end{align*}
\]

\[
\mathcal{C} \quad B \quad B' := \quad \begin{array}{c}
\mathcal{U} \\
B \quad B' \quad B \quad e
\end{array}
\]
Bit commitment is a cryptographic primitive involving two mistrustful parties, conventionally called Alice and Bob. Alice is supposed to submit an encoded bit of information to Bob in such a way that Bob has (in principle) no chance to identify the bit before Alice later decodes it for him, whereas Alice has (in principle) no way of changing the value of the bit once she has submitted it. In other words, Bob is interested in binding Alice to some commitment, whereas Alice would like to conceal her commitment from Bob.

In the first two sections of this Chapter we will describe the protocol: we will start from an historical perspective, from the first article published by Blum in 1983 [14] to the most recent developments of the last decade that focus on the impossibility of bit commitment in quantum theory [15, 16]. Then, we will rigorously define the protocol in the language of OPTs.

As every other cryptographic primitive, bit commitment does not need to be perfectly secure, i.e. probability of cheating equals to 0 for Alice and 1/2 for Bob (who can always randomly guess), to be efficient. In fact, even if a greater probability of error is admitted, iterating the protocol the error probability can be generally asymptotically reduced. A protocol that admits this possibility is called unconditionally secure, with a vast literature on the subject. However in our thesis we will only deal with perfectly secure bit commitment. Since we are considering the protocol within the operational framework, there are not yet the technical tools to analyse unconditionally secure bit commitment in OPTs and,
anyway, the perfectly secure protocol should be the starting point for a rigorous analysis. The exceptional thing is that in our analysis we will be anyway able to build a cheating scheme also for some unconditionally secure bit commitment protocol, as we will see in Section 4.2.

Finally, in the third part of the Chapter, we will study the proof of the impossibility of perfectly secure bit commitment in OP quantum theory done in Ref. [2]. We will review if all of the six axioms of quantum theory introduced in Section 1.2 are really necessary conditions. Some of them will be neglected and other will be replaced with weaker hypothesis. With the new sufficient conditions that we will find, the proof of impossibility of perfectly secure bit commitment can be extended to other theories than the quantum one. Some applications hypothesis are the fermionic theory and the real quantum theory and in particular the PR-box theory. To the latter will be focused the next two Chapter, in fact, in literature numerous bit commitment protocols (and more generally quantum key distributions protocols) have been studied in the context of non-local correlated box, i.e. PR-boxes. So our analysis will provide a solid (operational probabilistic) point of view from which study all these protocols that have been prosed in the past years.

2.1 From Coin Tossing to no-go Theorem

The bit commitment protocol was conceived for the first time by Blum in 1983 as a building block for secure coin tossing. To cite the abstract of the original work [14]:

Alice and Bob want to flip a coin by telephone. (They have just divorced, live in different cities, want to decide who gets the car.) Bob would not like to tell Alice HEADS and hear Alice (at the other end of the line) say “Here goes... I’m flipping the coin .... You lost!”.

A standard example to illustrate bit commitment is for Alice to write the bit down on a piece of paper, which is then locked in a safe and sent to Bob, whereas Alice keeps the key. At a later time, she will unveil it by handing over the key to Bob. However, Bob has a well-equipped toolbox at home and may have been able to open the safe in the meantime. So while this scheme may offer reasonably good practical security, it is in principle insecure. Yet all bit commitment schemes that have wide currency today rely on such technological constraints: not on strongboxes and keys, but on unproven assumptions that certain computations are hard to perform.

The first example of quantum bit commitment was first proposed by Bennet and Brassard in their famous paper of 1984 [17] as a primitive for implementing coin toss-
2.1. From Coin Tossing to no-go Theorem

In their scheme, Alice commits to a bit value by preparing a sequence of photons in either of two mutually unbiased bases, in a way that the resulting quantum states are indistinguishable to Bob. The authors show that their protocol is secure against so-called passive cheating, in which Alice initially commits to the bit value $k$ and then tries to unveil $1 - k$ later. However, they also prove that Alice can cheat with a more sophisticated strategy, in which she initially prepares pairs of maximally entangled states instead, keeps one particle of each pair in her laboratory and sends the second particle to Bob. For the first time, entanglement is recognized as a crucial factor in preventing perfectly secure bit commitment.

Subsequent proposals for bit commitment schemes tried to evade this type of attack by forcing the players to carry out measurements and communicate classically as they go through the protocol. At a 1993 conference Brassard et al. presented a bit commitment protocol [18] that was claimed and generally accepted to be unconditionally secure.

In 1996 Lo and Chau [19], and Mayers [20] realized that all previously proposed bit commitment protocols were vulnerable to a generalized version of the (EPR) attack that renders the BB84 proposal insecure, a result that they slightly extended to cover quantum bit commitment protocols in general. In OP terms, the determinant factor in the impossibility of bit commitment shifts from entanglement to the purification principle, this will be fully proved in Ref [2].

Their basic argument is the following. At the end of the commitment phase, Bob will hold one out of two quantum states $\psi_k$ as proof of Alice's commitment to the bit value $k \in \{0, 1\}$. Alice holds its purification $\Psi_k$, which she will later pass on to Bob to unveil. For the protocol to be concealing, the two states $k$ should be (almost) indistinguishable, $\psi_0 \approx \psi_1$. But Uhlmann's theorem then implies the existence of a unitary transformation $\mathcal{U}$ that (nearly) rotates the purification of $\psi_0$ into the purification of $\psi_1$. Since $\mathcal{U}$ is localized on the purifying system only, which is entirely under Alice's control, Lo-Chau-Mayers argue that Alice can switch at will between the two states, and is not in any way bound to her commitment. As a consequence, any concealing bit commitment protocol is argued to be necessarily non-binding.

Starting from 2000 the Lo-Chau-Mayers no-go theorem has been continually challenged, arguing that the impossibility proof does not exhaust all conceivable quantum bit commitment protocols. Several protocols have been proposed and claimed to circumvent the no-go theorem. These protocols seek to strengthen Bob’s position with the help of "secret parameters" or "anonymous states", so that Alice lacks some information needed to cheat successfully: while Uhlmann's theorem would still imply the existence of a unitary cheating transformation as described above, this transfor-
A rigorous definition is in order for a twofold reason. If it is true that it is necessary to define a framework where to operate, it is at the same time important to clearly remark the definition of bit commitment to which the statement “bit commitment is impossible” refers to.

Despite we already referred to the word protocol, we did not linger to define what we mean with it, therefore we will start by the notion of protocol and we will state the bit commitment in the OPT language and its key properties only thereafter.

2.2.1 The protocol
A protocol regulates the exchange of messages between participants, defining what are the honest strategies that they can adopt, so that at every stage it is clear what type of message is expected from the participants, although, of course, their content is not fixed. The expected message types can be either classical or quantum or a combination thereof.
2.2. A Formal Definition

In any bit commitment protocol, we can distinguish two main phases: the first is the commitment phase, in which Alice and Bob exchange classical and quantum messages in order to commit the bit. The second is the opening phase where Alice will send to Bob some classical or quantum information in order to reveal the bit value.

Commitment phase this phase can end either with a successful commitment, or with an abort, in which the two parties irrevocably give up the purpose of committing the bit (of course, in a well designed protocol, if both parties are honest the probability of abort should be vanishingly small). If no abort took place, the bit value is considered to be committed to Bob but, supposedly, concealed from him. Since bit commitment is a two-party protocol and trusted third parties are not allowed, the starting state necessarily has to be originated by one of the two parties. Moreover, since we can always include in the protocol null steps (in which no information, classical or quantum, is exchanged), without loss of generality, we can restrict our attention to protocols that are started by Alice.

Opening phase in the case of abort during the commitment, this is just a null step, whereas, in the case of successful commitment, at the opening Alice will send to Bob some classical or quantum information in order to reveal the bit value. Taking both Alice’s message and his own (classical and quantum) records, Bob will then perform a suitable verification measurement. His measurement will result in either a successful readout of the committed bit, or in a failure, e.g. due to the detection of an attempted cheat. Again, in a well-designed protocol the probability of failure should be vanishingly small.

2.2.2 OP bit commitment

Alice wants to commit a classical bit $b \in \{0, 1\}$ to Bob.

1. as we have seen before, the first phase is the commitment phase, in which Alice and Bob can perform any sequence of operations. Depending on whether Alice intended to commit $b = 0$ or $b = 1$, the state $\Psi_0$, $\Psi_1 \in St(AB)$ is selected, respectively. We will assume that Alice and Bob’s systems at the end of the commitment phase are $A$ and $B$ respectively, so that the two possible pure states that can be transmitted to Bob are $(e|_A \Psi_0$, $(e|_A \Psi_1 \in St(B)$;

2. between the commitment and the opening phase Alice and Bob can perform only local operations on their systems $A$ and $B$, respectively (together with every other ancillary system they control);
3. during the opening phase, Alice transmits her system $A$ to Bob who can perform any measurement on the joint system $AB$ to know the bit $b$ and to check if it is compatible with the commitment of Alice. In general, to do this Bob can perform a two outcome measurement (Positive Operator Valued Measure, POVM) $\{a_0, a_1\}$ on the joint system $AB$.

From now on, when we will refer to bit commitment we mean a protocol that is included in the previous scheme.

Already at the beginning of this Chapter we intuitively mention the two way of cheating that can occur. Being a two party protocol, we can have Alice’s cheating, i.e. she changes the bit after the commitment (the protocol is not binding), and Bob’s cheating, i.e. Bob discovers the bit committed before the opening (the protocol is not concealing). In this way we have now identified two key properties of any bit commitment protocol. However there is a third key property that is often omitted in the literature: the correctness of the protocol that guarantees the correct verification of the bit committed in the opening phase.

To recapitulate with a properer language, we will say that a bit commitment protocol is

- **binding**: if, for honest Bob, Alice should not be able to change the bit she committed. More precisely, assume that a possibly dishonest Alice committed $b$ but she wants to reveal $b' \neq b$, then it must be

  $$\Pr[\text{Bob accepts} | \text{Alice reveals } b'] < 1; \quad (2.1)$$

- **concealing**: if, for honest Alice, Bob should not be able to know the bit that Alice committed until she reveals it;

- **correct**: for honest Alice and Bob, if Alice commits $b$ and later reveals $b$ to Bob, then Bob accepts with probability grater then $1/2$:

  $$\Pr[\text{Bob accepts} | \text{Alice reveals } b] > \frac{1}{2}. \quad (2.2)$$

Furthermore, we will say that a bit commitment protocol is *perfect* if

- it is perfectly binding, namely for honest Bob, Alice cannot switch between $\Psi_0$ and $\Psi_1$ in such a way that Bob cannot detect the switch with certainty (i.e. Eq. (2.1) becomes $\Pr[\text{Bob accepts} | \text{Alice reveals } b'] = 0$). Namely it does not exists a reversible channel $\mathcal{U}$ such that

  $$(\mathcal{U} \otimes \mathcal{I}_B)|\Psi_0\rangle_{AB} = |\Psi_1\rangle_{AB}; \quad (2.3)$$
2.3. Lightening No Bit Commitment

- it is perfectly concealing, namely for honest Alice, Bob would not be able to perform some measurement of his system $B$ and gain at least partial information about which bit Alice committed

\[ (e|_A|\Psi_0)_{AB} = (e|_A|\Psi_1)_{AB} ; \quad (2.4) \]

- it is correct with probability one, namely Bob accepts with probability one. So Eq. (2.2) becomes

\[ \Pr[ \text{Bob accepts} \mid \text{Alice reveals } b] = 1 . \]

In the perfect implementation of the protocol the probability of cheating of Bob must be $1/2$, since in the worst case he can always make a random guess of the committed bit. The cheating probability of Alice is instead equal to zero.

