

Universidade de São Paulo
Instituto de Física

Teoria de Recursos para Contextualidade Generalizada

Rafael Wagner

Orientadora: Profa. Dra. Bárbara L. Amaral

Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre(a) em Ciências.

Banca Examinadora:

Profa. Dra. Bárbara Lopes Amaral - Orientadora (Instituto de Física - USP)

Prof. Dr. Frederico Borges de Brito (Instituto de Física de São Carlos - USP)

Prof. Dr. Ernesto Fagundes Galvão (Universidade Federal Fluminense - UFF, International Iberian Nanotechnology Laboratory - INL)

Assinatura do Orientador:



São Paulo
2020

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Wagner, Rafael

Teoria de Recursos para Contextualidade Generalizada. São Paulo, 2020.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Matemática

Orientador: Profa. Dra. Bárbara Lopes Amaral
Área de Concentração: Fundamentos da Teoria Quântica e Informação Quântica

Unitermos: 1. Contextualidade; 2. Teoria de recursos; 3. Teoria quântica.

USP/IF/SBI-071/2020

University of São Paulo
Physics Institute

Resource Theory for Generalized Contextuality

Rafael Wagner

Supervisor: Prof. Dr. Bárbara Lopes Amaral

Dissertation submitted to the Physics Institute of the
University of São Paulo in partial fulfillment of the re-
quirements for the degree of Master of Science.

Examining Committee:

Prof. Dr. Bárbara Lopes Amaral - Supervisor (Physics Institute - USP)

Prof. Dr. Frederico Borges de Brito (Physics Institute São Carlos - USP)

Prof. Dr. Ernesto Fagundes Galvão (Fluminense Federal University - UFF, International
Iberian Nanotechnology Laboratory - INL)

São Paulo

2020

*Aos meus pais,
Elaine e Mauro.*

Resumo

Contextualidade generalizada corresponde a uma hipótese sobre as descrições ontológicas de resultados experimentais. Um modelo ontológico que descreve tabelas de dados é considerado não contextual quando operações equivalentes do ponto de vista operacional possuem a mesma representação no modelo. Quando esses modelos não podem ser construídos, dizemos que há contextualidade. A teoria quântica é o mais importante exemplo de uma teoria física contextual. Neste trabalho descrevemos como contextualidade pode ser entendida dentro do formalismo de teorias de recurso com aplicações para tarefas quânticas, desenvolvendo um método geral para o estudo de descrições experimentais, chamadas de cenários. Em particular, discutimos três tarefas de informação distintas: comunicação com restrição de paridade, discriminação entre duas hipóteses e clonagem. Nós concluímos, usando ferramentas da teoria de recursos e de algoritmos para obtenção de desigualdades de não-contextualidade, que nessas tarefas existem vantagens devido à contextualidade da teoria quântica em relação a quaisquer protocolos clássicos análogos.

Palavras chave: contextualidade, teoria de recursos, teoria quântica, modelos ontológicos, clonagem quântica, discriminação de estados.

Abstract

Generalized contextuality constitutes the following definition: for a given set of experimental procedures, any ontological explanation for the probability distributions obtained will need to distinguish between operationally equivalent procedures. We name such models as contextual and consider a notion of classicality that does not allow that to happen, namely, a notion of classicality that arises from noncontextual ontological models for experiments. Quantum theory is the most important of physical theories that have such a property. In this work, we have studied the notion of generalized contextuality, rooted in the philosophy of science, but that has shown to be particularly useful for solving experimental loopholes present in the Kochen-Specker formulation of contextuality. We have also studied the resource theory framework, developing new tools for witnessing quantum contextuality in prepare-and-measure scenarios focusing on already known contextuality scenarios in the literature.

Keywords: contextuality, resource theory, ontological models, quantum cloning, minimum error state discrimination, parity oblivious.

ACKNOWLEDGEMENTS

I want to start thanking my supervisor Bárbara Amaral. 400 days ago, we started the project that concludes with this dissertation; I will be forever thankful for her guidance and patience. She has now become a model for me, both as a researcher and as a professor. I would also like to thank Roberto Baldijão, for all the help and attention. I would also like to thank Alisson Tezzin for all the discussions, the help with the algebraic formalism of quantum statistical mechanics, for the kindness and patience.

I would like to thank all my friends Renan, Thiago, Gabriel, Fernanda, Pedros, Mariana Cipolla, Leo, Marina, Giulio, Brenda, Silvio, Ari, for all the good moments, chats, and funny encounters. My friends Geeh, Xii, Theo, Mai, Cris, Cat, Maya, Cae, Lucas, for all the Star Wars previews and friendship in all these years. A super special thank to Cauê and Gustavo, for all the nights without sleep working on projects, for always being there whenever I needed, for travels to conferences. They were/are inspiring references for me and I am in great debt for their support.

An enormous thanks to Marlon, Andrea, and Paula, workers from my institute, for all the help during my entire graduate studies.

I would like to acknowledge the enormous help that I have received from my entire family—especially Elaine, Mauro, Isadora, and Dagma, for being the most amazing leftist family in the world. A special thanks to my cousin Adriana, for all the coffee, the great moments when we lived together, and the help of all these years I've worked with physics. I would like to thank my girlfriend Elisabetta Girardi for teaching me how to be pragmatic, study harder, help me with English and Italian, believe in me when I didn't, and for teaching me how to take punctuality seriously; without her, this work would have never been possible. Special thanks to the McKay family: Craig, Maggie, Neil and Ingrid for many good moments in Uppsala. Thank you to Dalila, Luquinhas, Lilico, Hugo, Eunice, Miguel, Angela, Tinilson, Ju, Ina, both Di, César, Gabi, Si, Carlos, Anna.

I thank many of the professors of the physics institute that inspired me so much throughout all these years: João Carlos Alves Barata, Walter Alberto de Siqueira Pedra,

Oscar Éboli, Gabriel Landi, and specially those that were incredible friends, that made remarkable impact on my personality such as professors Josif Frenkel and Ana Regina Blak, to whom I send the kindest hug for all the help.

I would also like to thank Cristhiano Duarte, Ernesto F. Galvão, Marcelo Terra Cunha, Frederico Borges de Brito and Ana Belén Sainz for incredible comments regarding my work, my writing and my ideas. Your comments were enlightening and inspiring, and hopefully they will lead to future improvements and interesting research.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

CONTENTS

List of Figures	iii
List of Tables	vi
1 Introduction	1
2 Generalized Contextuality Framework	4
2.1 Operational-probabilistic theories	5
2.2 Ontological models	18
2.3 Generalized contextuality	20
3 Noncontextuality inequalities	26
3.1 No-go theorems	27
3.2 Noncontextual measurement-assignment polytope	38
3.3 Prepare-and-measure noncontextuality inequalities	40
4 Resource Theory for Generalized Contextuality	44
4.1 Resource Theory framework	45
4.2 Quantifiers	52
4.3 Witnessing contextuality with the free operations	55
4.4 Box product	62
5 Quantum Information tasks	70
5.1 Parity Oblivious Multiplexing	70
5.2 Quantum State Discrimination	76
5.3 Quantum Cloning	88
6 Conclusions and Further Directions	95

Appendix A Glossary	107
Appendix B Contextuality for quantum theory	112
B.1 Preparation contextuality	112
B.2 Measurement contextuality	115
Appendix C PORTA	118
C.1 Vertex enumeration	118
C.2 Fourier-Motzkin elimination	120
Appendix D Linear programming	127
D.1 Noncontextuality from data-tables	127
D.2 Contextual Fraction	134
Appendix E Proofs for Noncontextual bounds	137
E.1 n -bit parity oblivious multiplexing	137

LIST OF FIGURES

2.1	Quantum theory as an operational theory.	7
2.2	Representation of a prepare-and-measure scenario: Given \mathbb{S} we experimentally choose to utilize the blue source S that can prepare $P_{[s_1 S]}$ with probability $p(s_1 S)$. $\mathbb{V}_S = \{s_1, s_2\}$ and $\mathbb{O}_M = \{-1, 0, +1\}$. In the figure, $p(+1 s_1, S, M) \neq 0$	10
2.3	Representation of a complete coarse-graining of the source events. Source event $[T S_T] = [0 S] + [1 S]$	11
2.4	Operational representation of the equivalence between M_1 and a coarse-graining of M_2 into \tilde{M}_1 . The first box represents a set \mathbb{P} of preparations of different colors; each button above with a different color represents a different preparation procedure. On the right side of the picture, each box is now a measurement procedure itself, and we represent the possible outcomes of the measurements M_2 and the coarse-graining \tilde{M}_1 . However, we note that in the light of the operational theory, we do not have the means to differentiate between M_1 and \tilde{M}_1	13
3.1	Overlapping probability distributions over ontic space Λ . If the ontic state is in the overlapping region a single-shot measurement could not distinguish between the two preparations. Here we abuse notation and denote the probability densities over $\Lambda = \mathbb{R}$ with the same symbol as the probability measures.	29
4.1	Example of a free operation: pre-processing box for the measurement procedures in an experimental scenario.	47

4.2 Solving the problem of operational inequalities using free operations, as described in (Mazurek et al., 2016), (Schmid and Spekkens, 2018). In figure 4.2a we consider that there are ideal procedures, that satisfy the operational equivalences described by the vertical and horizontal dashed lines, and that we would like to implement (denoted as $\{P_j^{\text{ideal}}\}_{j=1}^4$). Due to noise, the experiment implements the real procedures (denoted by $\{P_j^{\text{ideal}}\}_{j=1}^4$) that may fail to respect the operational equivalences. In figure 4.2b we use a free operation defining secondary procedures (denoted as $\{P_j^{\text{sec}}\}_{j=1}^4$) that respect the operational equivalences. 60

4.3 Representation of the polytope structure arising from the product scenario. We stress that it is not clear how should be the new form of the *quantum* set $Q(\mathbb{B})$, even though it is clear the polytope structure for both NC and the larger polytope of statistics. In this picture we have used the fact that $NC(\mathbb{B}) \subseteq Q(\mathbb{B})$. From the convex nature of $Q(\mathbb{B})$, the product scenario must have a quantum contextual set that is at least of the form given by curve 1, but it could also be given by curve 2 and the study of maximal violations for noncontextuality inequalities shall answer such questions (Ambainis et al., 2016),(Chailloux et al., 2016). 69

5.1 Operational description of 2-bit parity-oblivious multiplexing prepare-and-measure scenario. Alice generates a random string $x \in \{(0, 0), (1, 0), (0, 1), (1, 1)\}$, while Bob generates a number $y \in \{0, 1\}$. Corresponding to each of these outcomes we have some probability that Bob guesses the bit correctly. . . 72

5.2 Comparison between the optimal guessing strategy, $p(g = x_y)$, classical bound provided by noncontextual models of n -bit parity oblivious multiplexing communication protocol with respect to the optimal quantum strategy. We see that not only for the scenarios corresponding to 2 and 3 bits, we have that quantum theory is advantageous, as shown in (Spekkens et al., 2009a), also for any n -bit scenario, $n \geq 4$ we have contextual advantage. 75

5.3 Comparison between the quantum optimal noisy bounds for the average success probability s in equation 5.26 and the noncontextual bound given by the noncontextuality inequality from theorem 13. We are always considering the symmetries from (5.29a)-(5.29c). In the blue curve we have the behaviors that might be achieved by quantum theory and that are conjectured to be optimal. In the yellow region we have the noncontextuality facet associated with the noncontextual polytope $NC(\mathbb{B}_{sd}) \cap \{\text{symmetries}\}$. 85

5.4 Comparing the noncontextual and quantum bounds for the ideal case. The l_1 -distance will find the square region depicted in light blue. 87

5.5	Comparing the optimal quantum fidelity with the optimal fidelity that any noncontextual ontological model need to obey. We plot the optimal fidelities against the confusability $c = p_{b,a}$	93
5.6	Comparing the optimal quantum fidelity with the optimal fidelity that any noncontextual ontological model need to obey. In this case, we are considering that there exists noise, with visibility parameter given by $v = 0.015$. We see a region where optimal quantum fidelity is higher than the noncontextual bound, but that due to noise this is no longer the case for all $c = p_{b,a}$. For the noisy fidelity we have considered the one in (Lostaglio and Senno, 2020).	94
B.1	Representation of the mixed state and all the six decompositions we will consider in the proof of preparation contextuality.	112
D.1	Monte Carlo distribution: We have generated and calculated the contextual fraction for behaviors in \mathbb{B}_{s_i} . In the plot we have around 5 000 000 points.	135

LIST OF TABLES

5.1	Quantum Behavior for the simplest scenario.	74
5.2	Data-table for the final statistics obtained by quantum predictions.	74
5.3	Behavior representing the statistics of scenario \mathbb{B}_{sd}	81
5.4	Behavior representing the operational description of the success rate s , the confusability c and the noise ε of scenario \mathbb{B}_{sd} . Here we consider the validity of (5.29a)-(5.29c), but one could drop such constraints and simply work with the general behavior description of table 5.3.	81
5.5	Behavior representing the operational description of the primitives from minimum error state discrimination without considering the symmetries.	84
C.1	Vertex enumeration PORTA: efficiency	120
D.1	Table of vertices from the large polytope of all behaviors $C(\mathbb{B}_{si})$: Contextual vertices for the simplest scenario. Each of these vertices violate one of the tight noncontextuality inequalities defined by (C.14)-(C.22).	136

INTRODUCTION

When we talk of an interpretation, this means that, even if we find it hard or impossible to identify the ontological basis, the mere assumption that one might exist suffices to help us understand what the quantum mechanical expressions normally employed in physics, are actually standing for, and how a physical reality underlying them can be imagined.

([Hooft, 2016](#), Gerard't, pg. 34)

*The belief in an external world independent of the perceiving subject is the basis of all natural science.*¹

([Einstein, 1960](#), pg. 266)

Since the beginning of the twentieth century, experimental facts about the description of quantum mechanics were in conflict with current interpretations of what does it mean for a physical theory that aspires to describe reality, to be complete. Einstein, Podolsky, and Rosen ([Einstein et al., 1935](#)) argued that for a theory to be considered complete, it needs to have an accurate description of the *elements of reality* ([Howard, 1985](#)). Seen from this perspective, quantum mechanics would not be complete. Consequently, to describe the physical reality entirely, some new hidden structure in the formalism must exist since quantum mechanics correctly predicted all known low-energy experiments. The effort to formulating an interpretation of quantum theory by completing it with new variables has been named *realism*² since this was considered a “desperado” attempt to bring classical notions of reality into the quantum formalism. However, the realistic interpretation was not the only one considered; throughout the century, other explanations

¹Taken from ([Landsman, 2005](#), pg. 10); since the author is fluent in German he mentions the original sentence: *Der Glaube an eine vom wahrnehmenden Subjekt unabhängige Außenwelt liegt aller Naturwissenschaft zugrunde*, from ([Einstein, 1934](#)).

²Nowadays, realism is a term used for a much larger group of attempts at interpreting quantum theory. The one here could be better defined as *hidden-variable realism*. As an example, among many others, the author in ([Hooft, 2016](#)) construct a realist approach to quantum theory by relaxing the Free Will argument in Bell's theorem.

for the *strangeness* of quantum mechanics emerged. A wide variety of explanations were presented to comprehend the strange facts observed by the quantum theory - entanglement and steering being the most notable examples of strange effects that have baffled early realists - which created a conflict with the more classically minded interpretations, represented by the realistic approach; Nowadays, we study that spectrum of explanations in the field of research of *foundations of quantum theory*. The seminal works of (Kochen and Specker, 1975), (Bell, 1964) and (Bell, 1966), had a profound impact on that field as they have set rigorous mathematical proofs that ruled out some realistic interpretations of quantum mechanics. Their work led to a better understanding of the theory and helped to develop useful ideas that converged, for example, to the use of entanglement as a resource for quantum information (Nielsen and Chuang, 2002), (Wilde, 2013), (Horodecki et al., 2009).

In his work, Bell derived upper bounds on probabilistic correlations that any local hidden-variable model of quantum theory would need to satisfy, under the following assumptions, (Hermens, 2010, pg. 38), (Amaral, 2020, 10:30),

1. **Realism:** Measurement procedures only *reveal* pre-defined results. Therefore there exists a state of the system that determines the measurement events.
2. **Locality:** Any choice made by one side of a two-party experiment does not affect the other side. We could further specify this in terms of the following assumptions:
 - (a) **Outcome Independence:** Measurement events do not depend simultaneously on other measurement events made far away.
 - (b) **Context Independence:** Measurement events do not depend on the experimental setup of the party located far away.
3. **Free will:** Each party has free will to choose which measurement they will perform.

And under these considerations, Bell later showed that the quantum formalism allows for a violation of these bounds. These violations were the mathematical proof that the quantum formalism indeed permits an instance of nonlocality. In a later work, Kochen and Specker demonstrated an unusual feature about quantum probability theory that comprises Bell's earlier results: Quantum theory is intrinsically context-dependent. In other words, no hidden variable model, for which the value assignments do not depend on which context the measurements are performed, can adequately describe statistical data gathered by every quantum experiment. This theoretical fact, contextuality of quantum probability theory, discovered by Kochen and Specker (so-called KS-contextuality) between measurement procedures is a necessary condition for any interpretation of quantum data assuming realism. Nevertheless, there are experimentally restrictive constraints considered in KS-contextuality, e.g., the measurements are always represented by orthogonal projections, which are never experimentally implemented because of noise effects.

In (Spekkens, 2005), the author proposes a generalized notion of contextuality for operational theories suitable for noise, and that recovers KS-contextuality as a particular case. This new paradigm implemented a notion of contextuality not only for measurement procedures but also for preparations and transformations that one can perform in a quantum system; included other physical, operational models beyond quantum theory and relaxed the constraints on measurement procedures to allow noisy protocols that better describe non-idealized measurements in the laboratory.

Given that we have rigorously established that quantum theory constitutes an utterly different overview of physical reality, the following question remains: can we make direct technological usage of these discoveries? Following this path, researchers have shown entanglement to be a successful resource for applying several quantum information protocols, which paved the way for searching similar results associated with other effects such as quantum steering, Bell nonlocality, quantum thermodynamics, asymmetry, to name a few. The name *quantum resource theory* (Coecke et al., 2016), stands for the general framework of constructing tools built upon these quantum effects.

In this project, we want to study the resource theory for generalized contextuality developed in (Duarte and Amaral, 2018). The guiding question for this dissertation was:

*Can generalized contextuality be a resource for quantum information theory?
If so, would it be possible to formulate a framework to witness both the resource
and the advantage it provides for general experimental situations?*

By reviewing the literature, we answer the first question positively: Generalized contextuality represents a fundamental resource for the advantage of some quantum information protocols. We present proof of advantage for a large class of oblivious protocols and two quantum information protocols: binary minimum error state discrimination and state-dependent quantum cloning. For the second question, we developed two resource theory frameworks to witness nonclassicality: one framework that uses pre/post-processes and another one that might be useful to engineer complex scenarios with an embedded contextuality structure, but without exponentially increasing the computational complexity for obtaining noncontextuality inequalities. These two frameworks can witness the resource, but they do not serve as general proofs of advantage for protocols.

This dissertation's structure is as follows: We start providing a review of generalized contextuality in chapter 2, from the first breakthrough by Robert Spekkens, and up to more recent perspectives. We continue to describe experimental scenarios that fit the prepare-and-measure description, and to present, in chapter 3 an algorithm that obtains all the noncontextuality inequalities for any scenario with finite procedures. We then present the resource theory we work with, together with our main results, in chapter 4. We finish in chapters 5 and 6 by applying the resource theory framework and drawing our main conclusions.

GENERALIZED CONTEXTUALITY FRAMEWORK

For me, I am in a sense a realist, but what is important to my mind about Realism - maybe I don't even need to be committed to the term Realism - the thing that is unsatisfactory about Instrumentalism, to my mind, is that it fails to provide causal explanations; I am committed to providing causal explanations of quantum phenomena.

- (Spekkens, 2016, 19:55-20:15)

Our research's primary goal is to understand better how quantum contextuality *à la* Spekkens provides useful tools for quantum information protocols. Since the pioneering work of (Hardy, 2001), one of the paths towards understanding quantum physics at its core is to find axioms for generalized probabilistic theories, henceforth mentioned as GPT framework, that recover quantum probability theory. This perspective lies upon the assumption that physical processes can always be described operationally by *some* probabilistic formalism; from a very empiricist perspective, what we can infer from the laboratory is always, at most, how can we proceed experimentally to obtain raw statistical data in the form of probability of occurrence of the phenomena. In essence, we call this an operational theory framework.

To interpret experimental results, we attempt to explain how we perceive (philosophically) the world through the lenses of physical experience; we give a *meaning* to the acquired probabilities by the operational theory. To engage in such a meaning corresponds to engage with an ontological commitment (Bricker, 2016). This commitment is not always related with an underlying ontological explanation in the terms we will consider throughout this dissertation, but generally speaking, it is natural to give meanings to probabilities as if they show some aspect of the nature of the processes in question.

Hence, treating operationally any experimental procedure should give an overview of how we can generally approach *theory building*, since, *a priori*, we do not give any interpretation to the probabilities obtained - a job for the ontological models we build. From this, we can make powerful assertions about abstract ideas; rigorously by writing no-go

theorems, technically by deriving noncontextuality inequalities that can be implemented, and finally, foundationally by searching principles¹, or axioms, that are representatives of the theory.

We can view quantum mechanics as an operational theory, and one can consider the following question: How can we provide an ontological model that describes such an operational theory? Standard quantum mechanics² is just one possible ontological description. Re-framing this question using ontological models: *Consider any ontological description that can adequately describe the probabilistic features of quantum mechanics as evaluated/measured in the laboratory. By making some set of restrictions - in our case, an assumption that is both philosophically and experimentally motivated: Spekkens notion of contextuality - what can we infer?*

In what follows, we proceed to describe the abstract theory, properly defining what we mean by an operational theory and an ontological model. For a clear exposition of the concepts not directly related to the physics community, we refer the reader to appendix A. The notion of contextuality we will formulate is going to be described as a hypothesis over the interplay between the concepts mentioned above. We shall call this *Spekkens contextuality*, *generalized contextuality* or, if no confusion arises, just *contextuality*. We finish this chapter with comment on its philosophical roots: Leibniz principle of indiscernibles.

2.1 Operational-probabilistic theories

We call an operational-probabilistic framework, a set of laboratory prescriptions to obtain probabilistic data. We will see that even committing to such a description can be understood as committing oneself to some physically relevant assumptions. For example, we are already committing to the fact that the operational theory is capable of assigning probability distributions. Many of the quantities we will work with could be studied in more general terms as a somewhat *pure* operational theory. In this dissertation we only consider operational-probabilistic theories, therefore, for us, we drop the term probabilistic for the rest of this dissertation.

First of all, we need the concept of *state*, or *behavior*, which is related to the notion of a physical description of some part of nature that we want to delineate experimentally. For instance, take a particle that we characterize as having intrinsic angular momentum. The particle and its momentum are a small aspect of nature that we call a *system* and,

¹As an example, a known proposal for a principle that could represent a fundamental feature of quantum theory is the *exclusivity principle*, since this principle forbids a set of correlations larger than the quantum set (Amaral et al., 2014).

²Standard here does not mean the ideas present by Bohr, Heisenberg and Pauli, also commonly known as the Copenhagen interpretation. By standard we mean the set of rules used by physicists to obtain experimental results, such as the Born rule. For a discussion relating the different interpretations among the mentioned physicists, and also about the measurement problem, see (Janssen, 2008).

suppose that we would like to understand better how this system behaves when exposed by a magnetic field without making any further assumptions, such as its relationship with other systems, its inner complexity, and other possible degrees of freedom. Then we must, operationally, come up with a set of rules, or better, a list of instructions that we will use to obtain information from the system. We define two different procedures that achieve this goal: *preparations* and *operations*.

A preparation procedure will leave the system in its “primary configuration” — for example, a free particle in a configuration of fixed intrinsic angular momentum. In general terms, we consider a valid preparation procedure to be any initial configuration of the system being studied; e.g., once a given tissue has absorbed radiation, we can consider that tissue a prepared state of the system. It is now clear what is the point of such description: an *operation* will be generically any possible interaction, including non-local, strongly correlated ones, acting on a *prepared* system. In this work, we will consider operations as measurements and, in some sections, we refer to the notion of transformation procedures, also representing operations, since including transformations in the framework is particularly interesting to study abstract aspects of quantum theory (Baldijão et al., 2020), despite one can always - as can be seen by the generality of our considerations - consider a transformation as part of a measurement, or as part of a preparation procedure. Measurements are operations essentially different from the others since the scientist obtain information about the system.

Following (Amaral, 2015) and (Barrett, 2007) we make an assumption here that is fundamental for any operational theory framework:

Assumption 1. Every preparation or operation can be repeated to produce a probability of occurrence of outcomes when subject to the measurement operations.

This assumption has significant importance: it assumes that one could, for instance, perform enough laboratory experiments in order to build a data table, representing a set of probabilities of occurrences. For instance, consider an operational setting that wants to describe the state of some system experimentally: define the set \mathcal{S} of sources $S \in \mathcal{S}$ where each S represents a source of a finite set \mathbb{V}_S of different preparations, and we consider that \mathcal{S} is the set of all such possible sources for some operational theory. The assumption considers that a preparation event $[s|S]$, which means that the source S has prepared $s \in \mathbb{V}_S$, can be repeated indefinitely.

We also notice that assumption 1 does not express how many preparations or operations one should perform to reproduce the final probabilistic data representing the actual behavior of the system. That is a crucial point that can be considered an essential loophole of having probability distributions be the essential underlying features of the theory.

All the ingredients of the operational framework are now settled. We will rep-

represent the set of all possible measurements procedures by $\mathcal{M} = \{M_i\}_{i \in I}$, the set of all source settings $\mathcal{S} = \{S_j\}_{j \in J}$ and, to be complete, in the description given by (Spekkens, 2005), it is also mentioned the set of all transformation procedures $\mathcal{T} = \{T_q\}_{q \in Q}$, for I, J, Q arbitrary set of indices. In our work, we focus on prepare-and-measure experiments; therefore, scenarios where only preparations and measurements are performed. We denote the set of all possible outcomes within the scenario as $\mathcal{O}_M = \{k\}_{k \in K}$, where K is also an arbitrary set of indices. Furthermore, we denote the set $\mathcal{V}_S := \{s\}_{s \in S}$ of possible source values related to a setting S . Every pair $[s|S]$ represents a preparation procedure $P_{[s|S]}$, that we refer to as a preparation event. Whenever it is convenient, we will represent these procedures graphically, as depicted in 2.1.

Example 1. It is possible to consider quantum mechanics as an operational theory. In such a description, we have the generalized notions of states, transformations (also called channels in the quantum information literature), and measurements. States are trace-class positive operators $\rho \in \mathcal{B}(\mathcal{H}_{in})^{+3}$ with $\text{Tr}(\rho) = 1$, transformations are CP-maps $T : \mathcal{B}(\mathcal{H}_{in}) \rightarrow \mathcal{B}(\mathcal{H}_{out})$ and measurements are positive partitions of the unit $\mathcal{B}(\mathcal{H}_{out}) \supset \{E_k\}_{k \in K}$, for an arbitrary set of indices K .⁴

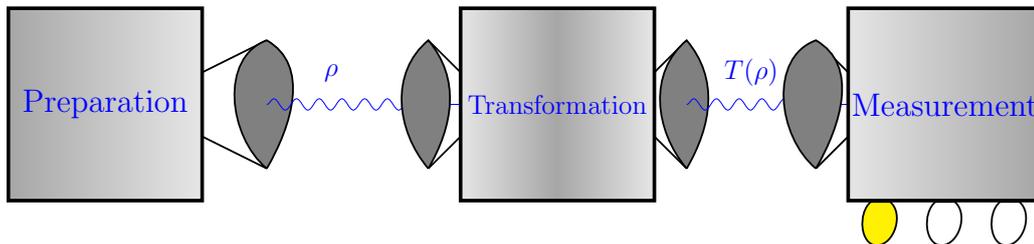


Figure 2.1: Quantum theory as an operational theory.

△

What we will mean by a *state*, or more commonly as a *behavior*, it was first introduced in (Hardy, 2001), (Barrett, 2007) as:

Definition 1. A *state*, also known in the literature by a *behavior*, a *data table*, a *black-box correlation* or a *prepare-and-measure statistics*, constitutes a list of probabilities of outcomes

$$B := \{p(k|M_i, T_q, P_j)\}_{k \in K, i \in I, q \in Q, j \in J} \quad (2.1)$$

given that the specific preparations followed by the transformations and measurements were performed on that system a sufficiently large number of times. ◇

³This is the set of bounded positive, $\rho \geq 0$, operators acting on the Hilbert space \mathcal{H}_{in} .

⁴In (Werner, 2019, Lecture 1, 8:04), professor R. Werner says: “I’ve learned this way of representing things from my thesis advisor Günther Ludwig, so we call them Ludwig boxes”, which is the best historical quote from the birth of such operational description we could find.

If we consider a prepare-and-measure experiment, then we write the behavior corresponding to the system as $B = \{p(k|M_i, P_j)\}_{k \in K, i \in I, j \in J}$. When the sets of operations are finite, we will change the notation from calligraphic, e.g., \mathcal{S} , to bold \mathbb{S} . We will mainly use the behaviors as described by (2.1), without the transformations, and we present a discussion about the recently introduced construction by (Kunjwal, 2019), following previous work (Kunjwal and Spekkens, 2015) that brings attention to the formulation in terms of sources \mathcal{S} . See, for example, the discussion following definition 17 and how graph invariants connecting KS-contextuality and robust noncontextuality inequalities for generalized contextuality can be formulated via preparation events (Kunjwal, 2019).

In definition 1 the notion of behavior of a system suffers from a practical difficulty. Suppose that in an experiment, we would like to describe the state of a qubit. There are infinitely many measurements that one can perform on a qubit relative to every point in the Bloch sphere. This implies that to describe all possible behaviors arising from the system would only be possible given that an infinite number of measurements have been performed, which is experimentally unreasonable. The *tomographic assumption* takes care of such a problem, and we will see that this gives rise to a meaningful discussion on the experimental perspective of testing noncontextuality for (quantum) models regarding theoretical loopholes.

Assumption 2. (Amaral, 2015) A set of operations is called *tomographic* if the list of probabilities for the outcomes of these operations completely specifies the state of the system. We call a set of measurements tomographically complete, or fiducial, if, in this set, there is the minimum number of measurement procedures necessary for the behavior to be completely specified.

This assumption supports the following observation: in our qubit case, one needs only to obtain the statistics of three different measurements along the defined axis that corresponds to the Pauli matrices. These three measurements fully characterize the state, and so the set $\{\sigma_x, \sigma_y, \sigma_z\}$ can be considered a fiducial set of measurements for the qubit. Hence, assumption 2 solves the operational problem of the necessity of an infinitely large number of measurements to find the correct finite set of measurement procedures that fully characterize the state. We will only consider procedures with a finite set of outcomes, and experimental scenarios that admit a finite set of tomographically complete procedures. The process of finding a state given such a finite set of measurements is called *tomography*.

Definition of the operational quantities

We consider a set \mathcal{S} to be the set of all source settings S . A source setting is composed of possible sources $s \in \mathcal{V}_S$, each representing a possible preparation event. To the pair $[s|S]$ we associate $(p(s|S), P_{[s|S]})$, the probability that one chooses $s \in \mathcal{V}_S$ therefore making the preparation procedure $P_{[s|S]}$. In the rest of this work, P always stand

for preparations, while p stands for probabilities. The conditional probability associated with choosing a source $s \in S$ does not necessarily need to be a free parameter, it may be controlled by the apparatus (such as flipping weighted coins to choose between sources) or it can also be a source of randomness unconstrained. Whenever we make changes in notation such as $\mathcal{V}_S \rightarrow \mathbb{V}_S$, it represents that we are considering finite sets of sources.

Example 2. Consider a common quantum information protocol where we have a laser pump and a beta barium borate (BBO) nonlinear crystal responsible for generating entangled photons. This is a setting $S \in \mathbb{S} \subseteq \mathcal{S}$.⁵ Each pair crystal + laser beam of photons represents a source S . The setting s corresponds to the laser states that can be prepared, such as a label representing different wavelengths, among all choices \mathbb{V}_S depending on the number of material resources one might have in the laboratory. This setting S is one of all possible settings \mathbb{S} available that can generate pairs of entangled photons. The pair $[s|S]$ is the operational event of choosing a specific crystal + laser preparation procedure $P_{[s|S]}$ with probability $p(s|S)$. \triangle

We cannot implement every element of \mathcal{S} since this set is infinite. Usually we speak about an operational framework via finite sets $\mathbb{S} \subseteq \mathcal{S}$, making use of assumption 2 by expecting \mathbb{S} to be tomographically complete. If our operational description is valid for \mathcal{S} , in what follows, we say that the assumptions are valid for the *theory*. If it is only valid for \mathbb{S} we call it a *model*. When \mathbb{S} is tomographically complete, we hypothesize that the assumptions made about the operational description are valid for the entire set \mathcal{S} .

For transformations, the situation is simple: there is no notion of “transformation event” since, in some sense, transformations are not fundamental in the operational framework. They can always be embedded in the preparation procedures if necessary, as mentioned before. For measurements, we have a similar situation, as described above. Hopefully, our intuitions are sufficiently clear to be summarized in a definition.

Definition 2. There are two notions of operational events: an operational event of a preparation procedure and an operational event of a measurement procedure. We denote them $[s|S]$ and $[k|M]$, respectively.

1. To $S \in \mathbb{S} \subset \mathcal{S}$ we associate $\{(p(s|S), P_{[s|S]})\}_{s \in \mathbb{V}_S}$ where \mathbb{V}_S is the finite set of possible settings in S . The preparation $P_{[s|S]}$ is obtained once $[s|S]$ happens.
2. The complete coarse-graining over $\{[s|S]\}_{s \in \mathbb{V}_S}$ corresponds to the effective event $[T|S_T]$. This preparation event we write as $P_{[T|S_T]}$, since by definition $|\mathbb{V}_{S_T}| = 1$.

$$[T|S_T] := \sum_s [s|S] \tag{2.2}$$

⁵Examples can be found in (Yu et al., 2019, Fig. 1, pg. 3), (Zhan et al., 2020, Fig. 4, pg. 5), (Bian et al., 2020, Fig. 2, pg. 3) and many others. This is a common preparation procedure in quantum optics.

3. To the event $[k|M]$ we mean that measurement procedure $M \in \mathbb{M} \subset \mathcal{M}$ was performed and we have obtained outcome $k \in \mathbb{O}_M$.⁶

◇

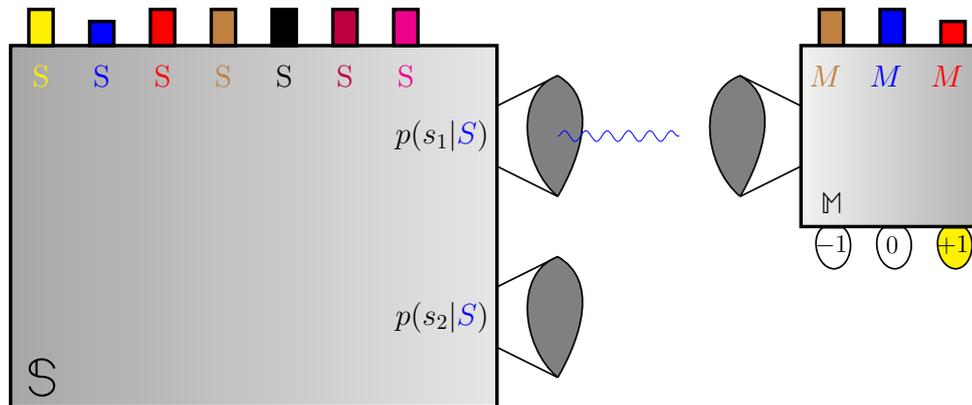


Figure 2.2: Representation of a prepare-and-measure scenario: Given S we experimentally choose to utilize the blue source S that can prepare $P_{[s_1|S]}$ with probability $p(s_1|S)$. $\forall_S = \{s_1, s_2\}$ and $\mathbb{O}_M = \{-1, 0, +1\}$. In the figure, $p(+1|s_1, S, M) \neq 0$.

This is the core of what we mean by an operational framework. There are left some comments on it: The completely coarse-grained event was defined by (Kunjwal, 2019), after working in the years before on the perspective of preparation events with application to new graph invariants, such as the Corr quantity.⁷ With these new techniques, the graph invariant for preparation procedures Corr was used to witness quantum contextuality (Mazurek et al., 2016). The notion of preparation event $[s|S]$ is relatively new in the generalized contextuality literature, and for applying many of the techniques that we present in this dissertation, it is still unknown the methodology to be used. For example, it is not clear how to apply the techniques we will discuss in chapter 3 to a prepare-and-measure scenario described in terms of preparation events.

Nevertheless, we think it is worth presenting this as a more general framework that we specify in future chapters for our case studies.

⁶The summation in (2.2) represents the logical “or”. So this means that $[T|S_T]$ is the procedure that prepares $[0|S_1]$ or $[1|S_1]$, for example.

⁷See also (Kunjwal and Spekkens, 2015). The definition and study of the quantity Corr can be found in (Kunjwal, 2019).

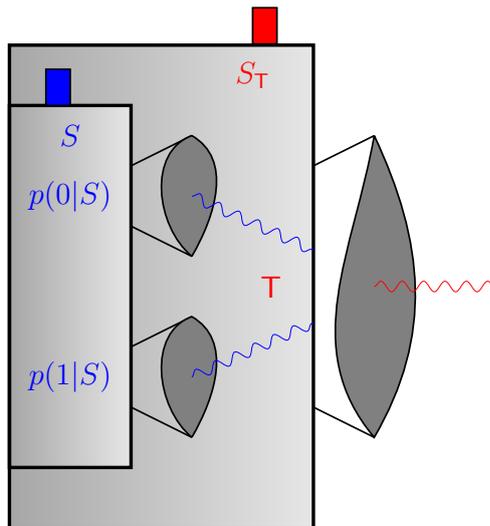


Figure 2.3: Representation of a complete coarse-graining of the source events. Source event $[T|S_T] = [0|S] + [1|S]$.

In what follows, we make observations relevant to someone starting in the field to help read the past fifteen years of literature. After the seminal paper, (Spekkens, 2005), presenting the notion of generalized contextuality, much of the way we approach the discussion has changed, and we try to give a brief account of these changes⁸. We also define the last ingredients needed for a complete picture in our theoretical treatment that will be necessary to define contextuality properly.

Operational equivalences: evolution of a concept

Since a behavior is described by how the probabilities that represent a system correspond to different preparations and measurements, it is possible that in a given system, we can obtain the same statistics \vec{B} for *different* experimental procedures. In what follows, we will try to formalize such a situation and argue about its importance. We follow the example given in (Spekkens, 2005) to study photon polarization scenarios:

Example 3. Take P to be a preparation defined by a homogeneous beam of photons and let M_1, M_2 be two measurement procedures described as polaroid oriented to pass vertically polarized light, M_1 , and in some different skewed direction, M_2 , both followed by a photodetector. Let now M_3, M_4 be a pair of birefringent crystals that are also oriented to transmit vertical, M_3 , and skewed light, M_4 , again followed by a photodetector. We notice that the behavior of the system does not change if we perform the procedure M_1 or M_3 for the same preparation P . The same happens for M_2 and M_4 . Such a description is deeply connected with the notion of generalized contextuality. By saying that the system does not change, we mean that we cannot perceive any difference in the

⁸A more state-of-the-art approach considers our boxes as processes in a certain category framework. See (Schmid et al., 2020c) and references therein for an overview of such ideas.

statistics obtained from the experimental procedures. Therefore, we have in mind the idea of empirical indistinguishability, where two different sets of procedures provide the same statistical description for the process. \triangle

We will define such property of the behavior in the operational theory by an *equivalence between measurements*. In (Spekkens, 2005), this notion is defined as follows:

Definition 3. Let $P, P' \in \mathcal{P}$. Then we say that P is operationally equivalent, or just equivalent to P' , if

$$\forall M \in \mathcal{M}, p(k|M, P) = p(k|M, P'). \quad (2.3)$$

Similarly, for every pair M, M' and T, T' we consider that they are operationally equivalent if, respectively,

$$\forall P \in \mathcal{P}, p(k|M, P) = p(k|M', P), \quad (2.4)$$

$$\forall P \in \mathcal{P}, \forall M \in \mathcal{M} p(k|M, T, P) = p(k|M, T', P). \quad (2.5)$$

We denote the equivalences between procedures by writing $P \simeq P'$. \diamond

Studying the works that came after the seminal paper of (Spekkens, 2005), we can see that there has been an evolution of such a concept of *operational equivalences*. Equation (2.4) has a conceptual issue: it does not allow the outcomes to vary. It seems like a small difference to write, for example, $p(k'|P, M') = p(k|P, M)$ but we give an example to clarify its importance.

Example 4. Let us consider two measurement procedures in the operational quantum formalism, with M_1 and M_2 given by

$$M_1 = \{|0\rangle\langle 0|, \mathbb{1} - |0\rangle\langle 0|\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}, \quad (2.6)$$

and

$$M_2 = \{|0\rangle\langle 0|, |1\rangle\langle 1|, |-1\rangle\langle -1|\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}, \quad (2.7)$$

where we have written the measurement procedures in the M_2 basis, each measurement procedure is here denoted as a set of operators. We are using a generalization of the quantum measurements: POVM's. Now we can see that, for instance, the events $[0|M_1]$ and $[0|M_2]$ should be operationally related since they both have the same probability of occurrence given a preparation of any state associated with the finite dimensional Hilbert space in case. Given any preparation P that we relate to the state ρ :

$$p(0|M_1) = \text{Tr}(|0\rangle\langle 0|\rho) = p(0|M_2)$$

We can recognize that with respect to (2.4) the two measurement procedures M_1 and M_2 are *not* operationally equivalent.

$$p(k = -1|M_2, P) \neq 0, p(k = -1|M_1, P) := 0.$$

Since we have understood that there are events that cannot be distinguished operationally among the measurement procedures we consider, we should improve the definition through defining *events* $[k|M]$ to be operationally equivalent instead of just the measurements M .

This, for example, accounts for measurement procedures with a different number of outcomes (in this formalism, a different number of possible events), and we can understand that correspondence noting that in the operational framework, we cannot distinguish between the event $[0|M_1]$ if we coarse-grain the measurement M_2 , as depicted in figure 2.4. △

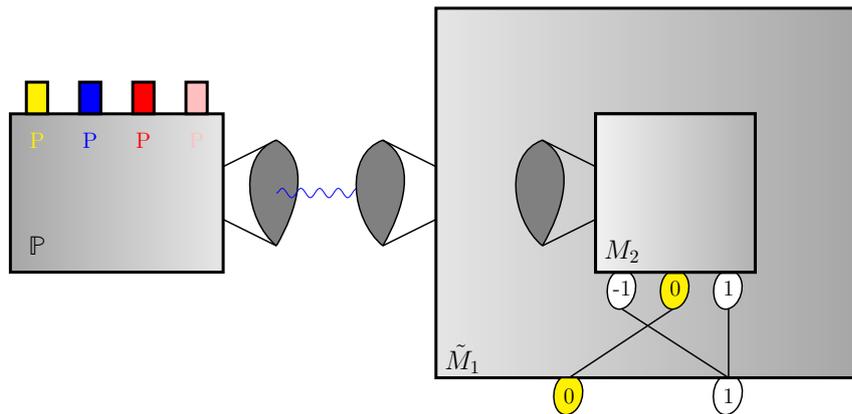


Figure 2.4: Operational representation of the equivalence between M_1 and a coarse-graining of M_2 into \tilde{M}_1 . The first box represents a set \mathbb{P} of preparations of different colors; each button above with a different color represents a different preparation procedure. On the right side of the picture, each box is now a measurement procedure itself, and we represent the possible outcomes of the measurements M_2 and the coarse-graining \tilde{M}_1 . However, we note that in the light of the operational theory, we do not have the means to differentiate between M_1 and \tilde{M}_1 .

This example shows that equation (2.4) is not the end of the story. We improve the definition by noticing that if we consider not only the measurement M but the *event* $[k|M]$, then our example has $[0|M_1] \simeq [0|M_2]$. Hence we make an improvement on definition 3:

Definition 4. Let the pair $[k|M]$ represent a measurement event. Then we say that a

pair of measurement events $[k|M]$ and $[k'|M']$ are operationally equivalent if

$$\forall P \in \mathcal{P}, \quad p(k|M, P) = p(k'|M', P), \quad (2.8)$$

and we say that $[k|M] \simeq [k'|M']$ or also $[k|M_i] \simeq [k'|M'_i]$, given a set of indices I . \diamond

Hence, equivalent measurement events are indistinguishable with respect to the probabilities arising in the operational framework. This change is present in the works (Duarte and Amaral, 2018), (Schmid and Spekkens, 2018) and (Schmid et al., 2018). The latest improvement for the notion was made in (Kunjwal, 2019) where we also recognize the operational equivalences between the preparation events. We provide a full definition with all such characteristics.

