Non-classicality as a computational resource

Lorenzo Catani

A thesis submitted to

University College London

for the degree of

Doctor of Philosophy

Department of Physics and Astronomy
University College London

October 19, 2018

I, Lorenzo Catani confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Non-classicality as a computational resource

Lorenzo Catani

Doctor of Philosophy of Physics

University College London

Prof. Dan Browne, Supervisor

Abstract

One of the main questions in the field of quantum computation is where the quantum computational speed-up comes from. Recent studies in the field of quantum foundations have suggested which are the features to be considered as inherently non-classical. One of the major contributions in this direction comes from a result known as Spekkens' toy theory, which is a model built to reproduce quantum theory as a classical phase-space-inspired theory with restrictions on what an observer can know about reality. The model reproduces many of the features of quantum mechanics, but it does not reproduce non-locality and contextuality. In this thesis we first complete Spekkens' toy theory with measurement update rules and a mathematical framework that generalises it to systems of any finite dimensions (prime and non-prime). We also extend the operational equivalence between the toy theory and stabilizer quantum mechanics to all odd dimensions via Gross' Wigner functions. We then use the toy theory to represent the non-contextual and classically simulatable part of the computation in state-injection schemes of quantum computation where contextuality is a resource. In the case of qubits, we show that the subtheories of quantum mechanics represented in the toy model can achieve the full stabilizer theory via state-injection and we associate different proofs of contextuality to different injection processes. Stepping back from Spekkens' toy theory, we conclude by focusing on single system protocols that compute non-linear functions (similarly to the popular CHSH game) which show quantum advantages even in absence of non-locality and contextuality (in its standard notions). We analyse their performances (formalised in Bell's and Tsirelson's bounds) in relation to Landauer's principle, which associates entropic costs to irreversible computations, and to a new notion of contextuality for sequences of transformations.

Impact Statement

We are probably living in the golden age of the foundations of quantum mechanics. Many research groups originated all over the world, conferences are organised every year and researchers do not have to hide with embarrassment while studying what is the nature of quantum reality. This is due, in addition to the universally fascinating questions charaterising the field, to the concrete applications that have resulted from these studies. The main example is Bell's theorem, stating that quantum mechanics is incompatible with the intuitive local realism underlying classical physics. It can be stated as the violation of a simple inequality, that has been proven to guarantee stronger cryptographic security and to certify whether numbers are intrinsically random (crucial matter, for example, for gambling companies). Much has still to be discovered and other applications, in particular in the directions of developing more powerful computers, are around the corner.

The results of this thesis, mainly concerning contextuality – a notion originating from a similar result to Bell's theorem – contribute to the understanding of what are the resources that justify why quantum computers are more powerful than classical ones. As such, they are beneficial towards the big goal of developing a quantum computer. The impact of this discovery would be huge as, for example, universal quantum computers would be able to break current cryptographic techniques that guarantee the security of our financial transactions. We also deal with examples of quantum computational advantages that involve single systems only, thus extending the area of application and impact of our results. Moreover, some of our mathematical techniques, in particular when dealing with systems of non-prime dimensions, will hopefully provide a new methodology to treat models of computation based on more complex systems than the usual two-level systems (bits/qubits). We believe that extending the knowledge on

the resources for quantum technologies would be a benefit in the academical environment, as it would trigger new research projects, open the possibility of new faculty courses and be one of the leading topics in top-level scientific journals. Lastly, understanding how the world works in light of the mysterious and bizarre phenomena described by quantum mechanics is already an extremely frequent theme in popular-science magazines, and our results would fit well in that context too.

List of Publications and Preprints

The work presented in this thesis contains material from the following publications:

1. L. Catani and D. E. Browne, *Spekkens' toy model and its relationship with stabiliser quantum mechanics*, New journal of physics, vol. 96, no. 5, p. 052112, 2017.

2. L. Catani, N. D. Silva, and D. E. Browne, *Spekkens' toy model in quantum computation and contextuality as a resource*, IOS Press Amsterdam, SIF Bologna: accepted for publication in Proceedings of the International School of Physics Enrico Fermi, course 197, foundations of quantum theory, edited by E. M. Rasel, W. P. Schleich and S. Woelk ed., 2017.

Other preprints by the author are:

1. L. Catani and D. E. Browne, *State-injection schemes of quantum computation in Spekkens' toy theory*, arXiv:1711.08676 [quant-ph], 2017. Accepted for publication in Physical Review A.

2. L. Henaut, L. Catani, D. E. Browne, S. Mansfield, and A. Pappa, *Tsirelson's bound and Landauer's principle in a single-system game*, arXiv:1806.05624 [quant-ph], 2018.

Acknowledgments

Many people deserve to be thanked for allowing me to achieve this important goal.

First of all, I am indebted to my supervisor, Dan Browne, for his expertise, patience and enthusiasm. For motivating me with his illuminating ideas and subtle comments. For knowing when it was time to leave me the independence, both to fail and to create, and for encouraging me when I was stuck. I have always had the constant feeling that I could rely on him and I could have not asked for a better guide.

My examiners, Jonathan Barrett and Lluis Masanes, need to be thanked for their careful reading of this thesis. Their insightful comments significantly improved this document. Any remaining inaccuracies result from me.

Besides my supervisor and examiners, I would like to thank the inspiring figures that allowed me to undertake my placement in Canada last year and start stimulating research projects with them: Robert Raussendorf for his knowledge and always inviting me to be part of his research group and Robert Spekkens for feeding my passion for quantum foundations with extremely motivating and deep conversations. I am also sincerely grateful to their research groups and close collaborators for widening my research from various perspectives, in particular Juan Bermejo-Vega, Piers Lillystone, Hammam Qassim and David Schmid. I also thank Matt Leifer for giving me the possibility to join his group in California, allowing me to follow my passion for foundational questions and work on ideas that I find inspiring and promising.

I want to express my sincere gratitude to the CDT in its entirety, starting from cohort one – Andrew, Cameron, Carlo, Claudia, Johnnie, Josh, Sherif, Tim and Tom – a group of friends I shared my academic pains and joys with over the last four years. I cannot forget Lopa for always being there and ready to cheer us up, whatever the issue. I also thank my second

supervisor Simone Severini for the never banal conversations and his humanity.

I sincerely thank the lads in office C25. Despite sharing a cage with no natural lights and fresh air, I have to admit that it has been the best place to spend my PhD years. Every joke, lunch and secret Santa made my days sweeter than expected. Carlo, both Toms, Alex and Mike, you will be missed. The thanks has to be extended to the whole circle of friends in the physics department here at UCL and my research group for being part of a delightful environment in which to do research.

Lastly, I owe a debt of gratitude to my friends from home, here in London, the ones back in Italy, and, more importantly, my family for always supporting me. I feel privileged to have you at my side. You have always been present in the bad moments and ready to celebrate my achievements that would not have been possible otherwise.

Contents

List of Figures

Stands at the sea,
wonders at wondering: I
a universe of atoms
an atom in the universe.

R. P. Feynman, *The Value of Science (1955)*

Chapter 1

Introduction

1.1 Overview

Quantum mechanics is the most precisely tested theory in the history of science. It provides theoretical predictions that agree with experiments with unprecedented success[1] and, after about a century of tests, it has never been experimentally falsified. It originated at the beginning of the twentieth century to explain observations of the black-body radiation that were inconsistent with the predictions of classical physics [2] and the photoelectric effect [3]. It then developed as a theory of atomic physics [4] and eventually underpinned all the modern physics except from gravity (for now!) [5].

In the last few decades quantum mechanics has been applied to other fields, such as information and computer sciences. The novel features of the theory such as superposition, entanglement and non-locality, originally perceived as problematic obstacles for a full understanding of the theory, over time became intended as resources for information processing and computing tasks that show better performances than approaches based only on classical resources. The first idea of a computer that exploits quantum phenomena was suggested by Richard Feynman in 1982 [6] in a lecture titled "Simulating Physics with Computers" and three years later developed by David Deutsch, who formalised the theoretical structure of a universal quantum computer [7]. The field of quantum computation became very popular after the invention of

[1] Up to eleven significant figures in the case of the anomalous magnetic moment of the electron for testing quantum electrodynamics [1].

Shor's algorithm in 1994 [8], which showed that a quantum computer would allow one to factor large numbers exponentially more quickly than the best known classical algorithms, thus being able to break current cryptographic techniques that guarantee the security of our financial transactions [9].

Other algorithms showing quantum advantages have been designed [10–13] as well as other evidences of the quantum computational supremacy over classical computation [14], thus motivating the widespread belief that the dream of a quantum computing era will eventually come true: world-leading companies, like IBM, Google and Microsoft, have started to invest considerable amount of funds in this area, many research groups have originated all over the world and an increasing number of universities have created programmes to develop quantum technologies (thus allowing people like the current author to carry out research in this area). As the state of the art stands at the moment, several theoretical models for realising quantum computers have been formulated [7, 15, 16], but their experimental realisations still remain challenging [17, 18]. Nevertheless, technologies based on quantum mechanics for specific tasks are already available on the market, in particular concerning cryptography [19].

Despite the central attention that quantum computation has gained in recent years, we cannot still understand which physical principles underlie its power. One of the main questions in the field concerns, therefore, the features that might be responsible for the supposed quantum computational speed-up. In this work we aim to gain insights into this issue by exploiting results coming from recent studies in the field of foundations of quantum mechanics, whose main goal is to find an uncontroversial understanding of the reality described by quantum mechanics. Despite being extremely effective, the mathematical formulation of quantum mechanics [20–25] is indeed highly abstract and does not suggest any clear interpretation of how nature behaves. This has lead to the formulation of many possible interpretations of the theory [26] and there is still no consensus among scientists on which one to adopt. The idea is therefore to reformulate quantum theory with a more intuitive framework, possibly starting from physically meaningful principles, similarly to the case of the theory of special relativity, which is singled out by the requirements that the speed of light in a vacuum – the maximum speed of any interaction – is independent of the motion of all observers and that the laws of physics have to work the same in any inertial reference frame [27]. A possible way to achieve this and obtain a better

understanding of the theory is to identify the features that are inherently non-classical and make quantum theory special. One of the main contributions in this direction comes from a result known as Spekkens' toy theory [28, 29], a model built to reproduce quantum theory as a classical phase-space-based theory with restrictions on what an observer can know about the ontic state (identified with a phase space point) describing the reality of a system. This model has shown that almost all phenomena and protocols that were considered to be non-classical, like entanglement [30] and teleportation [31], can actually be reproduced in this classical-like model: approximately everything but Bell non-locality [32] and contextuality [33, 34], that therefore emerge as inherently non-classical features.

The main goal of this work is to shed light on the physical sources of the quantum computational speed-up. We do that by using notions of non-classicality developed in studies of quantum foundations. We can concisely sum up the main contributions of this work as follows. We first complete the formulation of Spekkens' toy theory and state, by using the tool of Gross' Wigner functions, its equivalence for odd-dimensional systems with a well-known subtheory of quantum mechanics called stabilizer quantum mechanics. We then exploit the toy theory to support the statement that contextuality is a resource for universal quantum computation (UQC) in state-injection schemes of quantum computation. We finally address more restricted computational scenarios that show quantum advantages to compute non-linear functions, where there is no presence of non-locality and contextuality (in its standard notions [33,34]), providing instead an analysis in terms of information erasure, as famously studied by Landauer.

Before proceeding with a more detailed description of the structure of the thesis, we think it is fundamental to clarify what we mean by *classical physics* and *quantum physics*, in order to make sense of terms like "non-classicality", "classical computation" and "quantum computation".

Classical physics is usually referred to as the theory that describes the physical world at the macroscopic scales (above the atomic scale). It acquires a precise meaning depending on the context it is considered in, and approximately it includes all the theories that were established before the advent of quantum mechanics (*i.e.* Newtonian/Lagrangian/Hamiltonian mechanics, Maxwell electrodynamics, thermodynamics). In this thesis, when we refer to classical physics we will usually be referring to Hamiltonian mechanics [35](formulated in the formalism of phase

space and symplectic geometry). The reason for this choice is that Hamiltonian mechanics, among the formulations of classical mechanics, is the natural example of a hidden variable model as developed first by Einstein [36], and later by Bell [37] and Kochen-Specker [33]. In Hamiltonian mechanics the hidden variables – the states of physical reality that (pre)exist even without observers and experiments performed on the system – are points in the phase space, *i.e.* specifications of the position and momentum of a system at a given time, and any property of the system can be obtained from these specifications. In the framework of hidden variable models of Einstein, Bell, Kochen and Specker, that defines our notion of classicality, assumptions like non-contextuality and locality are included. The latter are not compatible with quantum mechanics, as proven by the well-known Bell and Kochen-Specker no-go theorems [32,33]. The precise notion of hidden variable models, more generally addressed as *ontological models*, will be given in subsection 3.1.2 as well as the notion of non-contextuality. Bell locality will be treated in subsection 4.1.1. We recall that, briefly said, non-contextuality – in its original formulation – refers to the fact that the outcome of a measurement does not depend on which other set of compatible measurements we perform with it [33]. Locality refers to the fact that no instantaneous (faster-than-light) interaction can take place between distant ("space-like") separated events.