As we have anticipated in Section 2.1, perfectly secure bit commitment protocol is impossible in quantum and classic theory. However, if the perfect concealing and perfect biding conditions are relaxed as follows

\[ (\mathcal{U} \otimes I_B)|\Psi_0)_{AB} \approx |\Psi_1)_{AB} \]
\[ (e|_A|\Psi_0)_{AB} \approx (e|_A|\Psi_1)_{AB} \]

interesting level of security relevant for concrete applications emerges. Within this scenario the main effort is in quantifying the cheating probabilities and their trade-off in operational terms. The goal is to achieve a protocol that is asymptotically binding and concealing (both the Alice and Bob cheating probability can be made arbitrarily small) against an adversary that has no restrictions on the computational resources. In this case one has unconditionally secure bit commitment.

2.3 Lightening No Bit Commitment

In this Section we will analyse the necessary assumptions to prove the impossibility of perfectly secure bit commitment. In the literature, a no-go theorem in quantum theory has been proven in Ref. [15, 16] but we found that not all of the axioms of quantum theory are necessary. So, using weaker assumption, we will generalize the impossibility of bit commitment to a larger set of operational probabilistic theories.

Clearly the causality Axiom [4] is essential, otherwise the very protocol could not be defined properly (there would not be a given order in the succeeding of the phases).
In our analysis we also assume Axiom 1, atomicity of composition, that is a sufficient condition to grant Corollary 1. Finally, instead of the purification Axiom 6, we take the following weaker assumption. Before stating it, we introduce the notion of dynamically faithful state.

**Definition 4** (Dynamically faithful state). A state $\sigma \in \text{St}(AC)$ is dynamically faithful for system $A$ if for any couple of transformations $\mathcal{A}, \mathcal{A}' \in \text{Transf}(A \to B)$ one has

$$\mathcal{A}|\sigma)_{AC} = \mathcal{A}'|\sigma)_{AC} \Rightarrow \mathcal{A} = \mathcal{A}'.$$

We are now in position to state the "new" axiom.

**Axiom 7.** For every system $A$ there exist a system $\tilde{A}$ and a pure state $\Psi^{(A)} \in \text{St}(A\tilde{A})$ that is dynamically faithful for system $A$. Furthermore, for every system $B$ and for every bipartite state $R \in \text{St}(B\tilde{A})$ such that

$$R|e_{\tilde{A}} = \Psi^{(A)}A_e\text{ e}_{\tilde{A}},$$

there exists a purification of $R$. If $\Phi, \Phi' \in \text{St}(C\tilde{A})$ are two purifications of $R$ then they are connected by a reversible transformation $\mathcal{U} \in \text{Transf}(C)$.

In quantum theory, the existence of mixed faithful states is a direct consequence of local discriminability. In a theory with both local discriminability and purification there exist also dynamically faithful states that are pure. In Axiom 7 we do not require local discriminability nor purification but the existence of dynamically faithful pure states. In addition we also require that a purification exists only for those states that have the same marginal of the dynamically faithful pure ones and that this purification is unique up to a reversible transformation on the purifying system.

An immediate consequence of Axioms 1 and 4 is the following lemma on the uniqueness of purification.

**Lemma 3** (Uniqueness of the purification up to channels on the purifying systems). Let $\Psi \in \text{St}(AB)$ and $\Psi' \in \text{St}(AC)$ be two purification of $\rho \in \text{St}(A)$. Then there exists a channel $\mathcal{C} \in \text{Transf}(B \to C)$ such that

$$\Psi'_{AC} = \mathcal{C} \Psi_{BC}.$$
Moreover, channel $C$ has the form
\[ B \xrightarrow{C} C = \begin{array}{c} \varphi_0 \end{array} \xrightarrow{U} \begin{array}{c} \psi \end{array} \xrightarrow{e} C. \]

for some pure state $\varphi_0 \in \text{St}(C)$ and some reversible channel $U \in \text{Transf}(B \to C)$.

**Proof.** Let $\eta$ and $\varphi_0$ be an arbitrary pure state of $B$ and $C$, respectively. Then, due to Axiom 1, $|\Psi'\rangle_{AC}|\eta\rangle_B$ and $|\Psi\rangle_{AB}|\varphi_0\rangle_C$ are still two pure states and so they are both purifications of $\rho$ with the same purifying system $BC$. Due to Axiom 7 we have
\[ \begin{array}{c} \eta \end{array} \xrightarrow{B} \begin{array}{c} \varphi_0 \end{array} \xrightarrow{\Psi'} \begin{array}{c} \psi \end{array} \xrightarrow{U} \begin{array}{c} \psi \end{array} \xrightarrow{\Psi} \begin{array}{c} \eta \end{array} \xrightarrow{B} \begin{array}{c} \varphi_0 \end{array}. \]

Applying the deterministic effect $e_B$ on system $B$ we obtain the thesis, with $C = (e|_B U|\varphi_0)$.

\[ \square \]

### 2.3.1 Reversible dilation of channels

Before starting the proof of the impossibility of bit commitment it is in order to derive some useful results about reversible dilations of channels.

**Lemma 4.** Let $R \in \text{St}(B\tilde{A})$ be a state such that
\[ R = \begin{array}{c} \psi^{(A)} \end{array} = \begin{array}{c} \psi^{(A)} \end{array} \xrightarrow{\psi^{(A)}} \begin{array}{c} \psi^{(A)} \end{array}. \]

where $\psi^{(A)} \in \text{St}(A\tilde{A})$ is a pure dynamically faithful state for system $A$. Then there exist a system $C$, a pure state $\varphi_0 \in \text{St}(BC)$, and a reversible channel $U \in \text{Transf}(ABC)$ such that
\[ R = \begin{array}{c} \varphi_0 \end{array} \xrightarrow{\psi^{(A)}} \begin{array}{c} \psi^{(A)} \end{array} = \begin{array}{c} \psi^{(A)} \end{array} \xrightarrow{U} \begin{array}{c} \psi^{(A)} \end{array} \xrightarrow{\psi^{(A)}} \begin{array}{c} \psi^{(A)} \end{array}. \]

Moreover the channel $\mathcal{V} \in \text{Transf}(A \to ABC)$ defined by $\mathcal{V} := U|\varphi_0\rangle_{BC}$ is unique up to reversible channels on $AC$. 

33
2.3. Lightening No Bit Commitment

Proof. Take a purification of $R$, say $\Psi_R \in \text{St}(C\tilde{B}\tilde{A})$ for some purifying system $C$ (existence of such a purification is granted by Axiom 7). One has

$$
\begin{array}{c}
\Psi_R \\
C \\
B \\
\tilde{A}
\end{array} \\
\begin{array}{c}
\tilde{C} \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
R \\
B \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\Psi^{(A)} \\
A \\
\tilde{A}
\end{array}
$$

that is, the pure state $\Psi_R$ and $\Psi^{(A)}$ have the same marginal on system $\tilde{A}$. Applying the uniqueness of purification expressed by Lemma 3 one then obtains

$$
\begin{array}{c}
\Psi_R \\
C \\
B \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\tilde{C} \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\varphi_0 \\
BC \\
A \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\Psi^{(A)} \\
A \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\varphi_0 \\
BC \\
A \\
\tilde{B} \\
\tilde{A}
\end{array}
$$

Applying the deterministic effect on system $C$ on both sides, one then proves Eq. (2.5). Moreover, if $\mathcal{V} := \mathcal{U}|\varphi'_0)_{BC}$ is channel such that Eq. (2.5) holds, then the pure states $\mathcal{V}|\Psi^{(A)})_{\tilde{A}}$ and $\mathcal{V}'|\Psi^{(A)})_{\tilde{A}}$ have the same marginal on system $B\tilde{A}$. Uniqueness of purification then implies

$$
\begin{array}{c}
\Psi^{(A)} \\
A \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\varphi' \\
AC \\
B \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\mathcal{V} \\
A \\
B \\
\tilde{B} \\
\tilde{A}
\end{array} = \\
\begin{array}{c}
\mathcal{V} \\
A \\
B \\
\tilde{B} \\
\tilde{A}
\end{array}
$$

for some reversible channel $\mathcal{W} \in \text{Transf}(AC)$. Since $\Psi^{(A)}$ is dynamically faithful for $A$, this implies $\mathcal{V}' = \mathcal{W}\mathcal{V}$. 

We now give the definition of dilatation and reversible dilatation.

**Definition 5 (Dilatation of a channel).** A dilatation of channel $\mathcal{C} \in \text{Transf}(A \rightarrow B)$ is a channel $\mathcal{V} \in \text{Transf}(A \rightarrow BE)$ such that

$$
\begin{array}{c}
A \\
\mathcal{C} \\
B
\end{array} = \\
\begin{array}{c}
A \\
\mathcal{V} \\
B \\
E \\
\tilde{E}
\end{array}.
$$

We refer to system $E$ as to the environment.
2.3. Lightening No Bit Commitment

**Definition 6** (Reversible dilation). A dilatation $\mathcal{V} \in \text{Transf}(A \rightarrow BE)$ is called reversible if there exists a system $E_0$ such that $AE_0 \simeq BE$ and

$$
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{V} \\
E \\
\hline
B
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \quad \varphi_0 \\
E_0 \\
\hline
B
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{U} \\
E \\
\hline
B
\end{array}
\end{array}
$$

for some pure state $\varphi_0 \in \text{St}(E_0)$ and some reversible channel $\mathcal{U} \in \text{Transf}(AE_0 \rightarrow BE)$.

According to the above definitions, we have the following dilatation theorem:

**Theorem 3** (Reversible dilatation of channels). Every channel $\mathcal{C} \in \text{Transf}(A \rightarrow B)$ has a reversible dilatation $\mathcal{V} \in \text{Transf}(A \rightarrow BE)$. If $\mathcal{V}, \mathcal{V}' \in \text{Transf}(A \rightarrow BE)$ are two reversible dilatations of the same channel, then they are connected by a reversible transformation on the environment, namely

$$
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{V}' \\
E \\
\hline
\epsilon \quad B
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{V} \\
E \\
\hline
\epsilon \quad B
\end{array}
\end{array}
$$

$$
\Rightarrow 
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{V}' \\
E \\
\hline
\epsilon \quad B
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{V} \\
E \\
\hline
\epsilon \quad B
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{W} \\
E \\
\hline
\epsilon \quad B
\end{array}
\end{array}
$$

for some reversible channel $\mathcal{W} \in \text{Transf}(E)$.

**Proof.** Let us store the channel $\mathcal{C}$ in the faithful state $\Psi^{(A)} \in \text{St}(A\tilde{A})$, thus getting the state $R_C$:

$$
\begin{array}{c}
\begin{array}{c}
\tilde{A} \quad R_C \\
\hline
B
\end{array}
\end{array}
:= 
\begin{array}{c}
\begin{array}{c}
\tilde{A} \quad \Psi \\
A \\
\hline
C \\
B
\end{array}
\end{array}
$$

Since $\mathcal{C}$ is a channel, it satisfy the normalization condition

$$
\begin{array}{c}
\begin{array}{c}
A \quad \mathcal{C} \\
B \\
\hline \epsilon
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
A \\
\hline \epsilon
\end{array}
\end{array}
$$

which implies

$$
\begin{array}{c}
\begin{array}{c}
\tilde{A} \quad R_C \\
\hline
B \quad \epsilon
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
\tilde{A} \quad \Psi^{(A)} \\
A \\
\hline
C \\
B \quad \epsilon
\end{array}
\end{array}
= 
\begin{array}{c}
\begin{array}{c}
\tilde{A} \quad \Psi^{(A)} \\
A \\
\hline \epsilon
\end{array}
\end{array}
$$
Now, applying Lemma 4 we obtain

$$R_C = \psi_0 \circ \Phi \circ U \circ \psi(A)$$

Since $\psi(A)$ is dynamically faithful for system $A$, this implies

$$\Phi_0 \circ \Phi \circ U \circ \psi(A) = \psi_0 \circ \Phi \circ U \circ \psi(A).$$

Therefore, $\mathcal{V} := U|_{\Phi_0 BC}$ is a reversible dilatation of $\mathcal{C}$, with $E_0 := BC$ and $E := AC$. Finally, the uniqueness clause in Lemma 4 implies uniqueness of the dilatation.