Definition 5. 1. Let $[k|M]$ a measurement event with $k \in \mathcal{O}_M$ and $M \in \mathcal{M}$. Then, we have that two measurement events are operationally equivalent iff

$$\forall [s|S], s \in \mathcal{V}_S, S \in \mathcal{S}, \quad p(k, s|M, S) = p(k', s|M', S) \quad (2.9)$$

where the full operational probabilities are described by the behaviors

$$B := \{p(k|M, S, s)p(s|S)\}_{s \in \mathcal{V}_S, k \in \mathcal{O}_M, M \in \mathcal{M}, S \in \mathcal{S}} = \{p(k, s|M, S)\}_{s \in \mathcal{V}_S, k \in \mathcal{O}_M, M \in \mathcal{M}, S \in \mathcal{S}},$$

and they were constructed by writing the probabilities in the operational description of the behavior in definition 1 with the notions in definition 2.

2. Let $[s|S]$ and $[s'|S']$ be two source events, we say that these are operationally equivalent iff

$$\forall [k|M], k \in \mathcal{O}_M, M \in \mathcal{M}, \quad p(k, s|M, S) = p(k, s'|M, S') \quad (2.10)$$

3. If we have a complete coarse-graining in the possible source events, we will effectively have only one preparation procedure, and we consider that two preparation procedures P and P' are equivalent iff,

$$\forall [k|M], k \in \mathcal{O}_M, M \in \mathcal{M}, \quad p(k|P, M) = p(k|P', M). \quad (2.11)$$

4. For transformation procedures we have the same definition as stated before, with the obvious changes exemplified by the last item.

\diamond

Prepare-and-measure scenario

The framework of operational theories is now defined, and we can construct the fundamental structure we work with, having in mind general experimental implementa-

tions. This fundamental structure we call a prepare-and-measure scenario.

Definition 6. We *specify* operational equivalences between preparation procedures P_j , with $j \in J$ a finite set of indices, as a set \mathbb{E}_P of pairs $(\alpha, \beta) \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$, with $|\mathbb{E}_P|$ being the number of equivalence relations of the type

$$\sum_{j=1}^{|\mathcal{S}|} \alpha_j^a P_j \simeq \sum_{j=1}^{|\mathcal{S}|} \beta_j^a P_j \quad (2.12)$$

with $a = 1, \dots, |\mathbb{E}_S|$, $0 \leq \alpha_j^a, \beta_j^a \leq 1$, and $\sum_j \alpha_j^a = \sum_j \beta_j^a = 1$ for any fixed a . The sums represent convex combination of preparations, and the alphas or betas are just the probabilities of the type $p(s|S)$. These convex combinations define new valid procedures. Similarly, we *specify* operational equivalences between measurement *events* $[k|M_i]$, for $k \in K$ and $i \in I$ finite sets of indices, as a set \mathbb{E}_M of pairs $(\alpha, \beta) \in \mathbb{R}^{2^{|\mathcal{O}_M|} \times 2^{|\mathcal{M}|}}$

$$\sum_{k=1}^{|\mathcal{O}_M|} \sum_{i=1}^{|\mathcal{M}|} \alpha_{[k|M_i]}^b [k|M_i] = \sum_{k=1}^{|\mathcal{O}_M|} \sum_{i=1}^{|\mathcal{M}|} \beta_{[m_k|M_i]}^b [k|M_i] \quad (2.13)$$

where these sums represent convex combinations of measurement events as before. \diamond

It is vital to notice that in this definition, we are not considering the inner structure of the equivalences in the source setting perspective. This means that there are different ways that the preparation procedures could be interpreted as sources, given some source setting, which could provide the same operational equivalences.

Example 5. Suppose that we have $S_1 = \{s_1^1, s_2^1\}$ and $S_2 = \{s_1^2, s_2^2\}$. Consider the following equivalence:

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 = \frac{1}{2}P_3 + \frac{1}{2}P_4 \quad (2.14)$$

which could come from letting $p(s_1^1|S_1) = p(s_2^1|S_1) = p(s_1^2|S_2) = p(s_2^2|S_2) = \frac{1}{2}$ and

$$p(s_1^1|S_1)P_{[s_1^1|S_1]} + p(s_2^1|S_1)P_{[s_2^1|S_1]} = p(s_1^2|S_2)P_{[s_1^2|S_2]} + p(s_2^2|S_2)P_{[s_2^2|S_2]} \quad (2.15)$$

but we could also have $S = \{s_1, s_2, s_3, s_4\}$, and $p(s_1|S) = \frac{1}{2}$ leading to the same operational equivalence,

$$p(s_1|S)P_{[s_1|S]} + p(s_2|S)P_{[s_2|S]} = p(s_3|S)P_{[s_3|S]} + p(s_4|S)P_{[s_4|S]}. \quad (2.16)$$

\triangle

The last example shows that different scenarios might have the same preparation equivalence. Making a connection that will help build the noncontextuality polytope in the next chapter, we define:

Definition 7. Consider a set \mathbb{S} such that we are only interested in the preparation procedure structure of the prepare-and-measure scenario. Then, we define the set \mathbb{P} to be the set of all preparation events:

$$P \in \mathbb{P} \iff P = P_{[s|S]} \quad (2.17)$$

for some event $[s|S]$. \diamond

This definition is not a fundamental one, but we present it mostly due to technical calculations that are to come. Different realizations of operational equivalences might lead to interesting new work, both theoretically and experimentally. We have followed (Duarte and Amaral, 2018) and (Kunjwal, 2019), and we notice that the work of (Kunjwal, 2019) lead to richer equivalences as in definition 5, but this last definition will be more practical in what follows on chapters 3, 5.

Definition 8. A prepare-and-measure scenario,

$$\mathbb{B} := (\mathbb{S}, \mathbb{M}, \mathbb{V}_S, \mathbb{O}_M, \mathbb{E}_S, \mathbb{E}_M) \quad (2.18)$$

consists of a finite set of source settings \mathbb{S} , a finite set of measurement procedures \mathbb{M} , a finite set of outcomes related to each measurement procedure \mathbb{O}_M , a finite set of values corresponding to a specific source \mathbb{V}_S and two finite sets of operational equivalences: one with respect to source settings \mathbb{E}_S ,

$$\sum_{s=1}^{|\mathbb{V}_M|} \sum_{j=1}^{|\mathbb{S}|} \alpha_{[s|S_j]}^b[s|S_j] = \sum_{s=1}^{|\mathbb{V}_M|} \sum_{j=1}^{|\mathbb{S}|} \beta_{[s|S_j]}^b[s|S_j] \quad (2.19)$$

and the other with respect to measurements \mathbb{E}_M . We assume that there is a list of instructions which assigns a probability distribution $p(s, m|S, M)$. The set of all conditional probabilities arising in such a way, we call a behavior B of the scenario \mathbb{B} . \diamond

The definition of a prepare-and-measure scenario takes into account the works of (Duarte and Amaral, 2018), (Hermens, 2019) and (Leifer, 2014). We have then a complete prescription of an operational prepare-and-measure scenario, which will be fundamental for our future considerations.

In our definition of a prepare-and-measure scenario it is also important to notice that the new procedures defined by convex combinations of the elements in \mathbb{M} , for example, are not defined in the scenario, but are only *hypothetical procedures* that could be performed (Chaturvedi et al., 2020). In this way we do not include the entire set of convex combinations of procedures in the prepare-and-measure scenario, which is a relevant specification for treating the numerical problems in chapter 3, since the convex hull of procedures would have an infinite number of new procedures.

Another important aspect in our scenarios is that we consider only the convex combinations that are fixed by the pair of weights $\gamma_P := (\alpha, \beta)$. Although this is not necessarily the most general definition of scenario that could be proposed, because we could define the operational equivalences in terms of equivalent classes, this also constitutes a specific choice to be made in order to study numerically some scenarios. This is why we refer to the scenarios \mathbb{B} as finitely defined, with fixed operational equivalences.

Example 6 (Simplest scenario). We will construct the simplest possible scenario in this framework, following (Pusey, 2018) and (Schmid et al., 2018). This scenario constitutes of four preparation events $P_1 = [s_1^1|S_1], P_2 = [s_2^1|S_1], P_3 = [s_1^2|S_2], P_4 = [s_2^2|S_2]$. The probabilities $p(s|S) = 1/2$ for all $s \in S$ and all S . We also have two binary measurements that are deterministic in its outcomes, $M_1, M_2 \in \mathbb{M}$, and $\mathbb{O}_{M_i} = \{0, 1\}$. We have no operational equivalences between measurements but we do have one operational equivalence between the complete coarse-graining of preparation procedures, that we write:

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 = \frac{1}{2}P_3 + \frac{1}{2}P_4 \quad (2.20)$$

corresponding to the vector (since $|\mathbb{E}_S| = 1$):

$$(\alpha, \beta) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4; \beta_1, \beta_2, \beta_3, \beta_4) = \left(\frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2} \right),$$

hence, the simplest scenario will be defined as

$$\mathbb{B}_{si} := \left(\{S_1, S_2, S_3, S_4\}, \{M_1, M_2\}, \{0, 1\}, \{s_1^1, s_2^1, s_1^2, s_2^2\}, \emptyset, \left\{ \left(\frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2} \right) \right\} \right) \quad (2.21)$$

or simply,

$$\mathbb{B}_{si} = \left(\{P_1, P_2, P_3, P_4\}, \{M_1, M_2\}, \{0, 1\}, \emptyset, \left\{ \left(\frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2} \right) \right\} \right). \quad (2.22)$$

It is also common in the literature, such as (Chaturvedi et al., 2020), that for brevity we write simply $\mathbb{B}_{si} = (4, 2, 2, \emptyset, \mathbb{E}_P)$ whenever everything in this notation is clear. \triangle

This scenario is specially important for our purposes, since is the simplest scenario allowing for quantum contextuality for preparation procedures, as we prove in the next chapter. But also, this same structure of preparation contextuality will appear in other experimental scenarios, see chapter 5. It is in general the scenario used as *safety checks* regarding developments in numerical estimation (Ambainis et al., 2016), or application of new tools such as the recently described ones of (Chaturvedi et al., 2020), where the authors studied \mathbb{B}_{si} and variations of this scenario by varying the convex weights (α, β) . The simplest scenario is also important from the point of view of Bell inequalities since

the scenario \mathbb{B}_{si} is isomorphic to the Bell scenario of CHSH⁹, see (Schmid et al., 2018) for a detailed discussion of this aspect.

Vector Characterization

As a side note, whenever our scenario is characterized by finite sets of the primitive operations, it is possible to define an order between the conditional probabilities. This defines a $n = |\mathcal{O}_M| \times |\mathcal{M}| \times |\mathcal{P}|$ dimensional vector $\vec{B} \in \mathbb{R}^n$. The notation $|\cdot|$ stands for cardinality of a set.

$$\vec{B} := (p(1|M_1, P_1), \dots, p(|\mathcal{O}_M||M_{|\mathcal{M}|}, P_{|\mathcal{P}|})) \in \mathbb{R}^n, \quad (2.23)$$

Since every p is a probability distribution we must have that a *generalized behavior*, i.e., any general element $\vec{B} \in C(\mathbb{B}) \subset \mathbb{R}^n$, needs to satisfy the following set of constraints

$$\sum_k p(k|M_i, P_j) = 1, \quad p(k|M_i, P_j) \geq 0, \quad (2.24)$$

and the equations associated with the operational equivalences defined in the specific scenario where \vec{B} lives. These equations are the H -representation for the polytope of the behaviors in \mathbb{R}^n .

2.2 Ontological models

The operational framework cannot provide a physical description of the *reality* embedded in experiments. As far as an abstract framework, it can only provide insights into the experiments: what the experimentalist can or cannot do. The operationalist perspective only provides lists and answers to questions of the form: *how are the probabilities describing the state of a system?* An ontological model is of fundamental difference; the main idea of an ontological model is to provide answers of the type: *why some operational theory prescription is experimentally successful?* But not only this, an ontological model supposes that there are aspects of reality, in a small part of nature (system), that have their existence *independently* of experimental verification (Spekkens, 2005, sec. II, pg. 2), having existing attributes, representatives of the reality as a thing.

The ontological model description of a system supposes the existence of states λ that fully characterize the representing the system. Here, λ is not a hidden variable on its own but an *ontic state*. Ontic states are the complete state of affairs concerning the

⁹This acronym stands for the experimental scenario described by Clauser-Horne-Shimony-Holt in (Clauser et al., 1969). In fact, the discussions from (Schmid et al., 2018) allow the conclusion that any scenario with one operational equivalence, such as the one described in the simplest scenario, is isomorphic to some Bell scenario. For a description of Bell scenarios see (Amaral and Terra Cunha, 2018) and (Santos, 2018).

system. Considering that a system exists independently of the way with respect to the experiments are performed, an ontological model hypothesizes about a complete set of variables that can fully describe and specify the state of the system. A *hidden variable* is not necessarily something that is experimentally hidden, but something that is imposed to specify the ontic state completely.

The philosophical notions and discussions on the ontological concept are presented in appendix A. Here we will consider the mathematical definitions connected to such philosophical ideas.

Definition 9. Every physical system has associated with it a set Λ of possible ontic states $\lambda \in \Lambda$. Each ontic state fully describes the system. \diamond

Since an ontological model provides meaning to operational probabilities, we have to understand how the operational approach and the ontological description are mathematically connected. Each operational primitive has an ontological counterpart. We will present here the ontological models' framework, following (Hermens, 2019) and (Leifer, 2014). An ontological model constitutes then constructing a pair (Λ, Σ) and the ontological counterparts for the operational primitives (Π, Θ) .

Definition 10. Let $B \in \mathbb{B}$ be a behavior described by the conditional probabilities $\{p(k|M_i, P_j)\}$. An ontological model that tries to explain the probabilities obtained by the operational framework is constituted by of a measurable space (Λ, Σ) were Λ is the set of all ontic states, and Σ is a σ -algebra related to the events of the system being in the ontic state λ , and also of a pair (Π, Θ) defined as the set of ontological counterparts of the operational primitives.

The set Π corresponds to the set of probability measures μ_{P_j} associated with the preparation procedures $\{P_j\}_{j \in J}$ in the operational framework. Hence, for each system that is prepared according to P_j , the ontic state λ is prepared and its epistemic state, that we interpret as the best prediction about the preparation of λ , will be some $\mu_{P_j} \in \Pi_{P_j} \subset \Pi$, where we allow the set Π_{P_j} to have several different possible measures. Each triple $(\Lambda, \Sigma, \mu_{P_j})$ constitutes a probability space.

The set Θ is the set of measurable functions, such that for every M there exists $\Theta_M \subset \Theta$ that is a subset of functions ξ_M associated to the measurement events of that operational description. For some state $\lambda \in \Lambda$ of the system, the probability that a measurement event $[k|M]$ occurs is given by $\xi_{[k|M]}(\lambda) = \xi(k|\lambda, M)$. We require that these functions are normalized with respect to the outcomes $k \in K$,

$$\sum_k \xi_{[k|M]}(\lambda) = 1, \forall \lambda \in \Lambda. \quad (2.25)$$

Finally, the probabilities of the operational framework are described in the ontological

model as

$$\forall [k|M_i], \forall P_j, p(k|M_i, P_j) = \int_{\Lambda} \xi_{[k|M_i]}(\lambda) d\mu_{P_j}(\lambda). \quad (2.26)$$

◇

We have then that our data $B \in \mathbb{B}$ admits an ontological model whenever the quadruple $(\Lambda, \Sigma, \Pi, \Theta)$ as in definition 10 exists, providing an ontological explanation for the behavior B . The interpretation of μ_{P_j} is important since these are intended to be only epistemic states, hence, they do not need to represent deterministic responses of the system to the ontic state λ , but instead, they represent only the amount of (probabilistic) knowledge we can have of the state λ via the model.

Given equation (2.26), we say that there exists an ontological model for the behavior. Changing $\mathbb{P} \rightarrow \mathcal{P}$, $\mathbb{M} \rightarrow \mathcal{M}$ represent that this ontological model could represent any procedure allowed in the entire *theory*: every possible procedure in the operational framework (Kunjwal, 2019, pg. 7).

Definition 11. Let $(\Lambda, \Sigma, \mu_{P_j})$ be a probability space. When Λ is a set of finite cardinality we consider $\Sigma := 2^\Lambda$ and we simply write $\mu_{P_j}(\{\lambda\}) \equiv \mu_{P_j}(\lambda)$ for the probability distribution μ_{P_j} , for all $\lambda \in \Lambda$. ◇

In the work of (Spekkens, 2005), the author provides the definitions by considering μ_{P_j} the probability density functions over the set of ontic states Λ . In (Kunjwal and Spekkens, 2015), the authors notice that for finitely defined prepare-and-measure scenarios, the set Λ always has finite cardinality, and therefore he only considers a discrete version of definition 10. In (Morris, 2009) and (Hermens, 2019), the authors notice that there exists a formal and general approach for defining the ontological description for the operational primitives for arbitrary sets Λ : for the measurement procedures, we associate Markov kernels, or also, fuzzy indicator functions; and for the epistemic states, they construct measures μ associated with each preparation procedure.

2.3 Generalized contextuality

The notion of contextuality in Spekkens formulation that we denote as generalized contextuality will be a hypothesis based upon the fact that one has equivalent ways of describing the probabilities obtained in the operational description. These equivalences are the operational equivalences of definition 5. They say that the implementations give rise to the same probability distributions, as far as the operational theory can tell.

Within an operational framework, noticing that there are no procedures that can be done to properly prescript operational differences between two procedures, say P_1 and P_2 , this should mean that these procedures are equivalent descriptions of physical

reality, assuming some sort of Leibniz principle (Spekkens, 2019)¹⁰. We will discuss this principle later with detail. However, note that here it is *fundamental* that no matter what measurements we perform, these measurements cannot distinguish between one or the other, concerning the way the system behaves, given by the set B . We define the generalized notion of contextuality as follows:

Definition 12. Given two different measurement events $[k|M], [k'|M']$ we consider the measurement noncontextuality hypothesis as,

$$[k|M] \simeq [k'|M'] \implies \xi_{[k|M]} = \xi_{[k'|M']} \quad (2.27)$$

◇

Definition 13. Given two different preparation events $[s|S], [s'|S']$ we consider the preparation noncontextuality hypothesis as,

$$[s|S] \simeq [s'|S'] \implies \mu_{[s|S]} = \mu_{[s'|S']} \quad (2.28)$$

For two equivalent events $[T|S_T], [T|S'_T]$ the noncontextuality hypothesis reads as

$$[T|S_T] \simeq [T|S'_T] \implies \mu_S = \mu_{S'} \quad (2.29)$$

Hence, for preparation procedures we will always write, when $P \simeq P'$

$$P \simeq P' \implies \mu_P = \mu_{P'}. \quad (2.30)$$

◇

Notice that in definition 5 all the operational equivalences are defined with respect to the sets \mathcal{M} and \mathcal{S} meaning that the noncontextuality hypothesis is taking into account the fact that no other primitive of the operational theory can be used in order to differentiate between the equivalent operations. We can also define what we mean by a *context* in the Spekkens formalism, such that we can compare it later with other notions of contextuality. We proceed here in the spirit of (Kunjwal, 2019).

Definition 14. A *generalized-context* is every distinction of labels between equivalent operational quantities. ◇

Example 7. Take $P_1 \simeq P_2$. This means that P_1 and P_2 are operationally equivalent, but since we make a distinction of labels between those two preparations, $P_1 \in \mathcal{C}_1^S$ and $P_2 \in \mathcal{C}_2^S$ for two different Spekkens-contexts \mathcal{C}_1^S and \mathcal{C}_2^S . △

¹⁰Or maybe also supposing some instance of a fine tuning argument. For a discussion of fine tuning arguments see (Landsman, 2015).

Example 8. For quantum theory, every state ρ represents a context because it represents an equivalence class of prepared states. Hence, we denote \mathcal{C}_ρ^S as the Spekkens-context associated with the class ρ . We then call the model contextual if there are two different contexts $\mathcal{C}_1^S, \mathcal{C}_2^S$ such that $\mu_\rho(\Omega|\mathcal{C}_1^S) \neq \mu_\rho(\Omega|\mathcal{C}_2^S)$ for any $\Omega \in \Sigma$. \triangle

Leibniz principle of indiscernibles

Although in the remaining chapters of this work we will not make direct use of the discussions present in this section, we find helpful to present an argument regarding the relevance of Leibniz principle, for different reasons: It is important for the completeness of our work; a dissertation about generalized contextuality that does not mention Leibniz principle is unimaginable, given the importance it plays for the contextuality literature. It is also relevant for new students that might read these notions for the first time, but that have no background on the connections between foundations of quantum theory and philosophy of science. Overall, this principle is constantly at stake when utilizing the contextuality hypothesis, posing itself as *the* essential methodology used for treating generalized contextuality in a generalized probability theory framework.

The notion of noncontextuality has its philosophical background in the work of Leibniz. His work attempted to prescribe, among other things, several general principles of knowledge. In (Leibniz, 1765), the author develops a complicated net of criticism to (Locke, 2001) and defines the notions of *innate ideas*. The perception of reality by sensations, the complexity of human thoughts, some in-depth discussions about essence of God, among many others, are present topics in (Leibniz, 1765). In his writings, Spekkens, in (Spekkens, 2019), argues that among all these Leibniz principles, one is of relevance for constructing physical theories: the principle of indiscernibles.

Leibniz had at least four principles in his philosophy. The principle of reason, the principle of the best, the principle of continuity, and the principle of the indiscernibles. This last principle concerns an axiomatic property of real objects. The idea is that he wanted to adequately define what is meant for two objects to be identical, as much, two objects are said to be indiscernible if and only if for every property that one object obeys, the other will also obey. We can write this idea as the equation

$$\forall F : (Fx \iff Fy) \implies x = y \quad (2.31)$$

where, given two objects x and y , no matter what property F is satisfied by x then y must also satisfy the same property, and vice-versa; we then conclude that the two objects are one and the same thing $x = y$.

We build the notion of Spekkens' noncontextuality, noticing that we can rewrite the principle of indiscernibles as follows: if two objects simultaneously have the same property - operational equivalent -, then these two objects must be indiscernible from one

another - at the ontological model. Indiscernible objects at the operational level should be explained in the same way at the ontological level.

Such a philosophical notion is actually so powerful that in (Spekkens, 2019) it is proposed as a methodological principle for theory building. Several essential points about the Spekkens notion of noncontextuality are present in (Spekkens, 2019). We focus here in the answer for the immediate criticism towards such methodology: It seems that such a method is based on some strong notion of empirical indiscernibility, that would be tied to momentary technological capabilities or even tied to the imaginative human capacity to propose better explanations, but not tied to truly intrinsic aspects of the physical system. Spekkens response to such criticism is that this notion of empirical indiscernibility is based on the generality of viable procedures - as well as viable ontological explanations - that are physically possible, and *not* to only those that are executable by human experimentation (in this dissertation, this reflects the difference between the sets \mathcal{M} and \mathbb{M}).

“Some commentators suggest that the Leibnizian methodological principle, though useful as a means of selecting among competing ontological theories, should not be taken as a constraint on theory construction. This is the position that is suggested by Maudlin’s discussion of the Newton-Leibniz debate. (...) Although he opines that “*one should be made at least uncomfortable by the postulation of empirically inaccessible physical facts*”, and consequently that “[*other things being equal*], *one would prefer a theory without them*”, he nonetheless grants that theories violating the Leibnizian methodological principle are still viable on the grounds that “*Man is not a measure of all things, and there is no reason to believe that all real properties must fall within the power of human observation.*” I do not find this argument persuasive because the Leibnizian methodological principle does not appeal to a parochial kind of empirical indiscernibility, judged relative to the particular in-born capabilities of humans or their particular technological capabilities at a given historical moment, but rather to the *in-principle* variety of empirical indiscernibility. This variety of indiscernibility must be understood as indiscernibility for *any* system that might be considered an agent within the universe. This is because (...) the only in-principle limits to human capabilities are the limits imposed by physics, and therefore the only limits on our capabilities are the limits on the capabilities of any system embedded in the universe and subject to its physical laws.” (Spekkens, 2019, Sec. 3, pg. 7)

We also note that, for Spekkens, this kind of observational indiscernibility leads to both an improvement to the ontological view, such as its understanding, and towards the proposal of *new* ontological models. However, instead of arguing that such (methodology) notion of noncontextuality implies that the quantum theory should be contextual,

implying that this principle does not lead to a better understanding to the physical reality of the quantum processes, Spekkens takes the position of maintaining the principle as a methodological tool and questioning the ontological models' framework itself.

“(...) skeptics of the principle could take the fact that quantum theory *does not* admit of a generalized-noncontextual ontological model as evidence *against* the principle of generalized noncontextuality and hence also as evidence against the Leibnizian methodological principle. To see that the argument need not lead one to having scepticism of the principle, it suffices to note that generalized noncontextuality is not the only assumption of the no-go theorem and consequently that one can preserve generalized noncontextuality (and hence the Leibnizian methodological principle) by giving up a different assumption. In my opinion, this is in fact the right attitude to take towards the no-go result. Specifically, I believe that it is the framework of ontological models that must be abandoned, and that it is a fruitful research program to seek an alternative to this framework that provides causal explanations of quantum statistics while strictly respecting the Leibnizian methodological principle.” (Spekkens, 2019, pg. 12)

It is also worth pointing that, in the light of ψ -epistemic considerations (a notion presented in several works (Spekkens, 2007), (Leifer, 2014) or also (Liang et al., 2011)) the usage of Leibniz methodological principle in the construction of the generalized notion of noncontextuality is based more upon an indistinguishability of the *knowledge* about physical responses than the empirical facts themselves.

“(...) to properly understand the principle of generalized noncontextuality as a special case of the Leibnizian methodological principle, it is necessary to reconceive the latter at the level of epistemology. That is, rather than conceiving of it as an inference from the indiscernibility of empirical facts to the identity of ontological facts, one must conceive of it as an inference from the indiscernibility of *states of knowledge about* empirical facts to the identity of *states of knowledge about* ontological facts.” (Spekkens, 2019, pg. 12)

Providing a physical hypothesis, the generalized noncontextual one, as a particular instance of the Leibnizian principle but equipped with a ψ -epistemic quantum foundational consideration, one constructs ontological descriptions for experimental investigation. Notice that in general, the literature considers the noncontextuality hypothesis as the arrows from definitions 12 and 13. From the above discussion, a stronger version of noncontextuality would be that equivalence in the ontological description, for all possible models, would imply operational equivalence at the operational level. This philosophical definition is known as indiscernibility of the identicals.¹¹ We will refer to this notion as

¹¹We can write this as the equation $x = y \implies \forall F : (Fx \iff Fy)$.

Leibniz-noncontextuality:

Definition 15. An ontological model is *Leibniz-noncontextual* if, for equivalent operational procedures, we have

$$P \simeq P' \iff \mu_P = \mu_{P'},$$

and similarly for measurement and preparation events. ◇

This means that equivalent operations are ontologically equivalent and that whenever two procedures have the same ontological explanations, they cannot be distinguished by operational procedures. Such a notion of noncontextuality was considered in (Lillystone et al., 2019). This notion of Leibniz-noncontextuality is not equivalent to new notions of classicality described by *Leibnizianity*. The notion of Leibnizianity is a recent notion regarding equivalences in causal-inferential models, for an introduction of these ideas see (Schmid et al., 2020d).

NONCONTEXTUALITY INEQUALITIES

What is proved by impossibility proofs is lack of imagination.

- (Bell, 1982, pg. 997)

Historically the discussion of classicality that emerged from the famous Einstein-Bohr debate (Landsman, 2005), arises from purely conceptual discussions, such as those present in (Einstein et al., 1935), to theoretically concrete grounds in the work of (Bell, 1964) by means of setting better logical definitions in the discussion, and from these definitions making use of the non-signaling hypothesis to build *inequalities*. Such seminal work represented an enormous breakthrough that converged to strong abstract results, opened a railway of experimental challenges, and nowadays, some of these experimental developments are already technological realities, for example in entanglement theory (Yin et al., 2020), (Ren et al., 2017), (Liao et al., 2017). In the field of quantum information, it is vital to develop both no-go theorems that will represent possible advantages of quantum-over-classical systems, and robust inequalities that can be experimentally verified and used in new technologies.

Following such a successful program, in this chapter we study noncontextuality by showing abstract results in the form of no-go theorems in section 3.1. Then we proceed to construct a full linear characterization of the noncontextuality scenario and develop all noncontextuality inequalities for general finite prepare-and-measure scenarios in sections 3.2 and 3.3. These inequalities are robust, meaning that they can be experimentally tested. We conclude with a numerical implementation of such construction. In chapter 4 we will make use of the resource theory framework to draw some conclusions of general polytope descriptions for prepare-and-measure scenarios. In chapter 5 we make use of noncontextuality inequalities obtained in this chapter to prove quantum advantages.

3.1 No-go theorems

The result of Bell's theorem fits into a class of abstract statements known as no-go theorems (see appendix A). Specifically, Bell's theorem implies that there can be no hidden-variable model that, simultaneously, respects the criteria present in chapter 1¹ and predicts the results of quantum mechanics. Not only that, but actually, Bell's theorem does not allow for an epistemic account of pre-determined values, meaning that although the values associated with measurement procedures have an *a priori* existence, the experimenter at best obtains partial knowledge of reality - epistemic states of reality.² Historically, Bell has constructed such no-go theorem providing an inequality that needs to be respected under these assumptions. But we should address that what goes under the name of *Bell inequalities* in the modern literature should be understood more as *robust Bell-type inequalities* and the inequalities that Bell first presented represent, in general, a no-go result (Schmid and Spekkens, 2018). The main difference is that Bell inequalities should be experimentally testable and, hence, make no reference to idealizations or the quantum formalism. In this work, we follow such convention and only consider robust non-contextuality inequalities. Kochen-Specker noncontextuality inequalities (Araújo, 2012), (Araújo et al., 2013) are, then, understood as certain types of no-go results such as Bell inequalities, since there are strong idealizations at play.

No-go results are fundamental results setting the grounds for two classes of further investigations: First, by drawing attention towards possible experimental resources, and second, by introducing new competing foundational ideas. For example, we have several competing notions of noncontextuality (generalized, extended, Kochen-Specker), such as competing ways for how to construct useful resource theories (for Bell nonlocality, as an example, we have competing ideas between two resource theories: one considering LOCC³ operations and the other considering LOSR⁴ operations, see (Schmid et al., 2020a)). Research in foundations of quantum theory usually uses no-go results as essential breakthroughs, such as Spekkens' proposal of abandoning the ontological models' framework, or at least, to state sound and thoughtful physical principles that would increase our comprehension of the theory.

As stated before, the generalized noncontextuality paradigm introduces new features for a system to produce *classical* probability distributions; preparation noncontextuality. Preparation noncontextuality is an important novelty, with respect to KS-noncontextuality. It could be the case that quantum theory is noncontextual regarding preparations, since the Kochen-Specker theorem only treats measurement procedures.

¹**Realism, Locality and Free will.**

²This is sometimes known as Bell's theorem for stochastic hidden variables, see (Hermens, 2010).

³Local Operations and Classical Communication.

⁴Local Operations and Shared Randomness. The difference is that instead of communicating the parties are only supposed to share a common source of classical variables.

The following no-go results will prove the opposite.

Some assumptions over noncontextual ontological models

There are two critical considerations heavily used in the demonstrations that follow: *Distinguishable procedures* and *Operational convexity implying ontological convexity*. The status of these features about ontological models can be considered as *definitions* for what preparation noncontextuality in the Spekkens sense means, or more generally, of what ontological models of operational theories should obey. Models that do not obey such characteristics, such as Meyer-Kent-Clifton models (Hermens, 2011)⁵ can then be viewed as preparation noncontextual models. These were first introduced in (Spekkens, 2005).

Definition 16 (Distinguishable procedures). Let $B \in \mathbb{B}$ be a behavior. If two preparation procedures, $P, P' \in \mathbb{P}$ are *distinguishable with certainty* in a single-shot measurement, then their associated probability distributions μ_P and $\mu_{P'}$ are non-overlapping:

$$\mu_P(\Omega)\mu_{P'}(\Omega) = 0, \quad \forall \Omega \in \Sigma, \quad (3.1)$$

for any noncontextual ontological model constructed for B . ◇

We make use of this definition in appendix B. Figure 3.1 clarifies the idea; suppose that one has two preparation procedures that are distinguishable and with associated probability distributions overlapping in a small region around $\lambda_0 \in \Lambda$. When the ontic state falls within the overlapping region, one could never distinguish between the two preparations since they both assign nonzero values to the same ontic state. Hence the distributions must satisfy the criteria (3.1); otherwise, the procedures could not be distinguishable.

⁵And references therein for an account of MKC-models. See also (Hermens, 2010) for a comprehensive discussion on the topic of nullifying the Kochen-Specker theorem. The original articles to propose noncontextual models for quantum theory, from arguments based on the finite precision loophole, were (Meyer, 1999) and (Clifton and Kent, 2000).

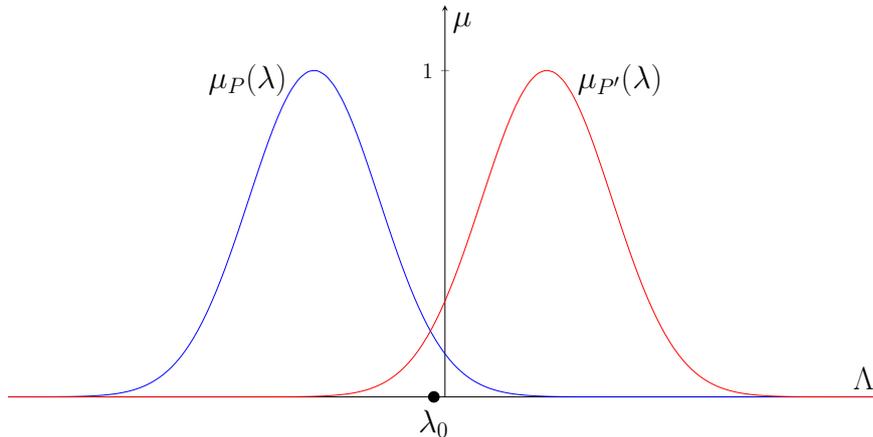


Figure 3.1: Overlapping probability distributions over ontic space Λ . If the ontic state is in the overlapping region a single-shot measurement could not distinguish between the two preparations. Here we abuse notation and denote the probability densities over $\Lambda = \mathbb{R}$ with the same symbol as the probability measures.

This discussion is related with the ψ -ontic/ ψ -epistemic debate over the reality of the quantum states ψ . We say that an ontological model for quantum theory is ψ -ontic whenever *any* two different pure quantum states $P_1 \rightarrow \rho_1, P_2 \rightarrow \rho_2, \rho_1 \neq \rho_2$, are also ontologically distinct. As stated in (Leifer, 2014), this definition captures the idea from figure 3.1, where distinct pure states have non-overlapping probability measures. Here, we suppose that definition 3.1 is always valid for noncontextual ontological models.

Definition 17 (Convex implies convex). A convex combination of procedures (preparations or measurements) is represented within an ontological model by the associated convex combination of probability distributions. \diamond

Such aspect is of direct importance given the formalism we introduced in chapter 2 where preparations $P_{[s|S]}$ correspond to source events $[s|S]$ occurring with probability $p(s|S)$. In the ontological model this should correspond to $p(s|S)\mu_{[s|S]}(\lambda)$. This means that:

1. We assume that the coarse-graining represented in the operational framework should be represented equivalently in the ontological model description.
2. We allow probabilistic mixtures of procedures to define new procedures within \mathcal{M} and \mathcal{S} , in the sense that the operational theory is convex, in the sense of definition 17.

In the class of operational mixtures we highlight coarse-grainings because of their experimental relevance.

Definition 18 (Coarse-graining of measurements). Let M be the coarse graining of \tilde{M} , represented by

$$[k|M] = \sum_{\tilde{k}} p(k|\tilde{k})[\tilde{k}|\tilde{M}] \quad (3.2)$$

defined for all k, \tilde{k} such that $p(k|\tilde{k}) \in \{0, 1\}$ and $\sum_k p(k|\tilde{k}) = 1$. Each M constructed in this way corresponds to a *new* procedure. We can represent this coarse-graining in the operational framework as

$$\forall [s|S], s \in \mathbb{V}_S, S \in \mathcal{S} : p(k, s|M, S) = \sum_{\tilde{k}} p(m|\tilde{k})p(\tilde{k}, s|\tilde{M}, S) \quad (3.3)$$

and in the ontological model as,

$$\xi_{[k|M]} = \sum_{\tilde{k}} p(k|\tilde{k})\xi_{[\tilde{k}|\tilde{M}]}. \quad (3.4)$$

◇

Notice how the convexity of operational quantities is translated into the ontological prescription by means of such a coarse-graining notion. This is an assumption consistent with quantum theory since the linearity present in the Born rule implies that convex operations are represented by convex quantum prescriptions. But we note that procedure M is not operationally equivalent with respect to independent measurement events, M defines a new procedure in equation (3.2). We assume the two procedures are then related at the operational level by equation (3.3), and equation (3.4) follows from convexity at the ontological level. These definitions do not suppose operational equivalent procedures, according to (Kunjwal, 2019).

Definition 19 (Coarse-graining of preparations⁶). Let P_j be the coarse-graining of $\{\tilde{P}_j\}_{j \in J}$, represented by

$$P_j = \sum_{\tilde{j}} p(j|\tilde{j})\tilde{P}_{\tilde{j}} \quad (3.5)$$

such that for all j, \tilde{j} : $p(j|\tilde{j}) \in \{0, 1\}$ and $\sum_j p(j|\tilde{j}) = 1$. Each P_j constructed in this way constitutes a *new* procedure. We represent this in the operational framework as

$$\forall [m|M], m \in \mathbb{O}_M, M \in \mathcal{M}, p(m|P_j, M) = \sum_{\tilde{j}} p(j|\tilde{j})p(m|\tilde{P}_{\tilde{j}}, M) \quad (3.6)$$

and in the ontological model as

$$\mu_{P_j} = \sum_{\tilde{j}} p(j|\tilde{j})\mu_{\tilde{P}_{\tilde{j}}}. \quad (3.7)$$

◇

⁶Here we abandon the notation present in (Kunjwal, 2019) for clarity and for the following developments of the chapter. For a definition in terms of preparation events see equations (16) and (17) of (Kunjwal, 2019).

Hence, definition 17 implies the following: Given a preparation procedure P that can be understood as the probabilistic mixture of $\{P_1, P_2\}$ by means of

$$P = pP_1 + (1 - p)P_2, \quad p \in [0, 1]$$

with the interpretation that P corresponds to preparing P_1 with probability p and P_2 with probability $1 - p$, we get that such preparation P is a *new* preparation, defined by means of $\{P_1, P_2\}$, that can be represented in the ontological model framework as

$$\mu_P = p\mu_{P_1} + (1 - p)\mu_{P_2}, \tag{3.8}$$

Definition 17 is of fundamental importance for the proof of contextuality and for experimental tests related to it. Actually, such a construction is not only fundamental for the notion of Spekkens contextuality but also, an ontological model that respect coarse-graining relations is also fundamental for the notion of KS-noncontextuality (Kunjwal, 2019, Appendix B, pg. 37).

Non-convexity at the ontological model implies failure of KS-contextuality

Consider the following criticism, (Hermens, 2011): Suppose we translate (3.8) into preparations of quantum states, and we prepare state ρ using $\{p, 1 - p\}$. Hence $\rho = p\rho_1 + (1 - p)\rho_2$, but, is ρ operationally equivalent to $p\rho_1 + (1 - p)\rho_2$? If we cannot control the random bit $\{p, 1 - p\}$, there can be no measurement capable of distinguishing between these two, but, if we have control, the situation is not the same; in other words, if the experimenter has control over the bit, he could check the value of the random variable, therefore distinguishing between these two preparations. Notice that this is not a criticism towards convex operational theories, like general GPT's, but towards the ontological identification of such convexity.

We could understand this as a criticism of the notion of generalized contextuality. Supposing that we do not have control over the bit $\{p, 1 - p\}$ would imply a contradiction, as we see above. On the other hand, suppose that we do have control over the bit, meaning that we knew that there is a weighted coin inside the preparation box of figure 2.3. Once we open the box, we could distinguish between the procedures, and therefore we would not have contextuality. Although we do not consider this debate finished, we present two different appealing responses.

In (Leifer, 2014, section 5.3, pg. 91-92), the author considers reasonable such a definition as 17 because the randomness present in the variable $\{p, 1 - p\}$ can be independent of the system under investigation, so that the ontic states would not need to have any dependence with the bit. Another appealing response is the one presented by Ravi Kunjwal in terms of the Kochen-Specker theorem.

Theorem 1. (*Kunjwal, 2019*) *If operational convexity does not imply ontological convexity, in the terms of definitions 17, and 18, then, the Klyachko-Can-Biniciöglu-Shumovsky scenario⁷ assumes a KS-noncontextual ontological model.*

Proof. We will follow the proof from (*Kunjwal, 2019*). Let $\mathbb{M} := \{M_i\}_{i=1}^5$, each measurement procedure heaving three outcomes $k \in \{0, 1, 2\} =: K$.

Now, suppose that M_i , for each $i \in \{1, \dots, 5\} = I_5$, corresponds to a coarse-graining of measurement procedures M'_i and M''_i , and

$$[0'|M'_i] \equiv [0|M_i], \tag{3.9}$$

$$[1'|M'_i] \equiv [1|M_i] + [2|M_i], \tag{3.10}$$

$$[0''|M''_i] \equiv [2|M_i], \tag{3.11}$$

$$[1''|M''_i] \equiv [0|M_i] + [1|M_i]. \tag{3.12}$$

Each binary-outcome measurement procedure M'_i, M''_i constitutes, in itself, *new procedures*⁸. These coarse-graining operations can be understood, operationally, as that for any preparation procedure $P \in \mathbb{P}$, the following set of equations is true,

$$p(0'|M'_i, P) \equiv p(0|M_i, P), \tag{3.13}$$

$$p(1'|M'_i, P) \equiv p(1|M_i, P) + p(2|M_i, P), \tag{3.14}$$

$$p(0''|M''_i, P) \equiv p(2|M_i, P), \tag{3.15}$$

$$p(1''|M''_i, P) \equiv p(0|M_i, P) + p(1|M_i, P). \tag{3.16}$$

And, as an assumption given by the theorem, we do not impose that these relations are respected for the ontological model constructions, i.e., for the functions $\xi_{[\cdot | M_i]}(\lambda)$, given $\lambda \in \Lambda$. The operational equivalences for the scenario KCBS are, $[0''|M''_i] \equiv [0'|M'_{i+5}]$ for all $i \in I_5$, with $+_5$ addition module 5. A Kochen-Specker noncontextual ontological model requires that we have,

$$\chi_{[0''|M''_i]}(\lambda) = \chi_{[0'|M'_{i+5}]}(\lambda), \forall \lambda \in \Lambda. \tag{3.17}$$

Where we use represent the functions $\xi \equiv \chi$ since for KS-models we have outcome determinism as an ontological requirement. An ontological model for the KCBS scenario must therefore specify the response functions associated with the measurement events in question.

⁷This is the experimental scenario described first in (*Klyachko et al., 2008*). See also (*Cabello et al., 2014*) for a description in terms of the graph approach to Kochen-Specker contextuality. The KCBS scenario is associated with the 5-cycle graph, and represents an example of existence of quantum correlations that *violate* KS-noncontextuality inequalities.

⁸See definitions 3.2.

Since there are no constraints from coarse-graining relations that must be respected by the functions χ we can assign arbitrary functions for $\{M'_i, M''_i\}_{i \in I_5}$, in such a way that these functions will respect only the KS-noncontextuality condition,

$$\forall \lambda \in \Lambda, \chi_{[0''|M'_i]}(\lambda) = \chi_{[0'|M'_{i+5}]}(\lambda). \quad (3.18)$$

Remember that for Kochen-Specker, the functions χ are idempotent. We can, therefore, construct an ontological model that is KS-noncontextual for the KCBS scenario, given that the functions χ do not respect the same coarse-graining relations of their associated operational events. \square

We think that this theorem is an interesting response to the criticism mentioned above, regarding definition 17. This is so because, supposing that we do not have convexity at the ontological level, many results that are already known from KS-noncontextuality would not be valid.

As it is demonstrated in (Morris, 2009), convexity for ontological models is *not* a fundamental restriction that a model should satisfy to represent a physically relevant theory. In (Morris, 2009, Section 2.4.3, pg. 39) the author constructs a non-convex ontological model for 2-level quantum system. Therefore, non-convex models are allowed by the definition of ontological models we give in chapter 2. The issue is that constructing a non-convex ontological model for a *convex* operational theory imply strange physical interpretations. Given his non-convex model, the author writes, after some early considerations, the following problematic feature of a non-convex model:

“(…), a non-convex model could potentially imply a dependence of the ontic state on whether or not a preparation takes place within a probabilistic ensemble of preparations or not. This does not seem like a desirable property for an ontological model.” (Morris, 2009, pg. 40)

We can argue that, convex ontological models are a necessary condition for the notion of generalized noncontextuality to hold. Without this feature one can construct noncontextual ontological models for quantum theory, but also heaving to deal with problems surrounding the interpretation concerning such non-convex ontological models. An argument linking convexity at the ontological level with experimental verification’s is still lacking in the literature. In other words, there is still the need that an experimental verification of quantum contextuality is such that no non-convex ontological model could reproduce the gathered statistics.

Preparation contextuality for quantum theory

The most fundamental no-go theorems for our work are the fact that quantum theory is both preparation and measurement contextual.