When talking about Spekkens' toy theory we will be more permissive with what a classical theory is, as we will endow the Hamiltonian formalism with a richer structure. We will add a restriction on which regions of the phase space represent states that can be known by an observer and this allows us to obtain features that are usually not considered classical, such as entanglement. We will also talk about classical computation when considering computation performed by using systems whose information is stored in bits (or more general dits), *i.e.* two(d)-level systems that cannot exhibit superpositions of logically exclusive states, and processed by logic gates that can generate any Boolean function of the bits (examples of universal sets of gates are AND,OR plus NOT gates, and the TOFFOLI gate). The final read-out just reveals the states of the processed bits. Classical computation will be contrasted with quantum computation, which is the computation performed using systems whose information is stored in qubits (or more general qudits) *i.e.* two(d)-level systems that can exhibit superpositions of logically exclusive states, processed by unitary reversible gates (examples of universal sets for

qubit gates are single qubit plus CNOT gates, and Hadamard plus Toffoli gates) and measured according to the formalism describing quantum measurements (reported below). We will also talk about classical simulations of quantum computations, meaning that the output statistics of the quantum computation can also be efficiently reproduced by a classical computer.

When referring to non-classicality, we therefore address features which do not arise in classical theories as defined above, even in cases where the classical theory is equipped with extra structures as in Spekkens' toy theory.

We can now provide the precise mathematical formulation of the theories of quantum and classical mechanics in terms of states, evolutions and measurements (see table 1.1). We follow the standard description of the postulates of quantum mechanics as described in [38] and we then formulate the analogue description in the classical case [35], thus comparing the two theories on the same ground. We assume the reader to be familiar with basic notions of linear algebra, vector spaces, symplectic geometry and phase spaces ([38] and [35, 39] are excellent references for these subjects).

Quantum mechanics [38].

- States. Associated to any isolated physical system is a Hilbert space $\mathcal{H}$ (complex vector space with inner product) known as state space of the system. The system is completely described by its density operator ρ (a positive trace one operator) acting on the state space of the system. We say that the system is in a pure state if its density operator is of the form $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a unit vector in the system's state space and $\langle\psi|$ is its transpose conjugate. The density operator of a mixed state can be written as $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where p_i denotes the probability of finding the system in the pure state $|\psi_i\rangle$. We recall that the inner product between two density operators in the Hilbert space is defined by their trace (note that density operators are Hermitian), $\mathrm{Tr}(\rho\sigma)$. The state space $\mathcal{H}$ of a composite physical system is the tensor product of the state spaces of the component physical systems, $\mathcal{H} = \bigotimes_i \mathcal{H}_i$.

- Evolutions. The evolution of a *closed* quantum system is described by a unitary transformation U, that acts on the state ρ as $U\rho U^\dagger$. We remind the reader that a unitary operator U is defined as $UU^\dagger = U^\dagger U = \mathbb{I}$, where $\mathbb{I}$ represents the identity operator. Unitary transformations are reversible, as the inverse $U^{-1} = U^\dagger$ always exists.

- Measurements. Quantum measurements are described by a collection $\{M_k\}$ of measurement operators acting on the state space of the system being measured. The index k refers to the possible outcomes of the measurement. The measurement operators satisfy the completeness equation $\sum_k M_k^\dagger M_k = \mathbb{I}$. If the outcome k occurs, the state of the system after measurement ρ' is given by $\rho' = \frac{M_k \rho M_k^\dagger}{p(k)}$, where ρ is the state before the measurement and $p(k)$ is the probability of obtaining the outcome k. For convenience let us define POVM (positive operator-valued measure) elements E_k as positive operators $E_k = M_k^\dagger M_k$, such that $\sum_k E_k = \mathbb{I}$. The probability $p(k)$ is given by the Born rule, $p(k) = \mathrm{Tr}(E_k\rho)$. If the POVM elements are orthogonal projectors $P_k \equiv E_k = E_k^2$ associated to each outcome k, then the measurement is known as a Von Neumann projective measurement (PVM). A PVM is associated to an observable, *i.e.* an Hermitian operator, O with spectral decomposition $O = \sum_k kP_k$, where k are the real eigenvalues of the projector P_k onto the eigenspace of O.

Classical mechanics [35].

- States. Associated to any physical system is a phase space Ω. The system is fully specified by a point λ in the phase space. It may be that the observer has partial knowledge of the physical state of the system and, in that case, the state of the system is associated to a probability distribution over the phase space. We recall that the phase space is a symplectic manifold, *i.e.* an even-dimensional manifold that locally has the structure of a symplectic vector space. The latter is a vector space equipped with a symplectic form, which means a bilinear real function $[\cdot,\cdot] : \Omega \times \Omega \to \mathbb{R}$ that is anti-symmetric, *i.e.* $[\lambda_1, \lambda_2] = -[\lambda_2, \lambda_1]$ and non-degenerate, *i.e.* $[\lambda_1, \lambda_2] = 0 \ \ \forall \lambda_2 \in \Omega$ implies $\lambda_1 = 0$. The state space Ω of a composite physical system is the Cartesian product of the individual phase spaces, $\Omega = \times_i \Omega_i$. We will mainly consider discrete phase spaces $\Omega = \mathbb{Z}_d^{2n}$, where d denotes the dimension and n is the number of systems we take into account. Every point λ in the phase space is denoted with respect to the conjugate variables position and momentum, $\lambda = (x_1, \ldots, x_n, p_1, \ldots, p_n)$. The standard symplectic form in $\Omega = \mathbb{Z}_d^2$ is $[\lambda_1, \lambda_2] = x_1 p_2 - x_2 p_1$, where $\lambda_1 = (x_1, p_1)$ and $\lambda_2 = (x_2, p_2)$.

- Evolutions. The physical evolution of a classical system is described by a symplectic diffeomorphism in the phase space $S : \Omega \to \Omega$. This is equivalent to state that the physical evolutions satisfy Hamilton equations [39]. We remind the reader that a symplectic transformation is a linear transformation that preserves the symplectic form, *i.e.* $[S(\cdot), S(\cdot)] = [\cdot,\cdot]$. A diffeomorphism is a map between manifolds which is differentiable and has a differentiable inverse.

- Measurements. Classical measurements have no special role in classical mechanics, unless we consider probability distributions for which measurement update rules are needed, and they are treated as any other physical evolution. The process of obtaining the outcome of a measurement just consists of revealing the value of the property being measured without disturbing the system and is therefore represented as a real (or integer) function on the phase space, $o : \Omega \to \mathbb{R}$.

	CLASSICAL MECHANICS	QUANTUM MECHANICS
STATES	Probability measures on phase space.	Density operators in Hilbert space.
EVOLUTIONS	Symplectic diffeomorphisms.	Unitaries.
OBSERVABLES	Real functions.	Hermitian operators.

Figure 1.1: **Classical and quantum mechanics.** States, evolutions and observables in classical and quantum mechanics. In order to avoid confusions, we here point out that in this work we do not consider evolutions of *open* quantum systems, that are described by completely-positive trace-preserving maps.

We conclude by also defining the notion of entanglement, arising in quantum mechanics, that is going to be useful in what follows. We define *entangled* quantum states – for simplicity we assume the bipartite scenario – as states ρ_{AB} that cannot be written as separable states, *i.e.* as convex mixtures of the form $\sum_i c_i \rho_A \otimes \rho_B$, where ρ_A and ρ_B denote states of the two parties A and B, respectively, and $c_i \in [0,1]$ is such that $\sum_i c_i = 1$.

1.2 Structure and outline of the thesis

The thesis is structured as follows:

- Chapter 2 primarily focuses on Spekkens' toy theory and overall presents the results cointained in [40], which is a joint work with Dan Browne. We start by introducing the original framework of the toy theory as developed firstly in [28] and later in [29]. More precisely, we rigorously define ontic and epistemic states, measurement observables and the rule to obtain the outcomes of a measurement observable on a system in a given state. We then describe stabilizer quantum mechanics [41], a subtheory of quantum

mechanics that only allows eigenstates of tensors of Pauli operators, Clifford unitaries and Pauli measurement observables. First derived for the purpose of constructing error-correcting codes, it now plays a role in many areas of quantum information theory. We also illustrate the tool we use to relate stabilizer quantum mechanics with Spekkens' toy theory: Wigner functions. These are a way of recasting quantum mechanics in the formalism of the phase space [42]. Wigner functions are defined as quasi-probability representations, which roughly means probability distributions with (sometimes negative) real values. In this chapter we mainly focus on the Wigner functions that are always non-negative – thus interpreted as actual probability distributions – for stabilizer states and observables in *odd* dimensions, also known as Gross' Wigner functions [43].

The original framework of the toy theory is constructed only for prime dimensional and infinite dimensional systems and formal rules for the update of states after measurements have not been written down. We here remedy this by deriving measurement update rules and extending the framework to derive models in all dimensions, both prime and non-prime. The distinction between the two cases has its roots in the mathematical difference between the set of integers modulo d, for d prime and non-prime. In the latter case $\mathbb{Z}_d$ is not a field, as the inverse of a number does not always exist. This fact implies that there are some problematic observables in the non-prime case (in the sense that they encode some degeneracy in the spectrum of possible outcomes) that need to be written in terms of the non-problematic ones in order to provide rules for the updating of states after measurements. Having developed the toy theory for all finite dimensions, we then focus on the general odd-dimensional case. In [29] it is proven that the toy theory is operationally equivalent to stabilizer quantum mechanics in odd prime dimensions. "Operationally equivalent" means that the two theories provide the same statistics of outcomes, given certain states, transformations and measurements. By exploiting Gross' theory of discrete Wigner function [43], that unlike most other studies concerns both prime and non-prime cases, and the now complete and general toy theory, we extend this equivalence to all odd dimensions (see figure 1.2). We also express the already found update rules in terms of Wigner functions and we use them to depict the elegant analogies between these three theories. The chapter ends with a discussion of the

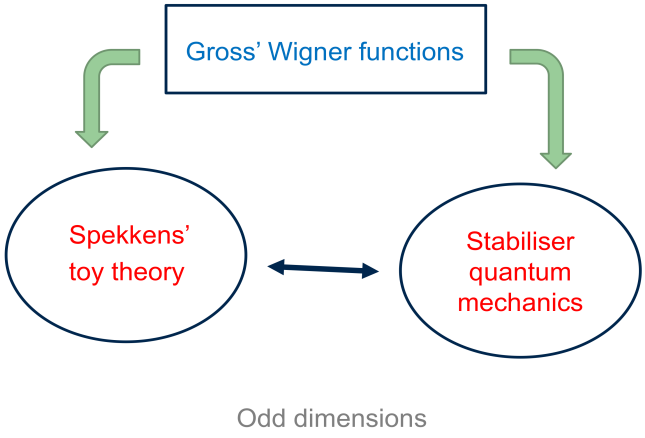


Figure 1.2: **Spekkens' toy theory, stabilizer quantum mechanics and Gross' Wigner functions in odd dimensions.** Spekkens' toy theory and stabilizer quantum mechanics are operationally equivalent theories in odd dimensions via Gross' Wigner functions.

possible applications of our achievements and a summary of the main results.

The question of which part of quantum mechanics is operationally equivalent to the toy theory in the even dimensional cases is still unanswered and motivates the subsequent chapter. It particularly applies to the dimension-two case, as nowadays qubits are the fundamental bricks in all the implementable models of quantum computation. A result that connects the toy theory and stabilizer quantum mechanics and analogue to the odd case cannot hold because of the unavoidable negativity of any Wigner function representation of qubit stabilizer quantum mechanics [44, 45]. This fact can also be seen as an expression of the contextuality present in qubit stabilizer quantum mechanics [46–48], not appearing in the odd dimensional qudit case.

- Chapter 3 predominantly illustrates an application of the toy theory for state-injection schemes of quantum computation, where contextuality is a resource and mainly reports the material contained in [49], which is a joint work with Nadish De Silva and Dan Browne, and [50], which is a joint work with Dan Browne.