Moreover, two reversible dilatations of the same channel with different environments are related as follows.

**Lemma 5.** Let $\mathcal{V} \in \text{Transf}(A \rightarrow BE)$ and $\mathcal{V}' \in \text{Transf}(A \rightarrow BE')$ be two reversible dilatations of the same channel $\mathcal{C}$, with generally different environment $E$ and $E'$. Then there is a channel $\mathcal{L}$ from $E$ to $EE'$ such that

$$\mathcal{V}' \circ E' = \mathcal{L} \circ E \circ \mathcal{V} \circ E.$$ 

The channel $\mathcal{L}$ has the form

$$\mathcal{L} = \eta_0 \circ \Phi \circ U \circ \psi(A),$$

for some pure state $\eta_0$ and some reversible transformation $\Phi \in \text{Transf}(EE')$.

**Proof.** Apply $\mathcal{V}$ and $\mathcal{V}'$ to the faithful state $\psi(A)$ and then use the uniqueness of purification stated in Lemma 3. □
2.3. Casually ordered channels and channels with memory

The last step before the proof of the theorem is the notion of casually ordered channels and channels with memory.

**Definition 7** (Casually ordered bipartite channel). A bipartite channel \( \mathcal{C} \) from \( A_1A_2 \) to \( B_1B_2 \) is casually ordered if there is a channel \( \mathcal{D} \) from \( A_1 \) to \( B_1 \) such that \( (e|_{B_2})\mathcal{C} = \mathcal{D} \otimes (e|_{A_2}) \). Diagrammatically,

\[
\begin{array}{c}
A_1 \\
\mathcal{C} \\
A_2 \\
\end{array} & \begin{array}{c}
B_1 \\
\mathcal{D} \\
B_2 \\
\end{array} = \begin{array}{c}
A_1 \\
\mathcal{D} \\
A_2 \\
\end{array} \begin{array}{c}
B_1 \\
\mathcal{I} \\
B_2 \\
\end{array} .
\]

(2.6)

Eq. (2.6) means that the channel \( \mathcal{C} \) does not allow for signaling from the input system \( A_2 \) to the output system \( B_1 \). In a relativistic context, this can be interpreted as \( B_1 \) being outside the causal future of \( A_2 \).

**Definition 8** (Sequence of two channel with memory). A bipartite channel \( \mathcal{C} \) from \( A_1A_2 \) to \( B_1B_2 \) can be realized as a sequence of two channels with memory if there exist two system \( E_1, E_2 \), called memory systems, and two channels \( \mathcal{C}_1 \in \text{Transf}(A_1 \rightarrow B_1E_1) \) and \( \mathcal{C}_2 \in \text{Transf}(A_2E_1 \rightarrow B_2E_2) \) such that \( \mathcal{C} = (e|_{E_2})(\mathcal{C}_2 \otimes \mathcal{I}_{B_1})(\mathcal{I}_{A_2} \otimes \mathcal{C}_1) \). Diagrammatically,

\[
\begin{array}{c}
A_1 \\
\mathcal{C} \\
A_2 \\
\end{array} = \begin{array}{c}
A_1 \\
\mathcal{C}_1 \\
A_2 \\
\end{array} \begin{array}{c}
B_1 \\
\mathcal{I} \\
B_2 \\
\end{array} \begin{array}{c}
E_1 \\
\mathcal{C}_2 \\
E_2 \\
\end{array} .
\]

(2.7)

For causally ordered bipartite channels the dilatation theorem implies the following result:

**Theorem 4** (Casual ordering is memory). A bipartite channel \( \mathcal{C} \) from \( A_1A_2 \) to \( B_1B_2 \) is casually ordered if and only if it can be realized as a sequence of two channels with memory. Moreover, the channels \( \mathcal{C}_1, \mathcal{C}_2 \) in Eq. (2.7) can be always chosen such that \( \mathcal{C}_2\mathcal{C}_1 \) is a reversible dilatation of \( \mathcal{C} \).

**Proof.** If Eq. (2.7) holds, the channel \( \mathcal{C} \) is clearly casually ordered, with the channel \( \mathcal{D} \) given by \( \mathcal{D} := (e|_{E_1})\mathcal{C}_1 \). Conversely, suppose that \( \mathcal{C} \) is casually ordered. Take a reversible dilatation of \( \mathcal{C} \), say \( \mathcal{V} \in \text{Transf}(A_1A_2 \rightarrow B_1B_2E) \), and a reversible dilatation...
of $\mathcal{D}$, say $\mathcal{V}_1 \in \text{Transf}(A_1 \to B_1 E_1)$. Now by the definition of casually ordered channels (Eq. (2.6)) we have

$$\begin{array}{c}
\begin{array}{c}
A_1 \vdash \mathcal{V} \vdash B_1 \\
A_2 \vdash E \vdash E
\end{array}
\end{array} = \begin{array}{c}
\begin{array}{c}
A_1 \vdash \mathcal{V}_1 \vdash B_1 \\
A_2 \vdash E_1 \vdash E
\end{array}
\end{array}.$$ 

This means that $\mathcal{V}$ and $\mathcal{V}_1 \otimes I_{A_2}$ are two reversible dilatations of the same channel. By the uniqueness of the reversible dilatation expressed by Lemma 5 we then obtain

$$\begin{array}{c}
\begin{array}{c}
A_1 \vdash \mathcal{V} \vdash B_1 \\
A_2 \vdash E \vdash E
\end{array}
\end{array} = \begin{array}{c}
\begin{array}{c}
A_1 \vdash \mathcal{V}_1 \vdash B_1 \\
A_2 \vdash E_1 \vdash E
\end{array}
\end{array}.$$ 

Once we have defined $E_2 := EE_1 A_2$ it only remains to observe that the above diagram is nothing but the thesis, with $\mathcal{C}_1 = \mathcal{V}_1$ and $\mathcal{C}_2 = \mathcal{L}$. By construction, $\mathcal{C}_2 \mathcal{C}_1$ is a reversible dilatation of $\mathcal{C}$.

The definition of casually ordered bipartite channel is easily extended to the multipartite case. Here we will only report the definition and the two main theorems of the theory. For their demonstrations we remand to the original article [2] since they are still valid also from our three Axiom as starting point.

**Definition 9** (Causally ordered channel). An $N$-partite channel $\mathcal{C}^{(N)}$ from $A_1 \ldots A_N$ to $B_1 \ldots B_N$ is causally ordered if for every $k \leq N$ there is a channel $\mathcal{C}^{(k)}$ from $A_1 \ldots A_k$ to $B_1 \ldots B_k$ such that

$$\begin{array}{c}
\begin{array}{c}
A_1 \vdash \ldots \vdash A_k \vdash \mathcal{C}^{(N)} \vdash \ldots \vdash A_{k+1} \vdash B_k \\
\vdots \vdash \ldots \vdash \vdots \vdash \ldots \vdash \vdots \\
A_N \vdash \vdots \vdash B_N \\
\vdots \vdash \ldots \vdash \vdots \vdash \ldots \vdash \vdots
\end{array}
\end{array} = \begin{array}{c}
\begin{array}{c}
A_1 \vdash \ldots \vdash A_k \vdash \mathcal{C}^{(k)} \vdash \ldots \vdash A_{k+1} \vdash B_k \\
\vdots \vdash \ldots \vdash \vdots \vdash \ldots \vdash \vdots \\
A_N \vdash \vdots \vdash B_N \\
\vdots \vdash \ldots \vdash \vdots \vdash \ldots \vdash \vdots
\end{array}
\end{array}.$$ 

The definition means that the output systems $B_1 \ldots B_k$ are outside the casual future of any input system $A_l$ with $l > k$. Causally ordered channels can be characterized as follows.
Theorem 5 (Causal ordering is memory for general N). An N-partite channels \( C^{(N)} \) from \( A_1 \ldots A_N \) to \( B_1 \ldots B_N \) is causally ordered if and only if there exist a sequence of memory systems \( \{E_k\}_{k=0}^{N} \) with \( E_0 = I \) and a sequence of channels \( \{V_k\}_{k=1}^{N} \), with \( V_k \in \text{Transf}(A_kE_{k-1} \rightarrow B_kE_k) \) such that

\[
\begin{array}{cccc}
A_1 & b_{(N)} & B_1 \\
\vdots & \vdots & \vdots \\
A_N & b_{(N)} & B_N \\
\end{array}
= \begin{array}{cccc}
V_1 & E_1 & A_2 & B_2 & \ldots & A_N & B_N \\
V_2 & E_2 & \ldots & E_{N-1} & V_N & E_N \\
& E & & & & \end{array}
\]

Moreover, \( V_N \ldots V_1 \) is a reversible dilation of \( C^{(N)} \).

We have also a uniqueness result:

Corollary 2 (Uniqueness of the reversible dilatation). Let \( \{V_k\}_{k=1}^{N}, V_k \in \text{Transf}(A_kE_{k-1} \rightarrow B_kE_k) \) be a reversible realization of the causally ordered channel \( C^{(N)} \) as a sequence of channels with memory, as in Theorem 5. Suppose that \( \{V'_k\}_{k=1}^{N}, V'_k \in \text{Transf}(A_kE'_{k-1} \rightarrow B_kE'_k) \) is another reversible realization of \( C^{(N)} \) as a sequence of channels with memory. Then there exist a channel \( R \) from \( E_N \) to \( E'_N \) such that

\[
\begin{array}{cccc}
\begin{array}{cccc}
V'_1 & E'_1 & A_2 & B_2 & \ldots & A_N & B_N \\
& V'_2 & E'_2 & \ldots & E'_{N-1} & V'_N & E'_N \\
& & E & & & \end{array}
= \begin{array}{cccc}
V_1 & E_1 & A_2 & B_2 & \ldots & A_N & B_N \\
V_2 & E_2 & \ldots & E_{N-1} & V_N & E_N \\
& E & & & & \end{array}
\end{array}
\]

2.3.3 No bit commitment

The bit commitment protocol defined in Section 2.2 is generally implemented with sequences of channels with memory, that can be used to describe sequences of moves of Alice or Bob. In this scenario, the memory systems are the private systems available to a party, while the other input-output systems are the systems exchanged in the communication with the other party.

We recall that for every theory that satisfy the starting hypothesis, this proof has an absolutely general validity: for every kind of input states and every possible strategy adopted, including both atomic and non-atomic transformations.