Theorem 2. (*Spekkens, 2005*) *The maximally mixed state of a qubit is preparation-contextual.*

Interestingly, one can obtain such a notion of preparation contextuality in more general situations. A stronger result that was later shown makes the preparation contextuality (and as a whole, the ontological models' framework) a much more useful concept. The fact that *any* mixed state of a qubit, and not only the maximally mixed one, has generalized contextuality.

Theorem 3. (*Banik et al., 2014*) *Any mixed state of a qubit is preparation contextual.*

Hence, employing these last theorems, we have that quantum theory for finite-dimensional Hilbert spaces has an experimentally testable preparation contextual structure - and this is the most crucial aspect for the construction of resource framework for generalized noncontextuality with applications in quantum information processes - apart from theoretical loopholes such as the the finite precision and the tomography loopholes.

Up to our knowledge, another significant result in the field is still lacking: a more general proof of quantum contextuality considering arbitrary separable Hilbert spaces.

Measurement contextuality for quantum theory

Since we already have a notion of measurement contextuality from (*Kochen and Specker, 1975*), the contribution of generalized contextuality is to obtain a no-go result for unsharp measurements. We have the 'unsharpness' in the proof provided by POVMs. In fact, we see in the proof of measurement contextuality for quantum theory how abstract, i.e., including a relatively wide range of possible generalized contexts, the notion of generalized contextuality is when we consider contexts such as:

“(...) the equivalence class of measurement procedures that contains M^9 also contains the “measurement” procedure \tilde{M} that completely ignores the system and just flips a fair coin to determine the outcome.” (*Spekkens, 2005*, pg. 9)

Where \tilde{M} completely ignores the system, be it quantum or classical, and it still represents a Spekkens-context. Therefore we see how *vast* is the notion of a generalized context. We will not focus on these details in this work, but arguably a fundamental question is whether such a notion for context is too broad. The essential result is the following:

Theorem 4. (*Spekkens, 2005*) *Quantum theory is contextual for unsharp measurements.*

In the proof, we have outcome determinism for sharp measurements but outcome *indeterminism* associated with a generalized quantum measurement: a POVM M . The full proof is given in appendix B.

⁹For the definition of the measurement M that Spekkens is mentioning here see the appendix B

Definition 20. (Morris, 2009) An ontological model respects outcome determinism when the functions $\xi_{[\cdot|M_i]}(\lambda) : \mathcal{O}_M \rightarrow [0, 1]$ associated with the measurements M_i are idempotent: $\xi_{[k|M_i]}^2 = \xi_{[k|M_i]}$. For idempotent effects we will often write ξ as χ . \diamond

Outcome determinism is so relevant for this discussion that we formulate it as an assumption.

Assumption 3. Any ontological model for quantum theory has to satisfy outcome determinism for sharp measurements.

One could criticize the proof of theorem 4 from two aspects (Spekkens, 2005):

1. Both sharp and unsharp measurements should respect outcome-determinism or,
2. that both, sharp and unsharp measurements, can be outcome indeterministic.

Outcome-determinism for unsharp measurements is inconsistent with assuming (generalized) measurement noncontextuality. By considering $M = \{\mathbb{1}/2, \mathbb{1}/2\}$ and generalized noncontextuality this imply that M *must* be represented in an ontological model by $\{1/2, 1/2\}$ which are not idempotent indicator functions, thus cannot be outcome-deterministic (Spekkens, 2005, pg. 12). That is the response given by Spekkens to the first criticism. Now the second one comes in the form of a theorem:

Theorem 5. (Spekkens, 2005) *Preparation noncontextuality implies outcome-determinism for sharp measurements in quantum theory.*

Proof. We will follow both (Schmid and Spekkens, 2018) and (Lostaglio and Senno, 2020). Let M correspond to a sharp measurement procedure, which using quantum theory we refer to as a set of projective measurements $M := \{E_k\}_{k \in K} \equiv \{|k\rangle\langle k|\}_{k \in K}$. Suppose that we have a corresponding basis of pure states $P_j := \rho_j \equiv |j\rangle\langle j|$, for $j \in J = K$. We have therefore that the probabilities $p(k|M, P_j) = \text{Tr}(E_k \rho_j) = \delta_{kj}$, with δ_{kj} the Kronecker delta. We denote these two objects with different notations because they are represented differently at the ontological models level.

Let (Λ, Σ) be the ontic space. To each ρ_j we associate μ_j over Λ , and to M we associate the object $\xi_{[\cdot|M]}(\lambda)$, for all $\lambda \in \Lambda$, as we have seen. Therefore we have that,

$$\int_{\Lambda} \xi_{[k|M]}(\lambda) d\mu_j(\lambda) = \delta_{kj}.$$

It follows then that,

$$\xi_{[k|M]}(\lambda) = \begin{cases} 1, & \lambda \in \text{supp}(\mu_k) \\ 0, & \lambda \in \text{supp}(\mu_{j \neq k}) \end{cases} \quad (3.19)$$

where this is true almost everywhere in the supports.¹⁰ Every ρ_j appears in *some* decomposition of the maximally mixed state $\frac{1}{d}\mathbb{1}$, where d is the dimension of the Hilbert space $\rho \in \mathcal{B}(\mathcal{H}) \simeq_{\mathcal{H}} \mathcal{B}(\mathbb{C}^d)$, with $\simeq_{\mathcal{H}}$ here meaning Hilbert space isomorphic. In a preparation noncontextual ontological model, every such decomposition is associated with the same measure $\mu_{\frac{1}{d}\mathbb{1}}$ over the ontic states. We conclude that, for all $\lambda \in \text{supp}(\mu_j)$ the same ontic state also satisfy $\lambda \in \text{supp}(\mu_{\frac{1}{d}\mathbb{1}})$. Since $\{\rho_j\}_{j \in J}$ is a basis, we get that

$$\frac{1}{d} \sum_{j \in J} \rho_j = \frac{1}{d} \mathbb{1}. \quad (3.20)$$

Preparation noncontextuality implies that,

$$\frac{1}{d} \sum_{j \in J} \mu_j = \mu_{\frac{1}{d}\mathbb{1}}. \quad (3.21)$$

And we might conclude that,

$$\Lambda = \text{supp} \left(\mu_{\frac{1}{d}\mathbb{1}} \right) = \text{supp} \left(\frac{1}{d} \sum_{j \in J} \mu_j \right) = \bigcup_{j \in J} \text{supp}(\mu_j). \quad (3.22)$$

We can then write the (measurable) function $\xi_{k|M}$ as,

$$\xi_{[k|M]}(\lambda) = \begin{cases} 1, & \lambda \in \text{supp}(\mu_k) \\ 0, & \text{otherwise} \end{cases} \quad (3.23)$$

Concluding that this function is now defined for all $\lambda \in \Lambda$, and that $\xi_{[k|M]}^2 = \xi_{[k|M]}$, concluding that we have outcome determinism for the ontological model. \square

Notice that in some sense, one should recover measurement contextuality in the Kochen-Specker perspective if one re-write the Kochen-Specker theorem in the ontological models' framework and assumes outcome determinism of measurements. A demonstration of the KS theorem in the ontological models' framework can found in (Leifer, 2014), (Kunjwal, 2016).

Transformation noncontextuality

Even though we are working with prepare-and-measure scenarios, there are no-go theorems applied to quantum transformations in the generalized formalism:

Theorem 6. (*Spekkens, 2005*) *Quantum theory is contextual with respect to transformation procedures.*

¹⁰This means, for example, that $\xi_{k|M}(\lambda) = 1$ almost everywhere on $\text{supp}(\mu_k)$, in the sense that this is true disregarding sets of measure zero.

We avoid explicit remarks about transformation procedures in the experimental scenarios because we assume a quantum result: The equivalence between Schrödinger and Heisenberg pictures. This means that one can treat either a preparation followed by a transformation $P + T \rightarrow \tilde{P}$ as a new preparation, or a transformation followed by a measurement $T + M \rightarrow \tilde{M}$ as a new measurement. This is possible because the quantum probabilities do not change by changing the pictures: suppose we consider a state prepared in $\rho \in \mathcal{B}(\mathcal{H})$ for some Hilbert space \mathcal{H} . Then the map $T_* : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ is the quantum channel, consisting of a completely positive trace-preserving map; we can think of $T_*(\rho) \in \mathcal{B}(\mathcal{H})$ as the effective preparation, considering the system as viewed in the Schrödinger picture. On the other hand, $T : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ can be viewed as a quantum channel acting on the POVM's, so that for every E_k of some POVM, the map $T(E_k) \in \mathcal{B}(\mathcal{H})$ can be considered as a new POVM element, since $T(E_k)$ is positive-definite for all k and

$$\sum_k T(E_k) = T\left(\sum_k E_k\right) = T(\mathbb{1}) = \mathbb{1},$$

where we required T to be unital¹¹. This is known as the Heisenberg picture. We then just make the following consideration about the maps discussed from the pictures above: $T = (T_*)^*$ ¹². This means that the maps in the Schrödinger and Heisenberg pictures are the adjoint with respect to each other and, since T is the adjoint of a CPTP map, we have that T must be unital. And also, we get that

$$p(k|P, T(M)) = \text{Tr}(\rho T(E_k)) = \langle \rho, T(E_k) \rangle_{HS} = \langle T_*(\rho), E_k \rangle_{HS} = \text{Tr}(T_*(\rho) E_k)$$

where the last term is just $p(k|T_*(P), M)$ and hence both pictures yield the same statistics in the quantum formalism¹³. Note that there are two very important remarks to make here. The first is that the noncontextual polytope can differ for the different pictures, since the equivalences in a prepare-and-measure scenario might change, for different choices of $P + T \rightarrow T_*(P)$ or $T + M \rightarrow T(M)$. Hence, if one constructs the operational description of a prepare-and-measure scenario, it is crucial *not* to change the pictures in later calculations/manipulations of the behaviors. Second is that such a description is firmly based upon quantum theory. Hence, for applications of generalized noncontextuality outside quantum theory, this consideration might also lead to errors in the analysis.

We conclude by noting that there are in the literature discussions of generalized noncontextuality beyond the prepare-and-measure scenarios, e.g., ([Mansfield and](#)

¹¹Such a request covers a lot of quantum transformations such as unitary transformations and decoherence processes ([Binder et al., 2018](#)).

¹²See ([Werner, 2019](#)) or ([Keyl and Werner, 2016](#)). Here we consider T^* as the adjoint with respect to the Hilbert-Schmidt inner product. The space $(\mathcal{B}(\mathcal{H}), \|\cdot\|_{HS})$ is a Hilbert space.

¹³We assume that $p(k|P, T, M) = p(k|T_*(P), M) = p(k|P, T(M))$ given that T is known *a priori*. We need to be extra careful here, since this is *not* necessary when dealing with a more general operational framework without having quantum theory in mind.

(Kashefi, 2018). In some cases, there are quite incredible results that arise from considering transformations and noncontextuality from it (Lostaglio, 2020). In other works, it seems that the notion of transformation noncontextuality is a notion too weak, making quantum theory contextual even for scenarios that were understood to be noncontextual for Kochen-Specker (Lillystone et al., 2019). This might suggest that for a notion of classicality, generalized noncontextuality is general and well suited for quantum theory but maybe too broad for some specific questions, such as being the fundamental resource for powering the speed-up of quantum computations.

3.2 Noncontextual measurement-assignment polytope

Knowing that noncontextual ontological models cannot describe quantum theory, we look now to develop noncontextuality inequalities that provide tools for experimental verification of this fact. We follow (Schmid et al., 2018), in order to present a method that obtains all the noncontextuality inequalities that define the universal noncontextual polytope for a prepare-and-measure scenario. The results of (Schmid et al., 2018) are of profound importance, since for *any* scenario, the full set of noncontextuality inequalities is provided. Obtaining the full set of noncontextuality inequalities for different notions, such as the Kochen-Specker notion, is known to be an extremely difficult task, and only classes of scenarios have a complete characterization of the noncontextual polytope in KS-contextuality.¹⁴

Remember that a behavior B can be understood as a vector $\vec{B} \in \mathbb{R}^n$ with components given by the resulting statistics $p(k|P_j, M_i)$, as stated in (2.23). Following (Schmid et al., 2018) we will ignore the source structure present in definition 8 so that we can write just a simplified description of prepare-and-measure scenarios:

$$\mathbb{B} := (\mathbb{P}, \mathbb{M}, \mathbb{O}_M, \mathbb{E}_P, \mathbb{E}_M), \quad (3.24)$$

so that given a scenario, defined by (3.24), the goal is to find all the inequalities that must be satisfied by a behavior \vec{B} that has a noncontextual ontological model. Hence, our goal is to find a set of inequalities of size $|H|$ such that for all \vec{B} that can be described by an universally noncontextual ontological model, its statistics must obey

$$\sum_{k,i,j} \gamma_{k,i,j}^h p(k|M_i, P_j) + \gamma_o^h \geq 0 \quad (3.25)$$

for all $h \in H$. If such a thing happens we shall say that $\vec{B} \in NC(\mathbb{B})$, where $NC(\mathbb{B})$ is

¹⁴But of course, there are general results of similar relevance such as the Lovász bound for the compatibility graph approach (Amaral and Terra Cunha, 2018). The Kochen-Specker scenarios that can be described by a cycle graph have a full description in terms of tight noncontextuality inequalities.

the noncontextual polytope, i.e, the set of all behaviors that satisfy (3.25). The set of all possible behaviors \vec{B} in the scenario \mathbb{B} , that we will call $C(\mathbb{B})$, also forms a polytope since every behavior must respect (2.24) and the equations that follow from the operational equivalences $\mathbb{E}_P, \mathbb{E}_M$.

$$\vec{B} \in C(\mathbb{B}) : \iff \begin{cases} p(k|M_i, P_j) \geq 0, \forall k, i, j \\ \sum_k p(k|M_i, P_j) = 1, \forall i, j \\ +\text{Operational equivalences} \end{cases} \quad (3.26)$$

We have then, by construction, that

$$NC(\mathbb{B}) \subset C(\mathbb{B}), \quad (3.27)$$

and we will later use the convex structure of these polytopes to develop tools for treating noncontextuality as a resource.

The first step towards (3.25) is to define the linear characterization of the non-contextual measurement-assignment polytope. If we restrict the attention just to the operational description of measurements we note that any noncontextual ontological model representing the statistics, as we defined in chapter 2, must have the functions $\xi_{[k|M_i]}$ satisfying the equations

$$\forall k, i : \xi_{[k|M_i]}(\lambda) \geq 0, \quad (3.28)$$

$$\forall i : \sum_k \xi_{[k|M_i]}(\lambda) = 1, \quad (3.29)$$

$$\forall b : \sum_{k,i} (\alpha_{[k|M_i]}^b - \beta_{[k|M_i]}^b) \xi_{[k|M_i]}(\lambda) = 0 \quad (3.30)$$

for any ontic state $\lambda \in \Lambda$ fixed. Hence, for every fixed ontic state we have that equations (3.28)-(3.30) define a polytope for the linear characterization

$$\vec{\xi}(\lambda) = \left(\xi_{[1|M_1]}(\lambda), \xi_{[2|M_1]}(\lambda), \dots, \xi_{[d|M_1]}(\lambda), \xi_{[1|M_2]}(\lambda), \dots, \xi_{[d|M_{|\mathbb{P}|}]}(\lambda) \right), \quad (3.31)$$

where just for now we set

$$d := |\mathbb{O}_M|.$$

Given such a description, the next step is to solve the *vertex enumeration problem* for that polytope. We use the command TRAF from PORTA for such an operation. As a step-by-step construction, we will consider two case studies for such a vertex enumeration task: the scenario from (2.21) and the scenario for the fair-coin-flip, which is a scenario with a non-trivial measurement assignment polytope. For the fair-coin-flip scenario we have $\mathbb{M} = \{M_1, M_2, M_3\}$, $\mathbb{P} = \{P_1, \dots, P_6\}$, $\mathbb{O}_M = \{0, 1\}$ and the following operational

equivalences for the preparation procedures:

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4 \implies \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0 \right) \in \mathbb{E}_P \quad (3.32)$$

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_5 + \frac{1}{2}P_6 \implies \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0; 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2} \right) \in \mathbb{E}_P \quad (3.33)$$

$$\frac{1}{2}P_3 + \frac{1}{2}P_4 \simeq \frac{1}{2}P_5 + \frac{1}{2}P_6 \implies \left(0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2} \right) \in \mathbb{E}_P \quad (3.34)$$

And for the measurement events we have the following equivalences,

$$\frac{1}{3}[0|M_1] + \frac{1}{3}[0|M_2] + \frac{1}{3}[0|M_3] \simeq \frac{1}{3}[1|M_1] + \frac{1}{3}[1|M_2] + \frac{1}{3}[1|M_3],$$

equivalent to the following element in \mathbb{E}_M

$$\left(\frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3}, 0; 0, \frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3} \right) \in \mathbb{E}_M. \quad (3.35)$$

We denote such a scenario as \mathbb{B}_{fcf} . Note that for the simplest scenario, we have that the vertices are just the deterministic assignments with 0's and 1's, but for the fair-coin-flip scenario, we have that there are indeterministic assignments for the vertices, as we can see from the results of (C.1)-(C.11). The numerical evaluation of the vertices is presented in detail in appendix C.

3.3 Prepare-and-measure noncontextuality inequalities

Once we were able to find the full set of vertices of the measurement-assignment polytope, it is now important that we describe every point in this polytope as a convex mixture of the vertices. This is possible since $NC(\mathbb{B})$ is a convex polytope. We define the measurement assignments of the vertices by the vectors $\tilde{\xi}(\kappa)$, where κ is the label for the vertices. The definition goes as:

$$\tilde{\xi}(\kappa) := \left(\tilde{\xi}_{[1|M_1]}(\kappa), \dots, \tilde{\xi}_{[d|M_1]}(\kappa), \dots, \tilde{\xi}_{[d|M_{|v|}]}(\kappa) \right) \quad (3.36)$$

For example, in \mathbb{B}_{s_i} one of the vertices is κ_1 , given by (C.1), which implies that

$$\tilde{\xi}(\kappa_1) = \left(\tilde{\xi}_{[0|M_1]}(\kappa_1), \tilde{\xi}_{[1|M_1]}(\kappa_1), \tilde{\xi}_{[0|M_2]}(\kappa_1), \tilde{\xi}_{[1|M_2]}(\kappa_1) \right) = (1, 0, 0, 1).$$

We write then that, for any $\vec{\xi}(\lambda)$, there exists a convex mixture of the vectors $\tilde{\xi}(\kappa)$, summed over all vertices κ , such that

$$\forall k, i : \xi_{[k|M_i]}(\lambda) = \sum_{\kappa} w(\kappa|\lambda) \tilde{\xi}_{[k|M_i]}(\kappa) \quad (3.37)$$

where $w(\kappa|\lambda)$ are convex weights. Considering that we can re-write, for every (fixed) ontic state λ the effects via the vertices of the noncontextual measurement-assignment polytope, we can write the ontological description of the operational probabilities $p(k|M_i, P_j)$ as,

$$p(k|M_i, P_j) = \int_{\Lambda} \xi_{[k|M_i]}(\lambda) d\mu_{P_j}(\lambda) = \int_{\Lambda} \left[\sum_{\kappa} w(\kappa|\lambda) \tilde{\xi}_{[k|M_i]}(\kappa) \right] d\mu_{P_j}(\lambda) \quad (3.38)$$

$$= \sum_{\kappa} \tilde{\xi}_{[k|M_i]}(\kappa) \underbrace{\left[\int_{\Lambda} w(\kappa|\lambda) d\mu_{P_j}(\lambda) \right]}_{=: \nu_{P_j}(\kappa)} = \sum_{\kappa} \tilde{\xi}_{[k|M_i]}(\kappa) \nu_{P_j}(\kappa) \quad (3.39)$$

For every vertex κ we note that $\nu_{P_j}(\kappa) \geq 0$, and if we sum over κ we get

$$\sum_{\kappa} \nu_{P_j}(\kappa) = \int_{\Lambda} \sum_{\kappa} w(\kappa|\lambda) d\mu_{P_j}(\lambda) = \int_{\Lambda} d\mu_{P_j}(\lambda) = 1.$$

Hence we have that ν_{P_j} is a probability distribution over the vertices. For the noncontextuality assumption over the operational equivalences of (2.12), we have that when μ_{P_j} satisfy the operational equivalences for preparation procedures, this implies that ν_{P_j} also respects them: For any κ ,

$$\sum_j (\alpha_j^a - \beta_j^a) \mu_{P_j} = 0 \implies \int_{\Lambda} \sum_j (\alpha_j^a - \beta_j^a) w(\kappa|\lambda) d\mu_{P_j}(\lambda) = 0$$

and hence $\nu_{P_j}(\kappa)$ satisfy the operational equivalences as well, for all vertices κ ¹⁵. The above calculations can be summarized in a lemma.

Lemma 1. (*Schmid et al., 2018*) *If any noncontextual ontological model over Λ exists for a scenario \mathbb{B} , there must also exist a noncontextual model with an ontic state space $\Lambda' := \bigcup_{\kappa} \{\kappa\}$ of finite cardinality. The latter model is constructed by identifying every ontic state λ with each extremal noncontextual measurement assignment, and then imagining*

¹⁵For the purposes we have in mind, it suffices to consider μ_{P_j} as a probability density. The RHS of this implication is then an equation valid for all values λ for the function μ_{P_j} and we suppose we can write $d\mu_{P_j} = \mu_{P_j}(\lambda)d\lambda$. Then the LHS holds.

every preparation as a probability distribution over those ontic states, as $\nu_{P_j}(\kappa)$.

We have then the final formulation of a universally noncontextual model over the set Λ' . A data-table $B := \{p(k|M_i, P_j)\}_{k,i,j}$ has a universally noncontextual ontological model if, and only if, there are $\{\nu_{P_j}(\kappa)\}_{j,k}$ such that,

$$\forall k, j : \nu_{P_j}(\kappa) \geq 0, \quad (3.40)$$

$$\forall j : \sum_{\kappa} \nu_{P_j}(\kappa) = 1, \quad (3.41)$$

$$\forall \kappa, a : \sum_j (\alpha_j^a - \beta_j^a) \nu_{P_j}(\kappa) = 0, \quad (3.42)$$

$$\forall k, i, j : \sum_{\kappa} \tilde{\xi}_{[k|M_i]}(\kappa) \nu_{P_j}(\kappa) = p(k|M_i, P_j) \quad (3.43)$$

were κ ranges over the vertices and the $\tilde{\xi}_{[k|M_i]}(\kappa)$ are the known measurement assignments over the vertices. Since the only unknown quantities from the equations (3.40)-(3.43) are the probabilities $\{\nu_{P_j}(\kappa)\}$, the final step towards the formulation (3.25) is to use Fourier-Motzkin elimination over $\{\nu_{P_j}(\kappa)\}_{j,k}$ so that we finally have linear inequalities of the form (3.25) over the statistics of the data-table. Hence the theoretical formulation is now complete and the method is settled since the equations (3.40)-(3.43) are necessary and sufficient conditions for generalized noncontextuality with respect to *any* prepare-and-measure scenario.

In appendix C.2 we provide the full construction of the tight noncontextuality inequalities present in (Schmid et al., 2018) for the simplest scenario \mathbb{B}_{si} , see for instance equations (C.14)-(C.22).

Note that the amount of inequalities grows extremely fast with the dimensionality of the scenario, and in some sense, it very easily enters the realm of being not useful anymore to derive the full set inequalities. Already for the scenario \mathbb{B}_{fcf} we find 1596 inequalities. In appendix D, we discuss how to use linear programming techniques to overcome such practical difficulty, again following (Schmid et al., 2018).

We conclude this chapter with the following lemma:

Lemma 2. (Pusey, 2018) *The scenario \mathbb{B}_{si} is indeed the simplest scenario that has non-trivial operational equivalences that presents contextual behaviors.*

Proof. Let us suppose any scenario with less operational structure than the scenario \mathbb{B}_{si} . For instance, consider a scenario that has only one measurement procedure. Then, there always exists a noncontextual ontological model for the behaviors in this scenario by letting the ontic states encode the outcomes.

$$\xi_{[k|M]}(\lambda) = \delta_{k\lambda} \quad (3.44)$$

If we let $\mu_{P_j}(\lambda) := p(k|M, P_j) = p(\lambda|M, P_j)$ then we obtain that

$$p(k|M, P_j) = \sum_{\lambda} \xi_{[k|M]}(\lambda) \mu_{P_j}(\lambda)$$

and any operational equivalences in the preparation procedures will be satisfied by $\mu_{P_j}(\lambda)$.

For two binary measurement procedures the operational probabilities for each preparation procedure P_j is defined by

$$\vec{P}_j := (p(0|M_0, P_j) - p(1|M_0, P_j), p(0|M_1, P_j) - p(1|M_1, P_j)),$$

where $j \in J$ labels all procedures. If $J = \{1, 2, 3\}$ or less, the convex hull of these two-dimensional vectors form a simplex. Every point in a simplex has one, and only one decomposition in terms of the vertices. Therefore, we can consider every ontic state λ as encoding the vertices in these convex hull, that we denote $\vec{\lambda}$, and the epistemic states as the unique decomposition of the behaviors in the simplex,

$$\vec{P}_j = \sum_{\lambda} \mu_{P_j}(\lambda) \vec{\lambda}.$$

For the effects $\xi_{[k|M]}(\lambda)$ we define such that the following equality holds,

$$\vec{\lambda} = (\xi_{[0|M_0]}(\lambda) - \xi_{[1|M_0]}(\lambda), \xi_{[0|M_1]}(\lambda) - \xi_{[1|M_1]}(\lambda)).$$

Such a model reproduces the operational probabilities and the operational equivalences for preparation procedures are satisfied, since they imply

$$\sum_j \alpha_j \mu_{P_j}(\lambda) \vec{\lambda} = \sum_j \beta_j \mu_{P_j}(\lambda) \vec{\lambda}$$

and the uniqueness of decomposition implies $\sum_j (\alpha_j - \beta_j) \mu_{P_j}(\lambda) = 0$. We then conclude that \mathbb{B}_{si} is indeed the simplest scenario with contextuality. \square

RESOURCE THEORY FOR GENERALIZED CONTEXTUALITY

What we are really looking for is a better understanding of the connections between information processing and physical principles in general.

- (Barrett, 2007, Jonathan, pg. 1)

As is natural to every quantum process with no direct classical analog, an exciting question for technological developments is: *How can we use these effects?* We have mentioned before, a possible approach to formally address this question is the so-called resource theory framework, where we choose an effect and build a formalism to understand how to use the chosen effect as a resource. The word *use* from our question usually is only limited by the amount of creativity in the development of the resource framework, or the no-go theorems one can prove therein.

We will see that the intuition behind *using* the resource will be made precise. But the idea is clear: some processes can only take place if the resource is present. Here, resource is truly anything that fits a mathematical description, having in mind that we would like to understand if it is possible to find processes where the resource is essential. Take the example of a car moving. This is a process that can only happen given that some resource exists: fuel. We want to show, using this formalism, that in a very similar way, some quantum information protocols have their success provided by a specific fuel: generalized contextuality.

In this chapter we will introduce our framework and prove the fundamental aspects that a resource theory should respect, in section 4.1. We then proceed to define the resource monotones in section 4.2, that shall quantify the resource. We then describe a new formalism to witness generalized contextuality in section 4.3, that is associated with quantum simulation and pre/post selection boxes. We will learn that this formalism is consistent with claims of experimental verification of quantum contextuality in the literature. We will also argue that the contextuality monotones might be useful to impose bounds of noise in experimental implementation of contextuality scenarios. We will finish

the chapter by describing a new resource theory map in section 4.4 that will be useful in chapter 5.

4.1 Resource Theory framework

The most general structure adapted to our study of the resource theory framework is the following:

Definition 21. A resource theory is defined by a quadruple $(\mathcal{C}, \mathcal{NC}, \otimes, \mathcal{F})$: A set \mathcal{C} of objects, a set of free objects \mathcal{NC} , a binary operation \otimes between the objects $C(\mathbb{B})$, a set \mathcal{F} of free operations acting on the objects. \diamond

Throughout this chapter we will mainly consider the discussions present in (Duarte and Amaral, 2018). For a mathematical description of general resource theories from the point of view of category theory see (Coecke et al., 2016). For a description of a general framework of convertibility relations between resources see (Fritz, 2015). For a study of quantum affine resource theories see (Gour, 2017).

As the notation of definition 21 suggests, we construct a resource theory for generalized contextuality where, given any prepare-and-measure scenario \mathbb{B} , the objects are the behaviors from $C(\mathbb{B})$. The free objects will be the behaviors in the noncontextual polytope associated with \mathbb{B} . We chose elements in $NC(\mathbb{B})$ to be the free objects of our resource theory because they can be interpreted as the objects that have no resource. There are noncontextual ontological models that reproduce the statistics of elements in $NC(\mathbb{B})$, by definition.

We will consider two different binary operations. The first one we define here, and for the second one we will devote a later section.

Definition 22. Given two behaviors $B_1 \in C(\mathbb{B}_1)$ and $B_2 \in C(\mathbb{B}_2)$, the juxtaposition of B_1 and B_2 , is the behavior obtained by *independently* choosing preparation and measurement procedures for B_1 and B_2 . That is, the preparations in $B_1 \otimes B_2$ correspond to a pair of preparations (P_{j_1}, P_{j_2}) , $j_1 \in J_1, j_2 \in J_2$ and analogously for measurement procedures. The data-table corresponding to the behavior $B_1 \otimes B_2$ is then given by the probabilities:

$$B_1 \otimes B_2 = \{p((k_1, k_2)|(M_{i_1}, M_{i_2}), (P_{j_1}, P_{j_2}))\} = \{p(k_1|M_{i_1}, P_{j_1})p(k_2|M_{i_2}, P_{j_2})\}.$$

Where we are considering the set of labels associated with the respective scenarios I_1, I_2 , for measurement procedures, K_1, K_2 , for outcomes, and J_1, J_2 , for preparation procedures, as described before. \diamond

By far, the most critical definition to be made within a resource theory framework is one of the free operations over the objects. These free operations induce order into the

set of objects. In essence, the way one defines such ordering, or equivalently, the free operations over the objects, will objectively influence how useful a resource theory might be. Our proposal for free operations will use the following maps.

Definition 23. A *stochastic map* from the left $q : I \rightarrow \tilde{I}$ between two finite set of indices is a map that satisfies the two constraints:

$$\sum_{i \in I} q(\tilde{i}|i) = 1 \quad (4.1)$$

$$q(\tilde{i}|i) \geq 0, \quad (4.2)$$

$\forall \tilde{i} \in \tilde{I}$, *i.e.*, the rows of q sum to one. Similarly, a stochastic map from the right is defined in the same way but switching i with \tilde{i} in the matrices $q(\tilde{i}|i)$ to $q(i|\tilde{i})$, *i.e.*, the columns sum to one. These free operations are defined with respect to the following stochastic matrices: \diamond

The interpretation is that a stochastic matrix Q can be understood as a (stochastic) dynamics over some matrix A . Consider that we have $A \rightarrow B$ in some Markov chain. Then for a left stochastic map we have $B = QA$ where B is the final state, and the columns add to one, and for a right stochastic map we have $B = AQ$, where the rows add to one. With this definition, we consider relevant the following definition for free operations,

Definition 24. Given a scenario \mathbb{B} we define the set of free operations \mathcal{F} as the set of maps $f : C(\mathbb{B}) \rightarrow C(f(\mathbb{B}))$ such that

$$f : \{p(k|M_i, P_j)\}_{k \in K, i \in I, j \in J} \mapsto \left\{ \sum_{i, j, k} q_O^i(\tilde{k}|k) p(k|M_i, P_j) q_M(i|\tilde{i}) q_P(j|\tilde{j}) \right\}_{\tilde{k} \in \tilde{K}, \tilde{i} \in \tilde{I}, \tilde{j} \in \tilde{J}}, \quad (4.3)$$

where $q_O^i : K \rightarrow \tilde{K}$, $q_M : \tilde{I} \rightarrow I$, $q_P : \tilde{J} \rightarrow J$ are stochastic maps between index sets corresponding to operational primitives in the different scenarios defined by \mathbb{B} and $f(\mathbb{B})$. \diamond

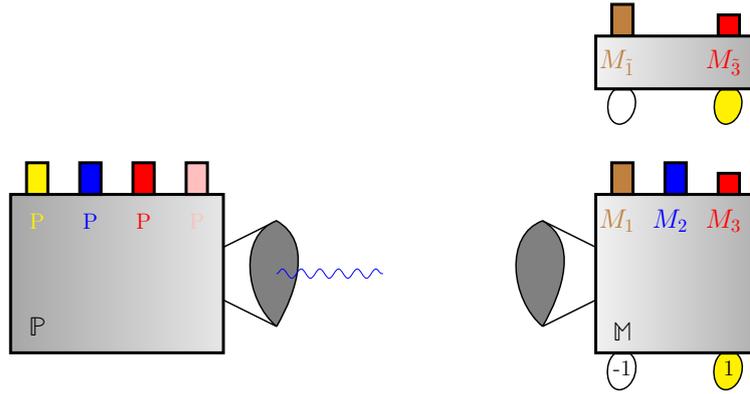


Figure 4.1: Example of a free operation: pre-processing box for the measurement procedures in an experimental scenario.

Suppose that we have the following: the experimental scenario with the primary primitives in figure 4.1 is composed of the boxes, \mathbb{P} and \mathbb{M} . The measurement box \mathbb{M} has three measurement procedures M_1, M_2 and M_3 . As an example, we can consider a free operation acting on the measurement procedures (pre-processing) as follows: for all outcomes of M_1 , we activate the measurement M_1 ; the same for M_3 . We then conclude that effectively, the procedure M_3 is never performed, so that the effective scenario under consideration has only two measurement procedures: $\{M_1, M_3\}$. This corresponds to a free operation f between two different scenarios, where $f(\mathbb{B})$ is the scenario with only two procedures.

Example 9 (Coarse-grainings are free-operations). Under this description, coarse-grainings can be understood as free operations. Let \tilde{M} be a measurement procedure, with measurement events $[\tilde{k}|\tilde{M}]$, $\tilde{k} \in \mathcal{O}_{\tilde{M}}$. Then, we define the events $[\tilde{k}|\tilde{M}]$ as a coarse-graining of other events $[k|M]$ as in equation (3.2),

$$[\tilde{k}|\tilde{M}] = \sum_k p(\tilde{k}|k)[k|M],$$

which can be understood as post-processing the outcomes of \tilde{M} (Kunjwal, 2019),

$$p(\tilde{k}|\tilde{M}, P) = \sum_k p(\tilde{k}|k)p(k|M, \tilde{P}).$$

For all procedures P . Therefore, in general we might let $q_O^i(\tilde{k}|k) = p(\tilde{k}|k)$, for each measurement procedure $i \in I$ in the scenario, and the other matrices q_M, q_P to be the identity. \triangle

Example 10 (Convex combinations are free operations). Let $\mathbb{P} := \{P_j\}_{j \in J}$ be a set of preparation procedures, and we suppose that we produce secondary procedures by convex

combinations of all elements from \mathbb{P} , such as in (Mazurek et al., 2016, Table III, pg. 16),

$$P_{\tilde{j}} = \sum_j p(j|\tilde{j})P_j,$$

where $p(j|\tilde{j})$ is the probability of implementing P_j given that the secondary procedure is $P_{\tilde{j}}$. As we can infer from the definition of a free operation, at the level of behaviors this relation reads

$$p(k|M, P_{\tilde{j}}) = \sum_j p(j|\tilde{j})p(k|M, P_j),$$

for all $[k|M]$. We let $q_P(j|\tilde{j}) = p(j|\tilde{j})$ and, in the case of convex combinations of the procedures, the procedures $\tilde{j} \in \tilde{J}$ satisfy that $|J| = |\tilde{J}|$. \triangle

It is important to stress that the scenarios \mathbb{B} and $f(\mathbb{B})$ are different, because the transformations f will in general change the operational equivalences. We write $\mathbb{E}_{f(P)}$ and $\mathbb{E}_{f(M)}$ for the new operational equivalences, and also $\mathbb{O}_{f(M)}$ the new outcomes. Since $f(\mathbb{B})$ is a *new* operational scenario, it is related with a *new* polytope $C(f(\mathbb{B}))$. In fact we can see by construction how the new operational equivalences should be uplifted in the new scenarios.

Let \mathbb{E}_P with respect to \mathbb{B} that can be written as, for a fixed element s ,

$$\sum_{j \in J} (\alpha_j^s - \beta_j^s)P_j = 0,$$

by defining, for every $j \in J$, the vectors $\tilde{\alpha}_j^s$ and $\tilde{\beta}_j^s$,

$$\alpha_j^s = \sum_{\tilde{j} \in \tilde{J}} \tilde{\alpha}_j^s q_P(j|\tilde{j}) \tag{4.4}$$

$$\beta_j^s = \sum_{\tilde{j} \in \tilde{J}} \tilde{\beta}_j^s q_P(j|\tilde{j}) \tag{4.5}$$

we get

$$\sum_{j \in J} (\alpha_j^s - \beta_j^s)P_j = \sum_{j, \tilde{j}} (\tilde{\alpha}_j^s - \tilde{\beta}_j^s)q_P(j|\tilde{j})P_j = \sum_{\tilde{j} \in \tilde{J}} (\tilde{\alpha}_j^s - \tilde{\beta}_j^s)P_{\tilde{j}} = 0$$

where the new set of preparations are defined within to the new scenario $f(\mathbb{B})$.

It is fundamental that a free operation does not create resource out of a resourceless object. In our formalism, this means that a free operation must send noncontextual behaviors in \mathbb{B} to noncontextual behaviors $f(\mathbb{B})$. We state such a result as a theorem:

Theorem 7. (Duarte and Amaral, 2018) *Let \mathbb{B} be any prepare-and-measure scenario. Every free operation $f \in \mathcal{F}$ sends noncontextual behaviors in $NC(\mathbb{B})$ to noncontextual behaviors in $NC(f(\mathbb{B}))$.*

Proof. Let $B \in NC(\mathbb{B})$. Hence, from definition 10 there exists an ontic space (Λ, Σ) and a pair (Π, Θ) , such that every function from Θ satisfies the operational equivalences \mathbb{E}_M and every epistemic state from Π satisfies, similarly, \mathbb{E}_P , by definition of the linear characterization of a noncontextual behavior.

For any behavior in $f(\mathbb{B})$, we get, for all $\tilde{i} \in \tilde{I}, \tilde{j} \in \tilde{J}, \tilde{k} \in \tilde{K}$ as in the definition 24, that the following holds¹,

$$p(\tilde{k}|M_{\tilde{i}}, P_{\tilde{j}}) = \sum_{i,j,k} q_O^j(\tilde{k}|k)p(k|M_i, P_j)q_M(i|\tilde{i})q_P(j|\tilde{j}) \quad (4.6)$$

$$= \sum_{i,j,k} q_O^i(\tilde{k}|k) \left(\sum_{\lambda \in \Lambda} \xi_{[k|M_i]}(\lambda) \mu_{P_j}(\lambda) \right) q_M(i|\tilde{i})q_P(j|\tilde{j}) \quad (4.7)$$

$$= \sum_{\lambda \in \Lambda} \left(\sum_{k,j} q_O^i(\tilde{k}|k) \xi_{[k|M_i]}(\lambda) q_M(i|\tilde{i}) \right) \left(\sum_j \mu_{P_j}(\lambda) q_P(j|\tilde{j}) \right) \quad (4.8)$$

$$= \sum_{\lambda \in \Lambda} \xi_{[\tilde{k}|M_{\tilde{i}}]}(\lambda) \mu_{P_{\tilde{j}}}(\lambda), \quad (4.9)$$

where we see that $\mu_{P_{\tilde{j}}}$ correspond to an epistemic state related to the new preparations from $f(\mathbb{B})$ that respects the new operational equivalences from (4.1). The same holds for the new effects $\xi_{[\tilde{k}|M_{\tilde{i}}]}$, where they correspond to valid effects since

$$\sum_{\tilde{k}} \xi_{[\tilde{k}|M_{\tilde{i}}]}(\lambda) = \sum_{\tilde{k}, k, i} q_O^i(\tilde{k}|k) \xi_{[k|M_i]} q_M(i|\tilde{i}) \quad (4.10)$$

$$= \sum_{\tilde{k}, i} \underbrace{\left(\sum_k q_O^i(\tilde{k}|k) \xi_{[k|M_i]}(\lambda) \right)}_{=: q_\lambda(\tilde{k}|i)} q_M(i|\tilde{i}) \quad (4.11)$$

$$= \sum_{\tilde{k}} \underbrace{\left(\sum_i q_\lambda(\tilde{k}|i) q_M(i|\tilde{i}) \right)}_{=: q_\lambda(\tilde{k}|\tilde{i})} = \sum_{\tilde{k}} q_\lambda(\tilde{k}|\tilde{j}) = 1. \quad (4.12)$$

We proceed with obtaining $\mathbb{E}_{f(P)}$, where we define the novel set of weights associated with

¹The tildes refer to the transformed set of labels, under the transformation given by the stochastic maps defined in the free operation.

the operational equivalences in the scenario $f(\mathbb{B})$ by $\tilde{\alpha}_{[\tilde{k}|M_{\tilde{i}}]}$ and $\tilde{\beta}_{[\tilde{k}|M_{\tilde{i}}]}$ with

$$\alpha_{[k|M_i]}^s = \sum_{\tilde{k}, \tilde{i}} q_O^i(\tilde{k}, k) \tilde{\alpha}_{[\tilde{k}|M_{\tilde{i}}]} q_M(i|\tilde{i}) \quad (4.13)$$

$$\beta_{[k|M_i]}^s = \sum_{\tilde{k}, \tilde{i}} q_O^i(\tilde{k}|k) \tilde{\beta}_{[\tilde{k}|M_{\tilde{i}}]} q_M(i|\tilde{i}) \quad (4.14)$$

so that for any s labeling the elements of the operational set \mathbb{E}_M we have

$$0 = \sum_{k, i} (\alpha_{[k|M_i]} - \beta_{[k|M_i]}) [k|M_i] = \sum_{k, i, \tilde{k}, \tilde{i}} (\tilde{\alpha}_{[\tilde{k}|M_{\tilde{i}}]} - \tilde{\beta}_{[\tilde{k}|M_{\tilde{i}}]}) q_O^i(\tilde{k}|k) q_M(i|\tilde{i}) [k|M_i] \quad (4.15)$$

$$= \sum_{\tilde{k}, \tilde{i}} (\tilde{\alpha}_{[\tilde{k}|M_{\tilde{i}}]} - \tilde{\beta}_{[\tilde{k}|M_{\tilde{i}}]}) [\tilde{k}|M_{\tilde{i}}] \implies 0 = \sum_{\tilde{k}, \tilde{i}} (\tilde{\alpha}_{[\tilde{k}|M_{\tilde{i}}]} - \tilde{\beta}_{[\tilde{k}|M_{\tilde{i}}]}) \xi_{[\tilde{k}|M_{\tilde{i}}]}(\lambda), \forall \lambda \quad (4.16)$$

since $\xi_{[\tilde{k}|M_{\tilde{i}}]}(\lambda)$ are the direct ontological descriptions for $[\tilde{k}|M_{\tilde{i}}]$. A very similar thing is valid for the epistemic states. Then $f(B)$ is in the noncontextual polytope $NC(f(\mathbb{B}))$. We conclude that any free operation f is such that

$$B \in NC(\mathbb{B}) \implies f(B) \in NC(f(\mathbb{B})), \quad (4.17)$$

as we wanted to show. \square

As a corollary, we have the following lemma:

Lemma 3. *Let $\mathcal{F} \ni f : \mathbb{B} \rightarrow f(\mathbb{B})$. If we consider contextual behaviors $B \in C(\mathbb{B}) \setminus NC(\mathbb{B})$, then,*

$$f(B) \in C(f(\mathbb{B})) \setminus NC(f(\mathbb{B})) \implies B \in C(\mathbb{B}) \setminus NC(\mathbb{B}). \quad (4.18)$$

Proof. This is not truly a lemma, but a different way of reading the definition of a free operation defined by the resource theory. Suppose that $f(B) \in C(f(\mathbb{B})) \setminus NC(f(\mathbb{B}))$. If B was a noncontextual behavior, we would have that $f(B)$ is also a noncontextual behavior, since f is a free operation. Therefore, we conclude that B cannot be a point in the noncontextual polytope $NC(\mathbb{B})$. \square

With this result we have constructed a resource theory for generalized contextuality. Remember that the resource theory is defined by the quadruple $(\mathcal{C}, \mathcal{NC}, \otimes, \mathcal{F})$. We know the set of objects as the set of all behaviors in a given finitely defined prepare-and-measure scenario,

$$\mathcal{C} := \{B \in C(\mathbb{B}) : \mathbb{B} \text{ is some prepare-and-measure scenario}\} \quad (4.19)$$

and free objects as,

$$\mathcal{NC} := \{B \in NC(\mathbb{B}) : \mathbb{B} \text{ is some prepare-and-measure scenario}\}. \quad (4.20)$$

A relation $(\text{map}) \otimes$ that acts between the objects, and a set of free operations over the allowed objects. Remember that all prepare-and-measure scenarios have their, possibly different, operational characterizations in terms of the vectors $\gamma^{a(b)}$, i.e., in terms of the operational equivalences between convexly generated hypothetical procedures. With this structure we have developed the background for producing reasonable questions of the type: Is this operation free? Does this data-table has the resource? Have I lost resource by performing a specific transformation? There exists a noncontextual ontological model for this data-table obtained by coarse-graining?