  The background section begins with a review of state-injection schemes of quantum computation [51], which are one of the leading models of fault tolerant universal quantum

computation. These schemes are composed of a "free" part,[2] which consists of quantum circuits that are efficiently simulatable by a classical computer – usually stabilizer circuits – and by magic resources – which are usually distilled from many copies of noisy states through magic state distillation [53] – that boost the computation to universal. Recent results have shown that the contextuality possessed solely by the magic state is a necessary resource for universal quantum computation [54–57]. As this chapter focuses on the role of contextuality in quantum computation, we accurately define the standard notion of contextuality [33] and its state-independent [47] and dependent manifestations [46]. These are the main versions found in the literature studying contextuality as a resource for quantum computation. We also briefly talk about a generalised notion of contextuality that extends the original notion of contextuality – which deals only with projective measurements – also to preparations, transformations and unsharp measurements [34] and its applications. We then describe the results on contextuality as a resource for state injection schemes of quantum computation, both regarding qudits of odd prime dimensions [54], where the free part is, as usual, stabilizer quantum mechanics, rebits (qubits with real density matrices) [55] and qubits [56, 57], where the free parts need to be more restrictive than stabilizer quantum mechanics in order to avoid manifestations of contextuality. Once again, all these results use the framework of the Wigner functions, which we treat in all its generality, following the work of Raussendorf *et al* in [57].

In order to see why Spekkens' toy theory well relates with the aforementioned works, we need to define and characterise the subtheories of the toy theory, in particular in the dimension two case where the full toy theory does not consistently match with any subtheory of quantum mechanics, which are operationally equivalent to subtheories of stabilizer quantum mechanics. We define subtheories of Spekkens' toy theory compatible with subtheories of stabilizer quantum mechanics as closed subtheories of quantum mechanics whose states and measurements are non-negatively represented by covariant Wigner functions. The property of covariance guarantees the preservation of the symplectic form, which is crucial in the definition of the toy theory and its epistemic restriction. We use Spekkens' subtheories to represent the non-contextual free

[2]The jargon adopted derives from the literature on resource theories [52].

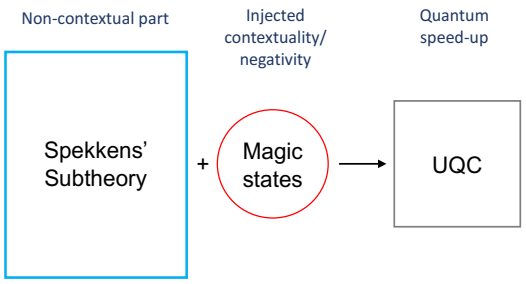


Figure 1.3: **Spekkens' theory and state-injection schemes with contextuality as a resource for quantum universality.** Spekkens' subtheories can be used to represent the non-contextual free part of state-injection schemes of quantum computation where contextuality is a resource injected through the magic states.

part of the known examples of state-injection schemes in [54, 55] where contextuality arises as a resource. These can be unified in the following framework (figure 1.3): $Spekkens'\ subtheory\ +\ Magic\ state(s)\ \rightarrow\ UQC$, as shown in figure 1.3. Furthermore, we prove that, in the case of qubits, stabilizer quantum mechanics can be obtained from a Spekkens' subtheory via state-injection, as all the objects not contained in the Spekkens' subtheory, namely non-covariant Clifford gates, can be state-injected via a circuit made of objects in the Spekkens' subtheory. This shows that within Spekkens' subtheories we possess the toolbox to perform state-injection of every object outside of them and suggests that there is no need to use bigger subtheories to reach universal quantum computation via state-injection. Before closing up and introducing the future challenges, we show a novel scheme of computation suggested by our approach which is based on the injection of particular states – CCZ states – and we also relate different proofs of contextuality to different state injections of non-covariant gates.

The results supporting the statement that contextuality is responsible for the quantum computational speed-up are not fully general, as they cannot be extended to any model of quantum computation. In the next chapter we therefore look at other scenarios showing quantum advantages where contextuality is not present, at least in its standard notions.

- Chapter 4 describes a single system protocol that computes non-linear functions and its sources of non-classicality. It mainly treats the content of [58], which is a joint work with Luciana Henaut, Dan Browne, Shane Mansfield and Anna Pappa.

The protocol considered in this chapter is inspired by the Clauser-Horne-Shimony-Holt (CHSH) game [59], which constitutes a way of recasting the celebrated Bell's scenario [32] into a two-player game for which quantum strategies can provide an advantage. In the CHSH game the probability of success of strategies involving only classical resources is bounded by a value, 0.75, known as Bell's bound; on the other hand, allowing the players to access quantum resources leads to a probability of success bounded by a value, $\cos^2 \frac{\pi}{8}$, known as Tsirelson's bound [60]. We describe the standard CHSH game, with a special focus on non-locality, which is the non-classical feature usually employed as the physical justification to why the quantum strategies can perform better than the Bell bound. Non-local correlations are strong correlations that cannot be found in any classical theory and are proven to be a resource for quantum technologies, such as device independent cryptography [61–63]. We treat the generalisation of the CHSH game, phrased in arithmetics modulo 2, to arithmetics modulo q, where q is any prime integer, about which tight Bell's and Tsirelson's bound have not been found yet [64–67]. We also discuss other protocols that are strictly related to the CHSH game, such as quantum random access codes [68] and parity oblivious multiplexing [69]. We conclude the background section illustrating Landauer's principle [70], that associates entropic costs to irreversible computations. This will serve us to analyse our protocol, where irreversible computation assumes a crucial role.

The protocol we present in this chapter, that we call CHSH* game, consists of an initial system in a fixed state, two gates controlled by classical bits and a fixed measurement. We consider the game of choosing gates to maximize the probability of computing the product (modulo 2) of the input bits. For qubit systems subject to unitary gates and projective measurements, we demonstrate that any strategy in our game can be mapped to a strategy in the CHSH game, which implies that Tsirelson's bound also holds in our setting. We then show that the optimal success probability depends on the set of operations allowed to the player (*e.g.* reversible versus irreversible and Clifford versus non-Clifford), the quantum or classical nature of the system and the dimension of the system. We also briefly describe the expected results when considering our scheme in arithmetics modulo q, where q is an arbitrary prime integer, by studying the case $q = 3$.

Finally, since the single-system protocol restricts the degrees of freedom to gates only, we analyse the bounds obtained in light of Landauer's principle, showing the entropic costs of the erasure associated with the game, which is a powerful tool for increasing the winning probability. This shows a connection between the reversibility in fundamental operations embodied by Landauer's principle and Tsirelson's bound, that arises from the restricted physics of a unitarily-evolving single-qubit system. Some considerations on the presence of a new notion of contextuality [71] in certain quantum strategies for the protocol are also discussed.

- In the Summary and Outlook chapter we recap the results shown throughout the dissertation, we illustrate their significance and we discuss the future challenges. To say it concisely, we complete Spekkens' toy model and study its relation with quantum mechanics, we use it to support the statement that contextuality is a resource for universal quantum computation in state-injection schemes of quantum computation and we analyse the sources of non-classicality in a protocol showing quantum advantages for computing non-linear functions, where neither standard contextuality nor non-locality are present. Among other implications, these results suggest that the source of quantum computational speed-up is scenario-dependent, as there are no resources which are strictly necessary in any context that provides quantum advantages. The inherent non-classical notions arising from the field of quantum foundations, like contextuality, do not always match the notions of non-classicality we have in quantum computation, *e.g.* non-efficient classical simulatability. Hence, the question of which form of contextuality is useful in this sense is still pending. It may be the case that we have to come up with a novel and inclusive notion of non-classicality, that manifests itself in different forms depending on the computational scenario.

Chapter 2

Spekkens' toy theory in all dimensions and its relationship with stabilizer quantum mechanics

A long tradition of research, starting from the famous "EPR paper" [36], has consisted of analysing quantum theory in terms of hidden variable models, with the aim of obtaining a more intuitive understanding of it. This has led to some crucial results in foundation of quantum mechanics, namely Bell's and Kochen-Specker's no-go theorems [32, 33]. Nowadays a big question is whether to interpret the quantum state according to the ontic view, *i.e.* where it completely describes reality, or to the epistemic view, where it is a state of incomplete knowledge of a deeper underlying reality which can be described by the hidden variables. In 2005, Robert Spekkens [28] constructed a non-contextual hidden variable model to support the epistemic view of quantum mechanics. The aim of the model was to replace quantum mechanics by a hidden variable theory with the addition of an epistemic restriction (*i.e.* a restriction on what an observer can know about reality). The first version of the model [28] was developed in analogy with qubits, with two-outcome observables. Despite the simplicity of the model, it was able to support many phenomena and protocols that were believed to be intrinsically quantum mechanical (such as entanglement and dense coding). Spekkens' toy model has influenced much research over the years: *e.g.* people provided a new notation for it [72], studied it from the

categorical point of view [73], used it for quantum protocols [74], exploited similar ideas to find a classical model of one qubit [75], and tried to extend it in a contextual framework [76]. Moreover, Spekkens' toy model addresses many key issues in quantum foundations: whether the quantum state describes reality or not, finding a derivation of quantum theory from intuitive physical principles and classifying the inherent non-classical features.

A later version of the model [29], which we will call Spekkens' Theory (ST), introduced a more general and mathematically rigorous formulation, extending the theory to systems of discrete prime dimension, where *dimension* refers to the maximum number of distinguishable measurement outcomes of observables in the theory, and continuous variable systems. Spekkens called these classical statistical theories with epistemic restrictions as *epistricted statistical theories*. By considering a particular epistemic restriction that refers to the symplectic structure of the underlying classical theory, the *classical complementarity principle*, theories with a rich structure can be derived. Many features of quantum mechanics are reproduced there, such as Heisenberg uncertainty principle, and many protocols introduced in the context of quantum information, such as teleportation. However, as an intrinsically non-contextual theory, it cannot reproduce quantum contextuality (and the related Bell non-locality)[1], which therefore arises as the signature of quantumness. Indeed, for odd prime dimensions and for continuous variables, ST was shown to be operationally equivalent to sub-theories of quantum mechanics, which Spekkens called quadrature quantum mechanics.

In the finite dimensional case quadrature quantum mechanics is better known as *stabilizer quantum mechanics* (SQM). The latter, as already mentioned in the introduction, is a sub-theory of quantum mechanics developed for the description and study of quantum error correcting codes [41], but subsequently playing a prominent role in many important quantum protocols. In particular, many studies of quantum contextuality can be expressed in the framework of SQM, including the GHZ paradox [46] and the Peres-Mermin square [47, 48] (see subsection 3.1.2). This exposes a striking difference between odd and even dimensional SQM. Even-dimensional SQM contains standard examples of quantum contextuality while odd-dimensional SQM exhibits no contextuality at all, necessary for its equivalence with Spekkens' Theory. While developed for qubits, SQM was rapidly generalised to systems of arbitrary di-

[1]Contextuality and non-locality will be precisely defined in subsection 3.1.2 and 4.1.1 respectively.

mension, [41]. However, for non-prime dimensions SQM remains poorly characterised and little studied (progress in this was recently reported in [77]).

Quasi-probability representations, such as the Wigner function, have been an important tool for the description of quantum systems for many years. Recently, negative quasi-probability representations and contextuality have been shown to have an important resource character in quantum computation [54–57, 78–83]. In particular, in certain fault tolerant quantum computation schemes, SQM plays a central role as both the set of operations that can be directly fault tolerantly realised, and the part of the computation which is efficiently simulatable by a classical computer [84]. As introduced in the previous section, such computation can be then boosted to quantum universality by "injecting" a resource state, known as a magic state. In the case of odd prime dimensions, Howard et al. [54] showed that the contextuality of the injected state is necessary for reaching universal quantum computation. Other similar results have been found in the case of qubits, at the cost of considering smaller subtheories than SQM for the classically simulatable non-contextual part of the computation [55–57]. The operational equivalence between SQM and ST in odd dimensions motivates the study of the role of ST in this research field, which is the main topic treated in the next chapter.

In spite of the importance of Spekkens' Theory, there remain some important aspects of it which have not yet been characterised and studied. First of all, all prior work on ST have only considered systems where the dimension is prime. Furthermore, while Spekkens' recent work strengthens the mathematical foundations of the model [29], one key part of the theory has not yet been described in a general and rigorous way. These are the measurement update rules, the rules which tell us how to update a state after a measurement has been made. In prior work, these rules, and the principles behind them have been described but not formalised.

In this chapter, we complete this step, deriving a formal description of the measurement rules for prime-dimensional ST. Having done so, we now have a fully formal description of the model, which can be used as a basis to generalise it. We do so, generalising the framework from prime-dimensions to arbitrary dimensions and finding that it is the measurement update rule, where the richer properties of the non-prime dimension can be seen, which provides the key to this generalisation.