Theorem 6 (No perfectly secure bit commitment). If a N-round protocol is perfectly concealing, then there is a perfect cheating.
2.3. Lightening No Bit Commitment

Proof. We first prove the impossibility for protocols that do not involve the exchange of classical information. Let \( \mathcal{A}_0, \mathcal{A}_1 \in \text{Transf}(A_1\ldots A_N \rightarrow B_1\ldots B_N F_N) \) two causally ordered N-partite channels (here the last output system of the causally ordered channels is the bipartite system \( B_N, F_N \)), representing Alice’s move to encode the bit value \( b = 0, 1 \), respectively. The system \( F_N \) is the system sent from Alice to Bob at the opening phase in order to unveil the value of the bit. If the protocol is perfectly concealing, then the reduced channels before the opening phase must be indistinguishable, namely \( (e|_{F_N}) \mathcal{A}_0 = (e|_{F_N}) \mathcal{A}_1 =: \mathcal{C} \). Now, due to Theorem 3, there exist two reversible dilatations \( \mathcal{V}_0 \in \text{Transf}(A_1\ldots A_N \rightarrow B_1\ldots B_N F_N G_0) \) and \( \mathcal{V}_1 \in \text{Transf}(A_1\ldots A_N \rightarrow B_1\ldots B_N F_N G_1) \) for \( \mathcal{A}_0 \) and \( \mathcal{A}_1 \), respectively. Since \( \mathcal{V}_0 \) and \( \mathcal{V}_1 \) are also two dilatations of the channel \( \mathcal{C} \), due to Lemma 5 there is a channel \( \mathcal{R} \) from \( F_N G_0 \) to \( F_N G_1 \) such that \( \mathcal{V}_1 = \mathcal{R} \mathcal{V}_0 \). Applying this channel to her private systems, Alice can switch from \( \mathcal{V}_0 \) to \( \mathcal{V}_1 \) just before the opening. Discarding the auxiliary system \( G_1 \), this yields channel \( \mathcal{A}_1 \).

The cheating is perfect, since Alice can play the strategy \( \mathcal{V}_0 \) until the end of the commitment and decide the bit value before the opening without being detected by Bob.

The above reasoning can be extended to N-round protocols involving the exchange of classical information. Indeed classical messages can be modeled by measure-and-prepare channels where the observation states are perfectly distinguishable. The fact that some systems can only be prepared in perfectly distinguishable states will be referred to as "communication interface" of the protocol. In this case, to construct Alice’s cheating strategy we can first take the reversible dilatations \( \mathcal{V}_0, \mathcal{V}_1 \) and the channel \( \mathcal{R} \) such that \( \mathcal{V}_1 = \mathcal{R} \mathcal{V}_0 \). In order to comply with the communication interface protocol, one can compose \( \mathcal{V}_0 \) and \( \mathcal{V}_1 \) with classical channels on all system that must be "classical" before the opening, thus obtaining two channels \( \mathcal{D}_0 \) and \( \mathcal{D}_1 \) that are no longer reversible but still satisfy \( \mathcal{D}_1 = \mathcal{R} \mathcal{D}_0 \). Discarding the auxiliary system \( G_1 \) and, of required by the communication interface, applying a classical channels on \( F_N \), Alice then obtains channel \( \mathcal{A}_1 \). Again, this strategy allows Alice to decide the value of the bit just before the opening without being detected. \( \square \)
In this Chapter we will analyse the probabilistic theory \cite{21, 22, 23} corresponding to the popular PR-boxes model introduced in Ref. \cite{7}.

In particular, in the first Section we will retrace the crucial steps and the underlying reasons that led to the development of the PR-box model. Then, in the second Section, after the formalization of the PR-boxes model in the language of operational probabilistic theories, we will report our results: the general POVM that grants perfect discriminability between any two pure states in the bi-partite case, the existence and uniqueness of the purification exclusively for the maximally mixed state (again in the bi-partite scenario) and finally some consideration on the general case of N-partite boxes.

3.1 Why PR-Boxes?

One of the most striking feature of quantum theory is certentantly non-locality. In fact, since the very beginning of the theory the incompleteness of the Copenhagen interpretation of quantum mechanics in relation to the violation of local causality was one of the main discussed aspect. In 1935 Einstein Podolsky and Rosen published the famous article of the EPR paradox \cite{24}. The thought experiment generated a great deal of interest in the following years. Their notion of a "complete description" was later formalized by the suggestion of hidden variables that determine the statistics of measurement results, but to which an observer does not have access.
3.1. Why PR-Boxes?

In 1964 John Bell proved that some predictions of quantum mechanics cannot be reproduced by any theory of local physical variables [3]. Although Bell worked within non-relativistic quantum theory, the definition of local variable is relativistic: a local variable can be influenced only by events in its backward light cone, not by events outside, and can influence events in its forward light cone only. Quantum mechanics, which does not allow us to transmit signals faster than light (super-luminal signalling), preserves relativistic causality. But quantum mechanics does not always allow us to consider distant systems as separate, as Einstein assumed.

Now quantum non-locality has been experimentally verified under different physical assumptions. Any physical theory that aims at superseding or replacing quantum theory should account for such experiments and therefore must also be non-local in this sense; quantum non-locality is a property of the universe that is independent of our description of nature.

So quantum non-locality is an essential feature of quantum theory but it often appear in a negative light. In 1994 Popescu and Rohrlich published a work [7] where they proposed to show quantum non-locality in a more positive light. They investigated the inversion of the logical approach to quantum mechanics, considering quantum non-locality as an axiom instead of as a theorem and wondering what non-locality together with relativistic causality would imply. They found that quantum theory is only one of a class of non-local theories consistent with causality, and not even the most non-local. In fact, in a certain sense, non-locality can be quantified.

In 1969, John Clauser, Michael Horne, Abner Shimony, and Richard Holt reformulated Bell’s inequality in a manner that best suites experimental testing, the homonymous CHSH inequality [5]. CHSH inequality, restricted to any classical theory, states that a particular algebraic combination of correlations lies between -2 and 2. This bound is obviously violated in quantum mechanics, where in fact the CHSH inequality allows a maximum value given by Cirel’son’s theorem as $2\sqrt{2}$ [4]. However, Popescu and Rohrlich wrote down a set of correlations that return a value of 4 for the CHSH expression, the maximum value algebraically possible, and that yet are non(super-luminal)-signalling. A question that now rises spontaneously is why does quantum theory not allow these strongly non-local correlations.

In the hope of making further progress with this question these correlations have been investigated in the context of a theory with well-defined dynamics. Abstractly this scenario may be described by introducing two observers that have access to a black box. Each observer selects an input from a range of possibilities and obtains an output. The box determines a joint probability for each output pair given
3.2 PR-Boxes as OPT

Each input pair. It is clear that a quantum state provides a particular example of such a box, with input corresponding to measurement choice and output to measurement outcome. More generally boxes can be divided into different types. Some will allow the observers to signal to one another via their choice of input, and correspond to two-way classical channels, as introduced by Shannon. Others will not allow signalling - it is well known, for example, that any box corresponding to an entangled quantum state will not. This is necessary for compatibility between quantum mechanics and special relativity. Among the non-signalling boxes, some will violate a Bell-type inequality, and we refer to any such a box as non-local. As we have described above, in terms of our boxes, there are some boxes that are non-signalling but are more non-local than any box allowed by quantum theory.

3.2 PR-Boxes as OPT

The model describes \( N \) correlated boxes (in the original paper it was \( N = 2 \)) in a causal context. Each box is represented by the same elementary system \( A \), with the \( N \) correlated boxes represented by the composite system \( A \otimes N \). We will consider the simplest situation where each box has both input and output as binary variables. On each elementary system \( A_i, i = 1, \ldots, N \) only two atomic binary observation tests are allowed, say \( B^{(0)} := \{ b^{(0)}, e_A - b^{(0)} \} \) or \( B^{(1)} := \{ b^{(1)}, e_A - b^{(1)} \} \), with \( b^{(0)}, b^{(1)} \in \text{Eff}(A) \) atomic effects, and \( e_A \) the deterministic effect of system \( A \). Notice that since the model is causal, the deterministic effect \( e_A \) is unique and it is \( e_A \otimes N = \otimes N e_A \).

The probabilistic model is typically presented in terms of the probability function \( P : (b_1, b_2, \ldots, b_N | B_1, B_2, \ldots B_N) \mapsto [0, 1] \), with \( B_i \in \{ B^{(0)}, B^{(1)} \} \) and \( b_i \in \{ b^{(0)}, b^{(1)} \} \), for every \( i = 1, \ldots, N \), which returns the probability of the outcomes \( b_1, b_2, \ldots, b_N \) given the observation tests \( B_1, B_2, \ldots B_N \). The constraint imposed on the function \( P \) is no-signalling, i.e.

\[
\sum_{b_k=b^{(0)},e-b^{(0)}} P(b_1, \ldots, b_N|B_1, \ldots B_N) = \]

\[
= P(b_1, \ldots, b_{k-1}, b_{k+1}, \ldots, b_N|B_1, \ldots B_{k-1}, B_{k+1}, \ldots, B_N). \tag{3.1}
\]

3.2.1 States, effects and transformations

We will start our analysis with the boxes of the elementary system \( A \) that are necessary local. They are described in terms of the following probability distribution

\[
p_{\alpha\beta}(a|x) = \begin{cases} 
1 & a = \alpha x \oplus \beta \\
0 & \text{otherwise}
\end{cases} \tag{3.2}
\]
with $\alpha, \beta = 0, 1$. This four probability distributions correspond to the four pure states of system $A$.

In fact the elementary system $A$ has dimension $\dim(A) = 3$, namely its states are described by vectors in $\mathbb{R}^3$ ($\text{St}_R(A) = \text{Eff}_R(A) = \mathbb{R}^3$). The four pure normalized states (3.2) can be represented by the following vectors

$$
\begin{align*}
\omega_0 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \omega_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \omega_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \omega_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},
\end{align*}
$$

where the correspondence between $\omega_n$ and the probability rule $p_{\alpha\beta}$ is given by $\alpha\beta = (\text{binary form of } n)$.

The convex set of states normalized is then represented by a square (see the square in the plane $z = 1$ in Fig. 3.1).

The set of effects of system $A$ is defined as the set of vectors $b$ such that $0 \leq \text{Tr}[b^T \omega] \leq 1$ for every state $\omega \in \text{St}(A)$, with $p_{y|x} = \text{Tr}[b(y)^T \omega_x]$ the rule providing the probability associated to an effect $b(y)$ on a state $\omega_x$. This leads to the truncated cone of effects in Fig. 3.1 with extremal points given by

$$
\begin{align*}
b^{(0)} &= \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad b^{(1)} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad b^{(3)} = \begin{pmatrix} 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}.
\end{align*}
$$

The deterministic effect (the effect $e_A$ such that $\text{Tr}[e_A^T \omega] = 1$ for every state $\omega \in \text{St}(A)$) is the vector $e_A = (0, 0, 1)^T$ (when no ambiguity arises we will simply denote the deterministic effect with $e$).
Our analysis now proceeds towards the composite system $A \otimes A$. In this case the probability function $p(a, b|x, y)$ form a table with $2^4$ entries, although these are not all independent due to the constraints of Eq. (3.1). The dimension of the set of boxes is found by subtracting the number of independent constraints from $2^4$, and turns out to be 8 (we notice that $8 = \dim \text{St}_1(A \otimes A)$, while $\dim \text{St}_R(A \otimes A) = 9$, so due to Theorem 2 it satisfies local discriminability). In this case we will have no more a square like in Fig. 3.1 but a polytope with 24 vertices. The vertices will be boxes that satisfy all of the constraints and saturates a sufficient number of the positivity constraints to be uniquely determined. These 24 bilocal pure states may be divided into two classes.