We also have that, by linearity of the map f and by convexity of the noncontextual polytope, the lemma:

Lemma 4. *For any $f \in \mathcal{F}$ we have that the following equality holds:*

$$f(\pi B_1 + (1 - \pi)B_2) = \pi f(B_1) + (1 - \pi)f(B_2) \quad (4.21)$$

for all $B_1, B_2 \in C(\mathbb{B})$ and $\pi \in [0, 1]$.

Proof. Let $B_1, B_2 \in C(\mathbb{B})$ be any two behaviors, described by

$$\begin{aligned} B_1 &:= (p^{(1)}(k|M_i, P_j))_{k \in K, i \in I, j \in J} \\ B_2 &:= (p^{(2)}(k|M_i, P_j))_{k \in K, i \in I, j \in J}. \end{aligned}$$

In this definition, K, I, J are the sets of labels for the procedures in \mathbb{B} . We use the notation $(p(s))_{s \in S} \equiv (p(s_1), \dots, p(s_{|S|}))$. Then, by definition, we have that each element of the behavior $f(\pi B_1 + (1 - \pi)B_2)$ is described by

$$\begin{aligned} &\sum_{i,j,k} q_O^j(\tilde{k}|k) (\pi p^{(1)}(k|M_i, P_j) + (1 - \pi)p^{(2)}(k|M_i, P_j)) q_M(i|\tilde{i}) q_P(j|\tilde{j}) = \\ &+ \pi \left(\sum_{i,j,k} q_O^j(\tilde{k}|k) p^{(1)}(k|M_i, P_j) q_M(i|\tilde{i}) q_P(j|\tilde{j}) \right) + \\ &+ (1 - \pi) \left(\sum_{i,j,k} q_O^j(\tilde{k}|k) p^{(2)}(k|M_i, P_j) q_M(i|\tilde{i}) q_P(j|\tilde{j}) \right) \end{aligned}$$

which proves the lemma. □

4.2 Quantifiers

Now that we have defined the resource theory structure, we can introduce some quantifiers to give meaning for questions of the type: *How much resource does a given object has?* As in our framework, this means to attempt at measuring how contextual is a given behavior from a scenario. The choice of quantifier might largely depend on the important aspects the researcher is interested in for specific purposes; here, we consider those most relevant for our formalism. Formally, a quantifier will be generally described as a function from the objects of the theory, in our case the behaviors, towards positive real numbers. The order \leq structure present in the set of real numbers will corresponds to the amount of resource, under the quantifier.

The class of functions that are helpful in this discussion for defining a quantifier, essentially, obey two inequalities:

1. Let \mathfrak{q} be a quantifier. Then we expect that for all objects $\mathfrak{q}(f(B)) \leq \mathfrak{q}(B)$, $f \in \mathcal{F}$.
2. Let \mathfrak{q} be a quantifier. Then for every pair of objects B_1, B_2 we have that $\mathfrak{q}(B_1 \otimes B_2) \leq \mathfrak{q}(B_1) + \mathfrak{q}(B_2)$.

The first requirement reflects the fact that a free operation should not increase the resource, in other words, a free operation cannot increase the amount of contextuality present in the behavior, in a sense provided by the quantifier \mathfrak{q} . The second requirement is that we cannot increase contextuality using the binary operation \otimes . This would be a rather strange situation since, *a priori*, the binary operation \otimes can refer to completely independent scenarios.

Definition 25. Consider \mathbb{B} to be any prepare-and-measure scenario, and consider the set of objects $C(\mathbb{B}) \subset \mathcal{C}$. Let $\mathfrak{q}_{\mathbb{B}} : C(\mathbb{B}) \rightarrow \mathbb{R}_+$ defines a family of functions, for each \mathbb{B} fixed, such that:

1. $\forall B_1 \in C(\mathbb{B}_1), B_2 \in C(\mathbb{B}_2)$ we have

$$\mathfrak{q}_{\otimes(\mathbb{B}_1, \mathbb{B}_2)}(B_1 \otimes B_2) \leq \mathfrak{q}_{\mathbb{B}_1}(B_1) + \mathfrak{q}_{\mathbb{B}_2}(B_2), \quad (4.22)$$

where $\otimes(\mathbb{B}_1, \mathbb{B}_2) \equiv \mathbb{B}_1 \otimes \mathbb{B}_2$ is the prepare-and-measure scenario defined by the product \otimes considered in the specific resource theory.

2. For any $f \in \mathcal{F}$, and every $B \in C(\mathbb{B})$, if we denote the image of the free operation f as $f(\mathbb{B})$ then,

$$\mathfrak{q}_{f(\mathbb{B})}(f(B)) \leq \mathfrak{q}_{\mathbb{B}}(B). \quad (4.23)$$

We call such map a *noncontextual monotone*, or also a *quantifier* for the resource theory. If the scenario in question is clear, we simplify the notation $\mathfrak{q}_{\mathbb{B}} \equiv \mathfrak{q}$. \diamond

Because of the polytope structure of $NC(\mathbb{B})$, arising from the linear characterization,

$$\text{behavior } B \equiv \text{point in } \mathbb{R}^N, \text{ for some } N \in \mathbb{N},$$

with respect to equations (3.40)-(3.43), the first obvious definitions of quantifiers that are meaningful relate the amount of resource with the convex-polytope structure. For Kochen-Specker contextuality a very interesting quantifier, the so called *contextual fraction*, from (Abramsky et al., 2017) has a direct analog for generalized contextuality.

Definition 26. (Duarte and Amaral, 2018) Let \mathbb{B} be any prepare-and-measure scenario. Let $f : C(\mathbb{B}) \rightarrow [0, 1]$ defined by,

$$f(B) := 1 - \max_{\omega \in [0,1]} \{ \omega \mid B = \omega B^{NC} + (1 - \omega)B', B^{NC} \in NC(\mathbb{B}), B' \in C(\mathbb{B}) \}. \quad (4.24)$$

We call f the contextual fraction. ◇

Since $C(\mathbb{B})$ and $NC(\mathbb{B}) \subseteq C(\mathbb{B})$ are convex polytopes, they are compact closed sets of points in $\mathbb{R}^{|J| \times |I| \times |K|}$. If $C(\mathbb{B}) = NC(\mathbb{B})$, $f(B) = 0$, for all $B \in C(\mathbb{B})$. Whenever $C(\mathbb{B}) \neq NC(\mathbb{B})$ we have that the max is unique and, therefore, f is well-defined.

Another important contextuality measure is defined as a class of measures, associated with distances between points.

Definition 27. Let D be any distance defined for $\mathbb{R}^{|K|}$, given a set of alphabet for outcomes K , that is associated with the scenario description $\mathbb{B} := (|J|, |I|, |K|, \mathbb{E}_P, \mathbb{E}_M)$. Then, given two data-tables $B, B' \in \mathbb{B}$ we define that

$$D(B, B') := \max_{i \in I, j \in J} D(p(\cdot | M_i, P_j), p'(\cdot | M_i, P_j)).$$

With this, we define the D -contextuality distance as:

$$d(B) := \min_{B' \in NC(\mathbb{B})} D(B, B') \quad (4.25)$$

of special importance is the case when D is the l_1 -distance:

$$D_1(x, y) = \sum_k |x_k - y_k|. \quad (4.26)$$

◇

Both the functions f and d are quantifiers for generalized contextuality.

Theorem 8. (Duarte and Amaral, 2018) Both the contextual fraction f and the l_1 -contextuality distance d are resource monotones with respect to the free operations.

The proof of this theorem can be found in (Duarte and Amaral, 2018, pg. 9-10, pg. 13-14).

Applying the monotones

Although we could not find any actual application for the monotones, such as the one known for the Kochen-Specker contextual fraction in measurement-based quantum computation, we mention a few initial attempts of using the monotones for practical applications.

For the monotone \mathbf{d} , we mention, in section 5.2, how we can think about \mathbf{d} when considering an idealized behavior for the quantum task of state discrimination. We could use this monotone for obtaining lower bounds for optimal quantum contextual behaviors. In the ideal case we treat, this would be a rather silly choice since calculus tools would suffice: the optimal quantum bounds are known. However, we still present it, since it could lead to a generalized version for searching optimal contextual behaviors for more complex experimental scenarios.

We also note that the monotone \mathbf{d} might be related in some way with the l_1 -measure over epistemic states μ . Let suppose we can write the epistemic states as probability densities μ_{P_j} over $\Lambda = \mathbb{R}$. We would then have that, allowing the probabilities $p(k|M_i, P_j)$ to have an ontological model,

$$\begin{aligned} |p(k|M_i, P_j) - p(k|M_i, P_{j'})| &= \left| \int_{\Lambda} (\mu_{P_j}(\lambda) - \mu_{P_{j'}}(\lambda)) \xi_{[k|i]}(\lambda) d\lambda \right| \\ &\leq \int_{\Lambda} |\mu_j(\lambda) - \mu_{j'}(\lambda)| \xi_{[k|i]}(\lambda) d\lambda, \end{aligned}$$

so

$$\sum_k |p(k|i, j) - p(k|i, j')| \leq \int_{\Lambda} |\mu_j(\lambda) - \mu_{j'}(\lambda)| d\lambda = \|\mu_j - \mu_{j'}\|_{l_1}. \quad (4.27)$$

We see that the LHS of equation (4.27) is $D_1(p(\cdot|M_i, P_j), p(\cdot|M_i, P_{j'}))$, which is related to the l_1 -distance monotone \mathbf{d} . The RHS is the l_1 -norm over the ontological quantities that are of importance for proving noncontextual bounds for quantum cloning, see (Lostaglio and Senno, 2020), and quantum state discrimination, see (Schmid and Spekkens, 2018). This induces us to speculate that the operational quantity \mathbf{d} could provide bounds for the ontological description of the epistemic states, which in itself might represent a clue for increasing our understanding of scenarios with preparation contextuality.

We describe how to obtain the contextual fraction \mathbf{f} for general scenarios in appendix D.2. In our research, we attempted to use this monotone as a tool for obtaining maximally contextual behaviors, i.e., behaviors B_{max} such that $\forall B \in Q(\mathbb{B}), \mathbf{f}(B) \leq \mathbf{f}(B_{max})$, where $Q(\mathbb{B})$ is the set of quantum behaviors in the generalized formalism, that reproduce the statistics of the general prepare-and-measure scenario \mathbb{B} . The idea was to develop an algorithm that would approach the maximally contextual behavior, from within the quantum set, by using the tools from (Ambainis et al., 2016) and (Schmid et al., 2018), hence obtaining a lower bound for B_{max} . Let $B_{facet} \in C(\mathbb{B})$ be a point that

lies in one of the facets of the convex polytope $C(\mathbb{B})$, but that is also a contextual behavior (not necessarily in $Q(\mathbb{B})$). Then, minimizing $\|B^Q - B_{facet}\|$ for $B^Q \in Q(\mathbb{B})$ corresponds to a quadratic semidefinite programming problem, once we fix either the set of quantum states associated with the quantum behavior, $\{\rho_j\}_{j \in J}$, or the set of measurement events, $\{E_i^k\}_{k \in K, i \in I}$. From a geometric perspective, it is clear that we are trying to find the quantum behavior closer to B_{facet} , but we need to find good candidates for B_{facet} , and this is where the contextual fraction can be relevantly applied. The solutions of the minimization 26, represented as B' for the contextual part, lie in the facets of the polytope $C(\mathbb{B})$. Therefore, we can try to use $f(B)$ as a step in an algorithm that would indicate both the growth in contextuality and new candidates for the minimization problem. We have found numerical evidence that this procedure finds the optimal quantum behavior in the simplest scenario, but there is still much to know for significant results in this direction.

4.3 Witnessing contextuality with the free operations

In light of the theory we presented, we propose a framework for witnessing contextuality that uses free operations. The essential aspects are, first, noticing the validity of lemma 3, and second, realizing that there exists a connection between pre/post-processing and simulation of quantum measurement and quantum correlations. An initial glance into such a field of research can be found in the following literature: (Guerini et al., 2017), (Oszmaniec et al., 2017), (Guerini, 2018), (Heinosaari et al., 2008).

We consider the following change in jargon: Let \tilde{M} be a generalized quantum measurement corresponding to a POVM, as usual. We say that the measurement $\tilde{M} = \{E_{\tilde{k}}\}_{\tilde{k} \in \tilde{K}}$ can be simulated by a set of POVMs $\mathcal{R} := \{R^i\}_{i \in I}$, with each $R^i = \{R_k^i\}_{k \in K}$ a POVM, if there exists pre/post processing maps, defined as q_M, q_O , such that,

$$E_{\tilde{k}} = \sum_{i,k} q_O^i(\tilde{k}|k) R_k^i q_M(i|\tilde{M}). \quad (4.28)$$

Notice that with respect to the jargon from the last section, equation (4.28) reads as “there exists a free operation that takes the POVM’s \mathcal{R} towards \tilde{M} ”. Whenever this happens, we say that \tilde{M} is \mathcal{R} -simulable. The idea behind this is that for any state ρ , the probabilities obtained by \tilde{M} are the same as the probabilities obtained by the simulated version, and therefore an experimenter would not need to have the specific procedure \tilde{M} available, but he/she could simply have \mathcal{R} and pre/post-process the data-table:

$$p(\tilde{k}|\rho) := \text{Tr}(E_{\tilde{k}}\rho) = \text{Tr}\left(\sum_{i,k} q_O^i(\tilde{k}|k) R_k^i \rho q_M(i|\tilde{M})\right), \quad (4.29)$$

which is extremely convenient for our purposes since we are analyzing *data-tables* in the

search for quantum contextuality. Next, we consider the following lemmas from (Guerini et al., 2017), and (Davies, 1976):

Lemma 5. (Guerini et al., 2017) *For any dimension of the system (Hilbert space of states), any 2-outcome POVM is projective-simulable. If the dimension of the system is two, binary-outcome simulability is equivalent to projective-simulability.*

Lemma 6. (Oszmaniec et al., 2017) *For two dimensional Hilbert spaces, projective measurements can simulate arbitrary two-outcome measurements.*

Now, this means that for any binary-outcome measurement \tilde{M} , there always exists a set \mathcal{R} of projective measurements that can simulate \tilde{M} . Hence, suppose we have a set of binary-outcome measurement procedures $\tilde{\mathbb{M}} = \{M_{\tilde{i}}\}_{\tilde{i} \in \tilde{I}}$. Each $M_{\tilde{i}}$ is simulable by a set of projective measurements $\mathcal{R}_{\tilde{i}}$. In terms of our scenario discussion we write that for any $\tilde{k}, E_{\tilde{k}}^{\tilde{i}} \in M_{\tilde{i}} \in \tilde{\mathbb{M}}$ we have that

$$E_{\tilde{k}}^{\tilde{i}} = \sum_{i,k} q_O^i(\tilde{k}|k) R_k^i q_M(i|\tilde{i}), \quad (4.30)$$

so that each $M_{\tilde{i}}$ is $\mathcal{R}_{\tilde{i}}$ -simulable for some set $\mathcal{R}_{\tilde{i}}$ of projective measurements. Hence, if we consider any scenario $\tilde{\mathbb{B}}$ containing $\tilde{\mathbb{M}}$ and another scenario \mathbb{B}_{proj} containing $\mathbb{M} = \bigcup_i \mathcal{R}_i$ we have that there exists a free operation $f \in \mathcal{F}$ such that

$$C(\mathbb{B}) \xrightarrow{f} C(\tilde{\mathbb{B}}). \quad (4.31)$$

And notice that, up to now, our discussion considered POVM simulation, so that the free operations $f \in \mathcal{F}$ are free operations over measurement procedures in the scenario, leaving the preparation procedures the same.

Clarification of the tilde notation

For clarifying the tildes: remember that in definition 24 we consider that the data-table $f(B)$ has the associated set of labels with the tildes. Therefore, if I labels the measurements in the scenario of the domain, \tilde{I} labels the measurements in the scenario from the codomain. Since we are looking at simulations of quantum processes, as we can see from equation (4.31), we have the following relationship between free operations and simulability:

$$\tilde{\mathbb{M}} \text{ is simulated by } \mathbb{M} \equiv \exists f \text{ free operation from } \mathbb{M} \text{ to } \tilde{\mathbb{M}}.$$

Therefore, the tildes are an attempt to maintain a coherent presentation with respect to the resource theory formalism.

The simulations witness contextuality in prepare-and-measure scenarios

Having the lemmas from before, we can notice that the following is true.

Lemma 7. *Let $M_1, \dots, M_{|I|}$ be a set of two-outcome measurement procedures. Then, there is always some stochastic maps $q_O : K \rightarrow \tilde{K} = K, q_M : \tilde{I} \rightarrow I$ that sends these POVMs $\{E_k^i\}$ to the following ones:*

$$E_{\tilde{1}}^{\tilde{1}} = \frac{1}{(1-v)^2 + 1} \begin{pmatrix} (1-v)^2 & 1-v \\ 1-v & 1 \end{pmatrix} \quad (4.32)$$

$$E_{\tilde{2}}^{\tilde{1}} = \frac{1}{(1+v)^2 + 1} \begin{pmatrix} (1+v)^2 & 1+v \\ 1+v & 1 \end{pmatrix} \quad (4.33)$$

$$E_{\tilde{1}}^{\tilde{2}} = \frac{1}{(1+v)^2 + 1} \begin{pmatrix} (-1-v)^2 & -1-v \\ -1-v & 1 \end{pmatrix} \quad (4.34)$$

$$E_{\tilde{2}}^{\tilde{2}} = \frac{1}{(1-v)^2 + 1} \begin{pmatrix} (-1+v)^2 & -1+v \\ -1+v & 1 \end{pmatrix} \quad (4.35)$$

where $v = \sqrt{2}$, for at least some quantum realization of the measurements $\{M_i\}_{i \in I}$. We call this set of measurements in the target scenario a quantum realization of $\tilde{\mathbb{M}} := \{M_{\tilde{1}}, M_{\tilde{2}}\}$. As we will see in section 5.1, this is a set of POVM's that obtain the optimal quantum contextual behavior for the simplest scenario.

Proof. This lemma follows directly from the fact that we might trivially consider $\tilde{\mathbb{M}} \cup \{M_i\}_{i=3, \dots, |I|}$ to simulate $\tilde{\mathbb{M}}$. Here, $\{M_i\}_{i=3, \dots, |I|}$ can be any sharp quantum realization of measurement procedures. \square

This is always possible and, at this point, seems to be a fairly trivial statement. We have then showed that any set of binary outcome measurement procedures \mathbb{M} have some quantum realization with the property of simulating the quantum realization of $\tilde{\mathbb{M}}$ from the above lemma. In the jargon of the resource theory for generalized contextuality, this shows that there exists a free operation $f \in \mathcal{F}$, acting only in the measurement procedures, with

$$\mathbb{M} \xrightarrow{f} \tilde{\mathbb{M}}. \quad (4.36)$$

But from these results, we can construct the following example:

Example 11. There is always a quantum behavior $B \in \mathbb{B}$ and a free operation $f \in \mathcal{F}$

towards \mathbb{B}_{s_i} when we have that the scenario \mathbb{B} to be of the form:

$$\mathbb{P} = \{P_1, P_2, P_3, P_4\} \quad (4.37)$$

$$\mathbb{M} = \{M_1, \dots, M_I\} \quad (4.38)$$

$$\mathbb{O}_M = \{0, 1\} \quad (4.39)$$

$$\mathbb{E}_M = \emptyset \quad (4.40)$$

$$\mathbb{E}_P : \iff \frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4 \quad (4.41)$$

such that $f(B) \in \mathbb{B}_{s_i}$ is a quantum *contextual* behavior. In other words, every prepare-and-measure scenario \mathbb{B} will present quantum contextual statistics for some data-table arising from such scenario. \triangle

Proof. Let a quantum realization of the measurement procedures $\mathbb{M} = \{M_i\}$ from the scenario \mathbb{B} be such that each measurement is a projective measurement. This means that, by lemma 7 we have that the quantum realization of $\tilde{\mathbb{M}}$ is \mathbb{M} -simulable, for some quantum realization of \mathbb{M} . We conclude that for the quantum measurements $E_{\tilde{k}}^{\tilde{i}}$ from $M_{\tilde{i}} \in \tilde{\mathbb{M}}$, we have that there exists maps q_O^i, q_M such that,

$$E_{\tilde{k}}^{\tilde{i}} = \sum_{i,k} q_O^i(\tilde{k}|k) E_k^i q_M(i|\tilde{i}). \quad (4.42)$$

Let the quantum realization of the preparation procedures in \mathbb{B} be (5.13)-(5.16). Then, if we define the free operation f using the maps in equation (4.42),

$$\begin{aligned} p(k|M_i, P_j) &= \text{Tr}(E_k^i \rho^j) \xrightarrow{f} \\ &= \sum_{i,k} q_O^i(\tilde{k}|k) \text{Tr}(E_{\tilde{k}}^{\tilde{i}} \rho^j) q_M(i|\tilde{i}) \\ &= \text{Tr}\left(\sum_{i,k} q_O^i(\tilde{k}|k) E_k^i q_M(i|\tilde{i}) \rho^j\right) \\ &= \text{Tr}\left(E_{\tilde{k}}^{\tilde{i}} \rho^j\right) = p(\tilde{k}|M_{\tilde{i}}, P_j), \end{aligned}$$

where $E_{\tilde{k}}^{\tilde{i}}$ are defined by (5.10)-(5.12). Hence, we get that f leads to a behavior $\tilde{B} \in \mathbb{B}_{s_i}$ that is described by table 5.2, which is a contextual behavior. Since f is a free operation we have that the quantum realization by means of the projections over \mathbb{B} cannot be noncontextual (see lemma 3). So we conclude that it is a quantum contextual behavior from \mathbb{B} . \square

This result can be stated as follows. If we consider the scenario $\mathbb{B} := (4, |I|, 2, \mathbb{E}_P, \emptyset)$ where \mathbb{E}_P is the same set of operational equivalences as in $\mathbb{B}_{s_i} := (4, 2, 2, \mathbb{E}_P, \emptyset)$. This scenario \mathbb{B} will have quantum contextual behaviors for any $|I| \geq 3$. We notice therefore that

the resource theory uses simple arguments to prove the following intuitive result: increasing the number of measurement procedures does not affect (preparation) contextuality.

We conclude that any scenario with the same operational structure for preparation procedures and any number of measurements will present generalized contextuality. The important aspect in the framework is not that by inserting any amount of measurements, a scenario with preparation contextuality will still have preparation contextuality for granted. The novelty is that since the class of simulations for quantum operations can be studied with semi-definite programming (Guerini et al., 2017) we can use them for witnessing contextuality in *any* scenario, by trying to simulate quantum contextual behaviors in scenarios with a more straightforward operational structure. We formulate this result as a proposition:

Proposition 1. *Let $\mathbb{B}, \mathbb{B}_{known}$ be two finitely defined prepare-and-measure scenarios, with \mathbb{B}_{known} a scenario that has known realizations of quantum contextual behaviors B_{known}^Q . Let $Q(\mathbb{B})$ be the set of quantum behaviors inside the polytope $C(\mathbb{B})$. Then, if there exist quantum preparations and measurements generating the statistics $B^Q \in Q(\mathbb{B})$ that is capable of simulating a contextual behavior B_{known} , then, we have that B^Q is a contextual behavior from \mathbb{B} . We would conclude, in this case, that the scenario \mathbb{B} presents quantum contextual correlations.*

Ontological relevance of free operations

One of the issues with verifying (quantum) contextuality experimentally can be described as *the problem of operational inequalities*. We describe the problem as follows: Any operational scenario \mathbb{B} has ideal operational equivalences \mathbb{E}_P or \mathbb{E}_M , but, in any experimental implementation of the procedures, due to errors the equivalences are not necessarily verified for the data-table. The procedures that satisfy the operational equivalences might be considered ideal procedures. The ones that are truly implemented in the laboratory we name as real procedures. We can state the problem as,

How can we witness generalized contextuality for the ideal data-table, obtained with respect to ideal procedures, if the experimental data-table relates to real procedures, and does not respect the operational equivalences of the scenario?

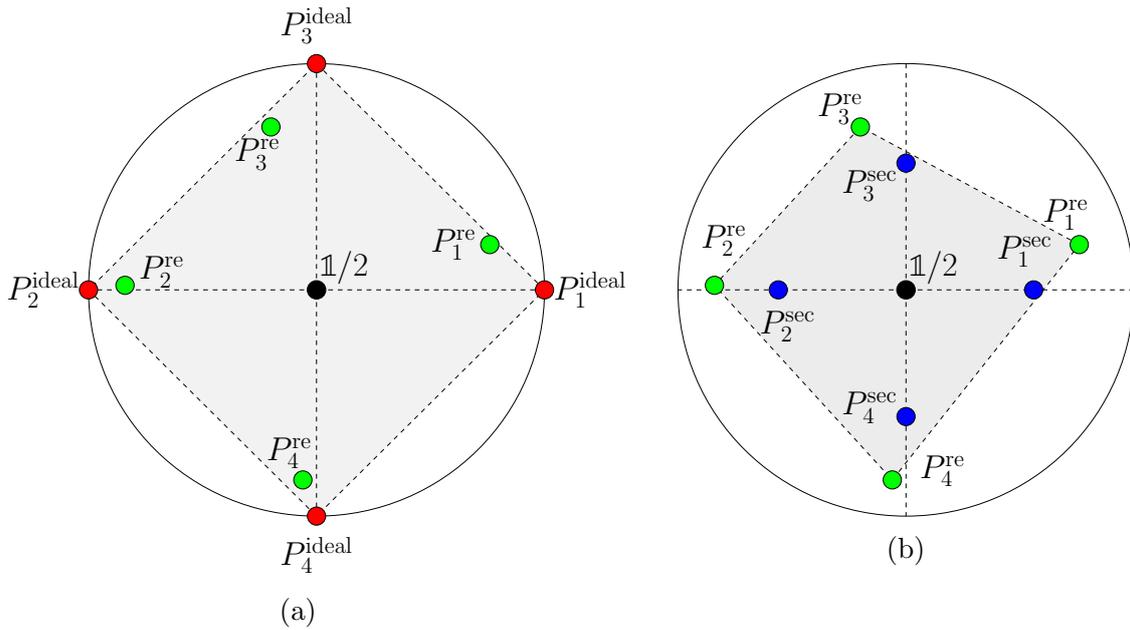


Figure 4.2: Solving the problem of operational inequalities using free operations, as described in (Mazurek et al., 2016), (Schmid and Spekkens, 2018). In figure 4.2a we consider that there are ideal procedures, that satisfy the operational equivalences described by the vertical and horizontal dashed lines, and that we would like to implement (denoted as $\{P_j^{\text{ideal}}\}_{j=1}^4$). Due to noise, the experiment implements the real procedures (denoted by $\{P_j^{\text{re}}\}_{j=1}^4$) that may fail to respect the operational equivalences. In figure 4.2b we use a free operation defining secondary procedures (denoted as $\{P_j^{\text{sec}}\}_{j=1}^4$) that respect the operational equivalences.

The proposal of (Mazurek et al., 2016) to solve this problem is to make a pre-processing of the real procedures, towards secondary procedures. We then understand the contextuality present in the secondary behavior (the pre-processed data-table obtained). In this way we can conclude the following, using the resource theory:

Theorem 9. *Let \mathbb{B} be a prepare-and-measure scenario, that has the ideal operational equivalences. Let $B_{re} \in \mathbb{B}_{re}$. This behavior is to be interpreted as the behavior obtained in the experiment, associated with the real procedures that were implemented. Then, if there exists a free operation $\mathcal{F} \ni f : \mathbb{B}_{re} \rightarrow f(\mathbb{B}_{re}) = \mathbb{B}$, we can conclude that,*

$$f(B_{re}) \in C(\mathbb{B}) \setminus NC(\mathbb{B}) \implies B_{re} \in C(\mathbb{B}_{re}) \setminus NC(\mathbb{B}_{re}) \quad (4.43)$$

In other words, generalized contextuality for the secondary behaviors imply generalized contextuality for the real behaviors.

Proof. The procedures of a real data-table $B_{re} \in \mathbb{B}_{re}$ that we consider due to errors in a GPT formalism lie inside the convex hull of the procedures associated with the ideal scenario (see figure 4.2). Any procedure inside this convex hull can be simulated by the procedures in the vertices. Therefore, there are free operations $f \in \mathcal{F}$, towards secondary

procedures, such that $f(B_{re}) \in \mathbb{B}$ lies in the ideal scenario \mathbb{B} . If $f(B_{re})$ is contextual, from lemma 3 we conclude contextuality of B_{re} . \square

It could be that, in fact, we could never access the contextuality of the real procedures, and only access the contextuality of the secondary procedures. One could then argue that the notion of generalized contextuality is consistent only with a notion of reality that depends on the pre-processing of the experimental data-table obtained. The resource theory demonstrates that this is not the case. In fact, using the same argument applied to the real procedures with respect to the ideal procedures, the resource theory guarantees that the ideal behaviors are contextual; notice from figure 4.2a that the real procedures can be simulated by the ideal procedures, meaning that there exists a free operation $\mathcal{F} \ni g : \mathbb{B} \rightarrow \mathbb{B}_{re}$, implying,

$$f(B_{re}) \in C(\mathbb{B}) \setminus NC(\mathbb{B}) \Rightarrow \underbrace{B_{re}}_{g(B_{ideal})} \in C(\mathbb{B}_{re}) \setminus NC(\mathbb{B}_{re}) \Rightarrow B_{ideal} \in C(\mathbb{B}) \setminus NC(\mathbb{B})$$

Our result strengthens the conclusions of (Mazurek et al., 2016), that generalized (quantum) contextuality was indeed verified, since we can conclude that we experimentally witnessing contextuality for the “real” procedures. There are other theoretical results reinforcing the claims of (Mazurek et al., 2016), such as (Schmid et al., 2020b). ²

Moreover, the resource theory is capable of describing bounds of how much noisy can we allow in our experiment, so that the secondary procedures will still witness contextuality of the real procedures.

Example 12. Let us consider the simplest scenario \mathbb{B}_{si} , and we define $Q(\mathbb{B}) \subseteq C(\mathbb{B})$ as the set of behaviors that are accessible with standard quantum theory. We know that, the optimal quantum contextual realizations, obtained with operators in $\mathcal{B}(\mathbb{C}^2)$, are such that the contextual fraction $f(B \in Q(\mathbb{B}_{si})) \approx 0.42$, see appendix D.2. We can conclude then that, one could use the contextual fraction as a tool to define a bound over the amount of noise allowed to still infer generalized contextuality, let the transformed behavior $f(B_{re})$ be called as the secondary behavior B_{sec} , then

$$f(B_{sec}) \in (0, 0.42] \implies B_{ideal} \in C(\mathbb{B}_{si}) \setminus NC(\mathbb{B}_{si}). \quad (4.44)$$

There are other proposals in the literature for bounding the amount of noise allowed to witness contextuality, see (Kunjwal, 2019).

\triangle

²For a detailed discussion of how to witness the set of tomographic complete procedures from the GPT framework, and obtaining bounds of experimental verification of generalized contextuality see (Mazurek et al., 2019).

4.4 Box product

We have already defined a binary operation for the resource theory in definition 22, we think that the product \otimes is reasonably interesting to work with and useful in any situation. However, we propose and study another binary operation between generalized scenarios that has a computational advantage and does not restrict the behaviors.

Definition 28. Given two behaviors $B_1 \in \mathbb{B}_1$ and $B_2 \in \mathbb{B}_2$ we define the binary operation $\boxtimes : \mathbb{B}_1 \times \mathbb{B}_2 \rightarrow \mathbb{B}$ as the union,

$$B_1 \boxtimes B_2 := \{p(k_1|M_{i_1}, P_{j_1})\} \cup \{p(k_2|M_{i_2}, P_{j_2})\} \quad (4.45)$$

With $i_1 \in I_1$, $|I_1| = |\mathbb{M}_1|$ and similarly for all other labels. The target scenario $\mathbb{B} \equiv \mathbb{B}_1 \boxtimes \mathbb{B}_2$ has the operational equivalences of both scenarios defined as, for $\{a\} := \{a_1\} \cup \{a_2\}$,

$$\gamma_P^a \in \mathbb{E}_P^{\mathbb{B}_1 \boxtimes \mathbb{B}_2}, \gamma_P^a := \begin{cases} (\alpha^{a_1}, \vec{0}; \beta^{a_1}, \vec{0}), & a = a_1 \\ (\vec{0}, \alpha^{a_2}, \vec{0}; \beta^{a_2}), & a = a_2 \end{cases}, \quad (4.46)$$

And the same definition for the operational equivalences for measurement events. As an operational constraint, the target scenario *does not* considers the probabilities of the form

$$p(k_1|M_{i_1}, P_{j_2}), p(k_2|M_{i_1}, P_{j_1}) \notin B_1 \boxtimes B_2, \quad (4.47)$$

for all the behaviors in the scenario $\mathbb{B}_1 \boxtimes \mathbb{B}_2$. \diamond

With the product given by equation (4.45) we are essentially thinking of the linear characterization of the behaviors B_i , $i = 1, 2$ from $C(\mathbb{B}_1) \subset \mathbb{R}^n, C(\mathbb{B}_2) \subset \mathbb{R}^m$, to the larger one,

$$B_1 \boxtimes B_2 \equiv (\vec{B}_1, \vec{B}_2) \in \mathbb{R}^{n+m}, \quad (4.48)$$

in terms of the polytope structure. With this definition, it is possible to prove that,

Theorem 10. *The binary operation \boxtimes preserves the resource:*

$$B_1 \in NC(\mathbb{B}_1), B_2 \in NC(\mathbb{B}_2) \Leftrightarrow B_1 \boxtimes B_2 \in NC(\mathbb{B}_1 \boxtimes \mathbb{B}_2).$$

Proof. Let $B_1 \in NC(\mathbb{B}_1), B_2 \in NC(\mathbb{B}_2)$. Hence, there are $(\Sigma^{(i)}, \Lambda^{(i)}, \Pi^{(i)}, \Theta^{(i)})$ where $\Pi^{(i)}$ and $\Theta^{(i)}$, $i = 1, 2$, respect the operational equivalences at the ontological model level respectively for each scenario. For sets of labels we define $K_i, I_i, J_i, \{a_i\}, \{b_i\}$ as before (see definition 28), for their respective operational primitives from \mathbb{B}_i . The scenarios are finite and the operational equivalences are fixed and finite as well, so each set ranges over

a finite set of labels.

$$\begin{aligned}
 K &:= K_1 \cup K_2, \\
 I &:= I_1 \cup I_2, \\
 J &:= J_1 \cup J_2, \\
 \{a\} &:= \{a_1\} \cup \{a_2\}, \\
 \{b\} &:= \{b_1\} \cup \{b_2\},
 \end{aligned}$$

By definition of noncontextuality at the ontological model level, we have the equations,

$$p(k_1|i_1, j_1) = \sum_{\lambda_1 \in \Lambda_1} \xi_{[k_1|i_1]}(\lambda_1) \mu_{j_1}(\lambda_1), \quad (4.49)$$

$$p(k_2|i_2, j_2) = \sum_{\lambda_2 \in \Lambda_2} \xi_{[k_2|i_2]}(\lambda_2) \mu_{j_2}(\lambda_2), \quad (4.50)$$

$$\sum_{j_1} (\alpha_{j_1}^{a_1} - \beta_{j_1}^{a_1}) \mu_{j_1}(\Omega_1) = 0, \quad \forall a_1, \forall \Omega_1 \in \Sigma_1 \quad (4.51)$$

$$\sum_{j_2} (\alpha_{j_2}^{a_2} - \beta_{j_2}^{a_2}) \mu_{j_2}(\Omega_2) = 0, \quad \forall a_2, \forall \Omega_2 \in \Sigma_2 \quad (4.52)$$

$$\sum_{k_1, i_1} (\alpha_{i_1}^{b_1} - \beta_{i_1}^{b_1}) \xi_{[k_1|i_1]}(\lambda_1) = 0, \quad \forall \lambda_1, b_1, \quad (4.53)$$

$$\sum_{k_2, i_2} (\alpha_{i_2}^{b_2} - \beta_{i_2}^{b_2}) \xi_{[k_2|i_2]}(\lambda_2) = 0, \quad \forall \lambda_2, b_2. \quad (4.54)$$

In (4.51)-(4.54), $\xi_{[\cdot|i_1]} \in \Theta^{(1)}$, $\mu_{j_1} \in \Pi^{(1)}$, and similarly for the rest of ontology measures. If we consider the product between the behaviors, $B_1 \boxtimes B_2$, we get that for the ontological description it is possible to set $\Lambda := \Lambda^{(1)} \sqcup \Lambda^{(2)}$ the disjoint union between the two sets. We then define $\tilde{\xi}_{[k|i]} : \Lambda \rightarrow [0, 1]$ as,

$$\tilde{\xi}_{[k|i]}(\lambda) := \begin{cases} \xi_{[k_1|i_1]}(\lambda_1), & \lambda = (\lambda_1, 1) \\ \xi_{[k_2|i_2]}(\lambda_2), & \lambda = (\lambda_2, 2) \end{cases}$$

$$\forall \lambda \in \Lambda, \sum_{k \in K} \tilde{\xi}_{[k|i]}(\lambda) = \begin{cases} \sum_{k \in K} \tilde{\xi}_{[k|i_1]}(\lambda), & \text{if } i \in I_1 \\ \sum_{k \in K} \tilde{\xi}_{[k|i_2]}(\lambda), & \text{if } i \in I_2 \end{cases} = \begin{cases} \sum_{k_1 \in K_1} \tilde{\xi}_{[k_1|i_1]}(\lambda), & \text{if } i \in I_1 \\ \sum_{k_1 \in K_1} \tilde{\xi}_{[k_1|i_2]}(\lambda), & \text{if } i \in I_2 \end{cases} = \begin{cases} 1, \\ 1 \end{cases}$$

So that the extended functions are normalized in the ontic space Λ . We have considered that, whenever $i \in I_1 \cap I_2$ any function $\xi_{[k_1|i_1]}$ or $\xi_{[k_2|i_2]}$ will serve, we then just need to pick one and use it for our noncontextual ontological model. This means that if we have two scenarios if the same procedures (not only same labels), $\{M_1, M_2\}, \{M_1, M_2\} \rightarrow \{M_{11}, M_{21}, M_{12}, M_{22}\} \equiv \{M_1, M_2\}$. Therefore we can recognize if two procedures are just simply the same. In this sense, we can have that the number of procedures in \mathbb{B}_{si} and $\mathbb{B}_{si}^{\boxtimes n}$

are the same, so that we simplify the scenario. This is a specific case when we consider how the different scenarios talk with one another.

For $\tilde{\mu}_j$, the ontic spaces are finite and we write $\tilde{\mu}_j(\{\lambda\}) \equiv \tilde{\mu}_j(\lambda)$. Let $j \in J$, we define that, if $j \in J_1$

$$\tilde{\mu}_j(\lambda) := \begin{cases} \mu_{j_1}(\lambda), & \text{if } \lambda = (\lambda_1, 1) \\ 0, & \text{if } \lambda = (\lambda_2, 2) \end{cases}$$

and similarly if $j \in J_2$. Again, when $j \in J_1 \cap J_2$ we choose one of the ontological descriptions as our fixed definition for the preparation procedure associated with it. With this definition we have that, for any $j \in J$, we will have that

$$\sum_{\lambda \in \Lambda} \tilde{\mu}_j(\lambda) = \sum_{\lambda=(\lambda_1,1) \in \Lambda} \tilde{\mu}_j(\lambda) + \sum_{\lambda=(\lambda_2,2) \in \Lambda} \tilde{\mu}_j(\lambda)$$

and, whenever $j \in J_1$ or $j \in J_2$ we recover the normalization condition from the already defined distributions in the parts. We then obtain that for any $p(k|i, j) \in B_1 \boxtimes B_2$ we will have that this probability comes from one of the two behaviors, in this sense,

$$\begin{aligned} \sum_{\lambda \in \Lambda_1 \sqcup \Lambda_2} \tilde{\xi}_{[k|i]}(\lambda) \tilde{\mu}_j(\lambda) &= \begin{cases} \sum_{\lambda \in \Lambda} \tilde{\xi}_{[k_1|i_1]}(\lambda) \tilde{\mu}_{j_1}(\lambda), & \text{if } k, i, j \in K_1, I_1, J_1 \\ \sum_{\lambda \in \Lambda} \tilde{\xi}_{[k_2|i_2]}(\lambda) \tilde{\mu}_{j_2}(\lambda), & \text{if } k, i, j \in K_2, I_2, J_2 \end{cases} \\ &= \begin{cases} \sum_{\lambda_1 \in \Lambda_1} \xi_{[k_1|i_1]}(\lambda_1) \mu_{j_1}(\lambda_1), & \text{if } k, i, j \in K_1, I_1, J_1 \\ \sum_{\lambda_2 \in \Lambda_2} \xi_{[k_2|i_2]}(\lambda_2) \mu_{j_2}(\lambda_2), & \text{if } k, i, j \in K_2, I_2, J_2 \end{cases} \\ &= \begin{cases} p(k_1|i_1, j_1), & \text{if } k, i, j \in K_1, I_1, J_1 \\ p(k_2|i_2, j_2), & \text{if } k, i, j \in K_2, I_2, J_2 \end{cases} \\ &= B_1 \boxtimes B_2 \end{aligned}$$

Notice that in the new scenario $B_1 \boxtimes B_2$ it is at play our operational constraint that the preparations of the parts do not interact with the measurements of one another. We will later discuss these assumptions and possible relaxations thereof that might be applicable and useful. The operational equivalences defined in the scenario $B_1 \boxtimes B_2$ are the ones from (4.46), so we need to study the following objects:

$$\sum_j (\alpha_j^a - \beta_j^a) \tilde{\mu}_j(\lambda), \quad \forall \lambda \in \Lambda_1 \sqcup \Lambda_2, a \in \{a_1\} \cup \{a_2\}, \quad (4.55)$$

where we can write an arbitrary order of $\{a\} = \{a_1\} \cup \{a_2\}$ as $a = 1, \dots, a_1, a_1+1, \dots, a_1+a_2$. So that for $\{a\}$ - see definition 28 if a clarification is needed, essentially $\{a\}$ labels the number of operational equivalences we have - it will be true, for all $\lambda \in \Lambda_1 \sqcup \Lambda_2$, the following holds,

$$\begin{aligned}
 \sum_j (\alpha_j^a - \beta_j^a) \tilde{\mu}_j(\lambda) &\stackrel{(4.51), (4.52)}{=} \underbrace{\sum_{j_1} (\alpha_{j_1}^{a_1} - \beta_{j_1}^{a_1}) \tilde{\mu}_{j_1}(\lambda)}_{=0} + \sum_{j=J_1+1}^{J_1+J_2} (\alpha_j^{a_1} - \beta_j^{a_1}) \tilde{\mu}_{j_2}(\lambda) \\
 &+ \sum_{j_1} (\alpha_{j_1}^{a_2} - \beta_{j_1}^{a_2}) \tilde{\mu}_{j_1}(\lambda) + \underbrace{\sum_{j=J_1+1}^{J_1+J_2} (\alpha_j^{a_2} - \beta_j^{a_2}) \tilde{\mu}_{j_2}(\lambda)}_{=0} \\
 &= \sum_{j_1} (\alpha_{j_1}^{a_2} - \beta_{j_1}^{a_2}) \tilde{\mu}_{j_1}(\lambda) + \sum_{j=J_1+1}^{J_1+J_2} (\alpha_j^{a_1} - \beta_j^{a_1}) \tilde{\mu}_{j_2}(\lambda) \stackrel{(4.46)}{=} 0.
 \end{aligned}$$

And for all $\lambda \in \Lambda_1 \sqcup \Lambda_2$ we also have,

$$\begin{aligned}
 \sum_{k,i} (\alpha_{[k|i]}^b - \beta_{[k|i]}^b) \tilde{\xi}_{[k|i]}(\lambda) &= \sum_{k_1, i_1} (\alpha_{[k_1|i_1]}^{b_1} - \beta_{[k_1|i_1]}^{b_1}) \tilde{\xi}_{[k_1|i_1]}(\lambda_1) + \sum_{k_2, i_2} (\alpha_{[k_2|i_2]}^{b_1} - \beta_{[k_2|i_2]}^{b_1}) \tilde{\xi}_{[k_2|i_2]}(\lambda_2) \\
 &+ \sum_{k_1, i_1} (\alpha_{[k_1|i_1]}^{b_2} - \beta_{[k_1|i_1]}^{b_2}) \tilde{\xi}_{[k_1|i_1]}(\lambda_1) + \sum_{k_2, i_2} (\alpha_{[k_2|i_2]}^{b_2} - \beta_{[k_2|i_2]}^{b_2}) \tilde{\xi}_{[k_2|i_2]}(\lambda_2) = 0
 \end{aligned}$$

Remembering that whenever we write the vectors $(\alpha^{a_1}, \vec{0}; \beta^{a_1}, \vec{0}) \in \mathbb{E}_P^{\mathbb{B}_1 \boxtimes \mathbb{B}_2}$ we have that the following is meant:

$$\gamma_P^{a_1} = (\alpha_1^{a_1}, \alpha_2^{a_1}, \dots, \alpha_{J_1}^{a_1}, \underbrace{0, \dots, 0}_{J_2 \text{ times}}, \beta_1^{a_1}, \dots, \beta_{J_1}^{a_1}, \underbrace{0, \dots, 0}_{J_2 \text{ times}}),$$

so that we get

$$\sum_{j=J_1+1}^{J_1+J_2} (\alpha_j^{a_1} - \beta_j^{a_1}) \tilde{\mu}_{j_2}(\lambda) = 0$$

by construction. We have the same for measurement procedures. This proves that the ontological model constructed is noncontextual for the behavior $B_1 \boxtimes B_2$ whenever B_1, B_2 are also noncontextual behaviors.