Having developed ST for all finite dimensions, we then focus on the general odd-dimensional

case, and prove that in all odd-dimensional cases Spekkens' Theory is equivalent to Stabilizer Quantum Mechanics. The bridge between SQM and ST is given by Gross' theory of discrete Wigner functions (GT) [43]. Unlike most other studies, Gross' treatment considered both prime and non-prime cases in its original formulation.

To summarise the contributions of this chapter, we provide a compete formulation of ST in *all* discrete dimensions, even and odd, endowed with the update rules for sharp measurements both for prime and non-prime dimensional systems. We extend the equivalence between ST and SQM via Gross' Wigner functions to *all* odd dimensions, and find the measurement update rules also for the Wigner functions. The above equivalence allows us to shed light onto a complete characterisation of SQM in non-prime dimensions. Finally, the elegant analogy between the three theories in odd dimensions: ST, SQM and GT, is depicted in terms of their update rules.

The remainder of the chapter is structured as follows. In the section 2.1, we first precisely and concisely describe the original framework of Spekkens' theory, in particular we define ontic and epistemic states, observables and the rule to obtain the outcome of the measurement of an observable given a state. We then describe more extensively stabilizer quantum mechanics and we define Wigner functions, with a particular focus on the ones used for systems of discrete odd dimensions (Gross' theory). In section 2.2 and 2.3 we state and prove the update rules in Spekkens' theory respectively for prime and non-prime dimensional systems. We prove these in two steps: first considering the case in which the state and measurement commute, and then the more general (non-commuting) case. The mathematical difference between the set of integers modulo d, for d prime and non-prime, results in having two levels of observables: the fundamental ones - the fine graining observables - and the ones that encode some degeneracy - the coarse-graining observables. The latter are problematic and are only present in the non-prime case. This is the reason why we need a different formulation in the two cases. The update rules for the coarse graining observables will need a step in which the coarse-graining observables are written in terms of fine graining ones. In section 2.4 we state the equivalence of ST and SQM via Gross' Wigner functions in all odd dimensions. We also express the already found update rules in terms of Wigner functions and we use them to depict the elegant analogies between these three theories. The chapter ends with a discussion of the possible applications of our achievements in section 2.5 and with a summary of the main results in section 2.6.

2.1 Background

This section does not contain original material. The references that have been used will be specified in the corresponding subsections.

2.1.1 Spekkens' theory

We start by reviewing and introducing Spekkens' theory for prime-dimensional systems. We take a slightly different approach to [28] and [29]. ST is a hidden variable theory, where the hidden variables are points in a phase space. The state of the hidden variables is called the *ontic state*. In Spekkens' model the ontic state is hidden and can never be known by an experimenter. The experimenter's best description of the system is the *epistemic state*, representing a probability distribution over the points in phase space.

For a single d-dimensional system, a phase space can be defined via the values of two conjugate fiducial variables, which we label X and P, in analogy to position and momentum. X and P can each take any value between 0 and $d-1$, and a single ontic state of the system is specified by a pair (x, p), where x is the value of X and p is the value of P. This phase space is equivalent to the space $\mathbb{Z}_d^2$. In figures 2.1 and 2.2 examples of states of one and two bits ($d = 2$), and one trit ($d = 3$) are depicted, where X and P are represented by the rows and columns in the phase spaces $\mathbb{Z}_2$, $\mathbb{Z}_2^2$ and $\mathbb{Z}_3$, respectively.

A collection of n systems is described by n pairs of independent conjugate variables X_j and P_j, with $j \in 0, \ldots, n-1$ a label indexing the systems. The phase space, denoted by Ω, is simply the cartesian product of single system phases spaces and thus $\Omega \equiv (\mathbb{Z}_d)^{2n}$.[2]

The ontic state of the n-party system represents a set of values for each fiducial observables X_j and P_j. In other words, an ontic state is denoted by a point in the phase space $\lambda \in \Omega$. We call X_j and P_j observables because they correspond to measurable quantities, and assume that these observables are sufficient to uniquely define the ontic state. We can refer to Ω as a vector space where the ontic states are vectors (bold characters) whose components (small letters) are

[2]The dimension d is any positive number, and we will not, in general, restrict it to odd or even, prime or non-prime, unless specified.

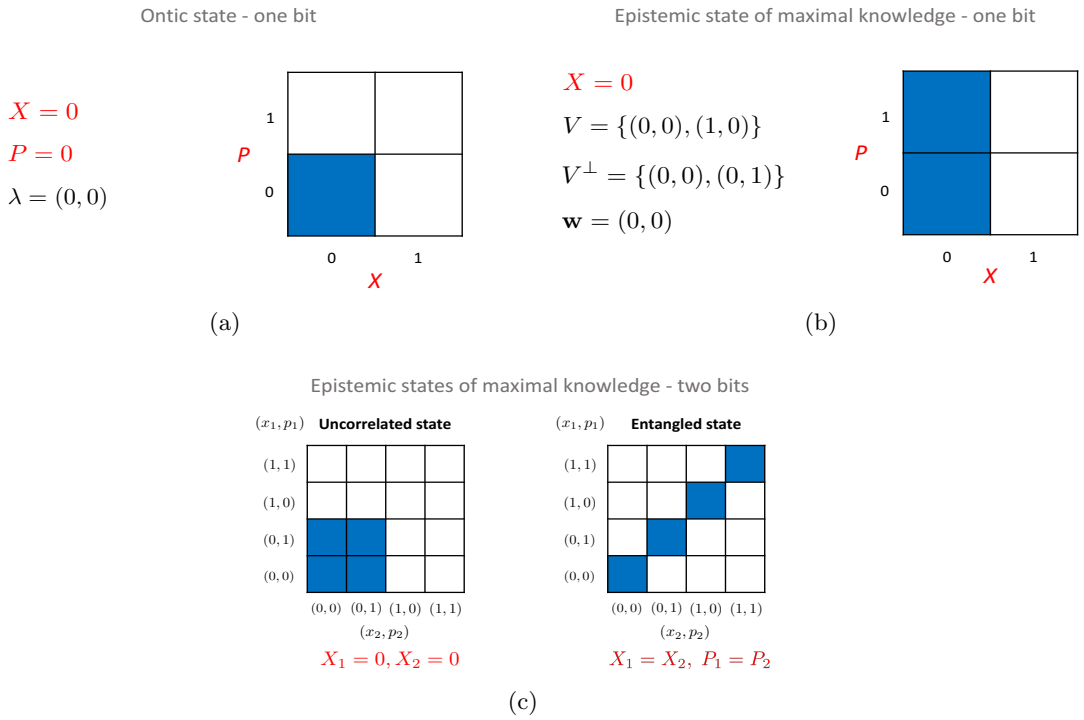


Figure 2.1: **Spekkens' toy states of one and two bits.** The figures 2.1a and 2.1b above show the elementary system of Spekkens' theory in two dimension: the bit. One possible ontic state of one bit is shown in 2.1a, where the observer both knows $X = 0$ and $P = 0$, so $\lambda = (0, 0)$. The epistemic restriction - classical complementarity principle - in this case corresponds to saying that at maximum the observer has "half" of the knowledge about the ontic state. For example a possible epistemic state is shown in figure 2.1b, where the observer only knows the variable $X = 0$, so $V = span\{(1, 0)\}$ and $\mathbf{w} = (0, 0)$. In this case the epistemic state $X = 0$ of one bit can be seen as the analogue of the quantum state $|0\rangle$ of one qubit. Figure 2.1c shows two kinds of two-bits epistemic states of maximal knowledge. The state on the left is a non-correlated state ($X_1 = 0 = X_2$), indeed we have the knowledge of the states of the individual subsystems, while the state on the right ($X_1 = X_2$ and $P_1 = P_2$) is perfectly correlated (*i.e.* entangled), indeed it would be impossible to know the states of the individual subsysytems, but we know exactly the correlation between them (in the case above we know that they have the same ontic states). This trade-off in choosing if knowing the correlation or the states of the individual subsystems is something which is not present in any purely classical theory.

the values of the fiducial variables:

$$\boldsymbol{\lambda} = (x_0, p_0, x_1, p_1, \ldots, x_{n-1}, p_{n-1}). \tag{2.1}$$

Not only are the fiducial variables important for defining the state space, they also generate

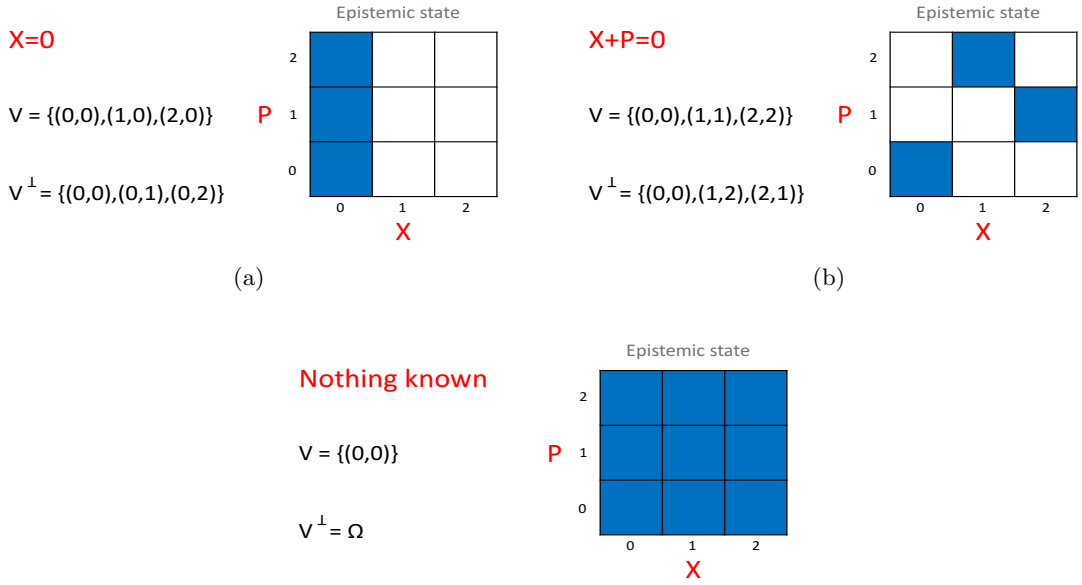


Figure 2.2: **Spekkens' epistemic states of one trit.** In the figures above we consider the case of one trit and we find the isotropic subspaces V and $V^{\perp}$ and the corresponding Spekkens epistemic state. In these cases the observables(linear functionals) are always of the form $aX+bP=0$, where $a,b\in\mathbb{Z}_3$. Moreover in the above examples we assume $\mathbf{w}=\mathbf{0}$. In figure 2.2a the observer only knows $X=0$ and this implies that the generator of V is $\mathbf{\Sigma}=(1,0)$. The subspace $V^{\perp}$ can be simply calculated from V by definition. In figure 2.2b the observer only knows that $X+P=0$ and this implies the generator of V to be $\mathbf{\Sigma}=(1,1)$. In figure 2.2c nothing is known. The subspace V is generated by $\mathbf{\Sigma}=(0,0)$ only. Here $V^{\perp}$ coincides with the whole phase space Ω. Note that it is not possible to have $V^{\perp}=(0,0)$, because this would correspond to have the knowledge of the ontic state.

the set of all general observables in the theory. A generic observable, denoted by Σ, is defined by any linear combination of fiducial variables:

$$\Sigma=\sum_m(a_mX_m+b_mP_m), \tag{2.2}$$

where $a_m,b_m\in\mathbb{Z}_d$ and $m\in 0,\dots,n-1$. The observables inhabit the dual space Ω^*, which is isomorphic to Ω itself. Therefore we can define them as vectors, in analogy with ontic states,

$$\mathbf{\Sigma}=(a_0,b_0,a_1,b_1,\dots,a_{n-1},b_{n-1}). \tag{2.3}$$

The formalism provides a simple way of *evaluating* the outcome σ of any observable measurement Σ given the ontic state λ, *i.e.* by computing their *inner product*:

$$\sigma = \mathbf{\Sigma}^T \boldsymbol{\lambda} = \sum_j (a_j x_j + b_j p_j), \tag{2.4}$$

where all the arithmetic is over $\mathbb{Z}_d$.

Spekkens' theory gains its special properties and, in particular, its close analogy with stabilizer quantum mechanics, via the imposition of an *epistemic restriction*, a restriction on what an observer can know about the ontic state of a system. The observer's best description is called the *epistemic state*, which is represented by a probability distribution $p(\lambda)$ over Ω (figure 2.2).