The local boxes, given by the following 16 probability distributions

$$p(a, b|x, y) = \begin{cases} 
1 & a = \alpha x \oplus \beta \\
1 & b = \gamma y \oplus \delta \\
0 & \text{otherwise}
\end{cases} \quad (3.4)$$

with $\alpha, \beta, \gamma, \delta = 0, 1$, and the non-local boxes, given by the 8 probability distributions

$$p_{a\beta\gamma}(a, b|x, y) = \begin{cases} 
1/2 & a \oplus b = xy \oplus \alpha x \oplus \beta y \oplus \gamma \\
0 & \text{otherwise}
\end{cases} \quad (3.5)$$

with $\alpha, \beta, \gamma = 0, 1$.

For convenience we can represent states and effects as $3 \times 3$ real matrices rather than as vectors in $\mathbb{R}^9$. The 16 local state of the bipartite system are nothing else that the factorized pure states

$$\Omega_{4i+j} := \omega_i \otimes \omega_j^T, \quad (3.6)$$

where $i, j \in \{0, 1, 2, 3\}$, and the 8 non-local states (playing the role of entangled states) are represented by the following matrices

$$\begin{align*}
\Omega_{16} &:= \frac{1}{2} \begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{17} &:= \frac{1}{2} \begin{pmatrix} -1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{18} &:= \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{19} &:= \frac{1}{2} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{20} &:= \frac{1}{2} \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{21} &:= \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{22} &:= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \\
\Omega_{23} &:= \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.
\end{align*} \quad (3.7)$$
From a generic probability rule $p_{\alpha\beta\gamma}$ we can identify one of the above matrix $\Omega_n$ by the equation: $n = 15 + [(3 + 3\alpha + 4\beta + 6\gamma) \mod 8]$.

Again, the probability associated to a bipartite effect $B_y$ on the state $\Omega_x$ is given by $p_{y|x} = \text{Tr}[B_y^T \Omega_x]$. Accordingly, the set of bipartite effects is easily derived via the consistency condition $\text{Tr}[B_j^T \Omega_i] \geq 0$ for every $j \in [0, 23]$. It follows that the only admissible extremal effects are the 16 factorized matrices

$$B_{4i+j} := b^{(i)} \otimes b^{(j)T}.$$  \hfill (3.8)

This is a relevant feature of PR-boxes model, whose strong correlation encapsulated in the eight non-local states $\Omega_x$, $x \in [16, 23]$ are incompatible with any in principle admissible non-factorized measurements. This feature has been firstly noticed in Ref. [25] and later in Ref. [26] where all possible theories compatible with the squit local system have been classified (among these theories is the dual version of PR-boxes that only have factorized states but eight non-local effects). Finally, the deterministic effect for the bipartite system is $e \otimes e^T$.

Analogously one defines the convex set of states and the convex set of effects for the arbitrary $N$-partite system $A \otimes N$. However, while for effect is nothing but a trivial generalization, for the states the discussion is not so straight and some interesting features arise. We will dedicate a following Section to discuss some of the most immediate aspect about the $N$-partite system, we will start our discussion from the tripartite boxes.

We can now turn our attention to the transformations of the theory. We focus here on the reversible transformations which are of interest for the present paper results. The set $\mathcal{U}(A)$ of reversible transformations of the system $A$ coincides with the finite group of symmetries of the square (the dihedral group of order eight $D_8$ containing four rotations and four reflections). In the chosen representation we have

$$\mathcal{U}(A) = \{ U^s_k : k = 0, \ldots, 3, s = \pm \}$$

$$U^s_k = \begin{pmatrix} \cos \frac{\pi k}{2} & -s \sin \frac{\pi k}{2} & 0 \\ s \sin \frac{\pi k}{2} & \cos \frac{\pi k}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \hfill (3.9)$$

The matrices $U^+_k$ and $U^-_k$ representing the four rotations and the four reflections of the square respectively. When we apply them to the four pure states of Eq. (3.3) we have

$$\omega_{j+k} = U^+_k \omega_j, \quad \omega_k = U^-_k \omega_{j+k}, \hfill (3.10)$$
for \( j \in [0, 3] \) and where the sum is \( \text{mod} \ 3 \), i.e. if \( j = 2 \) and \( k = 2 \) then \( j + k = 0 \).

We notice that these transformations are atomic. In Ref. [23] all atomic transformations of the squit system have been classified.

The set of reversible transformations \( \mathcal{U}(A \otimes A) \) of the composite system \( A \otimes A \) (which has been derived in Ref. [27]) is

\[
\mathcal{U}(A \otimes A) = \{ W^i(U^{s_1}_j \otimes U^{s_2}_k) \} \quad i = 0, 1, \ 0 \leq j, k \leq 3, \ s_1, s_2 = \pm,
\]

(3.11)

with \( W \) the swap map, namely the map that exchanges the two subsystems. This means that any reversible map corresponds to the tensor product of single system reversible transformations with, possibly, the application of the swap. As noticed in Ref. [27] reversible transformations cannot create entanglement. This result has been extended in Ref. [28] to the composition of an arbitrary number \( N \) of systems \( A \otimes^N \), showing that also in that case the set of reversible multipartite transformations is generated by local reversible operations and permutations of systems.

We notice that in the PR-boxes theory any non-local bipartite pure state can be reversibly mapped to any other non-local bipartite pure state. Moreover, this mapping can be done via the local application of a single system reversible map. For example, starting from the state \( \Omega_{16} \), one has

\[
\Omega_{16+k} = (U^+_k \otimes I)\Omega_{16}, \quad \Omega_{23-k} = (U^-_k \otimes I)\Omega_{16}.
\]

(3.12)

We finally remark that the set of reversible transformations of system \( A \), \( \mathcal{U}(A) \), in terms of description by the probability distributions in Eqs. (3.2), (3.4), (3.5) correspond to a local relabelling defined by the operations

\[
x \rightarrow x \oplus 1, \quad a \rightarrow a \oplus \alpha x \oplus \gamma, \\
y \rightarrow y \oplus 1, \quad b \rightarrow b \oplus \beta y \oplus \gamma.
\]

(3.13)

### 3.2.2 Discriminability between PR-boxes

An important question to address when dealing with the PR-boxes theory is the following: given two deterministic state of the theory is it possible to discriminate between them? We will address to this question only for the extremal points of the polytope, i.e. the pure states, in order to analyse a way to perfect discriminate between them.

For local boxes it is easy to show that there exist POVMs that are able to perfect discriminate between every two of the four possible state in Eq. (3.3) [23].

For what concern bipartite boxes, while it is trivial to find perfectly discriminable
POVMs for each pair of the 16 local boxes (since they are simply the tensor product of the 4 local boxes, so also the discriminable POVMs are just the tensor product of the perfectly discriminable POVMs of the local boxes) it is a bit more elaborate to explore the discriminability between non-local boxes.

To help in our analysis we introduce the following table regarding non-local boxes labelled by $(\alpha, \beta, \gamma)$ as described in Eq. 3.5:

| $x$ | $y$ | $a \oplus b$ |
|-----|-----|-------------|
| 0   | 0   | $\gamma$   |
| 0   | 1   | $\beta \oplus \gamma$ |
| 1   | 0   | $\alpha \oplus \gamma$ |
| 1   | 1   | $1 \oplus \alpha \oplus \beta \oplus \gamma$ |

Table 3.1: Correlations for a generic $(\alpha, \beta, \gamma)$ non-local box for all the possible given input combinations of $x$ and $y$.

It is now easy to verify that for every two different non-local boxes, i.e. for every choice of two different combinations of $(\alpha, \beta, \gamma)$: $c_1 = (\alpha_1, \beta_1, \gamma_1)$ and $c_2 = (\alpha_2, \beta_2, \gamma_2)$, there is always at least one input combination $(x, y)$, whose output relation is equal to 0 for $c_1$ and to 1 for $c_2$. If we denote with $e$ the deterministic effect of the bipartite system $A \otimes A$ and $a := b^{(3(1-x))} \otimes b^{(3(1-y))} + b^{(1+x)} \otimes b^{(1+y)}$, the POVM $\{a, e - a\}$ is able to perfectly discriminate between the two chosen non-local boxes.

Finally, the last remark regards the discrimination between one local and one non-local bipartite box. We can help ourselves with the two TABLEs of Appendix A that follows the notation of Eqs. 3.4, 3.5. For every pair of bipartite boxes (one local and one non-local) there is always a combination of $(x, y)$ such that in one case the outcome relation $a \oplus b$ is equal to 0 for one box and 1 for the other. So we can construct a perfectly discriminating POVM following the same strategy of above.

### 3.2.3 Purification in the PR-box (bipartite restriction) model

Until now nothing has been said about purification in the PR-box model. The following discussion represent an important result of the theory, valid in the general context of $N$-partite boxes (with $N$ finite), but we will operate in the significant restriction of admitting no more than bipartite correlated boxes. For the rest of the subsection when we will refer to PR-box model/theory we will mean the model under this limitation.
Let begin with some general definitions and preliminary considerations.

**Definition 10** (Transitivity). The ability to transform any pure state into any other by means of reversible transformations will be called transitivity, meaning that the action of the set of reversible transformations is transitive on the set of pure states.

Based on this definition, PR-box model clearly enjoys this property. In fact the transformations of Eq. (3.9) is transitive on the set of pure states of Eq. (3.3). Among the numerous consequences that transitivity implies one will be of our interest, the uniqueness of the maximally mixed state. So it is in order to properly define what a maximally mixed state is.

**Definition 11** (Maximally mixed state). If a state is invariant under the action of every reversible transformation, then it is a maximally mixed state.

We will not report here the demonstration of the uniqueness of the maximally mixed state from transitivity, for it we refer to Ref. [11].

From the transformations of Eq. (3.9) and the vector representation of the pure state in Eq. (3.3) it is not difficult to write down the maximally mixed state:

\[
\mu = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.14}
\]

We are now in position to state the main result of this Section.

**Theorem 7.** Given a system \( \otimes^N A \), the maximally mixed state \( \mu^{\otimes N} \in \text{St}(\otimes^N A) \) of Eq. (3.14) is the unique internal state that is purificable and its purification is unique up to a reversible transformation on the purifying systems.

**Proof.** The proof is divided in two part. In the first one we will prove the thesis for \( N = 1 \), then in the second part it is extended to an arbitrary number of systems.

Given the system \( A_1 \) we will consider a system \( A_2 \) as the purifying system. The pure states of \( A_1 \otimes A_2 \) are the 24 pure states \( \Omega_i \), for \( i = 0, \ldots, 23 \) of Eqs. (3.6), (3.7). We know that 16 of them (namely the local ones, that are expressed in Eq. (3.6)) are separable states. So if we consider the marginal state obtained by applying the deterministic effect \( e_{A_2} \) on the purifying system \( A_2 \) on these separable states we get one of the local 4 state represented in Eq. (3.3), that are pure. However, if we repeat
the same procedure on the 8 non-local pure states (represented in Eq. (3.7)) we get
the same state for all of them: the maximally mixed state $\mu \in \text{St}(A_1)$. So the unique
internal state that can be purificated is the maximally mixed state and since the 8
non-local bipartite pure states are all mapped to any other non-local bipartite pure
state by the application of a single system reversible map, as shown in Eq. (3.12), the
purification is unique up to a reversible transformation on the purifying system.