For the (\Leftarrow) part of the proof, suppose that the behavior $B_1 \boxtimes B_2$ has a noncontextual ontological model $(\Sigma, \Lambda, \Pi, \Theta)$. We know that this $\mathbb{B}_1 \boxtimes \mathbb{B}_2$ scenario has the same operational equivalences as both the scenarios \mathbb{B}_1 and \mathbb{B}_2 divided, by means of the weight vectors, e.g., $(\alpha_1^{a_1}, \alpha_2^{a_1}, \dots, \alpha_{j_1}^{a_1}, 0, \dots, 0)$. Hence, there exists an ontological model for B_1 inherited from $B_1 \boxtimes B_2$ using the operational equivalences:

$$\sum_j (\alpha_j^a - \beta_j^a) \mu_j(\lambda) = 0, \forall \lambda \implies \sum_{j_1} (\alpha_{j_1}^a - \beta_{j_1}^a) \mu_{j_1}(\lambda) = 0, \forall \lambda, \forall \gamma_P^a \in \mathbb{E}_P^{\mathbb{B}_1 \boxtimes \mathbb{B}_2},$$

where we can restrict $\{a\}$ to some set of labels $\{a_1\}$ and reduced vectors $\gamma_P^{a_1}$ by cutting

the zeros. We get the same for the behavior B_2 . The ontological description of the probabilities we get immediately:

$$p(k_1|i_1, j_1) := \sum_{\lambda \in \Lambda} \xi_{[k=k_1|i=i_1]}(\lambda) \mu_{j=j_1}(\lambda), \quad (4.56)$$

for any $k_1, i_1, j_1 \in K_1, I_1, J_1$, and similarly,

$$p(k_2|i_2, j_2) := \sum_{\lambda \in \Lambda} \xi_{[k=k_2|i=i_2]}(\lambda) \mu_{j=j_2}(\lambda). \quad (4.57)$$

Therefore, we get the wanted result,

$$B_1 \boxtimes B_2 \in NC(\mathbb{B}_1 \boxtimes \mathbb{B}_2) \Leftrightarrow B_1 \in NC(\mathbb{B}_1), B_2 \in NC(\mathbb{B}_2)$$

□

We notice that an important feature of such construction is that we do not require that the probability distributions associated with the behaviors B_1, B_2 need to be independent, for the construction of $B_1 \boxtimes B_2$ contrasted with definition 22. In particular, the product \boxtimes preserves the quantum contextual structure,

$$B_1 \boxtimes B_2 \in QC(\mathbb{B}_1 \boxtimes \mathbb{B}_2) \iff B_1 \in QC(\mathbb{B}_1), B_2 \in QC(\mathbb{B}_2), \quad (4.58)$$

where $QC(\mathbb{B}) := Q(\mathbb{B}) \setminus NC(\mathbb{B})$ and $Q(\mathbb{B})$ is the set of data-tables realizable by quantum operational prescriptions. More generally, we get that the quantum descriptions are preserved. Such a thing gives us the possibility to see noncontextuality as a building block feature, where one might have these blocks \mathbb{B} of scenarios that are defined operationally in a prepare-and-measure sense implying that scenarios with higher complexity can be constructed and still present quantum contextuality. For example, for any number N of scenarios, we will have $\mathbb{B}_1 \boxtimes \dots \boxtimes \mathbb{B}_N$ is again a prepare-and-measure scenario with unknown quantum contextuality, but, employing the last theorem, the scenario

$$\mathbb{B}_{si} \boxtimes \mathbb{B}_1 \boxtimes \dots \boxtimes \mathbb{B}_N$$

is guaranteed to allow for quantum contextual behaviors, inherited from the quantum contextuality of \mathbb{B}_{si} .

Notice that what we are effectively doing is the following procedure: let $B_1 \in NC(\mathbb{B}_1)$ and $B_2 \in NC(\mathbb{B}_2)$. Then, by a suitable choice of labeling the entries associated with the polytope, we have that $B_1 \boxtimes B_2 \rightarrow (\vec{B}_1, \vec{B}_2) \in NC(\mathbb{B}_1 \boxtimes \mathbb{B}_2)$. Such a construction is actually a well-known relation between convex polytopes.

Lemma 8. (*Henk et al., 2004*) *Let $P \subset \mathbb{R}^n$, $Q \subset \mathbb{R}^m$ be two convex polytopes. Then, the*

product defined by

$$P \times Q := \left\{ \begin{pmatrix} p \\ q \end{pmatrix} : p \in P, q \in Q \right\} \subset \mathbb{R}^{n+m}, \quad (4.59)$$

is again a convex polytope. Let $|V(P)|$ and $|V(Q)|$ represent the number of vertices of each of the convex polytopes P and Q , then, we also have that $|V(P \times Q)| = |V(P)| \cdot |V(Q)|$. Let $|F(P)|$ define the number of facets of the convex polytope P , and similarly for the convex polytope Q . Then, we have that $|F(P \times Q)| = |F(P)| + |F(Q)|$.

As we can see from lemma 8, we can know the number of vertices and the number of facets associated with the new larger polytope. As we see from the proof below, we can know these numbers because we can describe all the facets and all the vertices. This feature, looking at our noncontextual polytopes, represents a significant computational advantage for finding the tight noncontextuality inequalities associated with the larger scenarios $\mathbb{B}_1 \boxtimes \cdots \boxtimes \mathbb{B}_2$.

Proof. Let P and Q be two convex polytopes. Let the full characterization of these convex polytopes be defined by the tight inequalities $M_P p \leq b_P$ and $M_Q q \leq b_Q$, with \leq here representing element-wise ordering for real numbers.

$$p \in P \iff M_P p \leq b_P, \quad (4.60)$$

$$q \in Q \iff M_Q q \leq b_Q. \quad (4.61)$$

We call such representation of the convex polytopes as the H-representation. Now, the H-representation for the product polytope $P \times Q$ follows from (4.60)-(4.61), by making

$$M_{P \times Q} := \begin{pmatrix} M_P & \mathbb{0}_P \\ \mathbb{0}_Q & M_Q \end{pmatrix} \quad (4.62)$$

where $\mathbb{0}_P$ is a matrix with zeros and the same dimensionality as M_P , and similarly for $\mathbb{0}_Q$. We then have that

$$\begin{pmatrix} p \\ q \end{pmatrix} \in P \times Q \iff M_{P \times Q} \begin{pmatrix} p \\ q \end{pmatrix} \leq \begin{pmatrix} b_P \\ b_Q \end{pmatrix} \quad (4.63)$$

where (\Leftrightarrow) follows from (4.60)-(4.61). We have then proved that $|F(P \times Q)| = |F(P)| + |F(Q)|$ since each tight inequality defines a facet of the convex polytope. The equation for the vertices follows from the same reasoning. Let $\{p_v\}_v$ be the set of vertices defining P . The convex hull of $\{p_v\}_v$ can fully describe the convex polytope, known as the V-representation. Since there exists a fundamental duality theorem between the V and H representation of a polytope (Brøndsted, 2012), we might write that $P \times Q$ is the convex

hull of $\{(p_v, q_v)\}_v$, (Paffenholz, 2006, page 3):

$$P \times Q = \text{conv}(V(P \times Q)) = \text{conv}(\{(p_v, q_v) : p_v \in V(P), q_v \in V(Q)\}) \quad (4.64)$$

where we denote $\text{conv}(\cdot)$ as the convex hull for all the vertices. With this we show the equation from the lemma for the number of vertices in the product polytope. \square

Noticing that the map \boxtimes has the same structural effect over the polytopes that are constructed from the scenario, we conclude as a corollary that,

Lemma 9. *Let $\mathbb{B}_1, \mathbb{B}_2$ be two operational scenarios. Let $C(\mathbb{B}_i)$ and $NC(\mathbb{B}_i)$ for $i = 1, 2$ be, respectively, the polytope of all behaviors and the noncontextual polytope. Since by theorem 10 the map \boxtimes preserves the noncontextual structure,*

$$NC(\mathbb{B}_1 \boxtimes \mathbb{B}_2) = NC(\mathbb{B}_1) \times NC(\mathbb{B}_2) \quad (4.65)$$

where we are strongly using the update in the operational structure given by definition 28. Suppose that one knows the tight noncontextuality inequalities that define each part $NC(\mathbb{B}_1)$ and $NC(\mathbb{B}_2)$. Then, from lemma 8 we know the tight noncontextuality inequalities defining the polytope $NC(\mathbb{B}_1 \boxtimes \mathbb{B}_2)$.

As we will see in the next chapter, lemma 9 represents a great simplification associated with obtaining noncontextuality inequalities whenever the inner polytope structure of the scenario is understood. We can use the resource theory to obtain noncontextuality inequalities that are then robust, operationally described without assuming quantum theory and are related to quantum tasks.

Our results also suggest that generalized contextuality is a *good* resource to have in mind whenever thinking about complex scenarios. Suppose that we construct an operational task that has the following characterization: $\mathbb{B} := \mathbb{B}_{s_i} \boxtimes \mathbb{B}_1 \boxtimes \cdots \boxtimes \mathbb{B}_N$, for any $N \geq 1$. Since we know that \mathbb{B}_{s_i} has quantum contextual correlations, no matter how large N will be, and no matter how the scenarios $\{\mathbb{B}_i\}_{i=1, \dots, N}$ are described, because of theorem 10 we must have that \mathbb{B} have quantum contextual realizations. This means that no matter how complex our scenarios are, we can engineer new operational scenarios that are guaranteed to have the resource. In figure 4.3 we have tried to provide an intuitive view of this fact.

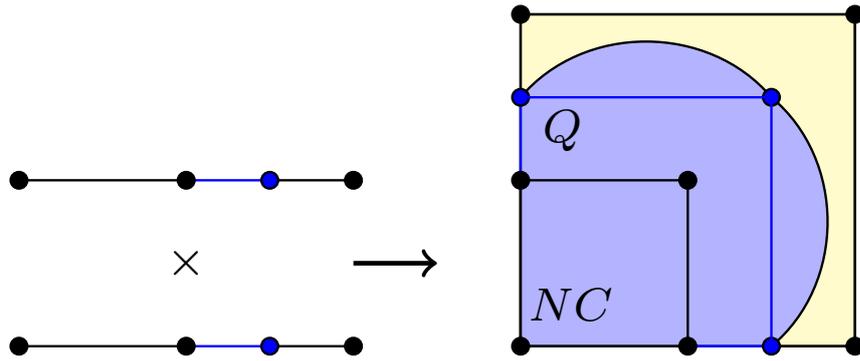


Figure 4.3: Representation of the polytope structure arising from the product scenario. We stress that it is not clear how should be the new form of the *quantum* set $Q(\mathbb{B})$, even though it is clear the polytope structure for both NC and the larger polytope of statistics. In this picture we have used the fact that $NC(\mathbb{B}) \subseteq Q(\mathbb{B})$. From the convex nature of $Q(\mathbb{B})$, the product scenario must have a quantum contextual set that is at least of the form given by curve 1, but it could also be given by curve 2 and the study of maximal violations for noncontextuality inequalities shall answer such questions (Ambainis et al., 2016), (Chailloux et al., 2016).

In figure 4.3 we have the visual description of the product between two polytopes, defined by the same set of vertices $\mathbb{B}_{1,2} = \text{conv}\{0, 1\}$, that imply the new set of vertices $\mathbb{B}_1 \boxtimes \mathbb{B}_2 = \text{conv}(\{(0, 0), (0, 1), (1, 0), (1, 1)\})$. If we suppose that in blue we have quantum behaviors, we an intuitive description of, supposing that we have quantum behaviors outside the noncontextual set, depicted by the blue vertices, they will still be outside the noncontextual vertex of the larger dimensional polytope, constructed with the box product. As we have shown in figure 4.3 we have no idea of how the quantum set will behave, despite the fact that this is a convex set, and we have depicted that letting the quantum set to be a little bit round.

QUANTUM INFORMATION TASKS

(...) quantum computing is transitioning from a research topic to a technology that unlocks new computational capabilities. We are only one creative algorithm away from valuable near-term applications.

- (Arute et al., 2019)

We now look further the applicability of the theoretical developments made so far. Contextuality has been shown to represent statistical contextual correlations for quantum information prepare-and-measure protocols. In what follows we present the protocols and their relevance for quantum information theory, as well as their quantum advantage witnessed by generalized contextuality.

In this chapter we will study a classical information task in section 5.1, the n -bit parity oblivious multiplexing task, and we will learn that quantum theory offers an essential improvement of its success rate. Then, we will proceed to study two quantum information tasks, quantum state discrimination in section 5.2 and state-dependent quantum cloning in section 5.3. These two tasks are known to be related. We will witness quantum contextuality in the experimental scenario describing these tasks, and understand the role generalized contextuality play.

5.1 Parity Oblivious Multiplexing

The first proof that quantum contextuality is advantageous was provided for an information protocol called n -bit parity-oblivious multiplexing (Spekkens et al., 2009a), which is a class of communication protocols, also called sometimes as oblivious communication (Saha and Chaturvedi, 2019; Saha et al., 2019). The protocol is described in the following way: consider that Alice has an n -bit string produced uniformly at random. She then sends a system to Bob that stores some information about her bit string, with the counterpart that there is a constraint to be respected; Bob then generates a random number between 1 and n , so that the task of Bob is, given that he generates number

$y \in \{1, \dots, n\}$, he guesses what is the bit in the y position of the string. The constraint imposed on the amount of information Alice can send gives the name of the protocol; Alice cannot inform the parity of the string to Bob. Mathematically this reads as follows (Spekkens et al., 2009a): let Par be the set defined by,

$$\text{Par} := \left\{ r \mid r \in \{0, 1\}^n, \sum_i r_i \geq 2 \right\} \quad (5.1)$$

For $n = 3$ we have for example $(1, 1, 0) \in \text{Par}$. For $n = 2$ the only string in Par is $(1, 1)$. Then, Alice cannot send information about the products of her bit string with elements of Par .

$$x \cdot r = \bigoplus_i x_i r_i \quad (5.2)$$

Where \oplus represents sum module 2. Putted in other words, the parity constraint implies that any measurement procedure performed by Bob cannot gather statistical information about the result of product $x \cdot r$; hence the statistics obtained from the measurements that give $x \cdot r = 1$ or $x \cdot r = 0$ need to be exactly the same:

$$\forall r \in \text{Par}, \forall M, \forall k : \sum_{x|x \cdot r=0} p(k|M, P_x) = \sum_{x|x \cdot r=1} p(k|M, P_x) \quad (5.3)$$

In other words, the restriction is not over *one* parity message. The restriction is over *all* possible parity checks, described in terms of the strings in the set Par , that could be performed by Bob, given an n -bit string.

For the 2-bit task we have the following operational restriction:

$$\forall M, \forall k : p(k|M, P_{(0,0)}) + p(k|M, P_{(1,1)}) = p(k|M, P_{(0,1)}) + p(k|M, P_{(1,0)}). \quad (5.4)$$

The description for an arbitrary operational protocol fits perfectly the prepare-and-measure scenario description from before. In fact, for a 2-bit protocol, the equivalent scenario corresponds to the simplest scenario from equation (2.21), with the noncontextual polytope given by (C.14)-(C.22).

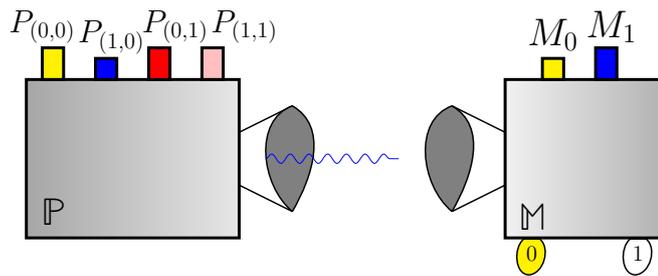


Figure 5.1: Operational description of 2-bit parity-oblivious multiplexing prepare-and-measure scenario. Alice generates a random string $x \in \{(0, 0), (1, 0), (0, 1), (1, 1)\}$, while Bob generates a number $y \in \{0, 1\}$. Corresponding to each of these outcomes we have some probability that Bob guesses the bit correctly.

From an operational perspective we have that the probability of success for Bob is given by,

$$p(g = x_y) = \frac{1}{2 \cdot 2^2} \sum_y \sum_x p(g = x_y | M_y, P_x), \quad (5.5)$$

This means that in general, the task given by Bob of predicting the y -th bit of Alice can be described as a function that takes operational behaviors $B \in \mathbb{B}_{si}$, for 2-bit oblivious communication, and sends it towards $p(g = x_y)$ by means of (5.5). The goal is then to understand what is the connection of noncontextuality with the best way one can construct probabilities and shuffle the behaviors from \mathbb{B}_{si} with the appropriate normalization given by (5.5), so that the success strategy is maximal.

Noncontextual bound

It is shown in (Spekkens et al., 2009a) that the best strategy for Bob to guess correctly the bit in a noncontextual ontological model is given by p such that $p \leq (n + 1)/2n$. For the 2-bit case we have that $p \leq 0.75$.

Theorem 11. (Spekkens et al., 2009a) *For any noncontextual ontological model describing the probability of success for n -bit parity oblivious multiplexing communication task, the upper bound over the guessing strategy is given by*

$$p(g = x_y) \leq \frac{n + 1}{2n}. \quad (5.6)$$

Then, by choosing the contextual behavior from table 5.2 with the labels $P_{(0,1)} = P_1, P_{(1,0)} = P_2, P_{(0,0)} = P_3, P_{(1,1)} = P_4$ we get that the probability of success for Bob in such quantum realization is of $p(q = x_y) = 0.8535 > 0.75$. What this shows is that the behavior 5.2 will present better strategies for the quantum task, by means of the quantum realizations, than any noncontextual one. That represents a proof that quantum contextuality is providing the advantage over classical analogs. A proof of theorem 11 can be found in appendix E. In what follows we provide a quantum contextual behavior

that provides $p(g = x_y) = 0.85$ and we also prove that this is the best bound that can be achieved by quantum theory.

Quantum realization in the simplest scenario

As a concrete useful example, the so called simplest scenarios \mathbb{B}_{si} from (2.21) represents operationally the 2-bit parity oblivious multiplexing protocol - were we drop the setting description, with definition 7, defining the scenario in the form of (3.24)¹

$$\mathbb{B}_{si} = (\mathbb{P}, \mathbb{M}, \mathbb{O}_M, \mathbb{E}_P, \mathbb{E}_M). \quad (5.7)$$

A possible quantum realization within the simplest scenario is given by (Pusey, 2018), were we simply consider that M_1, M_2 be given by,

$$M_1 = \frac{1}{\sqrt{2}}(\sigma_X + \sigma_Z), M_2 = \frac{1}{\sqrt{2}}(\sigma_X - \sigma_Z), \quad (5.8)$$

defined by the following effects $M_1 = \{E_1^1, E_2^1\}$, $M_2 = \{E_1^2, E_2^2\}$:

$$E_1^1 = \frac{1}{(1-v)^2 + 1} \begin{pmatrix} (1-v)^2 & 1-v \\ 1-v & 1 \end{pmatrix} = |v_1\rangle \langle v_1|, \quad (5.9)$$

$$E_2^1 = \frac{1}{(1+v)^2 + 1} \begin{pmatrix} (1+v)^2 & 1+v \\ 1+v & 1 \end{pmatrix} = |v_2\rangle \langle v_2|, \quad (5.10)$$

$$E_1^2 = \frac{1}{(1+v)^2 + 1} \begin{pmatrix} (-1-v)^2 & -1-v \\ -1-v & 1 \end{pmatrix} = |v'_1\rangle \langle v'_1|, \quad (5.11)$$

$$E_2^2 = \frac{1}{(1-v)^2 + 1} \begin{pmatrix} (-1+v)^2 & -1+v \\ -1+v & 1 \end{pmatrix} = |v'_2\rangle \langle v'_2|, \quad (5.12)$$

where $v = \sqrt{2}$. The choice is such that $|v_1\rangle, |v_2\rangle$ be the eigenvectors of $\frac{1}{\sqrt{2}}(\sigma_X + \sigma_Z)$ and $|v'_1\rangle, |v'_2\rangle$ the eigenvectors of $\frac{1}{\sqrt{2}}(\sigma_X - \sigma_Z)$. This POVM represent a quantum realization for the simplest scenario. For the preparations we can simply have, as is usual, the preparations to be equivalent to

$$\rho^1 = |0\rangle \langle 0| \quad (5.13)$$

$$\rho^2 = |1\rangle \langle 1| \quad (5.14)$$

$$\rho^3 = |+\rangle \langle +| \quad (5.15)$$

$$\rho^4 = |-\rangle \langle -| \quad (5.16)$$

¹Remembering that $\mathbb{E}_P : \iff \frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$, described by the linear characterization from (2.21)

such that this realization gives rise immediately to the correct operational equivalences for preparation procedures from \mathbb{B}_{si} . Such quantum realization gives rise to the following quantum behavior:

	ρ^1	ρ^2	ρ^3	ρ^4
E_1^1	$\left(\frac{1-v}{\sqrt{(1-v)^2+1}}\right)^2$	$\left(\frac{1}{\sqrt{(1-v)^2+1}}\right)^2$	$\left(\frac{2-v}{\sqrt{2((1-v)^2+1)}}\right)^2$	$\left(\frac{-v}{\sqrt{2((1-v)^2+1)}}\right)^2$
E_1^2	$\left(\frac{ -1-v }{\sqrt{(1+v)^2+1}}\right)^2$	$\left(\frac{1}{\sqrt{(1+v)^2+1}}\right)^2$	$\left(\frac{v}{\sqrt{2((1+v)^2+1)}}\right)^2$	$\left(\frac{-2-v}{\sqrt{2((1+v)^2+1)}}\right)^2$

Table 5.1: Quantum Behavior for the simplest scenario.

That for a sufficiently large number of repeated data collection give rise to the following probabilistic data-table. Here we omit probabilities corresponding to $1 - p$ for the events.

	ρ^1	ρ^2	ρ^3	ρ^4
E_1^1	0.1464	0.8535	0.1464	0.8535
E_1^2	0.8535	0.1464	0.1464	0.8535

Table 5.2: Data-table for the final statistics obtained by quantum predictions.

But since we have a full characterization of the noncontextual polytope for such scenario (Schmid et al., 2018) we notice that such a behavior is *contextual* by means of (C.21).

$$p_{12} + p_{24} - p_{22} - p_{13} = 0.8535 + 0.8525 - 0.1464 - 0.1464 = 1.4132 > 1.$$

Other oblivious tasks and discussion

Parity oblivious multiplexing represents one restriction of a large class of communication protocols known as *oblivious communication*. These class of communication is extremely fundamental for beyond quantum protocols. For the class of n -bit quantum parity oblivious scenarios there are known optimal quantum bounds (Ghorai and Pan, 2018; Chailloux et al., 2016), such that they achieve a probability of success given by

$$p(g = x_y) = \frac{1}{2} \left(1 + \frac{1}{\sqrt{n}}\right). \quad (5.17)$$

One can achieve such result by using the relation of parity oblivious scenarios and the Tsirelson bound for Kochen-Specker scenarios (Chailloux et al., 2016).

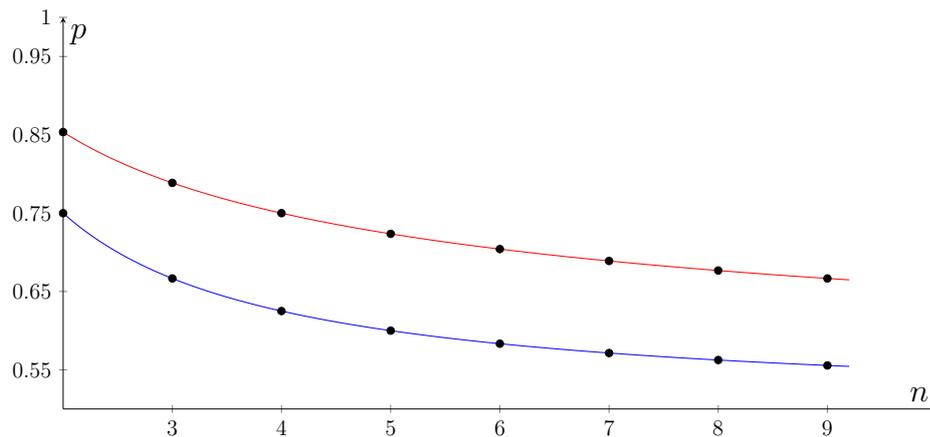


Figure 5.2: Comparison between the optimal guessing strategy, $p(g = x_y)$, classical bound provided by noncontextual models of n -bit parity oblivious multiplexing communication protocol with respect to the optimal quantum strategy. We see that not only for the scenarios corresponding to 2 and 3 bits, we have that quantum theory is advantageous, as shown in (Spekkens et al., 2009a), also for any n -bit scenario, $n \geq 4$ we have contextual advantage.

Since best guessing strategy can be achieved by finite dimensional systems, the question remains for what are the states that do achieve the optimal strategies. Because the structure of the quantum set for Spekkens scenarios is not fully understood there are still no direct algorithmic strategy to find maximal contextual behaviors. Nevertheless, there are partial results obtained specifically for oblivious communications tasks, where there has been considered the following approach (Ambainis et al., 2016):

1. First, since the guessing strategy is directly associated with contextuality simply construct an SDP² to find, dimensionally fixed, pure states and POVM's that maximise the guess $p(g = x_y)$ for n -bits. Since this strategy looks for quantum states within the scenario we will have a lower bound for quantum description.
2. Second, using an addaptation of NPA³ methods⁴ run an algorithm for as many hierarchies possible. Using NPA process we can find an upper bound, and the upper bound decreases as we increase the number of hierarchies considered.

This approach has been very successful for algorithmically finding close bounds for lower dimensional systems, very close to the optimal quantum guessing strategy. The cost is that see-saw lower bounds are very simple to obtain but NPA upper bounds are computationally demanding as the number of dimensions of the system increases.

²Semidefinite programming.

³NPA is an acronym that stands for Navascués, Pironio and Acín.

⁴NPA methods are convex optimization problems for finding the set of quantum correlations using a hierarchy of programs. For each step in the hierarchy, the algorithm approaches the quantum set from above: $Q^{k+1} \subset Q^k$, where Q^k is the quantum set after k steps, (Navascués et al., 2008).

5.2 Quantum State Discrimination

Another important quantum information task that is directly associated with generalized contextuality is quantum minimum error state discrimination (Schmid and Spekkens, 2018). The ideas for detecting specific states, that represent the distinction between two (or more) different hypothesis: system in state ρ_0 corresponding to hypothesis H_0 or system in state ρ_1 corresponding with hypothesis H_1 , and so on, are firstly reviewed, and also to some extent introduced, in the work of (Helstrom, 1969). In full generality, the estimation between states is a very difficult task that has been treated analytically for some specific cases (Helstrom, 1969), and that has been structured by means of linear programming and semi-definite programming techniques to obtain general optimal strategies, providing known necessary and sufficient conditions for reaching optimal discrimination between multiple hypothesis (Yuen et al., 1975; Barnett and Croke, 2008). One of many quantum detection paradigms corresponds to minimizing the error associated with the choice of the wrong hypothesis. We will focus here on the simplest version that has a known analytic solution: minimum error state discrimination between two non-orthogonal pure quantum states prepared with equal *a priori* probability. Later we relax the necessity of purity using techniques from the resource theory framework.

Quantum binary decision problem

We begin with a formal description of the quantum binary search as first provided by (Helstrom, 1969). It is normally said that the optimal quantum measurement for minimum error discrimination is the Helstrom measurement. This measurement procedures should maximize the probability of guessing correctly between the two hypothesis at stake. We describe the problem following the notation of (Schmid and Spekkens, 2018).

Let two nonorthogonal pure states $|\phi\rangle$ and $|\psi\rangle$ be among the possible preparations. Suppose that the system prepares one of the two states with equal probability, and the receiver would like to distinguish between one or the other. Let $|\bar{\phi}\rangle$ represent the antipodal point in the Bloch sphere representation associated with the equatorial description given by the space spanned by $\{|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|\}$. The states $|\phi\rangle$ and $|\bar{\phi}\rangle$ are orthogonal to one another, and a measurement over the ϕ basis can distinguish these two states completely using the projective measurements $M_\phi := \{|\phi\rangle\langle\phi|, |\bar{\phi}\rangle\langle\bar{\phi}|\}$. The same thing is true for the measurements over the ψ basis, $M_\psi := \{|\psi\rangle\langle\psi|, |\bar{\psi}\rangle\langle\bar{\psi}|\}$.

Whenever one implements M_ψ over a state $|\phi\rangle$ there exists a probability that a nonzero outcome is obtained, creating a sort of misleading impression that the prepared state was actually $|\psi\rangle$. The interpretation is that the state $|\phi\rangle$ passes the test of being $|\psi\rangle$ and, because of that interpretation we refer to this quantity as the *confusability*, or also the *cost*, c_q , where the subscript q is the notation for explicitly describing the quantity in terms of quantum theory. Later, we will provide a fully operational description for the

task and therefore drop the subscript q (see the beginning of section 3.1 for a clarification about such distinction in notation).

$$c_q := \text{Tr}(|\phi\rangle\langle\phi||\psi\rangle\langle\psi|) = |\langle\phi|\psi\rangle|^2. \quad (5.18)$$

If it is relevant we might write $c_{q\phi\psi}$ as the confusability associated with guessing ϕ when the true state is ψ . Although for our case this is irrelevant, such a notion is useful to describe the problem in more general terms. In the case that $c_q = 0$ we have that the two states are completely distinguishable. If $c_q \neq 0$ there can be no measurement procedure that will perfectly distinguish between the two. Let a discrimination measurement be defined as $M_d := \{E_{g_\phi}, E_{g_\psi}\}$ for,

1. If M_d results E_{g_ϕ} then we should guess that the prepared state is ϕ .
2. If M_d results E_{g_ψ} then we should guess that the prepared state is ψ .

Considering that the preparation performed is described as

$$P \simeq \frac{1}{2}P_\psi + \frac{1}{2}P_\phi \xrightarrow{QM} \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\phi\rangle\langle\phi|, \quad (5.19)$$

the quantum probability of guessing correctly is given by

$$s_q = \frac{1}{2}\text{Tr}(E_{g_\phi}|\phi\rangle\langle\phi|) + \frac{1}{2}\text{Tr}(E_{g_\psi}|\psi\rangle\langle\psi|). \quad (5.20)$$

The measurement scheme optimal for s_q , the Helstrom measurement, is the measurement M_d described as follows. To refer to the optimality we write the results as a theorem:

Theorem 12. (*Ryan O'Donnell, 2015*) *Let s_q be the optimal probability of guessing correctly between two states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$ is given by*

$$s_q = \frac{1}{2} + \frac{1}{2} \left(\frac{1}{2} \| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_1 \right) = \frac{1}{2} (1 + \sqrt{1 - c_q}) \quad (5.21)$$

where $\|\cdot\|_1$ denotes the trace norm.

Proof. First we notice that s_q can be rewritten as,

$$\begin{aligned}
 s_q &= \frac{1}{2}\text{Tr}(E_{g_\phi}|\phi\rangle\langle\phi|) + \frac{1}{2}\text{Tr}(E_{g_\psi}|\psi\rangle\langle\psi|) = \frac{1}{4}\text{Tr}(E_{g_\phi}|\phi\rangle\langle\phi|) + \frac{1}{4}\text{Tr}(E_{g_\psi}|\psi\rangle\langle\psi|) \\
 &+ \frac{1}{4}\text{Tr}(E_{g_\phi}|\phi\rangle\langle\phi|) + \frac{1}{4}\text{Tr}(E_{g_\psi}|\psi\rangle\langle\psi|) + \frac{1}{4}\text{Tr}(E_{g_\phi}|\psi\rangle\langle\psi|) \\
 &- \frac{1}{4}\text{Tr}(E_{g_\phi}|\psi\rangle\langle\psi|) + \frac{1}{4}\text{Tr}(E_{g_\psi}|\phi\rangle\langle\phi|) - \frac{1}{4}\text{Tr}(E_{g_\psi}|\phi\rangle\langle\phi|) \\
 &= \frac{1}{4}\text{Tr}((E_{g_\phi} + E_{g_\psi})(|\phi\rangle\langle\phi| + |\psi\rangle\langle\psi|)) + \frac{1}{4}\text{Tr}((E_{g_\phi} - E_{g_\psi})(|\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|)) \\
 &=: \frac{1}{2}T_1 + \frac{1}{2}T_2
 \end{aligned}$$

Where we notice that $T_1 := 1$ and that using the Hölder inequality for matrices,

$$\text{Tr}(AB) \leq \|A\|_p \|B\|_q, \quad \forall p, q \in [1, \infty], \quad \frac{1}{p} + \frac{1}{q} = 1 \quad (5.22)$$

for all A, B matrices, so that we have,

$$T_2 = \frac{1}{2}\text{Tr}((E_{g_\phi} - E_{g_\psi})(|\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|)) \leq \frac{1}{2}\|E_{g_\phi} - E_{g_\psi}\|_\infty \| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_1 \quad (5.23)$$

But we have that

$$\begin{aligned}
 \|E_{g_\phi} - E_{g_\psi}\|_\infty &= \max \sigma(E_{g_\phi} - E_{g_\psi}) = \max(\sigma(E_{g_\phi}) + \sigma(-E_{g_\psi})) \\
 &= \max(\sigma(E_{g_\phi})) + \max(-\sigma(E_{g_\psi})) = \max(\sigma(E_{g_\phi})) + \min(\sigma(E_{g_\psi})) \leq 1
 \end{aligned}$$

and that therefore $s_q \leq \frac{1}{2} + \frac{1}{4}\| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_1$. The important point is that this upper bound can be achieved. Let $\Delta := |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|$. Now, we consider $\sigma(\Delta) \subseteq [-\|\Delta\|_\infty, \|\Delta\|_\infty]$ separating the spectrum into nonnegative and negative eigenvalues. Take E_{g_ϕ} to be the projection into the eigenspace associated with the nonnegative eigenvalues and E_{g_ψ} the projection into the eigenspace associated with the negative ones. In this way we have that the eigenvalues of $E_{g_\phi}\Delta$ are all the positive eigenvalues of Δ and the eigenvalues of $E_{g_\psi}\Delta$ are the negative eigenvalues of Δ . This is the construction of the so-called Helstrom measurement. We notice that these measurements hits the bound since

$$\frac{1}{2}\|\Delta\|_1 = \frac{1}{2} \sum_{e \in \sigma(\Delta) \cap \mathbb{R}_0^+} e - \frac{1}{2} \sum_{e \in \sigma(\Delta) \cap \mathbb{R}^-} e = \frac{1}{2}\text{Tr}(E_{g_\phi}\Delta) - \frac{1}{2}\text{Tr}(E_{g_\psi}\Delta) = T_2. \quad (5.24)$$

Now, let $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle = \cos(\theta)$. We have that Δ has all eigenvalues equal to zero, but two, that are associated with eigenvectors of the form $|\Psi\rangle = a|\phi\rangle + b|\psi\rangle$. Let $|\Psi\rangle$ be an

eigenvector of Δ and then

$$\begin{aligned}\Delta |\Psi\rangle &= \Delta(a|\phi\rangle + b|\psi\rangle) = a\Delta|\phi\rangle + b\Delta|\psi\rangle \\ &= a|\phi\rangle - a\cos\theta|\psi\rangle + b\cos(\theta)|\phi\rangle - b|\psi\rangle \\ &= (a + b\cos(\theta))|\phi\rangle + (-a\cos(\theta) - b)|\psi\rangle,\end{aligned}$$

so that we have the eigenvalue problem to imply that, letting e to be the notation for the eigenvalues, $\Delta|\Psi\rangle = e|\Psi\rangle$

$$\begin{aligned}ea &= a + b\cos(\theta) \\ eb &= -a\cos(\theta) - b.\end{aligned}$$

Solving these equations for e we have that the only eigenvalue solutions are $e = \pm \sin(\theta)$ and therefore we have that $\|\Delta\|_1 = \sin(\theta) + \sin(\theta) = 2\sqrt{1 - c_q}$. \square

We have then that the Helstrom measurement as defined in the above theorem reaches the optimal quantum success probability given by

$$s_q = \frac{1}{2} (1 + \sqrt{1 - c_q}). \quad (5.25)$$

Since we are interested not only in no-go results that will represent formal proofs of contextual advantage for a task, but also in actual use of contextuality in experimental setups we need to have optimal quantum bounds that allow for noisy measurements. These robust bounds will express how much can a measurement be noisy and still represent a contextual advantage, so that experimental implementations of the protocol will actually be using the resource. In section 5.3 we will see that depending on the amount of noise in the system, the experimenter cannot guarantee contextual advantage of the protocol. We proceed then to analyse how we can introduce noise to the optimal bound given by equation (5.25).

First we consider that the measurements M_ϕ, M_ψ to be described by POVM elements $E_\phi, E_{\bar{\phi}}, E_\psi, E_{\bar{\psi}}$ that are not necessarily projections onto ϕ, ψ . Then, suppose that preparation of the state $|\phi\rangle$, say P_ϕ , is such that whenever we make a measurement procedure in the ideal case of the POVM E_ϕ we have, of course, 1. We attribute the noise factor as much as possible to the measurement procedures that fail to decide whether the correct hypothesis was chosen, and the remainder of the fluctuations we embed into the confusability. In this way, if for example P_ϕ prepares the state $|0\rangle\langle 0|$ we would have $p(\phi|M_\phi, p_\phi) = \langle 0|E_\phi|0\rangle = 1 - \varepsilon_\phi$, with ε_ϕ being the error estimation. In this way, we would attribute the confusability to $p(\phi|M_\phi, P_\psi) = c_\psi$. With this idea in mind, by means of a suitable description of the best choice of the two states in question ([Schmid and Spekkens, 2018](#), Appendix C) and imposing the symmetries we will discuss in detail when

a full description of the scenario is presented, we get the noisy optimal bound for the average success probability of guessing the correct state as

$$s_q = \frac{1}{2} \left(1 + \sqrt{1 - \varepsilon + 2\sqrt{\varepsilon(1 - \varepsilon)c_q(1 - c_q)} + c_q(2\varepsilon - 1)} \right) \quad (5.26)$$

where we notice that $\varepsilon \rightarrow 0$ reduces to the noiseless result given by equation (5.25).

Scenario description

Now that we have understood the quantum task of minimum error state discrimination, we would like to make a full description of the operational scenario in order to answer what is the nonclassical aspect of this task, if any does exist. The fact that one cannot distinguish between two nonorthogonal states cannot be viewed as essentially nonclassical, taking the notion of classicality to be generalized noncontextuality. Simply because as we have seen in figure 3.1 we might simply describe nonorthogonal states as overlapping epistemic states. In the region they overlap we could not fully distinguish between the two states of reality, but this event *can* still be described by a noncontextual ontological model, hence classical.

As the discussion we have so far suggests, the nonclassicality of the task does not arises as the impossibility to determine correctly the state, but in the dependence between the success of the task, that we denote operationally as s , and the epistemic distributions associated with the preparations.

Definition 29. Let \mathbb{B}_{sd} be the scenario defined by four preparation procedures P_1, \dots, P_4 that satisfy the operational equivalences,

$$\mathbb{E}_P : \iff \frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4 \quad (5.27)$$

and that the experimenter performs three binary-outcome tomographically complete measurement procedures $\mathbb{M} := \{M_1, M_2, M_d\}$. This is the operational scenario that fits the information task of minimum error state discrimination. \diamond

We notice that in quantum terms the we simply have the measurements $M_1 = M_\phi = \{E_\phi^\phi, E_\phi^{\bar{\phi}}\}$, $M_2 = M_\psi = \{E_\psi^\psi, E_\psi^{\bar{\psi}}\}$, $M_3 = M_d$, and that letting P_ϕ to be the preparation of the quantum state $|\phi\rangle\langle\phi|$, we have

$$\frac{1}{2}|\phi\rangle\langle\phi| + \frac{1}{2}|\bar{\phi}\rangle\langle\bar{\phi}| = \frac{\mathbb{1}}{2} = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\bar{\psi}\rangle\langle\bar{\psi}| \quad (5.28)$$

and since \mathbb{M} is chosen to be tomographically complete the probability, the Born rule can never distinguish these two new procedures that were convexly defined in equation (5.28).

Although the scenario \mathbb{B}_{sd} has been fully described there are some symmetries that can be imposed so that we reach the operational description given by the optimal bounds in equations (5.25) and (5.26). We describe these symmetries as:

$$s = p(g_\phi|M_d, P_\phi) = p(g_{\bar{\phi}}|M_d, P_{\bar{\phi}}) = 1 - p(g_\phi|M_d, P_\psi) = 1 - p(g_\psi|M_d, P_{\bar{\psi}}) \quad (5.29a)$$

$$c = p(\phi|M_\phi, P_\psi) = p(\psi|M_\psi, P_\phi) = p(\bar{\phi}|M_{\bar{\phi}}, P_{\bar{\psi}}) = p(\bar{\psi}|M_{\bar{\psi}}, P_{\bar{\phi}}) \quad (5.29b)$$

$$1 - \varepsilon = p(\psi|M_\psi, P_\psi) = p(\phi|M_\phi, P_\phi) = p(\bar{\psi}|M_{\bar{\psi}}, P_{\bar{\psi}}) = p(\bar{\phi}|M_{\bar{\phi}}, P_{\bar{\phi}}) \quad (5.29c)$$

These symmetries represent the fact that the average confusability does not vary if we are confusing in one way, expressed by $p(\phi|M_\phi, P_\psi)$, or in the related one, denoted by $p(\psi|M_\psi, P_\phi)$. It is possible not to impose these symmetries to consider the full generality of the process. For our discussion with the quantum case such description will be considered as a special case of interest for a cleaner comparison with the quantum optimal bounds given by (5.26) and (5.25). In this way, we will be able to describe the behavior with the following data-table:

	P_ϕ	$P_{\bar{\phi}}$	P_ψ	$P_{\bar{\psi}}$
E_ϕ^ϕ	$p(\phi M_\phi, P_\phi)$	$p(\phi M_\phi, P_{\bar{\phi}})$	$p(\phi M_\phi, P_\psi)$	$p(\phi M_\phi, P_{\bar{\psi}})$
E_ψ^ψ	$p(\psi M_\psi, P_\phi)$	$p(\psi M_\psi, P_{\bar{\phi}})$	$p(\psi M_\psi, P_\psi)$	$p(\psi M_\psi, P_{\bar{\psi}})$
$E_{g_\phi}^d$	$p(g_\phi M_d, P_\phi)$	$p(g_\phi M_d, P_{\bar{\phi}})$	$p(g_\phi M_d, P_\psi)$	$p(g_\phi M_d, P_{\bar{\psi}})$

Table 5.3: Behavior representing the statistics of scenario \mathbb{B}_{sd} .

And if we impose equations (5.29a)-(5.29c) we obtain the data-table representation of the operational results associated with minimum error state discrimination, including the noise model that respects and average relation given by equation (5.29c).

	P_ϕ	$P_{\bar{\phi}}$	P_ψ	$P_{\bar{\psi}}$
E_ϕ^ϕ	$1 - \varepsilon$	ε	c	$1 - c$
E_ψ^ψ	c	$1 - c$	$1 - \varepsilon$	ε
$E_{g_\phi}^d$	s	$1 - s$	$1 - s$	s

Table 5.4: Behavior representing the operational description of the success rate s , the confusability c and the noise ε of scenario \mathbb{B}_{sd} . Here we consider the validity of (5.29a)-(5.29c), but one could drop such constraints and simply work with the general behavior description of table 5.3.

Justifying the operational equivalences

Since we do not require quantum mechanics as an *a priori* underlying description of reality we need to justify that the operational equivalences defined by

$$\frac{1}{2}P_\phi + \frac{1}{2}P_{\bar{\phi}} \simeq P_{\frac{1}{2}\phi + \frac{1}{2}\bar{\phi}} \simeq P_{\frac{1}{2}\psi + \frac{1}{2}\bar{\psi}} \simeq \frac{1}{2}P_\psi + \frac{1}{2}P_{\bar{\psi}}. \quad (5.30)$$

are physically relevant. Although these operational equivalences are satisfied by the quantum minimum error state discrimination task, would it be reasonable to expect that general classical protocols would also respect such equivalences?

In (Schmid and Spekkens, 2018) the authors discuss that although noisy experiments will not necessarily satisfy neither the operational equivalences, neither the imposed symmetries; by means of the free operations of the resource theory developed in chapter 4, it is always possible to use post processing of the data-table obtained in order to force the symmetries and the operational equivalences in the scenario. In this way we are allowing for the experimentalist to verify contextuality, since the free operations cannot make any data-table contextual, by definition. It is also useful that the symmetries can be imposed with the free operations, since they greatly simplify the analysis for specific tasks. These were the first evidences that the resource theory have a great significance, but they do not notice that it could also be useful to witness contextuality itself as we have noticed in section 4.3.