The epistemic restriction of ST is called *classical complementarity principle* and it states that two observables can be simultaneously measured only when their Poisson bracket is zero. This is motivated by stabilizer quantum mechanics, since it captures the condition for two observables in there to *commute*. We shall adopt the quantum terminology and notation here, and say that if the Poisson bracket between two observables is zero they commute, *i.e.* $[\cdot, \cdot] = 0$. This can be simply recast in terms of the *symplectic inner product*:

$$[\Sigma_1, \Sigma_2] \equiv \mathbf{\Sigma_1}^T J \mathbf{\Sigma_2} = 0, \tag{2.5}$$

where $J = \bigoplus_{j=1}^{n} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_j$ is the usual invertible matrix used in symplectic geometry. Note that each observable Σ_j partitions Ω into d subsets, each of the form $(span\{\Sigma_j\})^{\perp} + w$, where w is any ontic state such that $\mathbf{\Sigma_j}^T \mathbf{w} = \sigma_j$, and σ_j takes values in $\{0, 1, \ldots, d-1\}$, one for each subset.

Let us now consider sets of variables that can be jointly known by the observer. Such variables commute, and represent a sub-space of Ω known as an *isotropic subspace*. We denote the subspace of the known variables as $V = span\{\Sigma_1, \ldots, \Sigma_n\} \subseteq \Omega$, where Σ_i denotes one of the generators (commuting observables) of V.

Sets of known commuting variables are important as these define the epistemic states within the theory. In particular, we can define an epistemic state by the set of variables V that are

known by the observer and also the values $\sigma_1, \ldots, \sigma_n$ that these variables take.

This means that $\mathbf{\Sigma}_j^T \cdot \mathbf{w} = \sigma_j$, where $w \in V$ is an ontic state that evaluates the known observables. We will call w a *representative ontic state* for the epistemic state. More precisely we can state the following theorem.

Proposition 1. *The set of ontic states consistent with the epistemic state described by* $(V, \mathbf{w})$ *is*

$$V^\perp + \mathbf{w}, \tag{2.6}$$

where the perpendicular complement of V *is, by definition,* $V^\perp = \{a \in \Omega \mid \mathbf{a}^T\mathbf{b} = 0 \; \forall \; b \in V\}$.

Proof. Let us start by considering the set of ontic states λ such that $\mathbf{\Sigma}_j^T \boldsymbol{\lambda} = 0 \; \forall j$. By definition of perpendicular complements, the ontic states λ belong to $V^\perp$. If we consider an ontic state w such that $\mathbf{\Sigma}_j^T \mathbf{w} = \sigma_j$, then $\mathbf{\Sigma}_j^T(\boldsymbol{\lambda} + \mathbf{w}) = \sigma_j$. Therefore the ontic states consistent with the epistemic state associated to $(V, \mathbf{w})$ are the ones of the kind $\lambda + w$, *i.e.* the ones belonging to $V^\perp + \mathbf{w}$. □

Note that the presence of $\mathbf{w} \neq \mathbf{0}$ simply implies a translation, that is why we can also call it *shift vector*.

By assumption the probability distribution associated to the epistemic state $(V, \mathbf{w})$ is uniform (indeed we expect all possible ontic states to be equiprobable), so the probability distribution of one of the possible ontic states in the epistemic state $(V, \mathbf{w})$ is

$$p_{(V,\mathbf{w})}(\lambda) = \frac{1}{N}\delta_{V^\perp + \mathbf{w}}(\lambda), \tag{2.7}$$

where the delta is equal to one only if $\lambda \in V^\perp + \mathbf{w}$ (note this means that the theory is a *possibilistic* theory) and N is a normalization factor. For epistemic states of maximal knowledge about the ontic state, the normalization N is equal to d^n. Like in quantum theory, duality in the description of states and measurements characterises ST. This means that we can represent the elements of a sharp measurement Π in an epistemic-state way, $(V_\Pi, \mathbf{r})$, where we can go from one element of the measurement to the other by simply shifting the representative ontic vector $\mathbf{r}$ (see figure 2.3).

The aim of Spekkens' theory is to show that epistemic states and measurements in the theory

are the analogue of quantum states and measurements in quantum theory. The analogue of unitary evolutions in quantum theory has to correspond to the allowed transformations in the toy theory. More precisely the allowed transformations are the ones that preserve the classical complementarity principle (the symplectic inner product), *i.e.* the symplectic affine transformations G in the phase space (in general a subset of the permutations in the phase space):

$$G(\lambda) = S\boldsymbol{\lambda} + \mathbf{a}, \tag{2.8}$$

where S is a symplectic matrix and $\mathbf{a} \in \Omega$ a translation vector. In figure 2.4 we report the table present in [29] that shows the achievements of ST in terms of the phenomena of quantum mechanics that it can and cannot reproduce. We refer again to figures 2.1a and 2.1b, that picture the notions defined so far in the two dimensional case. The notion of entanglement is depicted in figure 2.1c. Figure 2.2 provides three examples of states of one trit.

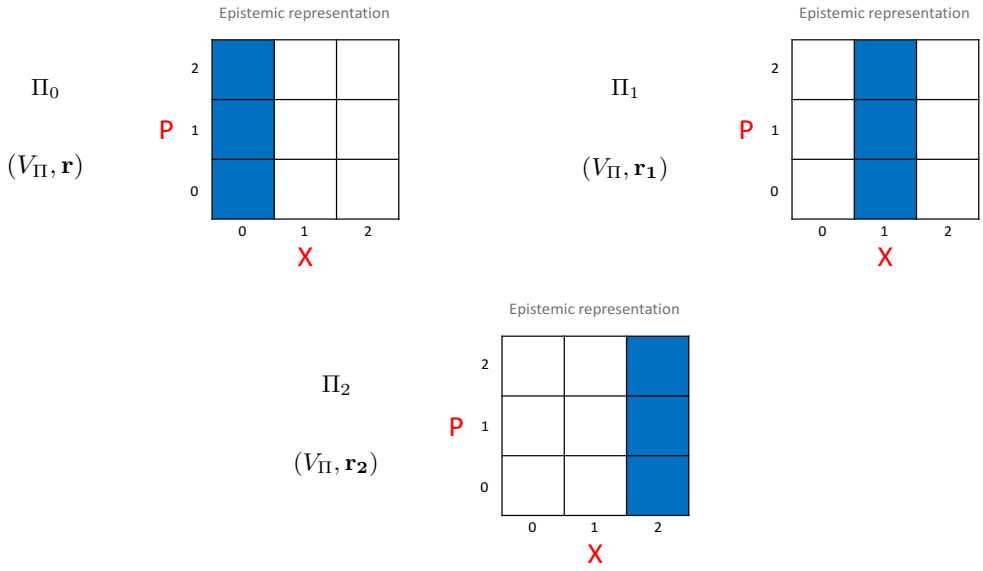


Figure 2.3: **Epistemic representation of a measurement.** The elements of the measurement Π can be represented as epistemic states. This duality is present also in quantum theory. The elements of the measurement Π_0, Π_1, Π_2 can be thought as the analogue of the projectors $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |2\rangle\langle 2|\}$. We can always go from one element to the other by shifting the representative ontic vector. In the above case we can go, for example, from Π_0 to Π_1 by adding to $\mathbf{r} = (0,0)$ the vector $(1,0)$, thus obtaining $\mathbf{r}_1 = (1,0)$. The measurement represented in the figure can be interpreted as asking the question "what is the value of the variable X?" about the ontic state of the system.

Phenomena arising in Spekkens' theory	Phenomena not arising in Spekkens' theory
Noncommutativity	Bell inequality violations
Coherent superposition	Noncontextuality inequality violations
Collapse	Computational speed-up (if it exists)
Complementarity	Certain aspects of items on the left
No-cloning	
No-broadcasting	
Interference	
Teleportation	
Remote steering	
Key distribution	
Dense coding	
Entanglement	
Monogamy of entanglement	
Choi-Jamiolkowski isomorphism	
Naimark extension	
Stinespring dilation	
Ambiguity of mixtures	
Locally immeasurable product bases	
Unextendible product bases	
Pre and post-selection effects	
Quantum eraser	
And many others...	

Figure 2.4: **Achievements of Spekkens' theory.** Almost all the phenomena of quantum mechanics are reproduced in the toy theory, apart from Bell non-locality and contextuality, that emerge as inherently non-classical features. The figure is taken from [29].

We can sum up our approach to Spekkens' model as follows:

1. Start from the intuitive (physically justified) formula (2.4) that relates observables Σ_j, ontic states λ and outcomes σ_j.

2. Epistemic restriction: the compatible observables are the ones whose symplectic inner product is zero.

3. Compute the shift vector $\mathbf{w}$. This allows us to shift back the set of points λ to obtain a subspace.

4. The set of ontic states compatible with the epistemic state $(V, \mathbf{w})$ is $V^{\perp} + \mathbf{w}$, where V is the isotropic subspace spanned by the observables Σ_j (the set of known variables).

We say that this approach is physically intuitive because we start with equation (2.4), which is physically motivated and states, observables and the corresponding outcomes are defined in terms of it. Equation (2.4) also allows us to see that the shift comes from the need to recover the subspace structure.

2.1.2 Stabilizer quantum mechanics

Stabilizer quantum mechanics is a subtheory of quantum mechanics where we only consider common eigenstates of tensors of Pauli operators, unitaries belonging to the Clifford group, and Pauli measurements. We first treat the formalism for systems of odd dimensions, as we will mainly deal with it in this chapter. Almost all the definitions will hold the same for the even case, apart from some crucial details.

A stabilizer state ρ can always be written as

$$\rho = \frac{1}{\mathcal{N}} \rho_1 \cdot \rho_2 \cdot \cdots \cdot \rho_N, \tag{2.9}$$

where $\mathcal{N} = Tr[\rho_1 \cdot \rho_2 \cdot \cdots \cdot \rho_N]$ and

$$\rho_j = (\mathbb{I}_d + g_j + g_j^2 + \cdots + g_j^{d-1}), \tag{2.10}$$

where $j \in \{1, \ldots, N \leq n\}$, n is the number of qudits ($N = n$ for pure states), $\mathbb{I}_d$ is the identity operator in dimension d and g_j is a generator of the stabilizer group, more precisely an element of the group generated by the Weyl operators or generalised Pauli operators that, for one qudit, read as:

$$\hat{W}(\lambda) = \chi(-2^{-1}px) Z(p) X(x), \tag{2.11}$$

where $\chi(a) = e^{\frac{2\pi i}{d} a}$ for any $a \in \mathbb{Z}_d$, x, p are the coordinates of the phase space point $\lambda = (x, p) \in \mathbb{Z}_d^2$, and X, Z are respectively the shift and boost operators (generalised Pauli X and Z operators) and the arithmetics is modulo d,

$$X(x) = \sum_{x' \in \mathbb{Z}_d} |x' - x\rangle \langle x'| \tag{2.12}$$

$$Z(p) = \sum_{x \in \mathbb{Z}_d} \chi(px) \left|x\right\rangle \left\langle x\right| . \tag{2.13}$$

Note that $2^{-1} = \frac{d+1}{2}$ is the multiplicative inverse of 2 modulo d. When considering more than one qudit, the Weyl operator is given by the tensor product of the single Weyl operators,

$$\hat{W}(\lambda) = \hat{W}(x_0, p_0, \ldots, x_{n-1}, p_{n-1}) = \hat{W}(x_0, p_0) \otimes \cdots \otimes \hat{W}(x_{n-1}, p_{n-1}), \tag{2.14}$$

where here $\lambda \in \mathbb{Z}_d^{2n}$. Weyl operators form a closed set under multiplication and an orthonormal basis in the space of operators on the Hilbert space with respect to the Hilbert-Schmidt product given by $d^{-n}\mathrm{tr}(.^{\dagger}.)$ [43]. We define the Pauli group $P_{n,d}$ as the group of Weyl operators on n systems of dimension d. We can write the stabilizer state ρ in a more compact way as

$$\rho = \frac{1}{\mathcal{N}} \prod_{j}^{N} \sum_{i}^{d-1} g_j^i. \tag{2.15}$$

However we will mostly use the following notation in terms of stabilizer generators,

$$\rho \to \langle g_1, \ldots, g_N \rangle . \tag{2.16}$$

The unitary evolution in SQM is due to Clifford group transformations $C_{n,d}$, *i.e.* the unitary transformations that map Pauli operators into Pauli operators:

$$C_{n,d} = \{U \mid UPU^{\dagger} \in P_{n,d} \ \ \forall P \in P_{n,d}\}. \tag{2.17}$$

Measurements in SQM are expressed as tensor products of generalised Pauli operators.