We now deal with the $N$-partite scenario. The more general state $\Psi \in \text{St}(\otimes^N A)$ has
the form:

\[
\Psi = \begin{array}{c}
\omega_1 \\
\vdots \\
\omega_{M/2} \\
\omega_{M/2+1} \\
\vdots \\
\omega_{N-M/2}
\end{array}
\begin{array}{c}
A_1 \\
A_2 \\
A_3 \\
A_4 \\
A_5 \\
A_6
\end{array}
\]

In order to be purificable, every monopartite state has to be pure or purificable
and every bipartite state to be pure (since we are admitting no more than bipartite
correlations, a bipartite state can be purificable only if it is pure). But if one of
the states $\omega_j$ for $j = 1, \ldots, N - M/2$ is pure then $\Psi$ can not be internal. Since
the only purificable internal state is $\mu$, we have that the only purificable internal $N$-
partite state is $\mu \otimes \ldots \otimes \mu$ (thanks to local discriminability the parallel composition
of internal states is still an internal state). Furthermore, for what we have seen before,
the purification of $\mu \otimes \ldots \otimes \mu$ is unique up to a reversible transformation. This
transformation is nothing else that the parallel composition of the maps $(U_k \otimes \mathcal{I})$
where $U_k$ are the maps of Eq. (3.12).

Remark: in the previous proof we noticed that the maximally mixed state of a system
$\otimes^N A$ is not the unique mixed state that is purificable. In fact, the more general state
3.2. PR-Boxes as OPT

\( \Phi \in \text{St}(\otimes^N A) \) that can be purificated is of the form: \( \Phi = \phi_1 \otimes \ldots \otimes \phi_N \) where

\[
\phi_i = \begin{cases} 
\mu & \text{for } j \in [0, 3] \\
\omega_j & \text{for } j \in [1, 15] \\
\Omega_k & \text{for } j \in [16, 23]
\end{cases} \quad \text{for } i = 1, \ldots, N,
\]

and its purification is still unique up to reversible transformations on the purifying systems.

### 3.2.4 N-partite PR-boxes

In the literature of PR-box theory a thorough and systematic study on \( k \)-partite correlated boxes, with \( k \geq 3 \), has never been made. This represents an important absence within the model since it prevents the theory to be complete. In this Section we show some important consequences that emerge by just considering tripartite correlated boxes with some speculations about the complete \( N \)-partite model.

In our analysis of PR-box model integrated with tripartite boxes we make use of the classification that has been made in Ref. [29]. In the article of Pironio et al., the no-signaling polytope is found to have 53856 extremal points, belonging to 46 inequivalent classes. The term inequivalent means that there not exist reversible local transformations that allow to move from a representative of one class to one of another class, while, inside the same class, all the extremal points are connected by local relabelling, see Eqs. (3.13) (since we are dealing with three parties boxes, also permutation of the parties is a local relabelling, i.e. \( x \rightarrow y, y \rightarrow z, \) and \( z \rightarrow x \) and so on for every possible permutation).

Firstly, it is no more granted that the maximally mixed state is the unique internal state purificable in the theory. In fact it could happen that between the 53856 pure tripartite states, there will be one whose marginal state is an internal state different from the maximally mixed one. This leads to think that increasing the number of correlated systems that we are considering the number of internal states that are purificable will also increase. Even if there are not academic works in this matter, it is a very likely and reasonable possibility and we address to future studies to investigate in this direction.

Secondly, even if the only internal purificable state would still be the maximally mixed one, the purification is no more unique. To see this it suffices to consider two
3.2. PR-Boxes as OPT

states \( \Psi_1, \Psi_2 \in \text{St}(A^\otimes 3) \) that have the following form:

\[
\Psi_1 = \begin{array}{c}
\Omega \\
\omega
\end{array} \quad \text{and} \quad \Psi_2 = \begin{array}{c}
\Phi^{(44)} \\\nA
\end{array},
\]

where \( \omega \in \text{St}(A) \) and \( \Omega \in \text{St}(A^\otimes 2) \) are two pure states and \( \Phi^{(44)} \) is a pure state, representative of the 44\(^{th} \) class described in Ref. [29]. If \( \Omega \) is one of the non-local bipartite states then they have the same marginal, namely

\[
\begin{array}{c}
\Omega \\
\omega
\end{array} = \begin{array}{c}
\mu \\
e
\end{array} = \begin{array}{c}
\Phi^{(44)} \\\nA
\end{array},
\]

where \( \mu \) is the maximally mixed state of Eq. (3.14) and the second equality derives straightforward once the probability rule of the 44\(^{th} \) class is written explicitly, as we will see in Eq. (4.4). So we found that \( \Psi_1 \) and \( \Psi_2 \) are two purification of the same state but since every reversible transformation \( U \in \mathcal{U}(A^\otimes 2) \) is the composition of local reversible maps, "correlation" can not be created and so

\[
\Psi_1 \neq \Psi_2 \quad \forall \, U \in \mathcal{U}(A^\otimes 2).
\]

Furthermore, restricting our attention to the pure bipartite states, we have noticed that when we pick up two of these states that have the same marginal, than there is always a local reversible map from one to the other and vice-versa. This is the case for the 8 non-local bipartite boxes, that are all connected by local transformations in \( \mathcal{U}(A) \) of Eq. (3.12). This mechanism will turn out to be exactly the one responsible for the impossibility of perfectly secure bit commitment, as we will see in detail in the next Chapter. Since in the tripartite scenario not all the tripartite non-local boxes are connected by local transformation (we remind that local relabelling is not enough to change from a class to another), it is reasonable to think that perfect bit commitment will be possible, or at least a completely new way of cheating has to be thought. With this purpose we will propose a scheme of bit commitment as the conclusive Section of the next Chapter.
In the past years numerous protocols have been proposed to realize bit commitment using PR-boxes. However, as outlined by A.J. Short, N. Gisin, and S. Popescu in Ref. [9], “it is surprising that the possibility that non-local correlations which are stronger than those in quantum mechanics could be used for bit commitment, because it is the very existence of non-local correlations which in quantum mechanics prevents bit commitment”. In that article they particularly referred to the protocol prosed by S. Wolf and J. Wullschleger [8] and showed that it was erroneous by argument of causality. After that also Buhrman et al. [10] proposed a bit commitment protocol in PR-box theory that was claimed to be unconditionally secure and where the counter-proof of Short, Gisin and Popescu did not work anymore.

From a OPT point of view, when dealing with PR-boxes, some issues arise since they still not have a complete and closed theory. In fact, $k$-partite boxes with $k \geq 3$ have been studied only roughly and a coherent and comprehensive theory has not been proposed. As we have seen in Section 3.2.3 simplistic generalizations are not adequate since admitting more than bipartite correlations alters significantly the theory. Nevertheless, in the literature not only PR-box model is generally considered admitting no more than bipartite states, but also local transformations (that are admissible in the theory) are ignored.

In this final Chapter we propose a proof of impossibility of perfectly secure bit commitment in PR-boxes (even if under two important limitations: pure input states and bipartite boxes, the proof includes almost all the protocols proposed in litera-
4.1 No-Perfectly Secure Bit Commitment

In this Section we provide an explicit proof of the impossibility of perfectly secure bit commitment in PR-box theory. However, two important limitation will be adopted. Even if we will consider arbitrary $N$-partite systems (with $N$ finite), we will not admit more than bipartite correlations (it will be taken for granted for the rest of the Section). Furthermore $\Psi_0$ and $\Psi_1$ will always be selected between the pure states. This is due to the fact that discrimination has never been studied in PR-box theory and the only results we rely on are those of Section 3.2.2 that refers to pure states. If it would turn out that the strategy exposed in Section 3.2.2 is the unique one to grant perfect discriminability between two arbitrary bipartite states then the protocol could be easily extended also to $\Psi_0$, $\Psi_1$ as arbitrary mixed states.

Actually we will prove our theorem in two different way. In the first proof, we will state the property of perfect bit commitment and we will show that inconsistencies arise. In the second, that we will call alternative proof, we will show that, with some shrewdness, the proof of Section 2.3 can be used also in this context.
4.1. No-Perfectly Secure Bit Commitment

4.1.1 First Proof

In this Section we make use of the definition of the protocol given in Section 2.2 and nomenclature of states and transformations given in Section 3.2. Before the main theorem a preliminary lemma is in order.

**Lemma 6.** If a bit commitment protocol is correct with probability one then the two input states $\Psi_0, \Psi_1 \in \text{St}(A^\otimes N B^\otimes N)$ shared by Alice ($A^\otimes N$) and Bob ($B^\otimes N$) are of the form

$$\Psi_i = \begin{cases} \begin{array}{c} \Omega_{k_1(i)} \\ \vdots \\ \Omega_{k_M(i)} \end{array} \\ \begin{array}{c} A_1 \\ \vdots \\ A_M \\ B_1 \\ \vdots \\ B_M \end{array} \end{cases}$$

for $i = 0, 1$ (4.1)

where $k_j(i) \in [16, 23]$ for $j = 1, \ldots, M$ and $k_j(i) \in [0, 3]$ for $j = M + 1, \ldots, 2N - M$.

**Proof.** The thesis follows immediately from the two assumptions we made: no more than bipartite correlated boxes and pure input states. In fact, for every two pure states a perfectly discriminating procedure always exists, as analysed in Section 3.2.2. For discriminate between two parallel compositions of pure states the parallel composition of the discriminating POVMs for each pair of states is sufficient.

In the previous lemma it would be possible that Alice and Bob have in control also non-factorized bipartite states, i.e. $\Omega_{k_{M+1}} \in \text{St}(A_{M+1}A_{M+3})$ for $k_{M+1} \in [16, 23]$ instead of $\omega_{k_{M+1}} \otimes \omega_{k_{M+3}} \in \text{St}(A_{M+1}A_{M+3})$ for $k_{M+1}, k_{M+3} \in [0, 3]$, however this does not carry any modification in the proof and so, to not make the notation even more troublesome, we will refer to Eq. (4.1) as the more general input states.

**Theorem 8.** Perfect bit commitment is impossible in PR-box theory.
Proof. If a bit commitment is perfect it means that it should be correct with probability one, perfectly concealing and perfectly binding.

If it is correct with probability one then, by the previous Lemma, the input states $\Psi_0, \Psi_1 \in \text{St}(A^\otimes N B^\otimes N)$ must have the form of Eq. (4.1).

If it is also perfectly concealing, then we have to impose the condition of Eq. (2.4), i.e.

\[ (e_{A^\otimes N}|\Psi_0)_{A^\otimes N B^\otimes N} = (e_{A^\otimes N}|\Psi_1)_{A^\otimes N B^\otimes N} \]

So, for the systems from $B_{M+1}$ to $B_N$ we have that $\omega_{k_{M+2s}(0)} = \omega_{k_{M+2s}(1)}$ for $s = 1, \ldots, N - M$. For the systems from $B_{M+1}$ to $B_N$ we can not make any further deductions since these 8 pure states have all the same marginal.

In any case if the protocol is correct with probability one and perfectly concealing it can not be perfectly binding.

In fact, for the non-local bipartite states there will be one of the local transformations of Eq. (3.9) such that

\[ \Omega_{k_i(0)} = \Omega_{k_i(1)} \]

as expressed by Eq. (3.12), where $U_{k_j} \in \mathcal{U}(A)$. Furthermore for all the other states, namely for the local states $\omega_{k_{M+2s+1}(i)}$ with $s = 0, \ldots, N - M - 1, i = 0, 1$, it is immediate that there will be local transformations $U_{k_{M+2s+1}} \in \mathcal{U}(A), s = 0, \ldots, N -
4.1. No-Perfectly Secure Bit Commitment

\[ M - 1 \] chosen from the 4 of Eq. (3.9) that will permit to Alice to switch unnoticed by Bob from \( \Psi_0 \) to \( \Psi_1 \) and vice-versa.

\[ \square \]

4.1.2 Alternative Proof

In Section 2.3 we pointed out the three sufficient conditions that a theory has to satisfy in order to ensure the impossibility of perfectly secure bit commitment: causality, atomicity of composition and the one required in Axiom 7.