Therefore, the operational equivalences are robust, in the sense that we can maintain them in real noisy experiments. By this does not mean that they are physically appealing: remember that we have introduced preparation $P_{\bar{\phi}}$ because we need operational equivalences to impose a noncontextuality assumption for the ontological model. It might be simply a unreasonable trick to create nonclassicality by force: Criticism in this spirit leads to completely new attempts to even define noncontextuality. But we would like to make the following interpretation (Lostaglio and Senno, 2020): let $P_{\bar{\phi}}$ be a preparation procedure such that,

$$p(\phi|M_{\phi}, P_{\bar{\phi}}) = 0 \tag{5.31}$$

This way we will have that, for all measurement events $[k|M]$ we will have operationally that,

$$p(k|M, P_{\phi}) + p(k|M, P_{\bar{\phi}}) = p(k|M, P_{\psi}) + p(k|M, P_{\bar{\psi}}) \tag{5.32}$$

and we can summarize this as follows: let the symmetries (5.29a)-(5.29c) be imposed, through considerations of the noise model and approximations of average confusability and success probability of the task. We are able to sustain these symmetries via the free operations. Now, since we can perceive from table 5.4, making $\varepsilon \rightarrow 0$, we have for any measurement event that equation (5.32) holds, which can be understood operationally by means of fair dice convex combination of the procedures associated with one another, having condition $p(\phi|M_{\phi}, P_{\bar{\phi}}) = \varepsilon \rightarrow 0$ as a reference. This establishes that the ideal scenario with $\varepsilon \rightarrow 0$ has a physical interpretation by acknowledging the existence of a preparation such that $p(\phi|M_{\phi}, P_{\bar{\phi}}) = \varepsilon \rightarrow 0$. As it was already pointed out, the noise model will most likely destroy the operational equivalences, but using the free operations we can restore them without loss of the argument.

We notice that it is important to proceed in this way and bring an intuitive physical interpretation here, differently to what we have done so far for the parity oblivious tasks because the operational equivalences there were *imposed* by the task itself. In the tasks of minimum error state discrimination and state-dependent quantum cloning we need to *introduce* these somewhat indirect operational descriptions in order to impose the noncontextual hypothesis. Doing so, we provide with a case study a methodology for application of generalized noncontextuality as a framework, for possibly new scenarios that still to be discovered to be a facet of nonclassicality, and also providing a rigorous technique for describing *what* is nonclassical in the task.

Noncontextual bound for the discrimination task

With the techniques we gathered so far we can obtain a relationship between (s, c, ε) , the three degrees of freedom defining the behavior in our scenario \mathbb{B}_{sd} with the symmetries that provides information about the interplay between the variables (s, c, ε) and our resource. The full set of facet-defining inequalities of the noncontextual polytope, as presented in chapter 3 and appendices D, C does that.

In (Schmid and Spekkens, 2018) the authors also prove the noncontextual bound for the ideal scenario using a diagrammatic intuitive description, but since we are mostly interested in the noisy version, that it is achieved using the techniques for general obtaining noncontextuality inequalities we want to show the validity of the following result.

Theorem 13. *For any noncontextual ontological model respecting the symmetries (5.29a)-(5.29c) over data-tables $B \in \mathbb{B}_{sd}$ we will have the trade-off between the average success rate of the task s , the average confusability c and the noise ε given by:*

$$s \leq 1 - \frac{c - \varepsilon}{2} \tag{5.33}$$

with the ideal scenario as a special case with condition $\varepsilon \rightarrow 0$.

Proof. The proof is greatly simplified once one notices that this trade-off is actually a tight noncontextuality inequality of the polytope $NC(\mathbb{B}_{sd})$. Therefore we might use the algorithmic approach, since the dimension of the problem does not make it unfeasible. Using this approach, with notation described by the algorithm of appendix C, we find the inequality:

$$\begin{aligned}
 (225) \quad & -x44 + x46 + x50 - x52 \leq 1 \\
 & -p_{\psi, \bar{\phi}} + p_{\psi, \psi} + p_{d, \phi} - p_{d, \bar{\phi}} \leq 1 \\
 & -(1 - c_{\bar{\phi}}) + (1 - \varepsilon_{\psi}) + s_{\phi} - (1 - s_{\bar{\phi}}) \leq 1 \\
 & c_{\bar{\phi}} + s_{\phi} + s_{\bar{\phi}} - \varepsilon_{\psi} \leq 2 \xrightarrow{(5.29a)-(5.29c)} \\
 & c - \varepsilon + 2s \leq 2 \implies s \leq 1 - \frac{c - \varepsilon}{2}
 \end{aligned}$$

The number (225) is the number given by inequality using PORTA. The labeling follows the direction that was chosen by the Fourier-Motzkin elimination protocol, where they have cancelled all odd variables, so we simply consider $x38 = p(\phi|M_{\phi}, P_{\psi}) = p(1|M_1, P_1)$ and we follow the labeling that is associated with table 5.3 and appendix C.

$$x44 \equiv p(1|M_2, P_2), x46 \equiv p(1|M_2, P_3), x50 \equiv p(1|M_3, P_1), x52 \equiv p(1|M_3, P_2) \quad (5.34)$$

We have also considered the situation that is depicted in appendix D of (Schmid and Spekkens, 2018), where they do not impose the symmetries, therefore letting the different noise and confusability quantities be labeled by the operational description that generates them, such as $c_{\phi} = p(\psi|M_{\psi}, P_{\phi})$, and similarly for the other quantities. In this way we obtain the more general table,

	$P_1 \equiv P_{\phi}$	$P_2 \equiv P_{\bar{\phi}}$	$P_3 \equiv P_{\psi}$	$P_4 \equiv P_{\bar{\psi}}$
$1 M_1 \equiv \phi M_{\phi}$	$1 - \varepsilon_{\phi}$	$\varepsilon_{\bar{\phi}}$	c_{ψ}	$1 - c_{\psi} - \varepsilon_{\phi} + \varepsilon_{\bar{\phi}}$
$1 M_2 \equiv \psi M_{\psi}$	c_{ϕ}	$1 - c_{\bar{\phi}}$	$1 - \varepsilon_{\psi}$	$c_{\phi} - c_{\bar{\phi}} + \varepsilon_{\psi}$
$1 M_3 \equiv g_{\phi} M_d$	s_{ϕ}	$1 - s_{\psi}$	$1 - s_{\bar{\phi}}$	$s_{\phi} - s_{\bar{\phi}} - s_{\psi}$

Table 5.5: Behavior representing the operational description of the primitives from minimum error state discrimination without considering the symmetries.

This is a tight inequality since the noncontextual behavior $\begin{pmatrix} 1, 0, 0, 1 \\ 1, 0, 1, 0 \\ 1, 1, 1, 1 \end{pmatrix}$ that can be a realization of table 5.3 achieves the equality, since $s_{\phi} = 1$, $s_{\bar{\phi}} = 0$, $c_{\bar{\phi}} = 1$, $\varepsilon_{\psi} = 0$. This behavior is also a vertex of the polytope $C(\mathbb{B}_{sd})$. We conclude that this is a tight noncontextual trade-off between the primitives that are associated with the minimum error state discrimination task, imposing a bound on the average success probability. \square

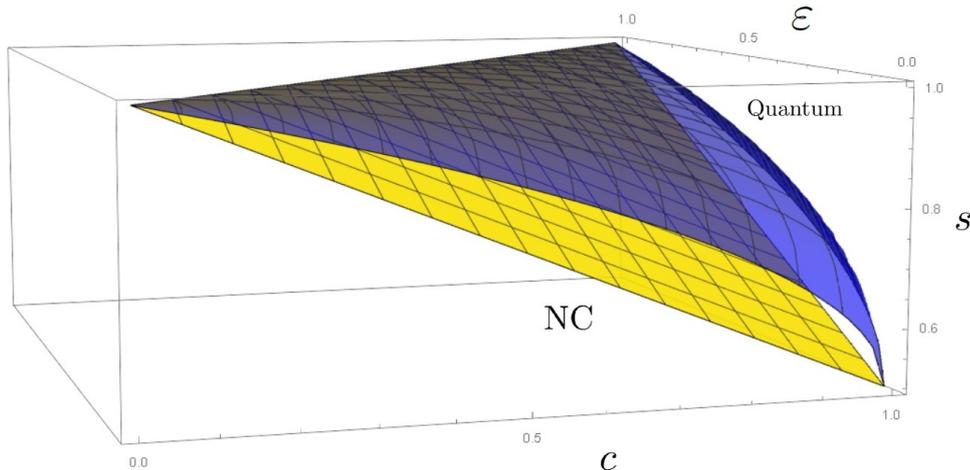


Figure 5.3: Comparison between the quantum optimal noisy bounds for the average success probability s in equation 5.26 and the noncontextual bound given by the noncontextuality inequality from theorem 13. We are always considering the symmetries from (5.29a)-(5.29c). In the blue curve we have the behaviors that might be achieved by quantum theory and that are conjectured to be optimal. In the yellow region we have the noncontextuality facet associated with the noncontextual polytope $NC(\mathbb{B}_{sd}) \cap \{\text{symmetries}\}$.

Proof. (of theorem 13 using the resource theory) We provide another demonstration that the noncontextual bound given is tight and associated with the noncontextual polytope. First, we notice that in the theorem we consider the scenario \mathbb{B}_{sd} described by the behaviors that are given by table 5.4. Notice that for any $s \in [0, 1]$ fixed we have that there is a free operation T that relates the first and the second rows of table 5.4. Therefore, the noncontextuality inequalities that use only the first two rows for this behavior cannot present nontrivial bounds for noncontextuality. If we then consider only the behavior of the first and third rows we will always be able to construct the second row. The structure is now the same as the one present in the scenario \mathbb{B}_{si} and we can effectively consider, that the *contextual* structure from the scenario $\mathbb{B}_{sd} \cap \{\text{symmetries}\}$ arises from the structure of the scenario \mathbb{B}_{si} .

If we write what we have just said, we have the following: Let us denote the set of behaviors of the same form as in 5.4 as $\mathbb{B}_{sd, sym} := \mathbb{B}_{sd} \cap \{\text{symmetries}\}$. Then, we consider two maps⁵ $T_1 : C(\mathbb{B}_{sd, sym}) \rightarrow C(\mathbb{B}_{si})$ and $T_2 : C(\mathbb{B}_{si}) \rightarrow C(\mathbb{B}_{sd, sym})$.

$$T_1 : q_M := \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \underbrace{\begin{pmatrix} 1-\varepsilon & \varepsilon & c & 1-c \\ c & 1-c & 1-\varepsilon & \varepsilon \\ s & 1-s & 1-s & s \end{pmatrix}}_{B \in \mathbb{B}_{sd, sym}} \xrightarrow{T_1} \underbrace{\begin{pmatrix} c & 1-c & 1-\varepsilon & \varepsilon \\ s & 1-s & 1-s & s \end{pmatrix}}_{B \in \mathbb{B}_{si}} \quad (5.35)$$

⁵Notice that, in fact, we are also mostly interested in behavior from \mathbb{B}_{si} that also satisfy a symmetry: $p_{21} = p_{24}$ and $p_{22} = p_{23}$, so that to be extra careful we indeed consider a smaller polytope than \mathbb{B}_{si} . For the purpose of the proof, which is finding the inequality $s \leq 1 - (c - \varepsilon)/2$ we not to focus so much in this fact.

For T_2 we consider similar stochastic maps, such that q_M copies the results of the first measurement procedure and uses permutations between the outcome results regarding preparation procedures in the same context.

$$\underbrace{\begin{pmatrix} c & 1-c & 1-\varepsilon & \varepsilon \\ s & 1-s & 1-s & s \end{pmatrix}}_{B \in \mathbb{B}_{si}} \xrightarrow{T_2} \underbrace{\begin{pmatrix} 1-\varepsilon & \varepsilon & c & 1-c \\ c & 1-c & 1-\varepsilon & \varepsilon \\ s & 1-s & 1-s & s \end{pmatrix}}_{B \in \mathbb{B}_{sd,sym}} \quad (5.36)$$

Now notice that any behavior of $C(\mathbb{B}_{si})$ that respects the noncontextuality assumption needs to be in $NC(\mathbb{B}_{si})$. Since we know that $NC(\mathbb{B}_{si})$ is given by (C.14)-(C.22) we can simply notice that (C.20) will give us, considering $1|M_1 \equiv g_\phi|M_d$ and $1|M_2 \equiv \phi|M_\phi$ we will have,

$$\begin{aligned} p_{23} + p_{14} - p_{12} - p_{22} &\leq 1 \\ p(1|M_2, P_3) + p(1|M_1, P_4) - p(1|M_1, P_2) - p(1|M_2, P_2) &\leq 1 \\ p(\phi|M_\phi, P_\psi) + p(g_\phi|M_d, P_{\bar{\phi}}) - p(g_\phi|M_d, P_{\bar{\phi}}) - p(\phi|M_\phi, P_{\bar{\phi}}) &\leq 1 \\ \xrightarrow{5.4} c + s - (1-s) - \varepsilon &\leq 1 \implies s \leq 1 - \frac{c-\varepsilon}{2}. \end{aligned}$$

We can conclude the following:

1. For any violation of the inequality (C.20) we have $B_{small} = T_1(B_{large})$ is a contextual behavior. Therefore, B_{large} can only be a contextual behavior as well, otherwise we would have a contradiction. We conclude then that violation of inequality (C.20), in the form of $s \leq 1 - (c - \varepsilon)/2$, implies contextuality in the scenario $\mathbb{B}_{sd,sym}$.
2. For any behavior B_{small} respecting the noncontextuality inequality we have that B_{small} is noncontextual⁶. Since T_2 is a free operation, $T_2(B_{small})$ is also noncontextual in the scenario $\mathbb{B}_{sd,sym}$. Therefore, if B_{large} respects the noncontextuality inequality (C.20) in the form of $s \leq 1 - (c - \varepsilon)/2$ we get that B_{large} is noncontextual, with respect this inequality.
3. We conclude that $s \leq 1 - (c - \varepsilon)/2$ is a noncontextuality inequality for any behavior $B_{large} \in \mathbb{B}_{sd,sym}$. Seing in the lights of the operational task, we get that this inequality is a noncontextual bound for the success probability s .

Therefore by using the resource theory framework we can understand better the inner structure of the noncontextual polytopes, obtaining noncontextual bounds for com-

⁶In fact, for B_{small} to be noncontextual it should respect all the noncontextuality inequalities for the simplest scenario. And indeed, if we write all the inequalities in terms of the elements of B_{small} we have that the only non-trivial inequalities are $s \leq 1 - (c - \varepsilon)/2$ and $s \leq 1 - (\varepsilon - c)/2$. The second inequality follows from the first if we consider $c \geq \varepsilon$, and this is fairly reasonable from an experimental perspective since, if the error is higher then the confusability one could actually reverse the roles, by making $\varepsilon \rightarrow c$.

plex scenarios from scenarios with a simpler structure. This is one example of such framework. Notice that without the imposition of the symmetries this is no longer the case, and we would need to obtain bounds using the full set of tight noncontextuality inequalities given in (Schmid and Spekkens, 2018). \square

We conclude that robust experimental implementations of minimum error state discrimination will represent advantage with respect to classical counterparts. The resource theory for generalized contextuality can be used to restore the operational equivalences and the symmetries imposed. Even without any consideration about the task, we can also use the resource theory to witness the existence of quantum contextual correlations via the resource theory, as we have discussed in section 4.3. We also notice that the l_1 -distance monotone can be used to provide information about the largest difference between the success probability given by the noncontextual bound, and the optimal quantum bound.

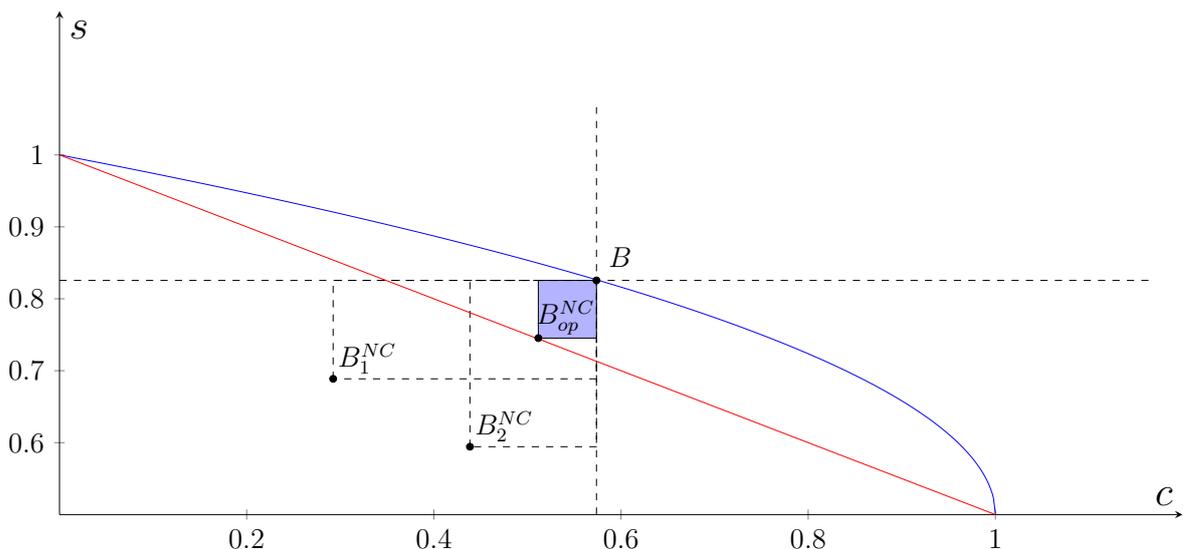


Figure 5.4: Comparing the noncontextual and quantum bounds for the ideal case. The l_1 -distance will find the square region depicted in light blue.

By definition, the l_1 -distance monotone will be calculated by means of,

$$d(B) := \min_{B^* \in NC(\mathbb{B})} \max_{i,j} \sum_k |p(k|M_i, P_j) - p^*(k|M_i, P_j)|.$$

and for the scenario $\mathbb{B}_{sd} \cap \{\text{symmetries}\}$ we will have that

$$\begin{aligned} d(B) &= 2 \min_{B^* \in NC(\mathbb{B})} \max_{i,j} |p_{i,j} - p_{i,j}^*| \\ &= 2 \min_{B^* \in NC(\mathbb{B})} \max\{|c - c^*|, |s - s^*|\} \end{aligned}$$

The optimal noncontextual ideal behavior B_{op}^{NC} will then be the one where $|c - c_{op}^{NC}| =$

$|s - s_{op}^{NC}|$ and $s_{op}^{NC} = 1 - c_{op}^{NC}/2$. In figure 5.4 we can see why this is the case. Let B_1^{NC} be some other noncontextual behavior, such that $|s - s_1^{NC}| < |c - c_1^{NC}|$. In this case, $\max\{|c - c^*|, |s - s^*|\} = |c - c_1^{NC}|$, but since $|c - c_{op}^{NC}| < |c - c_1^{NC}|$ it cannot be the final result of $\mathbf{d}(B)$. The same reasoning apply when we consider B_2^{NC} with $|s - s_2^{NC}| > |c - c_2^{NC}|$. The optimal case will then find the noncontextual behavior in the boundary of the noncontextuality inequality, such that the difference between the confusability and the success probability is the same. Since we know the quantum and the noncontextual bounds we can construct a linear optimization problem such that we obtain the highest quantum behavior violating the noncontextuality inequality the most: we simply optimize \mathbf{d} restricted to the blue curve in figure 5.4.

5.3 Quantum Cloning

Whenever speaking about quantum cloning, the first important result one remembers is the famous no-cloning theorem (Wootters and Zurek, 1982). An interesting extension to mixed state was given in (Barnum et al., 1996).

Theorem 14. (Wootters and Zurek, 1982) *Let $|\phi\rangle, |\psi\rangle$ be two different non-orthogonal quantum states. Then, these two states cannot be perfectly cloned.*

Therefore, a direct question we might ask is: how can quantum cloning serve as a case study for proving quantum contextual advantage? It is well known that classical states can be cloned at will, serving as the basis for most information processing in classical computers. As it is addressed in (Sainz, 2020), this question posits important steps for comprehending what generalized noncontextual bounds for prepare-and-measure scenarios actually mean.

As we have already noticed with the minimum error state discrimination, the nonclassicality arises when we notice operational, or in other words statistical equivalences in the experimental data. Such as saw in the last section, it is quite surprising that any advantage arises from a task that, *a priori*, has nothing to do with our common understanding of the word *classical*.

For this quantum task, we have a particular important detail to mention: we consider briefly the action of a transformation procedure. In fact, the idea is that given a preparation procedure P that we operationally describe, we will include in the scenario the procedure $T(P)$, that can be understood as considering the entire Ludwig box P followed by T as a preparation procedure. In quantum terms, considering a Schrödinger picture operational description. There is still much work towards a full comprehension of how, when and why we can consider the relationship between transformation procedures and prepare-and-measure scenarios. The paper of (Lostaglio and Senno, 2020) is one of the attempts to study such a situation. Another one that deals with anomalous weak

values and that it is interesting for quantum thermodynamics is the study presented in (Kunjwal et al., 2019).

Scenario description

The description of the state-dependent quantum cloning task is as follows: We consider two preparation procedures P_a, P_b , that are associated in the quantum description with preparing the two states $|a\rangle \langle a|$ and $|b\rangle \langle b|$ to be cloned. We then suppose that a new pair of preparation procedures, P_α, P_β that are the procedures $P_\alpha = T(P_a)$ and $P_\beta = T(P_b)$ associated with the preparation procedure of the cloned states. This would correspond to effectively using a cloning unitary $P_\alpha \stackrel{QM}{\equiv} |\alpha\rangle := U|a, 0\rangle$, where $|0\rangle$ is the initial state of some ancilla that we use for the cloning process. Given that we have these procedures, we would like to suppose that there are procedures P_{aa}, P_{bb} that give the ideal prepared clones, that we represent in quantum theory as $|a, a\rangle \equiv |aa\rangle$.

For each of these procedures we suppose that there exist some measurements that perfectly test whether the preparation was correct. Such as in the minimum error state discrimination task, we suppose that:

1. For each preparation P_s , for $s \in \{a, b, \alpha, \beta, aa, bb\} =: I_{qc}$ there exists measurement procedures M_s that test with certainty if preparation P_s was performed.
2. There exists complementary preparation procedures P_{s^\perp} , that are clear in light of our discussions from section 5.2.
3. For each $(s, s') \in \{(a, b), (\alpha, aa), (\beta, bb)\}$ we have the operational equivalences $\frac{1}{2}P_s + \frac{1}{2}P_{s^\perp} \simeq \frac{1}{2}P_{s'} + \frac{1}{2}P_{s'^\perp}$.

The intuition behind this operational description, that is ideal since we are not allowing for noise in the probabilities $p(1|M_s, P_s) = 1$, for all $s \in I_{qc}$, is similar to the one from \mathbb{B}_{sd} : we notice the fundamental description of the preparation and measure primitives, at the operational level, and then we also acknowledge some symmetries that are due to the so-called ‘‘purity’’ of the statistics. We here then consider the following description for the scenario \mathbb{B}_{qc} :

Definition 30. Let $\mathbb{P} := \{P_1, \dots, P_{12}\}$ a set of preparation procedures and $\mathbb{M} := \{M_1, \dots, M_6\}$ a set of binary-outcome measurement procedures. We consider that there are no operational equivalences between the measurement procedures but, we assume

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4 \tag{5.37}$$

$$\frac{1}{2}P_5 + \frac{1}{2}P_6 \simeq \frac{1}{2}P_7 + \frac{1}{2}P_8 \tag{5.38}$$

$$\frac{1}{2}P_9 + \frac{1}{2}P_{10} \simeq \frac{1}{2}P_{11} + \frac{1}{2}P_{12} \tag{5.39}$$

for the preparation procedures. We will refer to this scenario as \mathbb{B}_{qc} . \diamond

This operational scenario described in the terms of (Schmid et al., 2018) relates to the one in (Lostaglio and Senno, 2020) via the following definition:

Definition 31. Letting the following labels for the preparation procedures to act: $P_1 \equiv P_a, P_2 \equiv P_{a^\perp}, P_3 \equiv P_b, P_4 \equiv P_{b^\perp}$ and so on, we then suppose the label correlation

$$a, a^\perp, b, b^\perp, \alpha, \alpha^\perp, aa, aa^\perp, \beta, \beta^\perp, bb, bb^\perp \rightarrow 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \quad (5.40)$$

. Letting $M_1 \equiv M_a, M_6 \equiv M_{bb}$ following the ordering $a, b, \alpha, \beta, aa, bb$. And, by also imposing the symmetries associated with the purity, or also, the idealizations in the scenario, as

$$p(1|M_s, P_s) \equiv p_{s,s} = 1, \quad p_{s,s^\perp} = 0, \quad \forall s \in \{a, b, \alpha, \beta, aa, bb\} \quad (5.41)$$

\diamond

Quantum contextuality in the cloning scenario

Using the resource theory results we can infer the following:

Theorem 15. *The scenario \mathbb{B}_{qc} has quantum contextual realizations.*

Proof. First we notice that the scenario \mathbb{B}_{qc} can be written in the form of

$$\mathbb{B}_{qc} = \mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6, \quad (5.42)$$

where \mathbb{B}_6 is the scenario that constitutes the same operational structure for preparation procedures as the scenario \mathbb{B}_{si} but having 6 measurement procedures. Since we allow for ideal measurements for the purpose of quantum cloning task, we know that each scenario \mathbb{B}_6 has a quantum contextual behavior because there exists free operations $T : \mathbb{B}_6 \rightarrow \mathbb{B}_{si}$ with $T(B) \in C(\mathbb{B}_{si})$ quantum contextual correlation. Therefore B must be a quantum contextual correlation.

The operational structure for the scenario \mathbb{B}_{qc} given in terms of the weight vectors $\gamma_P^1, \gamma_P^2, \gamma_P^3$ are given by the following equations:

$$\gamma_P^1 = \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 \right) \quad (5.43)$$

$$\gamma_P^2 = \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0 \right) \quad (5.44)$$

$$\gamma_P^3 = \left(0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0 \right) \quad (5.45)$$

If we let $\gamma = (\frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2}) = (\vec{\alpha}, \vec{\beta}) = \gamma_{si}$ the weights for the simplest scenario we can see that, by using definition 28, we would have that the operational equivalences for $\mathbb{B}_6 \boxtimes \mathbb{B}_6$

$$\begin{aligned}\gamma_{P_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}}^1 &= \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0 \right) = \left(\vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1, \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1 \right) \\ \gamma_{P_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}}^2 &= \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2} \right) = \left(\vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2, \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2 \right)\end{aligned}$$

and, performing the same procedure again to obtain the weights for $\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6$ we will have,

$$\begin{aligned}\gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^1 &= \left(0, 0, 0, 0, \vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1; 0, 0, 0, 0, \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1 \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^2 &= \left(\vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1, 0, 0, 0, 0; \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^1, 0, 0, 0, 0 \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^3 &= \left(0, 0, 0, 0, \vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2; 0, 0, 0, 0, \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2 \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^4 &= \left(\vec{\alpha}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2, 0, 0, 0, 0; \vec{\beta}_{\mathbb{B}_6 \boxtimes \mathbb{B}_6}^2, 0, 0, 0, 0 \right)\end{aligned}$$

as we have explicitly from definition 28, this is how we construct the operational equivalences for the box product scenarios. This is then,

$$\begin{aligned}\gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^1 &= \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0 \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^2 &= \left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0; 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0, 0, 0 \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^3 &= \left(0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0; 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2} \right) \\ \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^4 &= \left(0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0 \right)\end{aligned}$$

and we notice that $\gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^1 = \gamma_{\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6}^4$, so we might as well drop one of these two, since we are clearly over counting. And we see that the weights for the box scenario $\mathbb{B}_6 \boxtimes \mathbb{B}_6 \boxtimes \mathbb{B}_6$ are the same weights for the scenario \mathbb{B}_{qc} , from equations (5.43)-(5.45). For the measurement procedures we have in each scenario \mathbb{B}_6 that $\mathcal{M}^{1,2,3} = \{M_1^{1,2,3}, \dots, M_6^{1,2,3}\}$ and we consider the relabel associated with the scenario \mathbb{B}_{qc} to be $M_i^{1,2,3} \equiv M_i$. We therefore have the dimensionality of the polytopes matching by,

$$\dim(\mathbb{B}_{qc}) = 2 \times 6 \times 12 = 3(2 \times 6 \times 4) = 3 \dim(\mathbb{B}_6) \tag{5.46}$$

Since we have that the map \boxtimes preserves the operational structure, and the noncontextuality polytope we conclude that the scenario \mathbb{B}_{qc} has quantum contextuality that is accessible. \square

There are then two conclusions we might draw from theorem 15: The first one is that we have only provided an operational demonstration of existence of quantum contextuality for the scenario. This corresponds then to evidence that the quantum cloning task might be advantageous with respect to classical analogs under these circumstances, but theorem 15 itself *does not* constitute a proof of quantum advantage. It constitutes a no-go theorem: there are quantum behaviors in \mathbb{B}_{qc} that cannot be achieved by classical probability, when classicality is understood as arising from noncontextuality hypothesis. The second conclusion we draw is that we only need to obtain the full set of tight inequalities for the scenario \mathbb{B}_6 , which represents a considerable simplification in terms of computational complexity. In this case, we are not even considering the symmetries arising from an idealization of the scenario, so there might have even more simplifications that appear in this case, such as it was when we were treating the ideal scenario for \mathbb{B}_{sd} .

Quantum advantage

To observe quantum advantages we must first consider *how* are we measuring the success of a given task. In the case of the cloning scenario, we would like to know if the preparation procedures P_α for example, for the cloning event $|a\rangle \rightarrow U(|a, 0\rangle)$ is as close as possible to the states P_{aa} associated by $|aa\rangle$. For such a purpose we use the fidelity (Nielsen and Chuang, 2002) as a measure of how much these two preparations are equal.

For treating the problem without assuming the validity of quantum theory we need an operational definition of the fidelity:

Definition 32. (Lostaglio and Senno, 2020) Let $p(1|M_{aa}, P_\alpha) \equiv p_{aa,\alpha}$ be our notation for the conditional probabilities in the behaviors corresponding to the scenario \mathbb{B}_{qc} from definition 31. Then, the global cloning fidelity operationally defined is, $F_g : \mathbb{B}_{qc} \rightarrow \mathbb{R}^+$ defined as,

$$F_g := \frac{1}{2}p_{aa,\alpha} + \frac{1}{2}p_{bb,\beta} \quad (5.47)$$

the average probability that the imperfect clones P_α, P_β pass the test for the ideal clones M_{aa}, M_{bb} . \diamond

The important considerations for quantum advantage need to be expressed in terms of the fidelity. For quantum theory it is known the optimal result given by,

Theorem 16. (Bruß et al., 1998) *The optimal quantum fidelity from the scenario \mathbb{B}_{qc} is described by the following equation,*

$$F_g^Q = \frac{1}{4} \left(\sqrt{(1 + p_{b,a})(1 + \sqrt{p_{b,a}})} + \sqrt{(1 - p_{b,a})(1 - \sqrt{p_{b,a}})} \right)^2 \quad (5.48)$$

where we have that $p_{b,a} = |\langle a|b\rangle|^2$ is again the confusability associated with the preparation $|a\rangle$ passing the test of M_b .

We have then to obtain a bound that any noncontextual model would be restricted to associated with the fidelity given operationally by definition 32. We consider only the ideal case, where we have the following result, proved in (Lostaglio and Senno, 2020).

Theorem 17. *Considering an ideal scenario, that assumes the operational description of definition 31, we have that any noncontextual model will have a fidelity bounded by the optimal F_g^{NC} ,*

$$F_g \leq F_g^{NC} = 1 - \frac{p_{b,a}}{2} + \frac{p_{bb,aa}}{2}. \quad (5.49)$$

This result shows that the quantum success arising from the operational task of state dependent quantum cloning measured by the fidelity is higher than any noncontextual explanation one could provide. Letting $p_{bb,aa} = p_{b,a}^2$ to be a symmetry induced by the quantum description, we have that the behavior of the fidelity under the noncontextuality assumption, compared with the optimal cloning one for the ideal description of definition 31 is always larger.

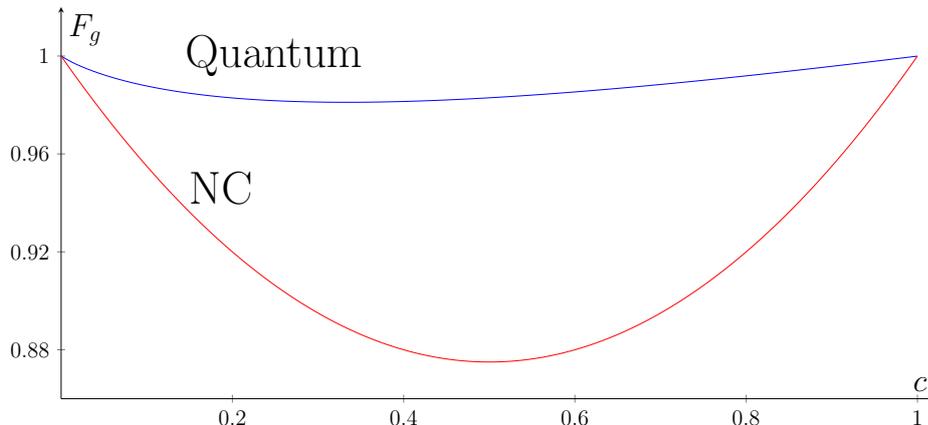


Figure 5.5: Comparing the optimal quantum fidelity with the optimal fidelity that any noncontextual ontological model need to obey. We plot the optimal fidelities against the confusability $c = p_{b,a}$.

We conclude that there exists quantum advantage for the cloning scenario with respect to noncontextual models. We also notice that there exists a threshold for robustness for a noisy experimental setting. In (Lostaglio and Senno, 2020) the author shows that for a white noise model,

$$\mathcal{N}(\rho) = (1 - v)\rho + \frac{v}{4}\mathbb{1} \quad (5.50)$$

where v is the visibility we get *regions* of advantage for the quantum fidelity. For a noise that is too high, the fidelity of quantum cloning becomes smaller than one that it is achieved by the noncontextual bound.

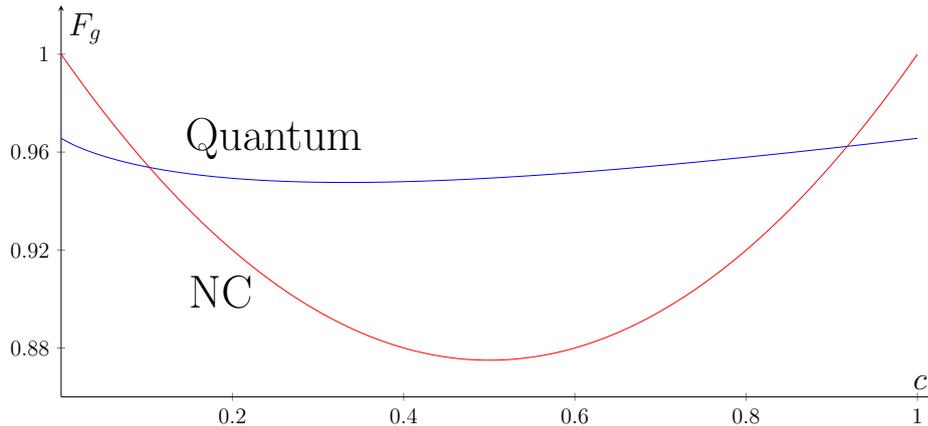


Figure 5.6: Comparing the optimal quantum fidelity with the optimal fidelity that any noncontextual ontological model need to obey. In this case, we are considering that there exists noise, with visibility parameter given by $v = 0.015$. We see a region where optimal quantum fidelity is higher than the noncontextual bound, but that due to noise this is no longer the case for all $c = p_{b,a}$. For the noisy fidelity we have considered the one in (Lostaglio and Senno, 2020).

We would like to conclude this section by pointing out that although the role that is played by the fidelity might be not directly related to a real measure of noncontextuality, both in the sense of a true resource theory monotone, but also in the sense of a mathematical distance in the space of quantum states, it is the common way to measure the success of cloning experimental tasks. Therefore, with respect to this function, and the operational description of the cloning scenario we see that contextuality poses itself as a resource for advantageous results.

It is also important to stress that in theorem 15 we have *not* relied on the fidelity as a witness of quantum contextuality. This is an operational account for the existence of quantum contextual behaviors. The direct implication is that, although one could criticize a utilization of the fidelity to express difference between contextual and noncontextual ontological models, the *structure* of the scenario \mathbb{B}_{qc} that was introduced in (Lostaglio and Senno, 2020) does present quantum contextuality, regardless if we take into account the specific measure of success for the quantum task.

CONCLUSIONS AND FURTHER DIRECTIONS

“There is hope that quantum mechanics will gradually lose its baffling quality... I have observed in teaching quantum mechanics, and also in learning it, that students go through an experience... The student begins by learning the tricks of the trade. He learns how to make calculations in quantum mechanics and get the right answers...it is comparatively painless. The second stage comes when the student begins to worry because he does not understand what he has been doing. He worries because he has no clear physical picture in his head... Then, unexpectedly, the third stage begins. The student suddenly says to himself, I understand quantum mechanics, or rather he says, I understand now that there isn’t anything to be understood... The duration and severity of the second stage are decreasing as the years go by. Each new generation of students learns quantum mechanics more easily than their teachers learned it...”

- Paul Dirac, quote from ([Susskind, 2016](#))

Although I agree with Paul Dirac, in the sense that students nowadays think about quantum mechanics in a much more natural way, it is by no means because scientists are more and more understanding that “there isn’t anything to be understood” but rather the complete opposite. More and more, we get used to quantum mechanics because of the invasion of new technologies in the life of all. The earlier we encounter quantum theory, the earlier we rephrase our imagination and our creativity in quantum terms, but never without truly searching for *why* quantum mechanics is the way it is. This is happening for practical reasons: we need better ways to understand quantum theory so that we will be more structured to enter the era of quantum computational advantage and quantum technological applications.

In this work, we have tackled the problem from a foundational perspective: to start, we devoted time to understanding the word *classical* means. We suppose a notion of classicality that is operationally relevant, generalized contextuality. By studying the resource theoretical approach, we have noticed the following pattern:

1. We first define a fragment of reality that is accessible by the experimenter. This is our description of a prepare-and-measure fragment.
2. This fragment provides an operational description of the quantum information task the experimenter will perform.
3. By imposing the assumption that there exists a noncontextual ontological explanation for the results of the experimental setting, we conclude that this information task is bounded by a parameter.
4. With the results of quantum theory, we are able to show that there are quantum procedures that violate the noncontextual bound.
5. Later we (generally) realize that the noncontextual bound was associated with a noncontextuality inequality for the noncontextual polytope associated with the experimental scenario.

In our work, we have attempted to answer the last item above. The idea was to provide general ways to witness quantum contextual behaviors present in general prepare-and-measure scenarios. However, not only that, to construct a framework so that the noncontextuality inequalities are understood: we based our approaches with known algorithmic results (Schmid et al., 2018) and the resource theory (Duarte and Amaral, 2018). We were not able to prove the existence of quantum advantage for general scenarios. However, we have provided a framework where one can engineer operational prepare-and-measure experimental settings with a quantum contextual structure *and* that have the noncontextuality inequalities given, using the box product.

With the resource theory binary map, we could understand why some scenarios in the literature, that are structurally related, all share quantum contextual behaviors. Although the complexity grows, some of the features of the noncontextuality polytope remain.

We have also pointed out that pre/post-processing the quantum behaviors might be associated with quantum simulation theory, and this can lead to a general technique for witnessing quantum contextuality for *any* scenario. These simulation problems are in general linear programming or semi-definite programming type problems.

Following this dissertation, there are many new directions to follow. To mention just a few, we could try to understand the connection between generalized contextuality and measurement-based quantum computation; Another direction possible is that many works indicate the relationship between protocols for measuring quantum thermodynamic quantities and negativity of quasi-probability distributions, which is an instance of contextual behavior (Lostaglio, 2020).

Other directions are to study the connection of the notion of transformation contextuality applied to experimental scenarios. There are structural results (Schmid et al.,

2020c) limiting the transformations under some circumstances, based on a categorical approach to quantum theory. In (Baldijão et al., 2020) the connection between scenarios with transformation procedures, prepare-and-measure scenarios, and the emergence of objectivity as an instance of noncontextuality was pointed out, and some general aspects of transformation contextuality were carefully considered. We believe that much more is still to be understood in that path.

A possible future direction can be associated with understanding the connection between generalized contextuality and other definitions of contextuality that try to encompass experimental robustness. In (Tezzin et al., 2020) a new interesting connection between the compatibility scenario approach and the contextuality-by-default was obtained, it would be interesting to search for similar connections thinking on generalized contextuality.

REFERENCES

- Abramsky, S., Barbosa, R. S., and Mansfield, S. (2017). Contextual fraction as a measure of contextuality. *Physical Review Letters*, 119:050504.
- Amaral, B. (2015). The exclusivity principle and the set of quantum correlations. *arXiv preprint arXiv:1502.03235*, Ph.D. Thesis.
- Amaral, B. (2020). SEMINÁRIOS VIRTUAIS - Prof. Bárbara Amaral (USP). *International Institute of Physics*, Youtube Channel - iiptv, last seen: 28, October 2020.
- Amaral, B., Cunha, M. T., and Cabello, A. (2014). Exclusivity principle forbids sets of correlations larger than the quantum set. *Physical Review A*, 89(3):030101.
- Amaral, B. and Terra Cunha, M. (2018). On graph approaches to contextuality and their role in quantum theory. *Springer Briefs in Mathematics*.
- Ambainis, A., Banik, M., Chaturvedi, A., Kravchenko, D., and Rai, A. (2016). Parity oblivious d-level random access codes and class of noncontextuality inequalities. *Quantum Information Processing*, 18(4):111.
- Araújo, M. (2012). Quantum realism and quantum surrealism. *arXiv preprint arXiv:1208.6283*, Master Thesis.
- Araújo, M., Quintino, M. T., Budroni, C., Terra Cunha, M., and Cabello, A. (2013). All noncontextuality inequalities for the n -cycle scenario. *Physical Review A*, 88:022118.
- Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G., Buell, D. A., et al. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510.
- Baldijão, R., Wagner, R., Amaral, B., Duarte, C., and Terra Cunha, M. (2020). Activating classicality in prepare-and-measure scenarios. (*In preparation*).

- Banik, M., Bhattacharya, S. S., Choudhary, S. K., Mukherjee, A., and Roy, A. (2014). Ontological models, preparation contextuality and nonlocality. *Foundations of Physics*, 44(11):1230–1244.
- Barnett, S. M. and Croke, S. (2008). On the conditions for discrimination between quantum states with minimum error. *arXiv preprint arXiv:0810.1919*.
- Barnum, H., Caves, C. M., Fuchs, C. A., Jozsa, R., and Schumacher, B. (1996). Noncommuting mixed states cannot be broadcast. *Physical Review Letters*, 76(15):2818.
- Barrett, J. (2007). Information processing in generalized probabilistic theories. *Physical Review A*, 75(3):032304.
- Bell, J. S. (1964). On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195.
- Bell, J. S. (1966). On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447.
- Bell, J. S. (1982). On the impossible pilot wave. *Foundations of Physics*, 12(10):989–999.
- Bian, Z., Majumdar, A. S., Jebarathinam, C., Wang, K., Xiao, L., Zhan, X., Zhang, Y., and Xue, P. (2020). Experimental demonstration of one-sided device-independent self-testing of any pure two-qubit entangled state. *Physical Review A*, 101:020301.
- Binder, F., Correa, L. A., Gogolin, C., Anders, J., and Adesso, G. (2018). Thermodynamics in the quantum regime. *Fundamental Theories of Physics (Springer)*.
- Boyd, S., Boyd, S. P., and Vandenberghe, L. (2004). Convex optimization. *Cambridge university press*.
- Bricker, P. (2016). Ontological Commitment. *Metaphysics Research Lab, Stanford University*, The Stanford Encyclopedia of Philosophy.
- Brøndsted, A. (2012). An introduction to convex polytopes. *Springer Science & Business Media*, volume 90.
- Bruß, D., DiVincenzo, D. P., Ekert, A., Fuchs, C. A., Macchiavello, C., and Smolin, J. A. (1998). Optimal universal and state-dependent quantum cloning. *Physical Review A*, 57(4):2368–2378.
- Cabello, A., Severini, S., and Winter, A. (2014). Graph-theoretic approach to quantum correlations. *Physical Review Letters*, 112(4):040401.
- Chailloux, A., Kerenidis, I., Kundu, S., and Sikora, J. (2016). Optimal bounds for parity-oblivious random access codes. *New Journal of Physics*, 18(4):045003.