Let us now consider even-dimensional SQM. All the above definitions still hold, apart from the fact that the exponent a in the phase factor $\chi(a)$ involves arithmetics modulo $2d$ instead of modulo d. This has some crucial implications, like the presence of contextuality in qubit SQM, that is going to be treated in the next chapter. As an example of even-dimensional SQM we can consider the popular case of qubits. The Pauli operators on the single system are, as usually

represented in the computational basis,

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.18}$$

A simple example of stabilizer state is the Bell state of two qubits: $|\psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, which is a $+1$ eigenstate of $XX, ZZ, -YY$ and $\mathbb{III}$, thus represented by the stabilizer generators $\langle XX, ZZ\rangle$. Notice that we drop the tensor product symbol in order to soften the notation. The Clifford gates are generated by $\langle H_i, S_i, CNOT_{i,j}\rangle$, where H_i is the Hadamard gate acting on the single qubit, S_i is the phase gate acting on the single qubit and $CNOT_{i,j}$ is the controlled not gate acting on two qubits. They are represented in the computational basis as

$$H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.19}$$

Notice that, as already mentioned, the stabilizer formalism here described is defined in terms of the set of integers modulo d, $\mathbb{Z}_d$. This set is a field only when d is a prime number, as the inverses do not exist for all its elements. This implies that some of the desired properties for stabilizer error correcting codes in non-prime dimensions no longer hold [77]. A fully satisfactory definition of the Pauli group in arbitrary dimensions is still to be found. In the case of prime power dimensions Gottesman has proposed to use the so called Galois finite fields [77]. We will not deal with these re-definitions of the Pauli groups in what follows and we will keep treating $\mathbb{Z}_d$ in any dimension, non-prime too.

The stabilizer formalism was developed by Gottesman in the late nineties in the field of quantum error correction [41]. The idea behind it is to calculate key properties of stabilizer codes (most of the utilised codes fall into this category [85–87]) by representing code-words through Pauli operators instead of state vectors, thus avoiding the exponential complexity that derives from the latter. For example, it is possible to detect a given Pauli error, if it anti-commutes with at least one stabilizer generator. SQM is also very important for its applications in the field of quantum computation [53, 77], as will be fully described in the next chapter (subsection

3.1.1).

2.1.3 Wigner functions

Wigner functions are a way of recasting quantum mechanics in the framework of the classical phase space [42, 44, 45]. Wigner functions are quasi-probability distributions, *i.e.* real valued functions that represent quantum states, transformations and measurements providing statistics for measurement outcomes that are consistent with quantum mechanics [88, 89]. The main difference between quasi-probability distributions and actual probability distributions is that the former can take negative values. Nevertheless their marginals represent probability distributions of measurement outcomes. The feature of negativity has often been associated with a signature of non-classicality [42, 82, 88]. The phase space formalism based on Wigner functions is very popular in the field of quantum optics [90] and it originated for the case of infinite dimensional systems (see figure 2.5).

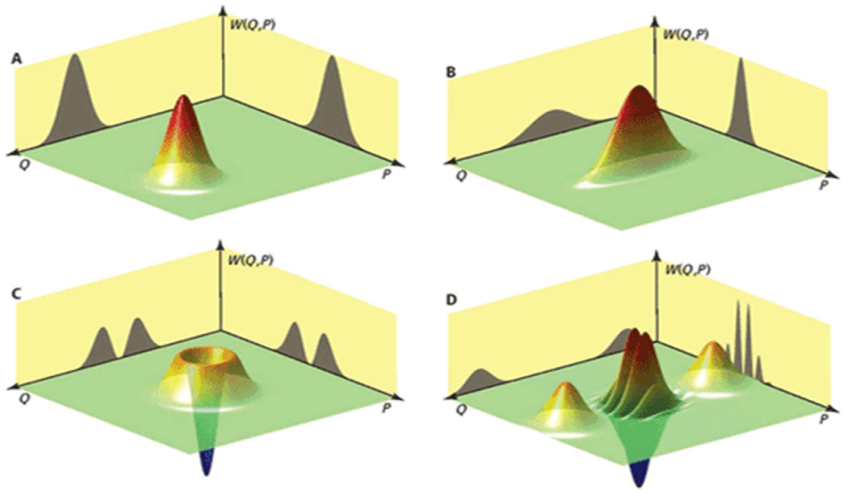


Figure 2.5: **Continuous Wigner functions.** The figure above, taken from *http://www.sciencemag.org/content/332/6027/313/F1.expansion.html*, shows four examples of continuous Wigner functions representing the state of a particle in the phase space, where the axes represent the position and momentum of the particle. Wigner functions are not proper probability distributions since they can be negative (like in the two figures at the bottom). However if we integrate out one of the variables we obtain a probability distribution (indicated by the shadows). The two figures in the top describe Gaussian states (notice the Gaussian shape of the probability distributions), which are very important states in the field of quantum optics [90].

Our work will focus on discrete Wigner functions. These were first proposed by Buot [91] and Berry [92] in the mid seventies, rediscovered for the case of the single qubit by Scully [93] and Feynman [94] and extended to prime-dimensional Hilbert spaces by Wootters [44] and Galetti [95] in the late eighties. In 2004 Gibbons et al. set a framework for Wigner functions based on the work by Wootters, where Wigner functions are defined by associating lines in a discrete phase space to projectors belonging to a fixed set of mutually unbiased bases [96]. Few years later, Galvao and collaborators [45,97] used the class of Wigner functions described by Gibbons and Wootters to study which states are non-negatively represented in qubit quantum mechanics. Among the same family of Wigner functions, in 2006, Gross [43] identified a particular Wigner function for qudits of odd dimensions that satisfies all the desired properties that were also present in the infinite dimensional case (listed below). Remarkably, this Wigner function turns out to be non-negative if and only if the pure quantum state it represents is a stabilizer state. Unfortunately a similar result does not hold for $n-$qubit SQM and the presence of negativity is unavoidable. Non-negativity of the Wigner function for states and measurements in qubit SQM holds only when a single qubit is considered, as shown by Galvao [45]. Non-negative Wigner functions can be interpreted as actual probability distributions and are important both from a foundational and computational point of view. If quantum states can be represented by probability distributions, then this is a strong argument in favour of the epistemic view of quantum mechanics [98], where quantum states are just the observer's partial knowledge about an underlying reality represented by the ontic states (phase-space points). The present work is an example of such approach, as we will use the Wigner function representations of SQM and its subtheories to prove their operational equivalence with ST and its subtheories. In addition, non-negative Wigner functions are also used to perform classical simulations of certain quantum computations, like in the case of odd dimensional qudit SQM [82].

In this chapter we will deal with odd dimensional qudit SQM and we now treat Gross' theory of discrete Wigner functions, that non-negatively describes it [43]. We will briefly mention the standard qubit Wigner function due to Wootters and Gibbons at the end of the subsection (equation (2.31)), but we refer the reader to the next chapter (subsection 3.1.4) for an extensive treatment of $n-$qubit Wigner functions and their usage in quantum computation.

We construct the Wigner function by defining the *characteristic function* associated with

an operator ρ as

$$\Xi_\rho(x,\xi) = \frac{1}{N}\mathrm{tr}(\hat{W}(x,\xi)^\dagger \rho), \tag{2.20}$$

where $\hat{W}$ denotes the Weyl operator as defined in equation 2.11 and the normalization N corresponds to d^n when considering pure states. The Wigner function W_ρ is defined as the symplectic Fourier transform of Ξ_ρ, [3]

$$W_\rho(\lambda) = \frac{1}{N^2}\sum_{\lambda'\in\Omega} \chi^*([\lambda,\lambda'])\mathrm{tr}(\hat{W}(\lambda')^\dagger \rho), \tag{2.21}$$

The above expression can be rewritten as

$$W_\rho(\lambda) = \frac{1}{N}\mathrm{tr}(A(\lambda)\rho), \tag{2.22}$$

where the Hermitian operators $A(\lambda) = \frac{1}{N}\sum_{\lambda'\in\Omega}\chi([\lambda,\lambda'])\hat{W}(\lambda')$ are called the *phase-space point operators*.

The above formulation of Wigner functions for the discrete case is easily extendible to the continuous case by considering the Weyl operators as $\hat{W}(\lambda) = e^{i(pX-xP)}$ and $\chi(x) = e^{ix}$, where $x, p \in \mathbb{R}$ and they are the usual momentum and position (see figure 2.5).

Some basic properties [43] of the discrete Wigner function W_ρ are:

- W_ρ is real and the following relations hold:

$$\sum_{\lambda\in\Omega} W_\rho(\lambda)W_\sigma(\lambda) = \frac{1}{N}\mathrm{tr}(\rho\sigma), \tag{2.23}$$

$$\sum_{\lambda\in\Omega} W_\rho(\lambda) = 1, \tag{2.24}$$

  where ρ, σ are two quantum states.

- The marginals of a Wigner function on the state ρ behave as a classical probability distribution:

$$\sum_{p\in\mathbb{Z}_d} W_\rho(x,p) = |\langle x|\,\rho\,|x\rangle|^2. \tag{2.25}$$

[3] Consider a function $f : \Omega \to \mathbb{C}$. The symplectic Fourier transform of f is $\tilde{f}(\lambda) = |\Omega|^{-\frac{1}{2}}\sum_{\lambda'\in\Omega}\chi^*([\lambda,\lambda'])f(\lambda')$, where $\lambda, \lambda' \in \Omega$.

- The Wigner function of many systems in a product state is the tensor product of the Wigner functions of each system, as a consequence of the factorability of the phase-point operators,

$$A(x_0, p_0 \dots x_{n-1}, p_{n-1}) = A(x_0, p_0) \otimes \cdots \otimes A(x_{n-1}, p_{n-1}). \tag{2.26}$$

- The Wigner function is covariant with respect to the Clifford gates U, *i.e.* for all the stabilizer states ρ,

$$W_{U\rho U^\dagger}(\lambda) = W_\rho(S\boldsymbol{\lambda} + \mathbf{a}), \tag{2.27}$$

  where S is a symplectic transformation and $\mathbf{a}$ is a translation vector.

We are interested in finding the Wigner function of a stabilizer state in any odd dimension and in rephrasing SQM in terms of Gross' Wigner functions. We recall that a stabilizer state is a joint eigenstate of a set of commuting Weyl operators. Two Weyl operators commute if and only if the corresponding phase-space points λ, λ' have vanishing symplectic inner product:

$$[\hat{W}(\lambda), \hat{W}(\lambda')] = 0 \text{ if and only if } [\lambda, \lambda'] = \boldsymbol{\lambda}^T J \boldsymbol{\lambda'} = 0. \tag{2.28}$$

This result derives from the product rule of Weyl operators:

$$\hat{W}(\lambda)\hat{W}(\lambda') = \chi([\lambda, \lambda'])\hat{W}(\lambda + \lambda').$$

We recall again that the square brackets denote the usual commutator when referring to operators and the symplectic inner product when referring to vectors. From this result, the sets of commuting Weyl operators and, as a consequence, the stabilizer states, are parametrized by the isotropic subspace M of Ω. Isotropic, as already mentioned when we introduced Spekkens' theory, means that it is composed by mutually commuting elements. In the case that M has dimension d^n, *i.e.* the eigenspaces are non-degenerate and uniquely map to the (pure) state vectors in the Hilbert space, it is *maximally isotropic*. For each M and each $w \in \Omega$ we can define a stabilizer state $\rho_{M,\mathbf{w}}$ as the projector onto the joint eigenspace spanned by $\{\hat{W}(\lambda) : \lambda \in M\}$, where $\hat{W}(\lambda)$ has eigenvalue $\chi([w, \lambda])$. Let M be maximally isotropic, then the Wigner function associated to the state $\rho_{M,\mathbf{w}}$ is always non-negative (necessary and sufficient condition in odd

dimensions, as proven in [43]) and it is of the kind

$$W_{\rho_{M,\mathbf{w}}}(\lambda) = \frac{1}{d^n}\delta_{M^C+\mathbf{w}}(\lambda), \tag{2.29}$$

where M^C is the symplectic complement of M. When we consider non-maximally isotropic subspaces M, *i.e.* mixed stabilizer states, the above expression (2.29) still holds, with the only exception of the normalization factor that is calculated in order to make the probability distribution uniform,

$$W_{\rho_{M,\mathbf{w}}}(\lambda) = \frac{1}{N}\delta_{M^C+\mathbf{w}}(\lambda), \tag{2.30}$$

where $N = \sum_\lambda \delta_{M^C+\mathbf{w}}(\lambda)$. However, notice that, unlike the case of pure stabilizer states, while it is true that mixed stabilizer states have non-negative Wigner function, it could be that mixed non-stabilizer states also are non-negatively represented [82].