Clearly PR-box theory is manifestly causal, since Eq. (3.1) imposes exactly the no-signalling constraint. Furthermore, also atomicity of composition is fulfilled by PR-box theory. The parallel composition of atomic operations is still atomic due to local discriminability, see Ref. [31]. To verify that also the sequential composition of atomic operation is still atomic it only need to consider all the atomic operations in the theory, see Ref. [23], and straightforwardly compute their composition.

For what concern Axiom 7 some considerations are in order. We required the existence of at least one dynamically faithful pure state, \( \Psi^{(A)} \), and the existence of a purification for every state \( \mathcal{R} \) that has the same marginal of \( \Psi^{(A)} \). Actually the latter assumption is excessive. If we look at Theorem 3, where this assumption needs to work, it would be enough that every state \( | \mathcal{R} \rangle_{BA} \) that is obtained from \( | \Psi^{(A)} \rangle_{AA} \) by a local channel \( \mathcal{C} \), i.e. \( | \mathcal{R} \rangle_{BA} := (\mathcal{C} \otimes \mathcal{I}_A) | \Psi^{(A)} \rangle_{AA} \), is purifiable. In fact, if \( \mathcal{R} \) is so defined, it certainly has the same marginal of \( \Psi^{(A)} \):

\[
\begin{array}{c}
\begin{array}{c}
\mathcal{R} \\
A
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
B \\
C
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
A \\
C
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
B \\
C
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
A \\
C
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
A \\
C
\end{array}
\end{array}
\end{array}
\]

Now if we limit to consider \( \mathcal{C} \) as a local reversible channel (since in the theory the reversible transformations are atomic), being \( \Psi^{(A)} \) pure by hypothesis, due to atomicity of composition also \( \mathcal{R} \) is pure and hence trivially purifiable. So, under this further constraint, we only need to check the existence of a dynamically faithful pure state within the theory.

We can help ourselves by the fact that the purification of an internal state is dynamically faithful (and obviously pure). This result was published in Ref. [2] for theories that satisfy local discriminability and purification but it can easily extended to PR-box theory given the existence of at least one internal state that is purifiable: \( \mu \).

Now the answer is immediate, all the non-local bipartite extremal point of the 8-dimensional polytope are dynamically faithful pure states. At this point it is easy
to see that we can just choose one of the pure states in Eq. (3.7), as the dynamically faithful pure state in Axiom 7.

Finally, it is straightforward to verify that the purification is unique up to a local reversible transformation on the purifying system, see Eq. (3.12).

The further limitation that we imposed, on the atomicity of $C$, practically requires that for every commitment protocol the two encodings $A_0, A_1 \in \text{Transf}(A_1...A_N \rightarrow B_1...B_NF_N)$ have the marginal atomic: $(e|F_N, A_0 = (e|F_N, A_1 := C$. But from Eq. (3.11) we know that every element of the set of reversible transformation of the composite system $A \otimes A$ is nothing else than the tensor product of local reversible transformations (with eventually the swap map). So if we limit to consider only atomic local transformations the initial new constraint is satisfied.

Finally, thanks to local discriminability, given two pure faithful states $\Psi^{(A)} \in \text{St}(A\bar{A})$ and $\Psi^{(B)} \in \text{St}(B\bar{B})$ for system $A$ and $B$, respectively, then also $\Psi^{(A)} \otimes \Psi^{(B)} \in \text{St}(A\bar{A}B\bar{B})$ is a pure dynamically faithful state for the compound system $AB$ (a rigorous proof can be found in Ref. [2]) and so the previous discussion is easily generalizable to any $N$-partite system.

In conclusion, the proof of impossibility of perfectly secure bit commitment of Section 2.3 can be easily extended to include PR-box theory (limited to no more than bipartite correlations and only reversible transformations).

### 4.2 Unconditionally Secure Bit Commitment

As outlined before, in Ref. [10] Buhrman et al. proposed a bit commitment protocol that was claimed to be unconditionally secure. They would like to show that super-strong non-local correlations in the form of non-local boxes enable to solve cryptographic problems otherwise known to be impossible. In particular, their result would imply that the no-signaling principle and secure computation are compatible in principle.

However, we now prove how it would be possible for Alice to perfectly cheat, i.e. with null probability of being detected by Bob, making use of the local reversible atomic transformations of Eq. (3.12).

In our analysis, we begin from dealing with only one bipartite PR-box and we suppose that Alice’s input is the committed bit (even if this is not a well defined bit commitment protocol it is as well an instructive example in order to simplify the following analysis).
4.2. Unconditionally Secure Bit Commitment

Alice and Bob share the non-local bipartite PR-box $\Omega_{18} \in \text{St}_\mathbb{R}(AB)$ (to which corresponds the probability rule $p_{000}$) but they have access only to system A and B, respectively. We can summarise the protocol as follows.

**COMMIT:**
- Alice selects her committed bit $x$, inputs it and obtain output bit $a$;
- Bob inputs a random bit $y$ and obtains output bit $b$.

**REVEAL:**
- Alice sends $x$ and $a$ to Bob;
- Bob checks to see if $a \oplus b = xy$. If this relation is true, Bob accepts $x$ as the revealed bit, otherwise he knows that Alice has cheated and rejects Alice’s revelation.

It is easy to see that Alice has a probability to cheat successfully equal to $\frac{1}{2}$.

With the help of local transformation Alice can make the probability of successful cheating equal to 1. In fact, it is sufficient to find $(\alpha, \beta, \gamma)$ of Eq. (3.5) such that, for given $x$ and $y$ and for the output couple $(a, b)$ (i.e. such that $p_{\alpha\beta\gamma}(a, b|x, y) \neq 0$) exists a generic function $f : \{0, 1\} \rightarrow \{0, 1\}$ such that, given $a' = f(a)$ and $x' = x \oplus 1$, $p_{000}(a', b|x', y) \neq 0$.

Mathematically, to find suitable $(\alpha, \beta, \gamma)$ it is sufficient to resolve

$$\begin{cases} a \oplus b = xy \oplus \alpha x \oplus \beta y \oplus \gamma, \\ a' \oplus b = x'y. \end{cases} \quad (4.3)$$

We find that $\beta = 1$ and $f(a) = a \oplus \alpha x \oplus \gamma$, for every choice of $\alpha$ and $\gamma$.

In conclusion, if Alice performs a local transformation (given by Eq. (3.9)) such that $p_{000} \rightarrow p_{\alpha 1\gamma}$ and then inputs her bit $x$ and gets output $a$, she can reveal $x'$ and $a' = a \oplus \alpha x \oplus \gamma$ to Bob, who will accept with probability 1.

We can now consider the unconditionally secure bit-commitment protocol proposed in Ref. [10] where $2n+1$ non-local bipartite PR-boxes in the state $\Omega_{18} \in \text{St}(AB)$ are shared between Alice and Bob.

The authors found that Alice probability of successfully cheating is at maximum equal to $1/2$ but can be asymptotically reduced if the protocol is repeated $k$ times. However, using local transformation Alice can cheat without being detected with probability 1.

We refer to the original article about the commitment protocol and we outline only
the "cheating procedure".

**COMMIT:**

- Alice wants to commit to bit $c$ but to send to Bob bit $c' = c \oplus 1$. So she chooses $x \in \{0, 1\}^{2n+1}$ by choosing the first $2n$ bits such that $|x_1 ... x_{2n}|_{11}$ is even where $x'_i = x_i \oplus 1$ for $i = 1, 2, ..., 2n$ (given a string of even length $x$, $|x|_{11}$ is the number of substring "11" in $x$ starting at an odd position) and then choosing $x_{2n+1} = c$ (analogously she can choose $x \in \{0, 1\}^{2n+1}$ such that $|x'_1 ... x'_{2n}|_{11}$ is odd and $x_{2n+1} = c \oplus 1$);

- Alice, for each of the $2n+1$ shared boxes, apply a local transformation changing the probability law from $p_{000} \rightarrow p_{\alpha \gamma}$, then se puts the bits $x_1, x_2, ..., x_{2n+1}$ into the boxes 1, 2, ..., $2n+1$. Let $a_1, a_2, ..., a_{2n+1}$ be Alice's output bits from the boxes;

- Alice computes the parity of all the "cheated" output bits $A' = \oplus_{i=0}^{2n+1} a'_i$ and send $A'$ to Bob, where $a'_i = a_i \oplus \alpha x'_i \oplus \gamma$ for $i = 1, 2, ..., 2n + 1$;

- Bob randomly chooses a string $y \in \mathbb{R} \{0, 1\}^{2n+1}$ and puts the bits $y_1, y_2, ..., y_{2n+1}$ into his boxes. We call the output bits from his boxes $b_1, b_2, ..., b_{2n+1}$.

Then the **REVEAL** phase:

- Alice sends $c'$, her string $x'$ (where $x'_i = x_i \oplus 1$ for $i = 1, 2, ..., 2n + 1$) and all her $2n + 1$ "cheated" outputs bits (i.e. $a'_i$) to Bob;

- Bob checks if Alice’s data is consistent: $\forall i \in \{0, 1\}^{2n+1}$, $x'_i \cdot y_i = a'_i \oplus b_i$ and $|x'_1 ... x'_{2n}|_{11} + |x'_2 ... x'_{2n+1}|_{11} + c'$ is even. Since he finds no error, he accepts $c'$ as the committed bit.

In fact, according to Eq. (4.3), every couple $(x'_i, a'_i)$ for $i = 1, 2, ..., 2n+1$ sent by Alice to Bob satisfies $x'_i \cdot y = a'_i \oplus b$ and by the proposed choices of $x \in \{0, 1\}^{2n+1}$ also the parity constraints are secured.

### 4.3 Bit Commitment in Tripartite Scenario

We have stressed that the proof of impossibility of perfectly secure bit commitment in PR-box theory was limited by considering no more than bipartite correlated boxes in $N$-partite systems. In fact, if we take in consideration just only tripartite boxes, the scenario changes considerably.

In this Section we highlight that a scheme of bit commitment protocol like the one...
in Ref. [10] that would make use of tripartite non-local boxes would not subject to the cheating by local reversible transformations and it would represent a possible perfectly (or at least unconditionally) secure bit commitment protocol.

In Section 3.2.4 we used the classification done in Ref. [29] where the tripartite correlated boxes were divided in 46 non equivalent classes, i.e. 46 classes whose states are not connected by local reversible transformations. In order to see it directly, it is sufficient to write explicitly the probability rules for those classes. As an example we write out the representatives of the probability rule for three non-local tripartite classes (number 44, 45, and 46 in Ref. [29]):

Class 44: \[ p(a, b, c|x, y, z) = \begin{cases} 1/4 & a \oplus b \oplus c = xyz \\ 0 & \text{otherwise} \end{cases} \]

Class 45: \[ p(a, b, c|x, y, z) = \begin{cases} 1/4 & a \oplus b \oplus c = xy \oplus xz \\ 0 & \text{otherwise} \end{cases} \]

Class 46: \[ p(a, b, c|x, y, z) = \begin{cases} 1/4 & a \oplus b \oplus c = xy \oplus xz \oplus yz \\ 0 & \text{otherwise} \end{cases} \]

and we note that the local relabelling of Eq. (3.13) plus the permutations of the parties do not permit to move from every representative of one class to any one of any other class.

Now to build our bit commitment protocol we decide to encode \( b = 0, 1 \) by choosing as input state \( \Psi_0, \Psi_1 \in \text{St}(A^3) \) one representative of the 44\(^{th} \) class and one of the 45\(^{th} \), say the ones in Eq. (4.4) for the sake of simplicity.