- Chaturvedi, A., Farkas, M., and Wright, V. J. (2020). Characterising and bounding the set of quantum behaviours in contextuality scenarios. *arXiv preprint arXiv:2010.05853*.
- Chauí, M. (1983). Vida e obra de Leibniz. *Leibniz*, Trad. Carlos Lopes de Mattos e Marilena Chauí. In *Newton e Leibniz. Coleção Os Pensadores* (Portuguese).
- Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A. (1969). Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880.
- Clifton, R. and Kent, A. (2000). Simulating quantum mechanics by non-contextual hidden variables. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 456(2001):2101–2114.
- Coecke, B., Fritz, T., and Spekkens, R. W. (2016). A mathematical theory of resources. *Information and Computation*, 250:59–86.
- Davies, E. B. (1976). Quantum theory of open systems. *Academic Press*.
- Duarte, C. and Amaral, B. (2018). Resource theory of contextuality for arbitrary prepare-and-measure experiments. *Journal of Mathematical Physics*, 59(6):062202.
- Einstein, A. (1934). Mein weltbild. *Nijmegen.-GJ Thieme*, German.
- Einstein, A. (1960). Ideas and opinions. *Crown Trade Paperbacks New York*.
- Einstein, A., Podolsky, B., and Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777.
- Forrest, P. (1996). The identity of indiscernibles. *Metaphysics Research Lab, Stanford University*, Stanford Encyclopedia of Philosophy, Section 1: Formulating the principle.
- Frembs, M., Roberts, S., and Bartlett, S. D. (2018). Contextuality as a resource for measurement-based quantum computation beyond qubits. *New Journal of Physics*, 20(10):103011.
- Fritz, T. (2015). Resource convertibility and ordered commutative monoids. *Mathematical Structures in Computer Science*, 27(6):850–938.
- Ghorai, S. and Pan, A. K. (2018). Optimal quantum preparation contextuality in an n -bit parity-oblivious multiplexing task. *Physical Review A*, 98(3).
- Giustina, M., Versteegh, M. A., Wengerowsky, S., Handsteiner, J., Hochrainer, A., Phelan, K., Steinlechner, F., Kofler, J., Larsson, J.-Å., Abellán, C., et al. (2015). Significant-loop-hole-free test of bells theorem with entangled photons. *Physical Review Letters*, 115(25):250401.

- Gour, G. (2017). Quantum resource theories in the single-shot regime. *Physical Review A*, 95(6).
- Guerini, L. (2018). Simulating quantum measurements and quantum correlations. *Universidade Federal de Minas Gerais*, Master Thesis.
- Guerini, L., Bavaresco, J., Terra Cunha, M., and Acín, A. (2017). Operational framework for quantum measurement simulability. *Journal of Mathematical Physics*, 58(9):092102.
- Hardy, L. (2001). Quantum theory from five reasonable axioms. *arXiv preprint arXiv:0101012*.
- Heinosaari, T., Reitzner, D., and Stano, P. (2008). Notes on joint measurability of quantum observables. *Foundations of Physics*, 38(12):1133–1147.
- Helstrom, C. W. (1969). Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252.
- Henk, M., Richter-Gebert, J., and Ziegler, G. M. (2004). 16 basic properties of convex polytopes. *Handbook of discrete and computational geometry*, pages 255–382.
- Hermens, R. (2010). Quantum mechanics: from realism to intuitionism. *Radboud University Nijmegen*, *arXiv preprint arXiv:1002.1410*.
- Hermens, R. (2011). The problem of contextuality and the impossibility of experimental metaphysics thereof. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 42(4):214–225.
- Hermens, R. (2016). Philosophy of quantum probability: An empiricist study of its formalism and logic. *Radboud University Nijmegen*, Ph.D. Thesis.
- Hermens, R. (2019). An operationalist perspective on setting dependence. *Foundations of Physics*, 49(3):260–282.
- Hooft, G. (2016). The cellular automaton interpretation of quantum mechanics. *Fundamental Theories of Physics (Springer)*, volume 185.
- Horodecki, R., Horodecki, P., Horodecki, M., and Horodecki, K. (2009). Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942.
- Howard, D. (1985). Einstein on locality and separability. *Studies in History and Philosophy of Science Part A*, 16(3):171–201.
- Janssen, H. (2008). Reconstructing reality. *University of Groningen*, Master Thesis.

- Keyl, M. and Werner, R. F. (2016). Channels and maps. *Quantum Information: From Foundations to Quantum Technology Applications*, pages 75–90.
- Klyachko, A. A., Can, M. A., Binicioğlu, S., and Shumovsky, A. S. (2008). Simple test for hidden variables in spin-1 systems. *Physical Review Letters*, 101(2).
- Kochen, S. and Specker, E. P. (1975). The problem of hidden variables in quantum mechanics. In *The logico-algebraic approach to quantum mechanics*, pages 293–328. The University of Western Ontario Series in Philosophy of Science (A Series of Books on Philosophy of Science, Methodology, and Epistemology Published in Connection with the University of Western Ontario Philosophy of Science Programme) vol 5a. Springer, Dordrecht.
- Krishna, A. (2015). Experimentally testable noncontextuality inequalities via Fourier-Motzkin elimination. *University of Waterloo*, Master Thesis.
- Kunjwal, R. (2016). Contextuality beyond the Kochen-Specker theorem. *arXiv preprint arXiv:1612.07250*, Ph.D. Thesis.
- Kunjwal, R. (2019). Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements. *Quantum*, 3:184.
- Kunjwal, R., Lostaglio, M., and Pusey, M. F. (2019). Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts. *Physical Review A*, 100(4).
- Kunjwal, R. and Spekkens, R. W. (2015). From the Kochen-Specker theorem to non-contextuality inequalities without assuming determinism. *Physical Review Letters*, 115(11):110403.
- Landsman, N. P. (2005). When champions meet: Rethinking the Bohr-Einstein debate. *arXiv preprint arXiv:0507220*.
- Landsman, N. P. (2015). The fine-tuning argument. *arXiv preprint arXiv:1505.05359*.
- Leibniz, G. W. (1765). *Nouveaux essais sur l’entendement humain*. (French) Paris: Hachette.
- Leifer, M. S. (2014). Is the quantum state real? An extended review of ψ -ontology theorems. *arXiv preprint arXiv:1409.1570*.
- Liang, Y.-C., Spekkens, R. W., and Wiseman, H. M. (2011). Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(1-2):1–39.

- Liao, S.-K., Cai, W.-Q., Liu, W.-Y., Zhang, L., Li, Y., Ren, J.-G., Yin, J., Shen, Q., Cao, Y., Li, Z.-P., et al. (2017). Satellite-to-ground quantum key distribution. *Nature*, 549(7670):43–47.
- Lillystone, P., Wallman, J. J., and Emerson, J. (2019). Contextuality and the single-qubit stabilizer subtheory. *Physical Review Letters*, 122:140405.
- Locke, J. (2001). *Essai sur l'entendement humain*. Vrin.
- Lostaglio, M. (2020). Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response. *arXiv preprint arXiv:2004.01213*.
- Lostaglio, M. and Senno, G. (2020). Contextual advantage for state-dependent cloning. *Quantum*, 4:258.
- Makhorin, A. (2014). GNU linear programming kit (glpk), reference manual, January 2014.
- Malament, D. B. (1996). In defense of dogma: Why there cannot be a relativistic quantum mechanics of (localizable) particles. *Springer, Perspectives on quantum reality*:1–10.
- Mansfield, S. and Kashefi, E. (2018). Quantum advantage from sequential-transformation contextuality. *Physical Review Letters*, 121:230401.
- Mazurek, M. D., Pusey, M. F., Kunjwal, R., Resch, K. J., and Spekkens, R. W. (2016). An experimental test of noncontextuality without unphysical idealizations. *Nature communications*, 7.
- Mazurek, M. D., Pusey, M. F., Resch, K. J., and Spekkens, R. W. (2019). Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories. *arXiv preprint arXiv:1710.05948*.
- Meyer, D. A. (1999). Finite precision measurement nullifies the Kochen-Specker theorem. *Physical Review Letters*, 83(19):3751–3754.
- Morris, R. (2009). Topics in quantum foundations: Ontological models, and distinguishability as a resource. *University of Waterloo*, Master Thesis.
- Navascués, M., Pironio, S., and Acín, A. (2008). A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013.
- Nielsen, M. A. and Chuang, I. (2002). Quantum computation and quantum information. *AAPT*.

- Oldofredi, A. (2019). No-go theorems and the foundations of quantum physics. *arXiv preprint arXiv:1904.10991*.
- Oszmaniec, M., Guerini, L., Wittek, P., and Acín, A. (2017). Simulating positive-operator-valued measures with projective measurements. *Physical Review Letters*, 119(19).
- Paffenholz, A. (2006). New polytopes from products. *Journal of Combinatorial Theory, Series A*, 113(7):1396–1418.
- Pusey, M. F. (2018). Robust preparation noncontextuality inequalities in the simplest scenario. *Physical Review A*, 98:022112.
- Ren, J.-G., Xu, P., Yong, H.-L., Zhang, L., Liao, S.-K., Yin, J., Liu, W.-Y., Cai, W.-Q., Yang, M., Li, L., et al. (2017). Ground-to-satellite quantum teleportation. *Nature*, 549(7670):70–73.
- Ryan O’Donnel, Z. J. (2015). Lecture 17: Discriminating two quantum states.
- Saha, D. and Chaturvedi, A. (2019). Preparation contextuality as an essential feature underlying quantum communication advantage. *Physical Review A*, 100(2).
- Saha, D., Horodecki, P., and Pawłowski, M. (2019). State independent contextuality advances one-way communication. *New Journal of Physics*, 21(9):093057.
- Sainz, A. B. (2020). What is non-classical about quantum no-cloning? *Quantum Views*, 4:39.
- Sankaranarayanan, S. (2018). 4 5 tutorial solving lps with gnu lp kit glpk 8 minutes. Youtube Channel - manoranjan sahu, last time seen: 28, October 2020.
- Santos, R. F. d. (2018). Contextualidade e Grafos. *Universidade Federal de Minas Gerais*, Master Thesis (Portuguese).
- Schmid, D., Fraser, T. C., Kunjwal, R., Sainz, A. B., Wolfe, E., and Spekkens, R. W. (2020a). Why standard entanglement theory is inappropriate for the study of bell scenarios. *arXiv preprint arXiv:2004.09194*.
- Schmid, D., Selby, J., Wolfe, E., Kunjwal, R., and Spekkens, R. W. (2020b). The characterization of noncontextuality in the framework of generalized probabilistic theories. *arXiv preprint arXiv:1911.10386*.
- Schmid, D., Selby, J. H., Pusey, M. F., and Spekkens, R. W. (2020c). A structure theorem for generalized-noncontextual ontological models.

- Schmid, D., Selby, J. H., and Spekkens, R. W. (2020d). Unscrambling the omelette of causation and inference: The framework of causal-inferential theories. arXiv preprint arXiv:2009.03297.
- Schmid, D. and Spekkens, R. W. (2018). Contextual advantage for state discrimination. *Physical Review X*, 8(1):011015.
- Schmid, D., Spekkens, R. W., and Wolfe, E. (2018). All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences. *Physical Review A*, 97(6).
- Spekkens, R. (2016). Robert Spekkens: Leibniz’s principle of the identity of indiscernible as... *Rotman Institute of Philosophy*, Youtube Channel - Rotman Institute of Philosophy, last seen: 28, October 2020.
- Spekkens, R. W. (2005). Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5).
- Spekkens, R. W. (2007). Evidence for the epistemic view of quantum states: A toy theory. *Physical Review A*, 75:032110.
- Spekkens, R. W. (2019). The ontological identity of empirical indiscernibles: Leibniz’s methodological principle and its significance in the work of Einstein. *arXiv preprint arXiv:1909.04628*.
- Spekkens, R. W., Buzacott, D. H., Keehn, A. J., Toner, B., and Pryde, G. J. (2009a). Preparation contextuality powers parity-oblivious multiplexing. *Physical Review Letters*, 102(1).
- Spekkens, R. W., Buzacott, D. H., Keehn, A. J., Toner, B., and Pryde, G. J. (2009b). Preparation contextuality powers parity-oblivious multiplexing: classical protocols and experimental details. *Physical Review Letters*, 101(2):020401.
- Susskind, L. (2016). Copenhagen vs Everett, teleportation, and ER=EPR. *Fortschritte der Physik*, 64(6-7):551–564.
- Tezzin, A., Wagner, R., and Amaral, B. (2020). Contextuality-by-default for behaviours in compatibility scenarios.
- Werner, R. (2019). Mathematical methods of quantum information theory. *Tobias Osborne*, Youtube Channel - T. Osborne, last seen: 28, October 2020.
- Wilde, M. M. (2013). *Quantum information theory*. Cambridge University Press.

- Wootters, W. K. and Zurek, W. H. (1982). A single quantum cannot be cloned. *Nature*, 299(5886):802–803.
- Yin, J., Li, Y.-H., Liao, S.-K., Yang, M., Cao, Y., Zhang, L., Ren, J.-G., Cai, W.-Q., Liu, W.-Y., Li, S.-L., et al. (2020). Entanglement-based secure quantum cryptography over 1,120 kilometres. *Nature*, pages 1–5.
- Yu, S., Budini, A. A., Wang, Y.-T., Ke, Z.-J., Meng, Y., Liu, W., Li, Z.-P., Li, Q., Liu, Z.-H., Xu, J.-S., Tang, J.-S., Li, C.-F., and Guo, G.-C. (2019). Experimental observation of conditional past-future correlations. *Physical Review A*, 100:050301.
- Yuen, H., Kennedy, R., and Lax, M. (1975). Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Transactions on Information Theory*, 21(2):125–134.
- Zhan, X., Wang, K., Xiao, L., Bian, Z., Zhang, Y., Sanders, B. C., Zhang, C., and Xue, P. (2020). Experimental quantum cloning in a pseudo-unitary system. *Physical Review A*, 101:010302.

Glossary

In this section we wanted to provide further clarifications to the concepts that are fairly used, or even fundamental, to those who intent to deepen the work presented here. Since there is a strong interplay between philosophy of science, quantum physics, quantum information theory and experimentally minded theories when we speak about quantum generalized contextuality, we believe that those starting this research field might feel the need of an introduction to concepts that are, for obvious reasons, meagre in standard courses of physics. For example, research in foundations of quantum mechanics tend to express ideas such as *realism*, which is perhaps the most important one in the current study of that area, but without reasoning *what type of realism* is one committed to it. This might be a lack of philosophical attention in early courses of physics, but it can lead to misinterpretations. For example, we know now that the EPR realism is different than the Einstein view about a realistic approach towards scientific reasoning. But also Bohr realism, neo-realism or realism that gathers independence of an experiment device and cosmological ideas are all different notions. In this work we wanted to bring light to *one* notion of realism due to Spekkens, that is committed to Leibniz and Reichenbach.

1. *Einstein Realism*: In the paper of 1935 Einstein, Podolsky and Rosen tried to express the negative perception towards the quantum mechanical description of physical reality that they had by an argument of reduction to absurd. In the paper they use arguments (there is a very nice analysis of the EPR¹ argument and a careful consideration of their assumptions in (Hermens, 2010)) that are usually described as the criticism of Einstein being mainly focused on non-locality. But as one can really see in Einstein's letters do Schrödinger (Howard, 1985) Einstein himself had a different (stronger) conception of what would be necessary for a physical theory to describe reality. In his mind, a realistic theory should be able to describe the fact - this is more close to his realistic approach towards nature - that there are independent and separable systems. We must notice that this is much stronger

¹This acronym is famously due to the paper (Einstein et al., 1935) by Einstein, Podolsky and Rosen.

than a notion of non-locality: he argues that two pairs of systems should have their existence in a separable manner and analysed paradoxes and contradictions growing from that notion of realism. The best source for that is (Howard, 1985) in which in fact we even learn that Einstein himself seemed a bit upset with the final printed paper of 1935.

2. *The Principle of Indiscernibles*: Leibniz had in his philosophical commitments a set of principles of knowledge that represented the nucleus of his theoretical construction. Using this principles he tried to make a study on the abstract aspects of knowledge, such as the ideas, the act of thinking, etc. One of these principles is called the principle of indiscernibles that can be put in words as: there exist no two beings that are identically equal in the universe; each being is intrinsically different than the other and this difference is the essence of that thing (Chauí, 1983). This can be putted logically as, given two objects x and y , for properties F we have that (Forrest, 1996)

$$x = y \rightarrow \forall F(Fx \iff Fy), \quad (\text{A.1})$$

were we have that, if there are two things in the universe that are identical then one need that these two things must have the same properties no matter which, otherwise they would be discernible. And we can make this construction equivalent to

$$\forall F(Fx \iff Fy) \rightarrow x = y \quad (\text{A.2})$$

which is the logical definition that Spekkens is committed towards his notion of noncontextuality. If no property can be used to distinguish either object from one another, hence these should be representing the exact same thing (Spekkens, 2019).

3. *Ontologic model*: We mean an ontological commitment towards one presumed operational theory. Given a list of instructions for the experimentalist to perform in the laboratory that will provide oneself with probabilities an ontological model will prescribe the existence of entities with respect to that given operational prescription that will propose the ontic states of the system that has been studied in the experiment, inferring with this at least something about that ontic state. The observed features, described here in the generalized probabilistic theories framework, are then inferences about the ontic states and, it's considered that the experiment represent some feature of the system that is related to the ontological commitment; in other words, the experiment will represent the entities that were considered to exist once one proposes a connection between the ontic states and the operationally obtained statistics. The number of ontic states, that will be the existing entities of the model, will depend on what does it take for the model to be considered as representative of physical reality (that it is not the notion present in the EPR paper

but the one present in philosophy: Being of a Thing.) For the purposes of our work an extremely succinct description can be found in (Spekkens, 2005, sec. II pg. 2)

“An ontological model is an attempt to offer an explanation of the success of an operational theory by assuming that there exist physical systems that are the subject of the experiment. These systems are presumed to have attributes regardless of whether they are being subjected to experimental test, and regardless of what anyone knows about them. These attributes describe the real state of affairs of the system.”

4. *Ontic*: Relates to a complete description, or specification, of the defining aspects and properties of something. This means that by ontic we mean that that something has all it’s real comprehension of physical purpose and intent known. Hence, by the ontic state of the system we consider that the state the system is correspond to the reality of the existing aspect. By knowing the ontic state of the system one could known every real aspect or infer everything of that specific system, it’s true being, meaning not only predictability with probability one but also (philosophical) understanding of the nature that is involved on the construction of that existence.

“In the present context, an ontic state refers to something that objectively exists in the world, independently of any observer or agent. In other words, ontic states are the things that would still exist if all intelligent beings were suddenly wiped out from the universe.” (Leifer, 2014, pg. 69)

5. *No-go theorem*: The physicist and philosopher Ronnie Hermens defines the term as follows:

“(..) what no-go theorems actually prove. The setup of such theorems is that, first, certain plausible assumptions for physical theories are formulated. Then, it is shown that no theory that satisfies these assumptions can reproduce the predictions of quantum mechanics. Consequently, if quantum mechanics is taken to provide the correct predictions, one of the assumptions has to be rejected. ” (Hermens, 2016, Introduction, pg. 4)

In his work, he proceeds to study what is the meaning of some important no-go results, such as the Kochen-Specker theorem, when dealing with different interpretations of quantum probability theory. We can therefore understand the content of no-go theorems formulated in terms of logical expressions leading to contradictions. For example, in (Hermens, 2011), the author draws the logical description of the assumptions at play in the Kochen-Specker theorem as,

$$\mathbf{QM} \wedge \mathbf{Realism} \wedge \mathbf{CP} \wedge \mathbf{FM} \rightarrow \perp, \quad (\text{A.3})$$

where **CP** means the **Correspondence Principle** stating that there exists a bijective correspondence between observables and self-adjoint operators, and **FM** means **Faithful Measurements** stating that measurements of the observables reveals their values at that time, and finally, the symbol \perp means contradiction. The non-contextuality assumption is then part of **CP**,

$$\mathbf{QM} \wedge \mathbf{Realism} \wedge \mathbf{IP} \wedge \mathbf{NC} \wedge \mathbf{FM} \rightarrow \perp, \quad (\text{A.4})$$

where **NC** is a formulation of **Noncontextuality** and **IP** stands for the **Identification Principle**. The idea is that, writing equations such as (A.4), makes clear the fact that we could give up some of the assumptions to avoid the contradiction.

In fact, this discussion is relevant for what we understand as *loopholes* as well, since stating every hidden assumption in the arguments of no-go theorems (or by simply considering idealizations in the arguments) make very complicated to implement experimentally (the validity of) these theorems, since some of the assumptions might not be satisfied entirely during the procedure; therefore not fully validating the logical contradictions such as (A.4). The most famous example is the experimental implementation of a loophole-free Bell test (Giustina et al., 2015). No-go theorems are not only present in quantum theory, but also in relativistic quantum mechanics or quantum field theory (the algebraic approach) there exists no-go theorems related to the impossibility of particle interpretations (Malament, 1996). For the philosophical discussions concerning no-go results, see (Oldofredi, 2019) for a recent reference.

6. *Ontological Commitment*: Regarding the set of choices that one makes when constructing a theory, there always exists a set of truth conditions that must be satisfy by any entity within the theory.

“On its face, the notion of ontological commitment for theories is a simple matter. Theories have truth conditions. These truth conditions tell us how the world must be in order for the theory to be true; they make demands on the world. Sometimes, perhaps always, they demand of the world that certain entities or kinds of entity exist. The ontological commitments of a theory, then, are just the entities or kinds of entity that must exist in order for the theory to be true. ” (Bricker, 2016, Introduction)

This clarifies the following: although many interpretations of quantum theory do not assume the existence of any underlying ontology, or ontological models in the description we consider for this work, every interpretation of the theory must have some ontological commitment regarding what are the truth conditions allowed. Denying

the ontological model framework as presented in this dissertation constitutes itself as an ontological commitment.

Contextuality for quantum theory

B.1 Preparation contextuality

In this section we will develop the proof given by (Banik et al., 2014) of preparation contextuality for quantum theory that was stated in theorem 3.

In this chapter we proceed to demonstrate theorem 3. The general idea is extremely similar to the one given by Spekkens and it is mostly based upon definitions 17 and 16.

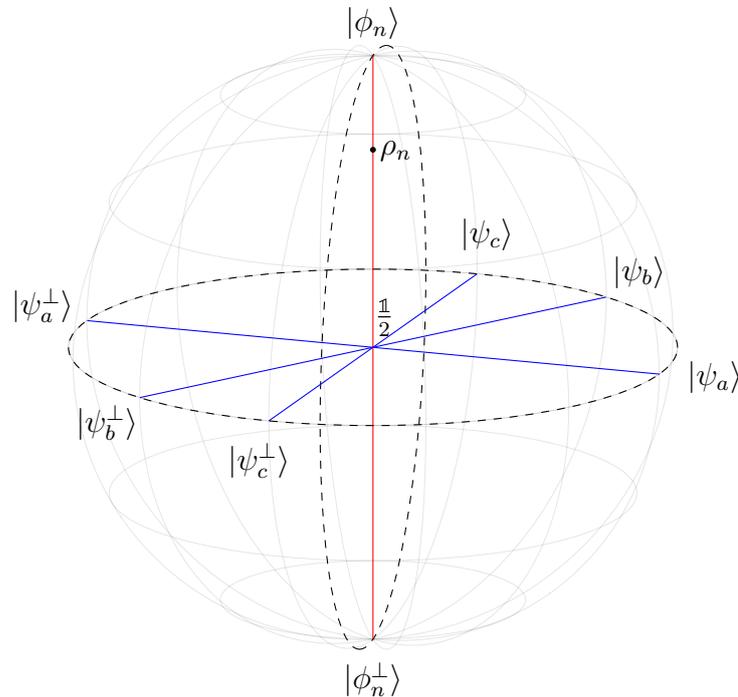


Figure B.1: Representation of the mixed state and all the six decompositions we will consider in the proof of preparation contextuality.

Any mixed state ρ_n from a qubit can be written as

$$\rho_n = \frac{1}{2}(\mathbb{1} + \vec{n} \cdot \vec{\sigma}) \tag{B.1}$$

were $0 \leq |\vec{n}| < 1$. We will follow (Banik et al., 2014) directly. Define $|\vec{n}| = q$. Then we will consider the six possible decompositions of the state ρ_n :

$$\rho_n = \frac{1-q}{2} |\phi_n^\perp\rangle\langle\phi_n^\perp| + \frac{1+q}{2} |\phi_n\rangle\langle\phi_n| \quad (\text{B.2})$$

$$= \frac{1-q}{2} (|\psi_a\rangle\langle\psi_a| + |\psi_a^\perp\rangle\langle\psi_a^\perp|) + q|\phi_n\rangle\langle\phi_n| \quad (\text{B.3})$$

$$= \frac{1-q}{2} (|\psi_b\rangle\langle\psi_b| + |\psi_b^\perp\rangle\langle\psi_b^\perp|) + q|\phi_n\rangle\langle\phi_n| \quad (\text{B.4})$$

$$= \frac{1-q}{2} (|\psi_c\rangle\langle\psi_c| + |\psi_c^\perp\rangle\langle\psi_c^\perp|) + q|\phi_n\rangle\langle\phi_n| \quad (\text{B.5})$$

$$= \frac{1-q}{3} (|\psi_a\rangle\langle\psi_a| + |\psi_b\rangle\langle\psi_b| + |\psi_c\rangle\langle\psi_c|) + q|\phi_n\rangle\langle\phi_n| \quad (\text{B.6})$$

$$= \frac{1-q}{3} (|\psi_a^\perp\rangle\langle\psi_a^\perp| + |\psi_b^\perp\rangle\langle\psi_b^\perp| + |\psi_c^\perp\rangle\langle\psi_c^\perp|) + q|\phi_n\rangle\langle\phi_n| \quad (\text{B.7})$$

where we have that $|\psi_n\rangle\langle\phi_n| = \frac{1}{2}(\mathbb{1} + \hat{n} \cdot \vec{\sigma})$ and we construct the vectors $|\psi_a\rangle, |\psi_b\rangle$ and $|\psi_c\rangle$ are chosen, similarly with the symmetric construction of all proofs in (Spekkens, 2005), such that every line between antipodal points has a distance of 60° with the other antipodal lines, from the labels b and c . Since we are considering antipodal points in the Bloch sphere this means that for each preparation associated with, say $|\psi_a\rangle\langle\psi_a|$ we have that the orthogonal one can be distinguished by a single-shot experiment, because of definition 16 we have that, for all $\lambda \in \Lambda$,

$$\mu_n(\lambda)\mu_{n^\perp}(\lambda) = 0 \quad (\text{B.8})$$

$$\mu_a(\lambda)\mu_{a^\perp}(\lambda) = 0 \quad (\text{B.9})$$

$$\mu_b(\lambda)\mu_{b^\perp}(\lambda) = 0 \quad (\text{B.10})$$

$$\mu_c(\lambda)\mu_{c^\perp}(\lambda) = 0 \quad (\text{B.11})$$

where we just define $\mu_n = \mu_{|\phi_n\rangle\langle\phi_n|}$, and etc. We have that the six decompositions are associated with six different contexts. Considering now definition 17 we have that the noncontextual assumption will lead us to the following relations:

$$\mu_{\rho_n}(\lambda) = \frac{1-q}{2}\mu_{n^\perp}(\lambda) + \frac{1+q}{2}\mu_n(\lambda) \quad (\text{B.12})$$

$$= \frac{1-q}{2}(\mu_a(\lambda) + \mu_{a^\perp}(\lambda)) + q\mu_n(\lambda) \quad (\text{B.13})$$

$$= \frac{1-q}{2}(\mu_b(\lambda) + \mu_{b^\perp}(\lambda)) + q\mu_n(\lambda) \quad (\text{B.14})$$

$$= \frac{1-q}{2}(\mu_c(\lambda) + \mu_{c^\perp}(\lambda)) + q\mu_n(\lambda) \quad (\text{B.15})$$

$$= \frac{1-q}{3}(\mu_a(\lambda) + \mu_b(\lambda) + \mu_c(\lambda)) + q\mu_n(\lambda) \quad (\text{B.16})$$

$$= \frac{1-q}{3}(\mu_{a^\perp}(\lambda) + \mu_{b^\perp}(\lambda) + \mu_{c^\perp}(\lambda)) + q\mu_n(\lambda) \quad (\text{B.17})$$

Since there are no solutions that are consistent with (B.8)-(B.17) we have that the representation of quantum theory of any mixed state is inconsistent with preparation non-contextual ontological model. Let's consider that Λ_ρ is the support of $\mu_{\rho_n}(\lambda)$. Then, considering all the ontic states $\lambda \in \Lambda_\rho$ we will have that, for example, take (B.8), this means that either $\mu_n(\lambda)$ or $\mu_{n^\perp}(\lambda)$ has to be equal to zero. Then there are in total 16 possible situations like that one, in order to satisfy (B.8)-(B.11). Let us consider each one:

1. For the case were $\mu_n(\lambda) = \mu_a(\lambda) = \mu_b(\lambda) = \mu_c(\lambda) = 0$ we will have that necessarily $\mu_{\rho_n}(\lambda) = 0$. But since $\lambda \in \Lambda_\rho$ we have a contradiction.
2. If we have that $\mu_n(\lambda) = \mu_{a^\perp}(\lambda) = \mu_b(\lambda) = \mu_c(\lambda) = 0$, then we have that from (B.13) it follows that $\mu_{\rho_n}(\lambda) = \frac{1-q}{2}\mu_a(\lambda)$ and from (B.16) we have that $\mu_{\rho_n}(\lambda) = \frac{1-q}{3}\mu_a(\lambda)$. Because we have that $q > 0$ the latter equations imply $\mu_{\rho_n}(\lambda) = 0$. Hence we get our contradiction. Notice that by the symmetry of the conditions (B.8)-(B.11) we have that same result for the other two symmetric cases b and c : $\mu_n(\lambda) = \mu_a(\lambda) = \mu_{b^\perp}(\lambda) = \mu_c(\lambda) = 0$ and also $\mu_n(\lambda) = \mu_a(\lambda) = \mu_b(\lambda) = \mu_{c^\perp}(\lambda) = 0$.
3. If $\mu_n(\lambda) = \mu_{a^\perp}(\lambda) = \mu_{b^\perp}(\lambda) = \mu_c(\lambda) = 0$, then from (B.15) we have that $\mu_{\rho_n}(\lambda) = \frac{1-q}{2}\mu_{c^\perp}(\lambda)$. And from (B.17) we have that $\mu_{\rho_n}(\lambda) = \frac{1-q}{3}\mu_{c^\perp}(\lambda)$, such that, as before, we get that $\mu_{\rho_n}(\lambda) = 0$ and hence, our contradiction. Similarly for when we have $\mu_a(\lambda) = 0$ and $\mu_b(\lambda) = 0$. Summing up we have then proved our result for 7 out of the 16 possibilities.
4. If $\mu_n(\lambda) = \mu_{a^\perp}(\lambda) = \mu_{b^\perp}(\lambda) = \mu_{c^\perp}(\lambda) = 0$ we get a contradiction directly following from (B.17).
5. When $\mu_{n^\perp}(\lambda) = \mu_a(\lambda) = \mu_b(\lambda) = \mu_c(\lambda) = 0$, then (B.16) and (B.12) imply that $\frac{1+q}{2}\mu_n(\lambda) = q\mu_n(\lambda) = \mu_{\rho_n}(\lambda)$, which is a contradiction as before.

6. If we consider $\mu_{n^\perp}(\lambda) = \mu_{a^\perp}(\lambda) = \mu_b(\lambda) = \mu_c(\lambda) = 0$, then we get from (B.13) and (B.16) that $\frac{1-q}{2}\mu_a(\lambda) + q\mu_n(\lambda) = \frac{1-q}{3}\mu_a(\lambda) + q\mu_n(\lambda)$. This last equation implies that $\mu_a(\lambda) = 0$ and we recover the case where we have that $\mu_{n^\perp}(\lambda) = \mu_a(\lambda) = \mu_b(\lambda) = \mu_c(\lambda) = 0$. This means that we get our contradiction. We have the same behavior happening when we exchange the orthogonality from a^\perp to the other letters b and c . Thus, we have proved our result for 12 out of the 16 possibilities.
7. If we consider $\mu_{n^\perp}(\lambda) = \mu_{a^\perp}(\lambda) = \mu_{b^\perp}(\lambda) = \mu_c(\lambda) = 0$ then we get from (B.15) and (B.17) that $\frac{1+q}{2}\mu_{c^\perp}(\lambda) + q\mu_n(\lambda) = \frac{1+q}{3}\mu_{c^\perp}(\lambda) + q\mu_n(\lambda)$ which implies that $\mu_{c^\perp}(\lambda) = 0$ and hence it follows the contradiction from the last case: see below the final item. Similarly we get other two possible cases by changing the letters.
8. We finally have the last possibility where we get that $\mu_{n^\perp}(\lambda) = \mu_{a^\perp}(\lambda) = \mu_{b^\perp}(\lambda) = \mu_{c^\perp}(\lambda) = 0$ this implies that, by equations (B.12) and (B.17) the relation $\mu_{\rho_n}(\lambda) = \frac{1+q}{2}\mu_n(\lambda) = q\mu_n(\lambda)$. This implies that $\mu_n(\lambda) = 0$ because $q < 1$. We have then that our contradiction follows from the first item. We have then got a contradiction from all 16 possibilities.

We then conclude that for ρ_n there is no noncontextual ontological model consistent with assignments of the ontic states λ and the quantum theory. This concludes the proof.

B.2 Measurement contextuality

Here we will prove that quantum theory is inconsistent with a noncontextual ontological model for unsharp measurements, as stated in (Spekkens, 2005) and, in our work, theorem 4.

Consider the following binary-outcome measurements M_a, M_b, M_c associated with the projection valued-measures $\{\Pi_a, \Pi_A\}, \{\Pi_b, \Pi_B\}, \{\Pi_c, \Pi_C\}$ defined by

$$\Pi_a := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \Pi_A := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{B.18}$$

$$\Pi_b := \begin{pmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{pmatrix} \quad \Pi_B := \begin{pmatrix} \frac{3}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix} \tag{B.19}$$

$$\Pi_c := \begin{pmatrix} \frac{1}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{3}{4} \end{pmatrix} \quad \Pi_C := \begin{pmatrix} \frac{3}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{1}{4} \end{pmatrix} \tag{B.20}$$

we get then that the following conditions are satisfied:

$$\Pi_a \Pi_A = 0 \tag{B.21}$$

$$\Pi_b \Pi_B = 0 \tag{B.22}$$

$$\Pi_c \Pi_C = 0 \tag{B.23}$$

$$\Pi_a + \Pi_A = \mathbb{1} \tag{B.24}$$

$$\Pi_b + \Pi_B = \mathbb{1} \tag{B.25}$$

$$\Pi_c + \Pi_C = \mathbb{1} \tag{B.26}$$

given that we assume sharp measurements respect outcome determinism in the ontological level, we shall have that there are idempotent functions χ_a, \dots, χ_C such that for all ontic states $\lambda \in \Lambda$ in our ontological model

$$\chi_a(\lambda) \chi_A(\lambda) = 0 \tag{B.27}$$

$$\chi_b(\lambda) \chi_B(\lambda) = 0 \tag{B.28}$$

$$\chi_c(\lambda) \chi_C(\lambda) = 0 \tag{B.29}$$

$$\chi_a(\lambda) + \chi_A(\lambda) = \mathbb{1} \tag{B.30}$$

$$\chi_b(\lambda) + \chi_B(\lambda) = \mathbb{1} \tag{B.31}$$

$$\chi_c(\lambda) + \chi_C(\lambda) = \mathbb{1} \tag{B.32}$$

We will construct the unsharp measurement M as the measurement procedure associated with the POVM

$$\left\{ \frac{1}{3} \Pi_a + \frac{1}{3} \Pi_b + \frac{1}{3} \Pi_c, \frac{1}{3} \Pi_A + \frac{1}{3} \Pi_B + \frac{1}{3} \Pi_C \right\} = \left\{ \frac{\mathbb{1}}{2}, \frac{\mathbb{1}}{2} \right\} \tag{B.33}$$

and by definition 17 in the ontological model we get that for any ontic state $\lambda \in \Lambda$,

$$\left\{ \frac{1}{3} \chi_a + \frac{1}{3} \chi_b + \frac{1}{3} \chi_c, \frac{1}{3} \chi_A + \frac{1}{3} \chi_B + \frac{1}{3} \chi_C \right\} \tag{B.34}$$

Now, consider the measurement \tilde{M} that flip a fair coin in order to determine the outcomes of the measurements. This measurement¹ does *not* have to be associated by any means with quantum theory, the only thing we need to notice is that the measurements M and \tilde{M} are operationally equivalent. We obtain that the probability associated with measurement effects of \tilde{M} is simply

$$\left\{ \frac{1}{2}, \frac{1}{2} \right\}$$

¹Which makes the argument even more beautiful.

were each one should be understood as the effects $\xi_{[k|\tilde{M}]}(\lambda)$ regarding measurement \tilde{M} for any ontic state $\lambda \in \Lambda$. But we note that, regarding the noncontextuality assumption we must have that since M and \tilde{M} are equivalent they're ontological model description should be the same implying that

$$\frac{1}{3}\chi_a + \frac{1}{3}\chi_b + \frac{1}{3}\chi_c = \frac{1}{2} \tag{B.35}$$

$$\frac{1}{3}\chi_A + \frac{1}{3}\chi_B + \frac{1}{3}\chi_C = \frac{1}{2} \tag{B.36}$$

1. From the equations (B.27)-(B.29) we can have eight possible cases: suppose we have that $\chi_a(\lambda) = \chi_b(\lambda) = \chi_c(\lambda) = 0$, then we must have that all other effects are equal to 1.
2. We get the same behavior by making $\chi_A(\lambda) = \chi_B(\lambda) = \chi_C(\lambda) = 0$.
3. Consider the case were $\chi_a(\lambda) = \chi_b(\lambda) = \chi_c(\lambda) = 0$. We have then that $\chi_B(\lambda) = \chi_A(\lambda) = \chi_c(\lambda) = 1$, hence we get that the assignment (B.34) would be $\{\frac{2}{3}, \frac{1}{3}\}$. If we have the other two cases (for two lowercase and one upper case) of $\chi_a(\lambda) = \chi_C(\lambda) = \chi_c(\lambda) = 0$ and also $\chi_A(\lambda) = \chi_b(\lambda) = \chi_c(\lambda) = 0$ we get the same result.
4. If instead we consider the 3 cases of the form $\chi_a(\lambda) = \chi_B(\lambda) = \chi_C(\lambda) = 0$ we get then $\chi_A(\lambda) = \chi_b(\lambda) = \chi_c(\lambda) = 1$ implying that we get the assignment (B.34) to be $\{\frac{1}{3}, \frac{2}{3}\}$.

and we get a contradiction because the constraints (B.27)-(B.32) imply that the set of possible values consistent with (B.35)-(B.36) are just $\{0, 1\}$, $\{1, 0\}$, $\{\frac{2}{3}, \frac{1}{3}\}$, $\{\frac{1}{3}, \frac{2}{3}\}$, and hence cannot be $\{\frac{1}{2}, \frac{1}{2}\}$ completing the proof.



PORTA

We shall present here the important features of the PORTA software in order to use it for computation of the noncontextual measurement-assignment polytope, such as obtaining the inequalities that characterize the generalized universally noncontextual polytope/model. We consider the most important and simplest of all scenarios for our work, and proceed following the work of (Schmid et al., 2018). For our purposes we will use two programs inside PORTA; the TRAF command and also the fmel command. Both are important and perform certain algorithms that can be found in other softwares such as QSKELETON and LRS for Fourier-Motzkin elimination, as well as SKELETON and CDDLIB for vertex enumeration.

C.1 Vertex enumeration

In order to perform vertex enumeration using the TRAF command from PORTA, consider the procedure to calculate, for example, the enumeration for the noncontextual measurement-assignment polytope \mathbb{B}_{si} . In such a scenario we have no operational equivalences and hence the enumeration is done over the following set of inequalities,

```
DIM = 4
```

```
INEQUALITIES_SECTION
```

```
(1) x1 >= 0
```

```
(2) x2 >= 0
```

```
(3) x3 >= 0
```

```
(4) x4 >= 0
```

```
(5) x1+x2 == 1
```

```
(6) x3+x4 == 1
```

```
END
```

where we have just four variables. Here the dimension of the problem is given by $|K| \times |J|$, i.e, the number of outcomes multiplied by the number of measurements in the scenario. The vector \vec{x} in the program corresponds to the vector $\vec{\xi}(\lambda)$ from (3.31), $x_1 = \xi_{[m_1|M_1]}(\lambda)$ and so on. This is the form of every TRAF operation on PORTA, were we just write down the dimension correctly and use the operations $\leq, \geq, =, =$ etc. The PORTA software reads the file as strings using commands CHAR perform Gauss elimination. This means that one cannot insert crazy notation such as $3.4*x2$, because the program will not recognize its format, and will return errors. Another important fact is that the command performs elimination on the equality's first, so if the equations are too restrictive the program will eliminate too many variables and not find any vertex solution. A final (but important) remark; in order to run TRAF for vertex enumeration it is important that the code be written in a .IEQ type file, otherwise as stated in the PORTA.C in the line

```
else if (is_set(Traf) && ieq_file)
```

returns reading errors. So, in order to run the command, go to the BIN directory in the PORTA-1.4.1 file run the command

```
traf name_of_the_file.ieq
```

For the simplest scenario \mathbb{B}_{si} we find the four vertices:

$$\kappa_1 \rightarrow (1, 0, 0, 1) \tag{C.1}$$

$$\kappa_2 \rightarrow (1, 0, 1, 0) \tag{C.2}$$

$$\kappa_3 \rightarrow (0, 1, 0, 1) \tag{C.3}$$

$$\kappa_4 \rightarrow (0, 1, 1, 0) \tag{C.4}$$

$$\tag{C.5}$$

and for the fair-coin-flip scenario \mathbb{B}_{fcf} we find, with notation for the vertex κ^* as

$$\left(\xi_{[0|M_1]}(\kappa^*), \xi_{[1|M_1]}(\kappa^*), \xi_{[0|M_2]}(\kappa^*), \xi_{[1|M_2]}(\kappa^*), \xi_{[0|M_3]}(\kappa^*), \xi_{[1|M_3]}(\kappa^*) \right),$$

were we will have the vertices of the noncontextual measurement-assignment polytope by

$$\kappa_1 \rightarrow \left(1, 0, \frac{1}{2}, \frac{1}{2}, 0, 1\right) \quad (\text{C.6})$$

$$\kappa_2 \rightarrow \left(1, 0, 0, 1, \frac{1}{2}, \frac{1}{2}\right) \quad (\text{C.7})$$

$$\kappa_3 \rightarrow \left(\frac{1}{2}, \frac{1}{2}, 1, 0, 0, 1\right) \quad (\text{C.8})$$

$$\kappa_4 \rightarrow \left(\frac{1}{2}, \frac{1}{2}, 0, 1, 1, 0\right) \quad (\text{C.9})$$

$$\kappa_5 \rightarrow \left(0, 1, 1, 0, \frac{1}{2}, \frac{1}{2}\right) \quad (\text{C.10})$$

$$\kappa_6 \rightarrow \left(0, 1, \frac{1}{2}, \frac{1}{2}, 1, 0\right) \quad (\text{C.11})$$

Since it is extremely simple to assign the vertices for a scenario with no operational equivalences, we have created the efficiency table

$ \mathcal{M} $	$ \mathcal{O}_M $	# κ 's	Time(s)	$ \mathcal{M} $	$ \mathcal{O}_M $	# κ 's	Time(s)
2	2	2^2	0	5	3	243	0
5	2	2^5	0	10	3	59049	16
10	2	2^{10}	1	11	3	177147	159
15	2	2^{15}	3	12	3	531441	1691
16	2	2^{16}	20	13	3	1594323	10870
17	2	2^{17}	58	8	5	390625	810
18	2	2^{18}	242				
19	2	2^{19}	1000				
20	2	2^{20}	3885				

Table C.1: Vertex enumeration PORTA: efficiency

From table C.1 we can have some knowledge about the efficiency of the PORTA program when performed by a normal PC. We have learned by the simulations that around $\text{Dim} > 43$ the program starts to run out of space to reallocate variables.

C.2 Fourier-Motzkin elimination

Fourier-Motzkin elimination, that from now on we refer to as FME, is an algorithm used in order to eliminate one variable from finite set of them, say x_1, \dots, x_n into x_2, \dots, x_n that are constrained through several inequalities. In our case, we want to use FME to eliminate unknown variables in order to obtain a set of known variables, representing the state of a system in an operational framework, that will represent the H-description of the generalized-noncontextual universal polytope. An explanation of the

algorithm can be found in (Krishna, 2015). The important message is that this algorithm is complex, and its computation complexity scale as $\mathcal{O}\left(\left(\frac{n}{4}\right)^{2^d}\right)$. where n is the initial amount of variables and d are the eliminated ones. This result draw attention at implementing the algorithm as a step in a larger program.