We can find the above form of the Wigner function in equation (2.29) by simply writing the characteristic function

$$\Xi_{\rho_{M,\mathbf{w}}}(\lambda) = \frac{1}{d^n}\chi([w,\lambda])\delta_M(\lambda),$$

and then make the same calculations of equation (2.21). Moreover it can be proven that the transformations that map between the non-negative Wigner functions above are the Clifford unitaries. GT is a faithful way of representing SQM.

Notice now that the Wigner function (2.30) has the same form of the probability distribution (2.7) associated to the epistemic state $(V, \mathbf{w})$ in Spekkens' theory. More precisely, they are equivalent if we assume $M = JV$,[4] indeed this transformation implies that $V^\perp = M^C$. The equivalence between GT and ST, using the matrix J as the bridge, also extends in terms of transformations and measurement statistics [29]. Once we will complete ST with measurement update rules for systems of any dimensions, this equivalence will also imply the equivalence between ST and SQM in odd dimensions. Therefore we can see the description based on known variables (Spekkens) and the description based on Wigner functions (Gross) as two equivalent descriptions of stabilizer quantum mechanics in odd dimensions.

In this subsection we have treated the Wigner function formalism developed by Gross, that is suitable for treating non-negatively SQM of odd dimensional qudits and showing its

[4]Note that the action of J is simply to map a variable into its conjugated.

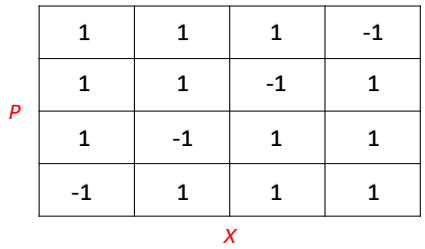


Figure 2.6: **Discrete phase space and Wigner function.** The figure above represents the discrete phase space of two qubits, where an example of Wigner function, the original Wootters Wigner function [44], for the Bell state $\frac{1}{\sqrt{2}}|00+11\rangle$ is calculated in each point. Note that the Wigner function assumes also negative values. The above Wigner function is not normalized, it should be divided by 8 in each term.

operational equivalence with Spekkens' toy theory. A formalism for qubit Wigner functions that shares analogous properties does not exist. We will treat several possible qubit Wigner functions in the next chapter. However, we want to conclude this section by reporting the standard Wigner function for *one* qubit first developed by Wootters [44, 96] and then largely used, in particular in the works by Galvao [45, 97] showing that it is always non-negative for qubit stabilizer states. It is defined as in equation (2.22) , with a different definiton of the phase point operators,

$$\begin{aligned} A(0,0) &= \frac{1}{2}(\mathbb{I}+X+Y+Z) \\ A(0,1) &= \frac{1}{2}(\mathbb{I}+X-Y-Z) \\ A(1,0) &= \frac{1}{2}(\mathbb{I}-X+Y-Z) \\ A(1,1) &= \frac{1}{2}(\mathbb{I}-X-Y+Z), \end{aligned} \tag{2.31}$$

where the operators $\mathbb{I}, X, Y, Z$ are the usual Pauli operators of equation (2.18) . When going to more than one qubit, by considering the Wigner function given by the tensor product of the phase point operators of equation (2.31), the Wigner function associated to the Bell state $\frac{1}{\sqrt{2}}|00+11\rangle$ does show negativity, as represented in figure 2.6.

The remainder of the chapter mainly presents the results cointained in [40], which is a joint work with Dan Browne.

2.2 Update rules - *prime* dimensional case

The formulation of ST in [29], made for prime (and infinite) dimensional systems and described in the previous section, does not provide a full treatment of the transformative aspect of measurements, *i.e.* how the epistemic state has to be updated after a measurement procedure. In the following we will provide a proper formalization of it, and in the next section we will generalise the formalism to all dimensions, non-prime too.

The set of integers modulo d shows different features depending on d being prime or not. In particular in the non-prime case it is not always possible to uniquely define the inverse of a number. The consequences of this will directly affect the update rules. In particular the possible observables sometimes will not show full spectrum: some outcomes will not be possible because they would derive from arithmetics involving numbers with not well-defined inverses. This will divide the set of possible observables in two categories depending on whether they have full spectrum or not. We start from the prime case where problematic observables are not present because inverses always exist.

In ST the measurement process corresponds to the process of *learning* some information (*aka* asking questions) about the ontic state of the system. According to the classical complementarity principle only the observables that are compatible (*i.e.* Poisson-commute) with the state of the system can be learned (jointly knowable). This means that the state after measurement will be given by the generators of the measurement and the generators of the state before the measurement, which are compatible with it.[5] It is then fundamental to understand how compatible sets of ontic states (the isotropic subspaces of known variables V and their perpendicular $V^\perp$) change when independent observables are added and removed from the set of known variables V.

2.2.1 Adding and removing generators to/from V

1. Let us start with the case of *adding* a generator Σ' to the set of generators of $V = span\{\Sigma_1, \ldots, \Sigma_n\}$. We assume that Σ' is linearly independent with respect to the set spanned by the Σ_j. Let us see what happens to $V^\perp$. The subspace V after the addition

[5] As an abuse of language we here talk of generators of a state meaning the orthogonal basis set that generates the subspace of known variables associated with the state.

becomes

$$V' = V \oplus span\{\Sigma'\}. \tag{2.32}$$

By definition the direct sum of two subspaces $A \oplus B$ returns a subspace such that for each $a \in A$ and $b \in B$, the sum $a + b$ belongs to $A \oplus B$. The direct sum of two subspaces is a subspace. We are interested in the orthogonal complement of a direct sum. It is well known that $(A \oplus B)^\perp = A^\perp \cap B^\perp$. This means that by adding a generator to V, its perpendicular $V^\perp$ is given by

$$V'^\perp = V^\perp \cap (span\{\Sigma'\})^\perp. \tag{2.33}$$

Note that $V'^\perp$ is smaller than $V^\perp$.

2. We now analyse what happens if we *remove* a generator, say Σ_n, from the set of generators of V. This means that now $V' = span\{\Sigma_1, \ldots, \Sigma_{n-1}\}$. The set $V^\perp$ is clearly contained in $V'^\perp$, since any vector orthogonal to all elements of V must also be orthogonal to all elements of V'. By definition, the set $V'^\perp$ is composed by all the ontic states λ such that $\mathbf{\Sigma}_j^T \boldsymbol{\lambda} = 0$ for all $j < n$, but $\mathbf{\Sigma}_n^T \boldsymbol{\lambda} \neq 0$. This means that we need to remove the constraint $\mathbf{\Sigma}_n^T \boldsymbol{\lambda} = 0$ to enlarge $V^\perp$ to $V'^\perp$, *i.e.* we simply need to add the ontic states $\lambda' = c\gamma$ to $V^\perp$, where $c \in \mathbb{Z}_d \neq 0$ and γ is a vector such that $\mathbf{\Sigma}_n^T \boldsymbol{\gamma} = 1$. Indeed this implies that

$$\mathbf{\Sigma}_n^T(\boldsymbol{\lambda} + \boldsymbol{\lambda}') = \mathbf{\Sigma}_n^T(\boldsymbol{\lambda} + c\boldsymbol{\gamma}) = 0 + c \neq 0.$$

In *prime* dimensions γ uniquely exists and it corresponds to $k^{-1}\Sigma_n$, where $k = \mathbf{\Sigma}_n^T \mathbf{\Sigma}_n$. Indeed the inverse of an integer $k \in \mathbb{Z}_d \neq 0$ always uniquely exists if d is a prime number. The formula for $V'^\perp$ then reads

$$V'^\perp = \bigcup_c (V^\perp + ck^{-1}\mathbf{\Sigma}_n) \equiv \bigcup_{\mathbf{w}_n \in V_n} (V^\perp + \mathbf{w}_n) = V^\perp \oplus V_n, \tag{2.34}$$

where the addition of $+\mathbf{w}_n$ means that the whole set $V^\perp$ is shifted by $\mathbf{w}_n$, and $V_n = span\{\Sigma_n\}$. The previous trick in general works as follows. Given the ontic state λ, the observable Σ and the outcome σ associated with them, *i.e.* $\mathbf{\Sigma}^T \boldsymbol{\lambda} = \sigma$, then it is possible

to shift the value σ by a constant k such that $\mathbf{\Sigma}^T\mathbf{\Sigma} = k$, by only adding Σ itself to the ontic state:

$$\mathbf{\Sigma}^T(\boldsymbol{\lambda} + \mathbf{\Sigma}) = \sigma + \mathbf{\Sigma}^T\mathbf{\Sigma} = \sigma + k. \tag{2.35}$$

Note that the above identity allows us to change the value of the outcome associated with an ontic state by a constant factor (that we can also choose) without affecting any commuting observable (in this case Σ).

2.2.2 Measurement update rules

We now want to find the update rules for the state $(V, \mathbf{w})$ of a prime dimensional system when we perform a measurement Π described by the subspace V_Π, with generators Σ'_i, and by the representative ontic vectors $\mathbf{r}_j$ associated to the different outcomes σ'_j. These vectors $\mathbf{r}_j$ can be obtained from each other by a shift vector (see figure 2.3), therefore we will consider the generic measurement element represented by $(V_\Pi, \mathbf{r})$ for stating the measurement update rules. The subspace of known variables V can be written in terms of the sets generated by the generators Poisson-commuting with all the Σ'_j, $V_{commute}$, and non-commuting ones, V_{other}. According to this definition $V_{commute}$ will always be a subspace. We cannot state the same for V_{other}, since the null vector does not belong to it. For this reason we augment V_{other} with the null vector in order to create a subspace. This implies that we can decompose V as

$$V = V_{commute} \oplus V_{other}. \tag{2.36}$$

We will now provide the update rules both for V and $\mathbf{w}$ in two steps: first considering the state and measurement to commute, and then the general (non-commuting) case.

Theorem 1. Commuting case. *Given the epistemic state* $(V, \mathbf{w})$ *and the measurement* Π *that commutes with it,* i.e. *the generators of* V *and* V_Π *all Poisson commute, the epistemic state* $(V', \mathbf{w}')$ *after the outcome* σ' *associated to the measurement element* $(V_\Pi, \mathbf{r})$ *has occurred, is described by*

$$V'^{\perp} = (V^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_\Pi^{\perp} + \mathbf{r} - \mathbf{w}'), \tag{2.37}$$

where $\mathbf{w}'$ *is given by equation*

$$\mathbf{w}' = \mathbf{w} + \sum_i {\boldsymbol{\Sigma}'}_i^T(\mathbf{r} - \mathbf{w})\boldsymbol{\gamma}_i, \tag{2.38}$$

where Σ'_i *are the generators of the measurement* Π *and the vectors* γ_j *are such that* ${\boldsymbol{\Sigma}'}_i^T\boldsymbol{\gamma}_j = \delta_{i,j}$.

Proof. When the state and measurement commute we have to add the generators of the measurement to the set of generators of V, as we have seen in the previous subsection 2.2.1 (learning stage). Therefore the update rule for the subspace V is (equation (2.32))

$$V \to V' = V \oplus span\{\Sigma'_0, \Sigma'_1, \dots \Sigma'_i, \dots\} = V \oplus V_\Pi. \tag{2.39}$$

In terms of perpendicular subspaces this implies that $V'^\perp = V^\perp \cap V_\Pi^\perp$.