Following the strategy in Section 3.2.2 we can construct a POVM that is able to discriminate between them, for example \( \{ a, e - a \} \) where \( a := b^{(0)} \otimes b^{(3)} \otimes b^{(0)} + b^{(0)} \otimes b^{(1)} \otimes b^{(2)} + b^{(2)} \otimes b^{(1)} \otimes b^{(0)} + b^{(2)} \otimes b^{(3)} \otimes b^{(2)} \) and \( e = e_A \otimes e_A \otimes e_A \). So the protocol is correct with probability one.

Furthermore \( \Psi_0 \) and \( \Psi_1 \) are not connected by any local reversible transformation (as pointed out before) and so the protocol is also perfectly binding. At the same time the two input states have the same marginal on Bob system, and so it is also perfectly concealing. Namely, it is not difficult to derive from Eq. (4.4) that

\[
\begin{pmatrix}
\Psi_0 \\
\Lambda \\
\Lambda \\
\Lambda \\
e
\end{pmatrix}
= \begin{pmatrix}
\mu \\
\Lambda \\
\Lambda \\
\Lambda \\
e
\end{pmatrix}
= \begin{pmatrix}
\Psi_1 \\
\Lambda \\
\Lambda \\
\Lambda \\
e
\end{pmatrix}
\]

In conclusion, by simply choosing as our input states two suitable non-local tripartite boxes we build a protocol that is correct with probability 1, perfectly concealing and
seems to be also perfectly binding (we are considering only reversible transformations). Naturally, since the theory is not complete, we would not like to fall in the same mistake of claiming our bit commitment protocol perfectly (or, when applied in Ref. [10], unconditionally) secure in PR-box theory but, less then unexpected turns in the theory, PR-box theory could really represent the first example of a theory with entanglement and bit commitment.
Conclusions

In this thesis we have formalized the bit commitment protocol in the operational language to investigate its feasibility in a more general context than quantum theory. In this way we are willing to make the first step in the understanding of the relation that exists between bit commitment and the operational axioms of quantum theory. In particular we focused on the study of BC in PR-box theory, a theory that is more non-local than the quantum one but where the purification property does not hold.

In performing this analysis we were forced to investigate new aspects of PR-box theory, since it is far away to be closed and complete. Some of the most remarkable results that we achieved are the following. Firstly, we described a strategy that grants to always find a POVM able to perfectly discriminate between any two bipartite pure states. In addition, we proved that the maximally mixed state $\mu$, is purificable. Furthermore, if the theory is limited to no more than bipartite correlated boxes, then $\mu$ is the unique internal state that is purificable and its purification is also unique up to reversible transformations on the purifying system. Then, we were able to show how simplistic generalizations to the arbitrary $N$-partite case are not appropriate. For example just admitting tripartite boxes we showed how the purification of the maximally mixed state is not more unique, and how it is not even granted that $\mu$ is still the unique internal state that is purificable.

After the presentation of the PR-box theory in Chapter 3 in Chapter 4 we presented the results about the impossibility of bit commitment in PR-boxes. In the literature of BC performed on non-local boxes, these have been rarely con-
sidered as part of a coherent theory and in fact tripartite correlated boxes or local reversible transformations (that are admissible in the theory) have always been neglected.

By simply taking in consideration the reversible transformations we proposed a proof of impossibility of perfectly secure bit commitment in a PR-box theory limited by the following constraint:

1. no more than bipartite correlated boxes admitted;

2. only reversible transformations considered;

3. only pure input states.

Even under these three important limitation our scenario is still enough general to include all the protocol proposed in literature. Furthermore we were also able to adapt the solid proof in Ref. [2] (in same context as above) to obtain an identical result for the impossibility of bit commitment in PR-boxes. In addition, even if we dealt only with perfectly secure bit commitment, we explicitly described a scheme in which Alice is able to cheat perfectly in the protocol proposed in Ref. [10], that was claimed to be unconditionally secure.

Finally we relaxed the limitations that we imposed on the theory and we proposed a protocol that seems to be perfectly secure. However, since the theory is not complete, we address future studies to investigate this and the other questions that remain unanswered in this thesis.

First of all, if the discriminating strategy could include not only pure states but all of them, then the proof of impossibility of perfectly secure BC could be extended to non-pure input states, too. Anyhow, the great unknown variable is still the integration of generic $N$-partite non-local correlated boxes in the theory. The consequences could be very surprising, for example it would even be possible that the number of internal states that are purificable asymptotically increases increasing $N$, and so that the purification principle could hold in the limit $N \to \infty$.

In conclusion we presented some results on the impossibility of perfectly secure bit commitment in PR-boxes and we precisely pointed out their limit of validity, that, even if under considerable assumptions, still have an important comparison with the literature on the subject. However, as we have outlined many times in this work, to achieve definitive results, other progresses in the fundamental aspects of the theory are absolutely necessary.
Table A.1: Outcome relations for the 16 bilocal-boxes.

| \((\alpha\beta\gamma\delta)\) | \(a \oplus b\) |
|-------------------------------|----------------|
| 0000                          | 0              |
| 0001                          | 1              |
| 0010                          | \(y\)          |
| 0011                          | \(y \oplus 1\) |
| 0100                          | 1              |
| 0101                          | 0              |
| 0110                          | \(y \oplus 1\) |
| 0111                          | \(y\)          |
| 1000                          | \(x\)          |
| 1001                          | \(x \oplus 1\) |
| 1010                          | \(x \oplus y\) |
| 1011                          | \(x \oplus y \oplus 1\) |
| 1100                          | \(x \oplus 1\) |
| 1101                          | \(x\)          |
| 1110                          | \(x \oplus y \oplus 1\) |
| 1111                          | \(x \oplus y\) |
| $(\alpha\beta\gamma)$ | $a \oplus b$ |
|------------------------|-------------|
| 000                    | $xy$        |
| 001                    | $xy \oplus 1$  |
| 010                    | $xy \oplus y$  |
| 011                    | $xy \oplus y \oplus 1$  |
| 100                    | $xy \oplus x$  |
| 101                    | $xy \oplus x \oplus 1$  |
| 110                    | $xy \oplus x \oplus y$  |
| 111                    | $xy \oplus x \oplus y \oplus 1$  |

Table A.2: Outcome relations for the 8 bipartite nonlocal-boxes.
Bibliography

[1] G.M. D’Ariano, P. Perinotti, and A. Tosini, Information and disturbance in operational probabilistic theories, pre-print arXiv:1907.07043 [quant-ph], (2019)

[2] G. Chiribella, G.M. D’Ariano, and P. Perinotti, Probabilistic Theories with Purification, Physical Review A 81, 062348 (2010).

[3] J. Bell, On the Einstein Podolsky Rosen paradox, Physics Physique 1 (3): 195–200 (1964)

[4] B.S. Cirel’son, Quantum generalizations of Bell’s inequality, Lett. in Math. Phys. 4, 83 (1980)

[5] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett., 23 (15): 880–4 (1969)

[6] W. van Dam, Nonlocality & Communication Complexity, Ph.D. thesis, University of Oxford, Department of Physics, (2000)

[7] S. Popescu and D. Rohrlich, Quantum Nonlocality as an Axiom, Foundations of Physics 24, 379 (1994).

[8] S. Wolf and J. Wullschleger Oblivious transfer and quantum non-locality, (quant-ph/0502030) (2006)

[9] A.J. Short, N. Gisin, and S. Popescu, The Physics of No-Bit-Commitment: Generalized Quantum Non-Locality Versus Oblivious Transfer, Quantum Information Processing volume 5, pages131–138 (2006)

[10] H. Buhrman, M. Christandl, F. Unger, S. Wehner, and A. Winter, Implications of Superstrong Nonlocality for Cryptography, Proceedings of the Royal Society A, 462(2071), pages 1919-1932, (2006)

[11] G.M. D’Ariano, G. Chiribella, and P. Perinotti, Quantum Theory From First Principles, An informational Approach, Cambridge University Press, (2017)
[12] G.M. D’Ariano, F. Manessi, P. Perinotti, and A. Tosini, *Fermionic computation is non-local tomographic and violates monogamy of entanglement*, EPL (Euro-physics Letters), **107**(2), 20009, (2014).

[13] G.M. D’Ariano, F. Manessi, P. Perinotti, and A. Tosini, *The Feynman problem and fermionic entanglement: Fermionic theory versus qubit theory*, International Journal of Modern Physics A Vol. **29** 1430025 (2014).

[14] S. Blum, *Coin Flipping by telephone - A Protocol for Solving Impossible Problems*, SIGACT News **15**, 23 (1983).

[15] G.M. D’Ariano, D. Kretschmann, D. Schlingemann, and R.F. Werner, *Reexamination of quantum bit commitment: The possible and the impossible*, Phys. Rev. A **76**, 032328 (2007).

[16] G. Chiribella, G.M. D’Ariano, P. Perinotti, D. Schlingemann, and R.F. Werner, *A short impossibility proof of Quantum Bit Commitment*, Phys. Lett. A **377** (2007).

[17] C.H. Bennet and G. Brassard, *Quantum Cryptography: Public Key Distribution and Coin Tossing*, International Conference on Computers, Systems & Signal Processing (1984).

[18] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois in Proceedings of the 34th Annual IEEE Symposium on the Foundations of Computer Science, edited by Leonidas Guibas (IEEE Computer Society Press, Los Alamitos, 1993), p. 362.

[19] H.K. Lo and H.F. Chau, *Is Quantum Bit Commitment Really Possible?*, Phys. Rev. Lett. **78**, 3410 (1997).

[20] D. Mayers, *Unconditionally Secure Quantum Bit Commitment is Impossible*, Phys. Rev. Lett. **78**, 3414 (1997).

[21] J. Barret, *Information processing in generalized probabilistic theories*, Phys. Rev A **75**, 032304 (2007).

[22] J. Barret, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, *Non-local correlations as an information-theoretic resource*, Phys. Rev. A **71**, 022101 (2005).

[23] G.M. D’Ariano and A. Tosini, *Testing axioms for quantum theory on probabilistic toy-theories*, Quantum Information Processing **9**, 95 (2010).
[24] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Physical Review 47 (10): 777–780 (1935)

[25] A.J. Short and J. Barret, *Strong nonlocality: a trade-off between states and measurements*, New Journal of Physics 12, 033034 (2010).

[26] M. Dall’Arno, S. Brandsen, A. Tosini, F. Buscemi, and V. Vedral, *No-Hypersignaling Principle* Phys. Rev. Lett. 119, 020401 (2017).

[27] D. Gross, M. Müller, R. Colbeck, and O.C.O. Dahlsten, *All Reversible Dynamics in Maximally Nonlocal Theories are Trivial* Phys. Rev. Lett. 104, 080402 (2010).

[28] S.W. Al-Safi and A.J. Short, *Reversible dynamics in strongly non-local Boxworld systems* Journal of Physics A: Mathematical and Theoretical 47, 325303 (2014).

[29] S. Pironio, J.-D. Bancal, and V. Scarani, *Extremal correlations of the tripartite no-signaling polytope*, J. Phys. A: Math. Theor. 44, 065303 (2011)

[30] H. Barnum, O.C.O. Dahlsten, M. Leifer, and B. Toner, *Nonclassicality without entanglement enables bit commitment* Proceedings of IEEE Information Theory Workshop, pp. 386-390, 2008

[31] G.M. D’Ariano, F. Manessi, and P. Perinotti, *Determinism without causality*, Physica Scripta, T163:014013 (2011)