Simplest scenario

Proceeding with the calculations of the full set of tight inequalities of the polytope $NC(\mathbb{B}_{si})$. Since we know the vertices (C.1)-(C.4) we can write a PORTA program, in order to use command FMEL, by just writing all equations (3.40)-(3.43).

```

DIM = 32

ELIMINATION_ORDER
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

INEQUALITIES_SECTION
(1) x1 >= 0
(2) x2 >= 0
(3) x3 >= 0
(4) x4 >= 0
(5) x5 >= 0
(6) x6 >= 0
(7) x7 >= 0
(8) x8 >= 0
(9) x9 >= 0
(10) x10 >= 0
(11) x11 >= 0
(12) x12 >= 0
(13) x13 >= 0
(14) x14 >= 0
(15) x15 >= 0
(16) x16 >= 0
(18) x1+x2+x3+x4 == 1
(19) x5+x6+x7+x8 == 1
(20) x9+x10+x11+x12 == 1
(21) x13+x14+x15+x16 == 1
(22) +1/2x1+1/2x5-1/2x9-1/2x13== 0
(23) +1/2x2+1/2x6-1/2x10-1/2x14== 0
(24) +1/2x3+1/2x7-1/2x11-1/2x15== 0
(25) +1/2x4+1/2x8-1/2x12-1/2x16== 0

```

```

(26) +1x3+1x4== x17
(27) +1x1+1x2== x18
(28) +1x2+1x4== x19
(29) +1x1+1x3== x20
(30) +1x7+1x8== x21
(31) +1x5+1x6== x22
(32) +1x6+1x8== x23
(33) +1x5+1x7== x24
(34) +1x11+1x12== x25
(35) +1x9+1x10== x26
(36) +1x10+1x12== x27
(37) +1x9+1x11== x28
(38) +1x15+1x16== x29
(39) +1x13+1x14== x30
(40) +1x14+1x16== x31
(41) +1x13+1x15== x32
END

```

Above we have considered that the variables are representatives of the following, with respect to (3.40)-(3.43):

$$\begin{array}{llll}
x1 = \nu_{P_1}(1) & x2 = \nu_{P_1}(2) & x3 = \nu_{P_1}(3) & x4 = \nu_{P_1}(4) \\
x5 = \nu_{P_2}(1) & x6 = \nu_{P_2}(2) & x7 = \nu_{P_2}(3) & x8 = \nu_{P_2}(4) \\
x9 = \nu_{P_3}(1) & x10 = \nu_{P_3}(2) & x11 = \nu_{P_3}(3) & x12 = \nu_{P_3}(4) \\
x13 = \nu_{P_4}(1) & x14 = \nu_{P_4}(2) & x15 = \nu_{P_4}(3) & x16 = \nu_{P_4}(4) \\
x17 = p(0|M_1, P_1) & x18 = p(1|M_1, P_1) & x19 = p(0|M_2, P_1) & x20 = p(1|M_2, P_1) \\
x21 = p(0|M_1, P_2) & x22 = p(1|M_1, P_2) & x23 = p(0|M_2, P_2) & x24 = p(1|M_2, P_2) \\
x25 = p(0|M_1, P_3) & x26 = p(1|M_1, P_3) & x27 = p(0|M_2, P_3) & x28 = p(1|M_2, P_3) \\
x29 = p(0|M_1, P_4) & x30 = p(1|M_1, P_4) & x31 = p(0|M_2, P_4) & x32 = p(1|M_2, P_4)
\end{array}$$

Hence, the format of the code is now clear, we want to eliminate all the variables that have to do with the ν 's so that we only have the set of equations/inequalities with respect to the probabilities, and hence, with respect to the data-table form the scenario. By running the FMEL we have that the program stops, because the line (25) represents a redundancy in the construction. After one just delete line (25) the program gives back the following result:

```
DIM = 32
```

```
INEQUALITIES_SECTION
```

```

( 1)  -x17-x18 == -1
( 2)      -x19-x20 == -1
( 3)      -x21-x22 == -1
( 4)      -x23-x24 == -1
( 5)      -x25-x26 == -1
( 6)      -x27-x28 == -1
( 7)  -x18      -x22      +x26      -x29 == -1
( 8)      -x29-x30 == -1
( 9)      -x20      -x24      +x28      -x31 == -1
(10)      -x31-x32 == -1

( 1) -x22 <= 0
( 2)  -x24 <= 0
( 3)  -x26 <= 0
( 4)  -x28 <= 0
( 5)  -x30 <= 0
( 6)  -x32 <= 0
( 7)  -x26 -x30 <= 0
( 8)  -x26 -x30 <= 0
( 9)  -x26 -x32 <= 0
(10)  -x28-x30 <= 0
(11)  -x28 -x32 <= 0
(12)  -x28 -x32 <= 0
(13)  +x24 -x28 -x32 <= 0
(14) +x22 -x26 -x30 <= 0
(15) -x22+x24 -x28 -x32 <= 0
(16) +x22-x24-x26 -x30 <= 0
(17)      +x32 <= 1
(18)      +x30 <= 1
(19)      +x28 <= 1
(20)      +x26 <= 1
(21)      +x24 <= 1
(22) +x22 <= 1
(23)  -x26      +x32 <= 1
(24)      -x28+x30 <= 1
(25)      +x28-x30 <= 1
(26)      +x26      -x32 <= 1
(27) -x22 +x26 +x30 <= 1
(28)  -x24 +x28 +x32 <= 1
(29) -x22-x24 +x28 +x32 <= 1
(30) -x22-x24 +x28+x30 <= 1

```

```

( 31) -x22-x24+x26      +x32 <= 1
( 32) -x22-x24+x26  +x30   <= 1
( 33) -x22+x24   -x28+x30   <= 1
( 34) -x22+x24+x26      -x32 <= 1
( 35) +x22-x24-x26      +x32 <= 1
( 36) +x22-x24   +x28-x30   <= 1
( 37) +x22+x24-x26   -x30   <= 1
( 38) +x22+x24-x26      -x32 <= 1
( 39) +x22+x24   -x28-x30   <= 1
( 40) +x22+x24   -x28   -x32 <= 1
( 41)                +x28   +x32 <= 2
( 42)                +x28   +x32 <= 2
( 43)                +x28+x30   <= 2
( 44)                +x26      +x32 <= 2
( 45)                +x26   +x30   <= 2
( 46)                +x26   +x30   <= 2
( 47) -x22+x24+x26  +x30   <= 2
( 48) +x22-x24   +x28   +x32 <= 2

```

END

Note that from (1)-(10) in the equations section we have the equations telling us that the probabilities for the same outcomes sum to one and, importantly, we have the operational equivalences from the equality's for preparations (choosing to write $p(1|M_i, P_j) = p_{ij}$).

$$p_{11} + p_{12} = p_{13} + p_{14} \tag{C.12}$$

$$p_{21} + p_{22} = p_{23} + p_{24} \tag{C.13}$$

And also, from the inequalities section we have that from (1)-(28) and (41)-(48) we have inequalities that arise from the characteristics of the behaviors being probabilistic and not from the noncontextuality assumptions. Hence we have the set of twelve inequalities (29)-(40) that we can write as the probabilities:

$$\begin{aligned}
 (29) : & \quad p_{23} + p_{24} - p_{22} - p_{12} \leq 1 \\
 (30) : & \quad p_{23} + p_{14} - p_{22} - p_{12} \leq 1 \\
 (31) : & \quad p_{13} + p_{24} - p_{22} - p_{12} \leq 1 \\
 (32) : & \quad p_{13} + p_{14} - p_{22} - p_{12} \leq 1 \\
 (33) : & \quad p_{22} + p_{14} - p_{12} - p_{23} \leq 1 \\
 (34) : & \quad p_{22} + p_{13} - p_{12} - p_{24} \leq 1 \\
 (35) : & \quad p_{12} + p_{24} - p_{22} - p_{13} \leq 1 \\
 (36) : & \quad p_{12} + p_{23} - p_{22} - p_{14} \leq 1 \\
 (37) : & \quad p_{12} + p_{22} - p_{13} - p_{14} \leq 1 \\
 (38) : & \quad p_{12} + p_{22} - p_{13} - p_{24} \leq 1 \\
 (39) : & \quad p_{12} + p_{22} - p_{14} - p_{23} \leq 1 \\
 (40) : & \quad p_{12} + p_{22} - p_{23} - p_{24} \leq 1
 \end{aligned}$$

But we notice that using (C.12) and (C.13) we get that the highlighted inequalities (29),(32),(37),(40) are also somewhat “not tight” since they are just trivial:

$$\underbrace{p_{23} + p_{24} - p_{22}}_{p_{21}} - p_{12} \leq 1 \implies p_{21} - p_{12} \leq 1$$

Always true since $0 \leq p_{ij} \leq 1$. Doying the same for the other inequalities, we have that

$$\begin{aligned}
 \underbrace{p_{13} + p_{14} - p_{12}}_{p_{11}} - p_{22} \leq 1 & \implies p_{11} - p_{22} \leq 1 \\
 p_{22} + \underbrace{p_{12} - p_{13} - p_{14}}_{-p_{11}} \leq 1 & \implies p_{22} - p_{11} \leq 1 \\
 p_{12} + \underbrace{p_{22} - p_{23} - p_{24}}_{-p_{21}} \leq 1 & \implies p_{12} - p_{21} \leq 1
 \end{aligned}$$

Hence, we have that the full set of tight inequalities that represent the noncontextual polytope $NC(\mathbb{B}_{si})$ is, by the method given in (Schmid et al., 2018):

$$0 \leq p_{ij} \leq 1, \forall M_i, P_j \tag{C.14}$$

$$p_{12} + p_{22} - p_{14} - p_{23} \leq 1 \tag{C.15}$$

$$p_{12} + p_{22} - p_{13} - p_{24} \leq 1 \tag{C.16}$$

$$p_{22} + p_{13} - p_{12} - p_{24} \leq 1 \tag{C.17}$$

$$p_{12} + p_{23} - p_{22} - p_{14} \leq 1 \tag{C.18}$$

$$p_{22} + p_{14} - p_{12} - p_{23} \leq 1 \tag{C.19}$$

$$p_{23} + p_{14} - p_{12} - p_{22} \leq 1 \tag{C.20}$$

$$p_{12} + p_{24} - p_{22} - p_{13} \leq 1 \tag{C.21}$$

$$p_{13} + p_{24} - p_{22} - p_{12} \leq 1 \tag{C.22}$$

were we have that (C.20) has corrected the typo with respect to (Schmid et al., 2018, Eq. 29g, pg. 7).

Linear programming

D.1 Noncontextuality from data-tables

By the calculations performed in appendix C and remarks from section 3.3 it is not a computationally useful way to study the contextuality of a quantum scenario by fully deriving all the noncontextuality inequalities. A much better method is obtained by means of the *Farkas dual*. In (Schmid et al., 2018) they have developed a way to analyse the noncontextuality assumption of an ontological model only by looking at the data-table \vec{b} . This is done by first writing the equations (3.41)-(3.43) as a matrix equation:

$$M \cdot \vec{x} = \vec{b}^* \quad (\text{D.1})$$

where \vec{b}^* is a possible behavior of the scenario \mathbb{B} that admits a noncontextual model $\vec{B}^* \in NC(\mathbb{B})$ were the distinction between \vec{b} and \vec{B} will be made clear in a second at equation (D.6). The condition of equation (3.31) is then just $\vec{x} \geq 0$. We then formulate a primal LP as

$$\begin{aligned} &\exists \vec{x} \text{ such that} \\ &M \cdot \vec{x} = \vec{b}^*, \\ &\text{and } \vec{x} \geq 0 \end{aligned} \quad (\text{D.2})$$

for any data-table \vec{b}^* given. This LP is checking whether any solution can be found, such that $\vec{B}^* \in NC(\mathbb{B})$. When such primal LP is unfeasible we can obtain a certificate by means of the Farkas dual. This certificate is obtained by solving

$$\begin{aligned} &\min_{\vec{y}} \vec{y} \cdot \vec{b}^* \text{ such that} \\ &\vec{1} \geq \vec{y} \cdot M \geq \vec{0}. \end{aligned} \quad (\text{D.3})$$

Hence we have the following method:

1. If the primal LP is feasible \implies Farkas dual is also feasible with $\vec{y} \cdot \vec{b}^* \geq 0 \implies$ there exists a noncontextual ontological model for \vec{B}^* .
2. If the primal LP is unfeasible \implies from the Farkas dual we get that $\vec{y} \cdot \vec{b}^* < 0 \implies$ there is no noncontextual ontological model for \vec{B}^* .

Simplest scenario

Considering the scenario as \mathbb{B}_{si} we get that the matrix M constructed following equations (3.41)-(3.43) acts on a 16 dimensional vector \vec{x} associated with the $\nu_{P_j}(\kappa)$ since we have 4 vertices and 4 preparation procedures. Now, it will really make a difference for the matrix M the way one constructs the vector \vec{x} ! Consider the following manner:

$$\vec{x} := (\nu_{P_1}(\kappa_1), \nu_{P_1}(\kappa_2), \nu_{P_1}(\kappa_3), \nu_{P_1}(\kappa_4), \nu_{P_2}(\kappa_1), \nu_{P_2}(\kappa_2), \dots, \nu_{P_4}(\kappa_4)) \quad (\text{D.4})$$

$$M = \begin{pmatrix}
 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
 \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\
 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\
 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 \\
 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0
 \end{pmatrix} \tag{D.5}$$

were we have used the convention for the behavior as

$$\vec{b} = (1, 1, 1, 1, 0, 0, 0, 0, p(0|M_1, P_1), p(0|M_1, P_2), p(0|M_1, P_3), p(0|M_1, P_4), \\
 p(1|M_1, P_1), p(1|M_1, P_2), \dots, p(1|M_2, P_4)) = (1, 1, 1, 1, 0, 0, 0, 0, \vec{B}) \tag{D.6}$$

Notice that we have highlighted the lines in matrix M that we can see from (C.1)-(C.4) as the lines depicted in M being the ‘‘columns’’ from (C.1)-(C.4). This is because we take as an example, from (3.43) we have that for $p(0|M_1, P_1)$ we have

$$1 \cdot \nu_{P_1}(\kappa_1) + 1 \cdot \nu_{P_1}(\kappa_2) + 0 \cdot \nu_{P_1}(\kappa_3) + 0 \cdot \nu_{P_1}(\kappa_4) = p(0|M_1, P_1)$$

because we fix the measurement event $[0|M_1]$ which is the first entry of the vectors from (C.1)-(C.4).

LP for the Farkas dual

By following (D.3) we have only to write a linear program for the Farkas lemma. There are many different ways to proceed in writing an LP as (D.3); normally a very good resource for linear programming is the MATLAB/OCTAVE programming language, that offers several possibilities and it is extremely simple to write the code. Also MATHEMATICA is a good resource for such type of calculation. But, despite the fact that these two resources are incredible for all sorts of analytic/LP calculations, in our work, whenever possible we shall make use of free software: in what follows we use GLPK - which is built in to other programming languages - and if necessary, PYTHON, so that we are able to produce data statistics for the LP's. In order to use CPLEX/GLPK there are helpful resources (Makhorin, 2014),(Sankaranarayanan, 2018) in order to use them. Here we have used a very rudimentary C program to write the GLPK file below:

```

var x1;
var x2;
var x3;
var x4;
.
.
.
var x24;
minimize obj:1.0000 * x1+1.0000 * x2+1.0000 * x3+1.0000 * x4+1.0000 *
    x9+1.0000 * x11+1.0000 * x14+1.0000 * x16+1.0000 * x17+1.0000 * x20+1.0000
    * x22+1.0000 * x23+0.0000 * x24;
c1:1.00 * x1 + 0.50 * x5 + 1.00 * x9 + 1.00 * x12 + 0.00 * x24 <= 1;
c2:1.00 * x1 + 0.50 * x6 + 1.00 * x9 + 1.00 * x11 + 0.00 * x24 <= 1;
c3:1.00 * x1 + 0.50 * x7 + 1.00 * x10 + 1.00 * x12 + 0.00 * x24 <= 1;
c4:1.00 * x1 + 0.50 * x8 + 1.00 * x10 + 1.00 * x11 + 0.00 * x24 <= 1;
c5:1.00 * x2 + 0.50 * x5 + 1.00 * x13 + 1.00 * x16 + 0.00 * x24 <= 1;
c6:1.00 * x2 + 0.50 * x6 + 1.00 * x13 + 1.00 * x15 + 0.00 * x24 <= 1;
c7:1.00 * x2 + 0.50 * x7 + 1.00 * x14 + 1.00 * x16 + 0.00 * x24 <= 1;
c8:1.00 * x2 + 0.50 * x8 + 1.00 * x14 + 1.00 * x15 + 0.00 * x24 <= 1;
c9:1.00 * x3 + -0.50 * x5 + 1.00 * x17 + 1.00 * x20 + 0.00 * x24 <= 1;
c10:1.00 * x3 + -0.50 * x6 + 1.00 * x17 + 1.00 * x19 + 0.00 * x24 <= 1;
c11:1.00 * x3 + -0.50 * x7 + 1.00 * x18 + 1.00 * x20 + 0.00 * x24 <= 1;
c12:1.00 * x3 + -0.50 * x8 + 1.00 * x18 + 1.00 * x19 + 0.00 * x24 <= 1;
c13:1.00 * x4 + -0.50 * x5 + 1.00 * x21 + 1.00 * x24 <= 1;
c14:1.00 * x4 + -0.50 * x6 + 1.00 * x21 + 1.00 * x23 + 0.00 * x24 <= 1;
c15:1.00 * x4 + -0.50 * x7 + 1.00 * x22 + 1.00 * x24 <= 1;
c16:1.00 * x4 + -0.50 * x8 + 1.00 * x22 + 1.00 * x23 + 0.00 * x24 <= 1;

```

```

c17:-1.00 * x1 + -0.50 * x5 + -1.00 * x9 + -1.00 * x12 + -0.00 * x24 <= 0;
c18:-1.00 * x1 + -0.50 * x6 + -1.00 * x9 + -1.00 * x11 + -0.00 * x24 <= 0;
c19:-1.00 * x1 + -0.50 * x7 + -1.00 * x10 + -1.00 * x12 + -0.00 * x24 <= 0;
c20:-1.00 * x1 + -0.50 * x8 + -1.00 * x10 + -1.00 * x11 + -0.00 * x24 <= 0;
c21:-1.00 * x2 + -0.50 * x5 + -1.00 * x13 + -1.00 * x16 + -0.00 * x24 <= 0;
c22:-1.00 * x2 + -0.50 * x6 + -1.00 * x13 + -1.00 * x15 + -0.00 * x24 <= 0;
c23:-1.00 * x2 + -0.50 * x7 + -1.00 * x14 + -1.00 * x16 + -0.00 * x24 <= 0;
c24:-1.00 * x2 + -0.50 * x8 + -1.00 * x14 + -1.00 * x15 + -0.00 * x24 <= 0;
c25:-1.00 * x3 + 0.50 * x5 + -1.00 * x17 + -1.00 * x20 + -0.00 * x24 <= 0;
c26:-1.00 * x3 + 0.50 * x6 + -1.00 * x17 + -1.00 * x19 + -0.00 * x24 <= 0;
c27:-1.00 * x3 + 0.50 * x7 + -1.00 * x18 + -1.00 * x20 + -0.00 * x24 <= 0;
c28:-1.00 * x3 + 0.50 * x8 + -1.00 * x18 + -1.00 * x19 + -0.00 * x24 <= 0;
c29:-1.00 * x4 + 0.50 * x5 + -1.00 * x21 + -1.00 * x24 <= 0;
c30:-1.00 * x4 + 0.50 * x6 + -1.00 * x21 + -1.00 * x23 + -0.00 * x24 <= 0;
c31:-1.00 * x4 + 0.50 * x7 + -1.00 * x22 + -1.00 * x24 <= 0;
c32:-1.00 * x4 + 0.50 * x8 + -1.00 * x22 + -1.00 * x23 + -0.00 * x24 <= 0;
solve;
display x1;
display x2;
display x3;
display x4;
.
.
.
display x24;
end;

```

Check ([Sankaranarayanan, 2018](#)) about the format of the file above. We have considered the following behavior from ([Schmid et al., 2018](#)):

$$\vec{B}_e = (1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0) \quad (\text{D.7})$$

that maximally violates inequality (C.22). If we run the above program

```
glpsol --math Name_of_the_file.txt
```

we get

```

GLPSOL: GLPK LP/MIP Solver, v4.65
Parameter(s) specified in the command line:
--math LP.txt

```

```
Reading model section from LP.txt...
83 lines were read
Generating obj...
Generating c1...
Generating c2...
.
.
.
Generating c32...
Model has been successfully generated
GLPK Simplex Optimizer, v4.65
33 rows, 24 columns, 140 non-zeros
Preprocessing...
32 rows, 24 columns, 128 non-zeros
Scaling...
  A: min|aij| = 5.000e-01 max|aij| = 1.000e+00 ratio = 2.000e+00
Problem data seem to be well scaled
Constructing initial basis...
Size of triangular part is 32
*   0: obj = 0.000000000e+00 inf = 0.000e+00 (12)
*  14: obj = -1.000000000e+00 inf = 0.000e+00 (0)
OPTIMAL LP SOLUTION FOUND
Time used: 0.0 secs
Memory used: 0.2 Mb (207206 bytes)
Display statement at line 59
x1.val = 0
Display statement at line 60
x2.val = 0
Display statement at line 61
x3.val = 0
Display statement at line 62
x4.val = 1
Display statement at line 63
x5.val = 0
Display statement at line 64
x6.val = 0
Display statement at line 65
x7.val = 0
Display statement at line 66
x8.val = -2
Display statement at line 67
```

```

x9.val = 0
Display statement at line 68
x10.val = 1
Display statement at line 69
x11.val = 0
Display statement at line 70
x12.val = 0
Display statement at line 71
x13.val = 0
Display statement at line 72
x14.val = 0
Display statement at line 73
x15.val = 1
Display statement at line 74
x16.val = 0
Display statement at line 75
x17.val = 0
Display statement at line 76
x18.val = 0
Display statement at line 77
x19.val = 0
Display statement at line 78
x20.val = 0
Display statement at line 79
x21.val = 0
Display statement at line 80
x22.val = -1
Display statement at line 81
x23.val = -1
Display statement at line 82
x24.val = 0
Model has been successfully processed

```

Hence obtaining as a result the value $\vec{y} \cdot \vec{b} = -1 < 0$ showing that we indeed have that the primal LP is unfeasible and hence that \vec{B}_e is a behavior that is outside the noncontextual polytope. As a result for \vec{y} we get

$$\vec{y} = (0, 0, 0, 1, 0, 0, 0, -2, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, -1, -1, 0) \quad (\text{D.8})$$

which is the same result we have obtained by means of using the LINPROG function from MATLAB/OCTAVE, and also using the MINIMIZE from MATHEMATICA. We reinforce that

the way one defines the vector \vec{b} and the ordering of the vertices (C.1)-(C.4) influences on how the vectors \vec{y}, \vec{x} and also the matrix M will look like, so that the presentation here are only mere examples and not fixed, uniquely defined, quantities. The important aspect is that no matter *how* the construction is made, the linear program will present a final result independent of such format; in other words, the value -1 should be always obtained. This is so because we can understand the certificate as a noncontextual inequality, which is a fixed constraint over the noncontextual polytope.

D.2 Contextual Fraction

Because of the necessary and sufficient linear conditions that define noncontextuality in a finite prepare-and-measure scenario it is possible to use convex optimization (Boyd et al., 2004) to obtain some of the monotones that are associated with the convex structure of the scenario. In particular, the set of linear conditions:

$$\begin{aligned} \forall k, j : \nu_{P_j}(\kappa) &\geq 0, \\ \forall j : \sum_{\kappa} \nu_{P_j}(\kappa) &= 1, \\ \forall \kappa, a : \sum_j (\alpha_j^a - \beta_j^a) \nu_{P_j}(\kappa) &= 0, \\ \forall k, i, j : \sum_{\kappa} \tilde{\xi}_{[k|M_i]}(\kappa) \nu_{P_j}(\kappa) &= p(k|M_i, P_j) \end{aligned}$$

can be used to apply the contextuality monotones that are associated with the convexity structure of the *noncontextual polytope*: the contextual fraction is our case study here due to its importance for application in quantum computing, see (Abramsky et al., 2017), (Frembs et al., 2018).

We propose the following, in light of (Abramsky et al., 2017), suppose that \vec{B} is a possible behavior for the operational scenario given. We then relax the equality between the operational probabilities and the ontological distributions, since the data-table might be contextual, but we let the vector $\vec{\nu}$ measure the best noncontextual description possible, and we maximize over the weight $\frac{1}{|\mathcal{P}|} \vec{\nu} \cdot \vec{1}$, where we have divided by the number of preparation procedures so that the contextual fraction would be in the range $[0, 1]$. We construct then the LP,

$$\begin{aligned}
 &\text{Maximize} && \frac{1}{4} \vec{v} \cdot \vec{1} \\
 &\text{Subject to} && A_{eq} \vec{v} = \vec{0} \\
 &&& A_{ineq} \vec{v} \leq \vec{B} \\
 &&& \vec{v} \geq \vec{0}
 \end{aligned}$$

If we let r be the result of the LP, the contextual fraction is then $f(B) = 1 - r$, since we are finding the optimal noncontextual fraction in this manner. Each $\vec{v} \cdot \vec{1} = c$ represents a hyperplane for $c \in \mathbb{R}$, and the optimal result is then the closest point with the noncontextuality inequality that represents a facet of the noncontextual polytope.

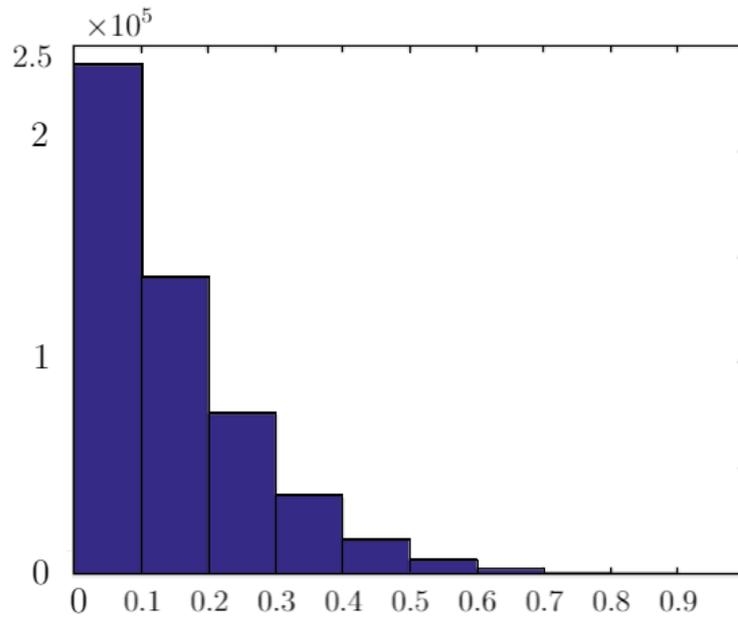


Figure D.1: Monte Carlo distribution: We have generated and calculated the contextual fraction for behaviors in \mathbb{B}_{si} . In the plot we have around 5 000 000 points.

We notice that in the scenario \mathbb{B}_{si} we have exactly eight contextual vertices, that are the points with $f(B_v) = 1$. To find all the vertices for $C(\mathbb{B})$ we simply need to find the deterministic assignments: all the deterministic assignments can be found by noticing

the imposed restrictions over the probabilities for preparation procedures.

$$\begin{aligned}
 p_{i1} + p_{i2} &= p_{i3} + p_{i4} \Rightarrow \\
 1 + 1 &= 1 + 1 \\
 1 + 0 &= 1 + 0 \\
 0 + 1 &= 1 + 0 \\
 1 + 0 &= 0 + 1 \\
 0 + 1 &= 1 + 0 \\
 0 + 0 &= 0 + 0
 \end{aligned}$$

From which we construct the following table, we have highlighted the vertices for which we have contextual fraction equal to one, and all the other are noncontextual vertices for which the contextual fraction returns zero:

	(1, 1, 1, 1)	(1, 0, 0, 1)	(1, 0, 1, 0)	(0, 1, 0, 1)	(0, 1, 1, 0)	(0, 0, 0, 0)
(1, 1, 1, 1)	(1, 1, 1, 1) (1, 1, 1, 1)	(1, 0, 0, 1) (1, 1, 1, 1)	(1, 0, 1, 0) (1, 1, 1, 1)	(0, 1, 0, 1) (1, 1, 1, 1)	(0, 1, 1, 0) (1, 1, 1, 1)	(0, 0, 0, 0) (1, 1, 1, 1)
(1, 0, 0, 1)	(1, 1, 1, 1) (1, 0, 0, 1)	(1, 0, 0, 1) (1, 0, 0, 1)	(1, 0, 1, 0) (1, 0, 0, 1)	(0, 1, 0, 1) (1, 0, 0, 1)	(0, 1, 1, 0) (1, 0, 0, 1)	(0, 0, 0, 0) (1, 0, 0, 1)
(1, 0, 1, 0)	(1, 1, 1, 1) (1, 0, 1, 0)	(1, 0, 0, 1) (1, 0, 1, 0)	(1, 0, 1, 0) (1, 0, 1, 0)	(0, 1, 0, 1) (1, 0, 1, 0)	(0, 1, 1, 0) (1, 0, 1, 0)	(0, 0, 0, 0) (1, 0, 1, 0)
(0, 1, 0, 1)	(1, 1, 1, 1) (0, 1, 0, 1)	(1, 0, 0, 1) (0, 1, 0, 1)	(1, 0, 1, 0) (0, 1, 0, 1)	(0, 1, 0, 1) (0, 1, 0, 1)	(0, 1, 1, 0) (0, 1, 0, 1)	(0, 0, 0, 0) (0, 1, 0, 1)
(0, 1, 1, 0)	(1, 1, 1, 1) (0, 1, 1, 0)	(1, 0, 0, 1) (0, 1, 1, 0)	(1, 0, 1, 0) (0, 1, 1, 0)	(0, 1, 0, 1) (0, 1, 1, 0)	(0, 1, 1, 0) (0, 1, 1, 0)	(0, 0, 0, 0) (0, 1, 1, 0)
(0, 0, 0, 0)	(1, 1, 1, 1) (0, 0, 0, 0)	(1, 0, 0, 1) (0, 0, 0, 0)	(1, 0, 1, 0) (0, 0, 0, 0)	(0, 1, 0, 1) (0, 0, 0, 0)	(0, 1, 1, 0) (0, 0, 0, 0)	(0, 0, 0, 0) (0, 0, 0, 0)

Table D.1: Table of vertices from the large polytope of all behaviors $C(\mathbb{B}_{si})$: Contextual vertices for the simplest scenario. Each of these vertices violate one of the tight noncontextuality inequalities defined by (C.14)-(C.22).

We have seen that the quantum behavior B^Q from (5.2) violates a noncontextuality inequality for \mathbb{B}_{si} . If we calculate the contextual fraction $f(B^Q) \approx 0.413$, which is the same amount of how much the behavior violates the inequality (C.21), as we have seen in chapter 5.

Proofs for Noncontextual bounds

E.1 n -bit parity oblivious multiplexing

In this section we shall prove theorem 11. We will follow (Spekkens et al., 2009b) closely. In the text we have already described the operational equivalences for n -bit parity as

$$\forall r \in \text{Par}, \forall M, \forall k : \sum_{x|x \cdot r=0} p(k|M, P_x) = \sum_{x|x \cdot r=1} p(k|M, P_x), \quad (\text{E.1})$$

but this is truly achieved after some steps. The fact is that the operational equivalences are written as

$$\forall k \in \mathcal{O}_M, \forall M \in \mathcal{M}, \forall r \in \text{Par} : \sum_{x|x \cdot r=0} p(P_x|k, M) = \sum_{x|x \cdot r=1} p(P_x|k, M),$$

from Bayes theorem we have that

$$p(P_x|k, M) = \frac{p(k|P_x, M)p(P_x)}{p(k)},$$

since $p(P_x)$ is uniform, because it is the probability associated with the distribution of the classical inputs generated by Alice, it does not depend on x . Hence, we get the relation we have provided in the main text. Any message m that can be sent by Alice will be read from Bob as measurement outcomes $k|M$, given that they don't communicate with each other. Hence, we can write that for any message m that Alice sends to Bob, we must have the operational equivalences given by,

$$\forall r \in \text{Par} : \sum_{x|x \cdot r=0} p(m|P_x) = \sum_{x|x \cdot r=1} p(m|P_x), \quad (\text{E.2})$$

Lemma 10. *We can describe the probabilities $p(m|P_x)$ using the following equation,*

$$p(m|P_x) = p(0)p_0(m) + \sum_{i=1}^n p(i) [p_{i,0}(m)\delta_{x_i,0} + p_{i,1}(m)\delta_{x_i,1}], \quad (\text{E.3})$$

where $p(i)$ is a normalized probability distribution over $\{0, \dots, n\}$. We have that $p_0(m)$, $p_{i,0}(m)$ and $p_{i,1}(m)$ are normalized probability distributions over the message m .

1. Here we have that $i \in \{0, 1, \dots, n\}$ represents the same variables as y , but we here use the symbol i with generality and the symbol y to be the one that was chosen to be guessed. Intuitively we could interpret any oblivious multiplexing protocol as Alice generating i from distribution $p(i)$; for $i = 0$ she sends the message m chosen from $p_0(m)$, otherwise she sends a message m depending on the value of the i -th bit of the message, by choosing between two different distributions, $p_{i,0}$ and $p_{i,1}$.
2. Remember that P_x is chosen associated with x , which is generated at random so we also have that the preparation P_x is performed uniformly at random. Therefore the above mentioned Bayes rule works.

Using this representation we can choose the distributions that achieve the maximum probability of Bob guessing correctly. Notice that from (E.3) we have that whenever $i = 0$ there is no information gained by Bob of what is the value of the entry of x .

1. x is the message Alice has such that Bob want's to guess the y -th coordinate.
2. m is the message that Alice actually sends to Bob.

Therefore the optimal strategy needs to set $p(0) = 0$. Again looking at (E.3) any information about x_i will be associated with the degree of distinction Bob can make between the probabilities $p_{i,0}(m)$ and $p_{i,1}(m)$. Bob will be able to fully distinguish between these two probability distributions whenever they are completely non-overlapping, otherwise there could be some arbitrary aspect on the choice, that we have already tackled in figure 3.1. This means that for any i and m we must have $p_{i,0}(m)p_{i,1}(m) = 0$. The optimal strategy for Bob is then clear: Given any y , find out weather the message m is in the support of $p_{y,0}$ or $p_{y,1}$, using a scheme that has outcomes $b = 0$ or $b = 1$ respectively.

The reason this is optimal is intuitive. The message Alice sends m contains information about x_y only if Alice generates $i = y$ and so Bob will guess correctly with probability one. When she does not generate $i = y$ Bob has equal probability of guessing between 0 or 1, according with the results of his measurements (either he finds in which support is, $p_{y,0}$ or $p_{y,1}$, or he does not find), hence, 1/2 of chance.

Then we get that for y chosen uniformly at random the probability that $i = y$ is $1/n$ so that the best probability for Bob to guess correctly will be

$$p(g = x_y) = \frac{1}{n}(1) + \left(1 - \frac{1}{n}\right) \frac{1}{2} = \frac{1+n}{2n},$$

and this is the first important result. It remains to prove that equation (E.3) is valid.

Proof. (of equation (E.3)) Let $r \in \{0, 1\}^n$ and $\chi_r : \{0, 1\}^n \rightarrow [-1, 1]$ defined as

$$\chi_r(x) = (-1)^{x \cdot r},$$

we have that,

$$\sum_{x \in \{0,1\}^n} \chi_r(x) \chi_{r'}(x) = \sum_{x \in \{0,1\}^n} (-1)^{x \cdot (r \oplus r')} = 2^n \delta_{r,r'}.$$

This can be shown by induction. For $n = 2$ we see that $r = r'$ imply that

$$r \oplus r' \in \{(0, 0)\} \implies x \cdot (r \oplus r') \in \{0\},$$

and when $r \neq r'$ we can have all possible outcomes,

$$r \oplus r' \in \{(0, 1), (1, 0), (1, 1)\} \implies x \cdot (r \oplus r') \in \{0, 1\}$$

and the sum gets $\forall(r \oplus r') :$

$$\sum_{x \in \{0,1\}^2} (-1)^{x \cdot (r \oplus r')} = (-1)^{(0,1) \cdot (r \oplus r')} + (-1)^{(1,0) \cdot (r \oplus r')} + (-1)^{(0,0) \cdot (r \oplus r')} + (-1)^{(1,1) \cdot (r \oplus r')} = 0$$

Now if we consider for $n + 1$ we get

$$\begin{aligned} \sum_{x \in \{0,1\}^{n+1}} (-1)^{x \cdot (r \oplus r')} &= \sum_{x \in \{(0,s):s \in \{1,0\}^n\}} (-1)^{x \cdot (r \oplus r')} + \sum_{x \in \{(1,s):s \in \{1,0\}^n\}} (-1)^{x \cdot (r \oplus r')} \\ &= 2^n \delta_{r,r'} + 2^n \delta_{r,r'} = 2 \cdot 2^n \delta_{r,r'} = 2^{n+1} \delta_{r,r'} \end{aligned}$$

This is so because, suppose that the first entry of $(r \oplus r')_1 = 1$, this can only happens if $r \neq r'$. In this case we get that, letting $\tilde{r} \oplus \tilde{r}' = a(r \oplus r')$ were a is the operator that

annihilates the first entry,

$$\begin{aligned} & \sum_{x \in \{(0,s):s \in \{1,0\}^n\}} (-1)^{x \cdot (r \oplus r')} + \sum_{x \in \{(1,s):s \in \{1,0\}^n\}} (-1)^{x \cdot (r \oplus r')} = \\ & \sum_{x \in \{(0,s):s \in \{1,0\}^n\}} (-1)^{(0 \times 1) \oplus s \cdot a(r \oplus r')} + \sum_{x \in \{(1,s):s \in \{1,0\}^n\}} (-1)^{(1 \times 1) \oplus s \cdot a(r \oplus r')} = \\ & \sum_{s \in \{1,0\}^n} (-1)^{s \cdot a(r \oplus r')} - \sum_{s \in \{1,0\}^n} (-1)^{s \cdot a(r \oplus r')} = 0. \end{aligned}$$

Remember that $x \cdot r := \bigoplus_i x_i r_i$. We then proceed to note that χ_r is an orthonormal basis for the space of all functions on $\{0, 1\}^n$ since this space has 2^n cardinality and the orthonormal set χ_r has the same cardinality. Therefore,

$$p(m|P_x) = \sum_r \hat{p}(m, r) \chi_r(x).$$

We see that

$$\begin{aligned} 2^n \hat{p}(m, r) &= \sum_x \chi_r(x) \chi_r(x) \hat{p}(m, r) = \sum_x \chi_r(x) p(m|P_x) = \sum_x (-1)^{x \cdot r} p(m|P_x) \\ &= \sum_{x|x \cdot r=0} p(m|P_x) - \sum_{x|x \cdot r=1} p(m|P_x). \end{aligned}$$

and because of (E.2) we get $\hat{p}(m, r) = 0$ for all $r \in \text{Par}$. The only strings r that will have $\hat{p}(m, r) \neq 0$ are those with a number symbols different of zero (Hamming weight) of either 0 or 1. Let $\hat{p}_0(m)$ be the Fourier coefficient associated with the string with only zeros, and those with a single value 1 at position i by $\hat{p}_i(m)$. Therefore,

$$p(m|P_x) = \hat{p}_0(m) + \sum_{i=1}^n \hat{p}_i(m) (-1)^{x_i}.$$

Write $(-1)^{x_i} = \delta_{x_i,0} - \delta_{x_i,1}$ and $1 = \delta_{x_i,0} + \delta_{x_i,1}$. We might rewrite the above equation defining

$$\begin{aligned} a_{i,0}(m) &= 2\hat{p}_i(m), \text{ if } \text{sgn}(\hat{p}_i(m)) \geq 0, \\ a_{i,1}(m) &= 0, \text{ if } \text{sgn}(\hat{p}_i(m)) \geq 0, \\ a_{i,0}(m) &= -2\hat{p}_i(m), \text{ if } \text{sgn}(\hat{p}_i(m)) < 0, \\ a_{i,1}(m) &= 0, \text{ if } \text{sgn}(\hat{p}_i(m)) < 0. \end{aligned}$$

Such that $\hat{p}_i(m) \geq 0 \implies a_{i,0}(m)\delta_{x_i,0} + a_{i,1}(m)\delta_{x_i,1} = 2\hat{p}_i(m)\delta_{x_i,0}$, and also that when $\hat{p}_i(m) < 0$ we have that $a_{i,0}\delta_{x_i,0} + a_{i,1}\delta_{x_i,1} = -2\hat{p}_i(m)\delta_{x_i,1}$. This means that,

$$\sum_{i=1}^n [a_{i,0}(m)\delta_{x_i,0} + a_{i,1}(m)\delta_{x_i,1}] = \sum_{i|\hat{p}_i(m) \geq 0} 2\hat{p}_i(m)\delta_{x_i,0} - \sum_{i|\hat{p}_i(m) < 0} 2\hat{p}_i(m)\delta_{x_i,1}$$

and since

$$\begin{aligned} \sum_{i=1}^n \hat{p}_i(m)(-1)^{x_i} &= \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)(-1)^{x_i} - \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)(-1)^{x_i} \\ &= \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)\delta_{x_i,0} - \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)\delta_{x_i,1} + \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)\delta_{x_i,0} - \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)\delta_{x_i,1} \\ &= \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)\delta_{x_i,0} - \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)(1 - \delta_{x_i,0}) + \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)(1 - \delta_{x_i,1}) - \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)\delta_{x_i,1} \\ &= 2 \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m)\delta_{x_i,0} - 2 \sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m)\delta_{x_i,1} + \underbrace{\left(\sum_{i|\hat{p}_i(m) < 0} \hat{p}_i(m) - \sum_{i|\hat{p}_i(m) \geq 0} \hat{p}_i(m) \right)}_{\text{Implicit in } a_0(m)}. \end{aligned}$$

So that we get

$$p(m|P_x) = \hat{p}_0(m) + \sum_{i=1}^n \hat{p}_i(m)(-1)^{x_i} = a_0(m) + \sum_{i=1}^n [a_{i,0}(m)\delta_{x_i,0} + a_{i,1}(m)\delta_{x_i,1}]$$

with $a_{i,0}(m), a_{i,1}(m) \geq 0$, remains to show that $a_0(m)$ is also nonnegative. Let $z(m) \in \{0, 1\}^n$ be the bit string defined as

$$z(m) = \begin{cases} 1, & \text{sgn}(\hat{p}_i(m)) \geq 0 \\ 0, & \text{sgn}(\hat{p}_i(m)) < 0 \end{cases}, \quad (\text{E.4})$$

so we have that, by definition, for $x = z(m)$ we have that

$$a_{i,0}(m)\delta_{x_i,0} + a_{i,1}(m)\delta_{x_i,1} = 0, \forall i \implies p(m|P_{z(m)}) = a_0(m),$$

and since $p(m|P_{z(m)})$ is a probability we get $a_0(m)$ is nonnegative. The final form of (E.3) is achieved by noticing that, for any x ,

$$1 = \sum_m p(m|P_x) = \sum_m a_0(m) + \sum_{i=1}^n \sum_m a_{i,x_i}(m),$$

define $A_0 := \sum_m a_0(m)$ and $A_{i,x_i} := \sum_m a_{i,x_i}(m)$. We have that $A_0 + \sum_{i=1}^n A_{i,x_i} = 1$ for all x so that $\sum_{i=1}^n A_{i,x_i}$ does not depend on x , therefore it does not depend on x_i and so

$A_{i,0} = A_{i,1}$, for all i . Then we have that the form of (E.3) is achieved by letting

$$\begin{aligned} p(0) &= A_0 \\ p(i) &= A_{i,0} = A_{i,1} \\ p_0(m) &= \frac{a_0(m)}{p(0)} \\ p_{i,b}(m) &= \frac{a_{i,b}(m)}{p(i)} \end{aligned}$$

if $p(0), p(i) \neq 0$. This finishes the proof. \square

We have proved then the following lemma.

Lemma 11. *For any classical protocol the best guessing strategy is given by $(1+n)/2n$.*

To prove theorem 11 we need to prove that any noncontextual ontological model cannot do better. Here we again follow closely the lines of (Spekkens et al., 2009a). Let's define $P_{s,b}$ the procedure obtained by choosing a random x such that $x \cdot s = b$ and implementing P_x . For all measurement procedures that can be performed we have that

$$p(k|M, P_{s,b}) = \frac{1}{2^{n-1}} \sum_{x|x \cdot s=b} p(k|M, P_x) \quad (\text{E.5})$$

From this and (E.1) we get that

$$\forall s \in \text{Par}, \forall k, \forall M : p(k|M, P_{s,0}) = p(k|M, P_{s,1})$$

For any ontic state λ that is prepared by $P_{s,b}$ we will have that,

$$p(\lambda|P_{s,b}) = \frac{1}{2^{n-1}} \sum_{x|x \cdot s=b} p(\lambda|P_x)$$

and we again conclude that

$$\forall s \in \text{Par} : p(\lambda|P_{s,0}) = p(\lambda|P_{s,1}).$$

This means that, using Bayes rules again we get that the operational equivalences are also obey at the ontological level,

$$\forall s \in \text{Par} : \sum_{x|x \cdot s=1} p(P_x|\lambda) = \sum_{x|x \cdot s=0} p(P_x|\lambda)$$

We conclude from that the following: knowledge of the ontic state of the system would not improve the classical bound of $(1+n)/2n$ because the ontic state itself *also* does not provide parity information.