Let us initially assume the measurement to consist only of one generator Σ'. Let us recall that the outcome associated with Σ' is σ'. We assume $\mathbf{w}$ is not compatible with this outcome, *i.e.* ${\boldsymbol{\Sigma}'}^T\mathbf{w} = \sigma' + x$, for some shift $x \in \mathbb{Z}_d$, and we want to find $\mathbf{w}'$ such that

$${\boldsymbol{\Sigma}'}^T\mathbf{w}' = \sigma'. \tag{2.40}$$

The identity (2.35) we used in the previous section does the job. More precisely,

$$\mathbf{w}' = \mathbf{w} - x\boldsymbol{\gamma},$$

where the vector γ is such that ${\boldsymbol{\Sigma}'}^T\boldsymbol{\gamma} = 1$. The above expression can be also written as

$$\mathbf{w}' = \mathbf{w} - k^{-1}x\boldsymbol{\Sigma}',$$

where $k = {\boldsymbol{\Sigma}'}^T\boldsymbol{\Sigma}'$. The inverse of k always exists because we are in the prime dimensional case. Without referring to x we can restate the update rule for the representative ontic vector as

$$\mathbf{w} \to \mathbf{w} + k^{-1}(\sigma' - {\boldsymbol{\Sigma}'}^T\mathbf{w})\boldsymbol{\Sigma}' = \mathbf{w} + k^{-1}{\boldsymbol{\Sigma}'}^T(\mathbf{r} - \mathbf{w})\boldsymbol{\Sigma}'. \tag{2.41}$$

Note that if we consider more than one generator of the measurement, we simply have to sum

over all those generators in the second term. This immediately follows from considering the whole measurement Π as a sequence of measurements given by each generator Σ'_i and apply every time the rule (2.41). In addition, we need also to require that the γ_j's are such that $\mathbf{\Sigma'}_i^T \boldsymbol{\gamma}_j = \delta_{i,j}$, so that $\mathbf{\Sigma'}_i^T \mathbf{w'}$ gives σ' without unwanted additional terms. We state again that the above formula always holds for prime dimensional systems. We cannot claim the same in non-prime dimensions. The correct update rule for the subspace $V'^{\perp}$ is found by combining the update rules for V and $\mathbf{w}$ as in (2.37). This correction simply sets the subspaces to the same origin in order to correctly compute their intersection, as schematically shown in figure 2.8. At the end we obtain for the epistemic state $(V', \mathbf{w'})$ that $V'^{\perp} + \mathbf{w'} = (V^{\perp} + \mathbf{w}) \cap (V_{\Pi}^{\perp} + \mathbf{r})$. We recall that the probability associated to each ontic state consistent with the epistemic state is uniform, *i.e.* given by $\frac{1}{|V'^{\perp}+\mathbf{w'}|} = \frac{1}{|V'^{\perp}|} = \frac{1}{|(V^{\perp}+\mathbf{w})\cap(V_{\Pi}^{\perp}+\mathbf{r})|}$, where $|\cdot|$ indicates the size of the subspace.

□

Figure 2.7 shows a basic example of theorem 1.

Theorem 2. Non-commuting case. *Given the epistemic state* $(V, \mathbf{w})$ *and the measurement* Π *that does not commute with it,* i.e. *some of the generators of* V_{Π} *do not Poisson commute with the generators of* V*, the epistemic state* $(V', \mathbf{w'})$ *after the outcome* σ' *associated to the measurement element* $(V_{\Pi}, \mathbf{r})$ *has occurred, is described by*

$$V'^{\perp} = (V_{commute}^{\perp} + \mathbf{w} - \mathbf{w'}) \cap (V_{\Pi}^{\perp} + \mathbf{r} - \mathbf{w'}), \tag{2.42}$$

where $V_{commute}^{\perp}$ *is given by*

$$V_{commute}^{\perp} = V^{\perp} \oplus V_{other}. \tag{2.43}$$

The representative ontic vector $\mathbf{w'}$ *is given by*

$$\mathbf{w'} = \mathbf{w} + \sum_i \mathbf{\Sigma'}_i^T (\mathbf{r} - \mathbf{w}) \boldsymbol{\gamma}_i, \tag{2.44}$$

where Σ'_i *are the generators (even the non-commuting ones) of the measurement* Π *and the vectors* γ_j *are such that* $\mathbf{\Sigma'}_i^T \boldsymbol{\gamma}_j = \delta_{i,j}$.

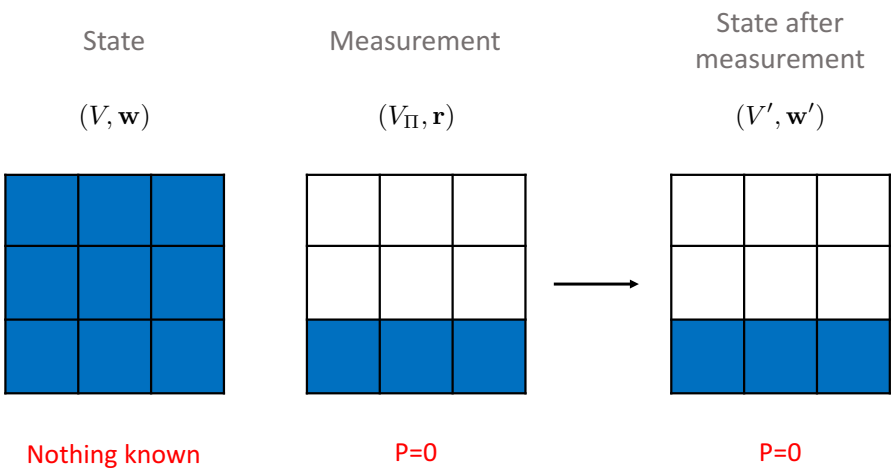


Figure 2.7: **Update rules in the prime commuting case.** The figure above shows a simple one-trit example of theorem 1 regarding the update rule to predict the state after a sharp measurement that commutes with the original state. The state after measurement is given by $V'^{\perp} + \mathbf{w}' = (V^{\perp} + \mathbf{w}) \cap (V_\Pi^{\perp} + \mathbf{r})$. In the above case the shift vectors are all $(0, 0)$, the perpendicular subspaces are $V^{\perp} = \Omega$, $V_\Pi^{\perp} = span\{(1, 0)\}$, and $V'^{\perp} = V_\Pi^{\perp}$. Note that with "measurement" we are here representing one element of the measurement. The other elements can be obtained by simply shifting $\mathbf{r}$ as seen in figure 2.3. The final state is associated to each element of the measurement, each one with a corresponding probability of happening. The same reasoning holds for figures 2.9 and 2.12.

Proof. Let us assume that Σ'_j, for $j \in \{0, \dots, m-1\}$, do not commute with the generators of V. In addition to the learning stage of the previous commuting case, we also have a removal stage of the disturbing part of the measurement. We have already seen that we can split the subspace V in $V = V_{commute} \oplus V_{other}$. Therefore we can reduce to the commuting case if we only consider $V_{commute}$ instead of the whole V. The update rule for the subspace V then becomes

$$V \to V' = V_{commute} \oplus span\{\Sigma'_0, \Sigma'_1, \dots \Sigma'_i, \dots\} = V_{commute} \oplus V_\Pi.$$

In terms of the perpendicular subspaces note that we can both write

$$V'^{\perp} = (V^{\perp} \oplus V_{other}) \cap V_\Pi^{\perp},$$

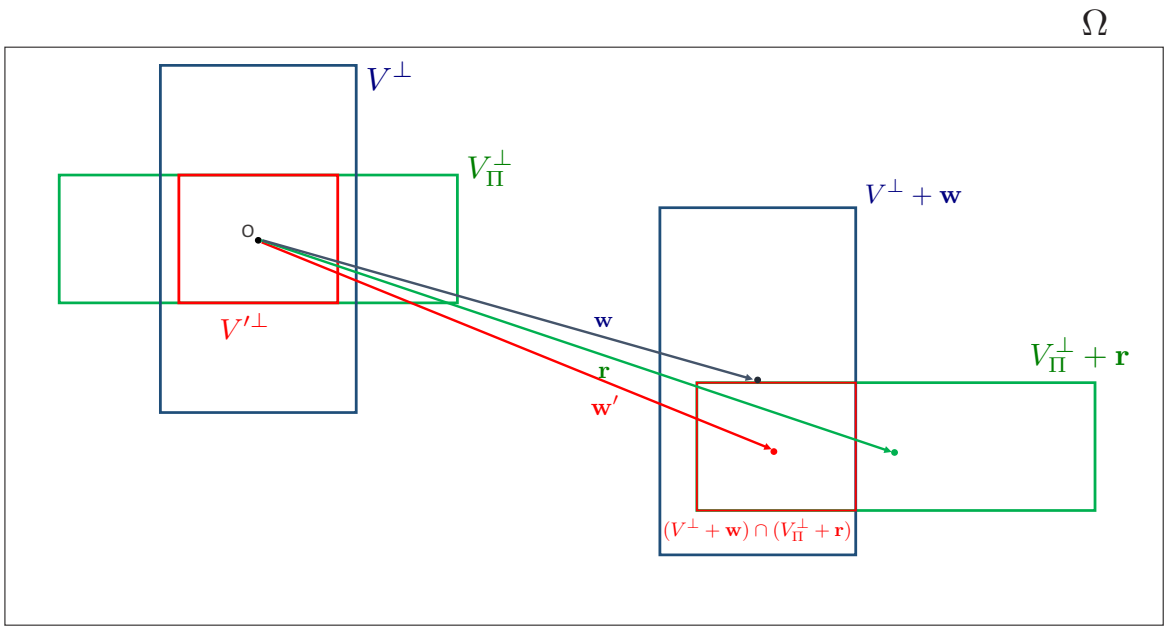


Figure 2.8: **Update rules via Venn diagrams.** The figure above schematically shows the subspaces $V^\perp, V_\Pi^\perp, V'^\perp$ and the shifted ones (after applying the corresponding representative ontic vectors $\mathbf{w}, \mathbf{r}, \mathbf{w}'$). In particular this picture explains the expression $V'^\perp = (V^\perp + \mathbf{w} - \mathbf{w}') \cap (V_\Pi^\perp + \mathbf{r} - \mathbf{w}')$ as a result of combining the update rules for the epistemic subspaces and the representative ontic vectors. It is important to notice that to obtain the correct intersection we have to shift the subspaces $V^\perp + \mathbf{w}$ and $V_\Pi + \mathbf{r}$ back to the same origin (this is the role of $\mathbf{w}'$). Indeed note that $V^\perp \cap V_\Pi^\perp$ is different from $(V^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r})$.

and

$$V'^\perp = V_{commute}^\perp \cap V_\Pi^\perp,$$

from the usual property that the perpendicular of a direct sum is the intersection of the perpendicular subspaces. The update rule for the representative ontic vector is the same as in the previous case (equation (2.38)). The correct update rule for the subspace $V'^\perp$ is found by combining the update rules for V and $\mathbf{w}$ as in the previous case (2.37), where $V^\perp$ is replaced by $V_{commute}^\perp$. At the end we obtain for the epistemic state $(V', \mathbf{w}')$ that $V'^\perp + \mathbf{w}' = (V_{commute}^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r})$.

□

Figure 2.9 shows a basic example of theorem 1.

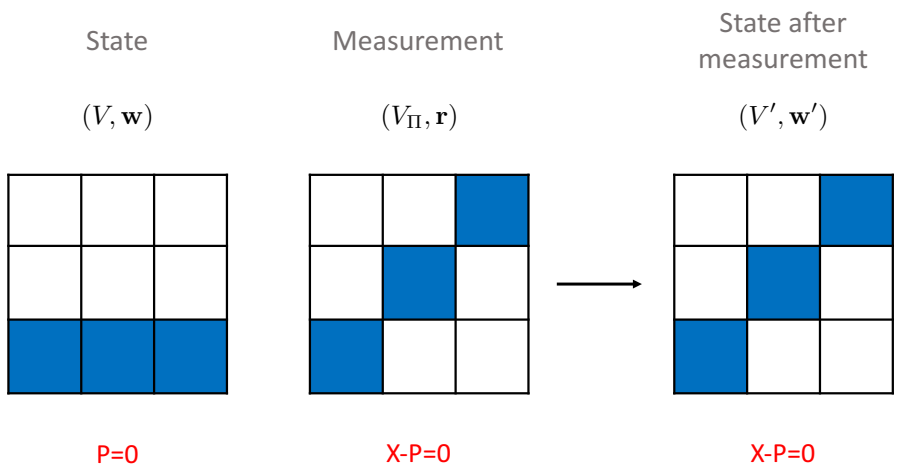


Figure 2.9: **Update rules in the prime non-commuting case.** The figure above shows a simple one-trit example of theorem 2 regarding the update rule to predict the state after a sharp measurement that does not commute with the original state. The state after measurement is given by $V'^{\perp} + \mathbf{w}' = (V^{\perp}_{commute} + \mathbf{w}) \cap (V^{\perp}_{\Pi} + \mathbf{r})$. In the above case the shift vectors are all $(0, 0)$, the perpendicular subspaces are $V^{\perp}_{commute} = \Omega$, $V^{\perp}_{\Pi} = span\{(1, 1)\}$, and $V'^{\perp} = V^{\perp}_{\Pi}$.

2.3 Update rules - *non prime* dimensional case

It is quite common in studies of discrete theories, like Spekkens' model and SQM, to only consider the prime dimensional case because of the particular features of the set of integers modulo d, $\mathbb{Z}_d$, when d is non-prime, like the impossibility of uniquely define inverses of numbers. For example in our present case, figure 2.10 shows the peculiar properties of the observable $3X$ in $d = 6$, which has not full spectrum of outcomes. The general formulation of Spekkens' model of section 2.1.1 does not change; not even the rules for calculating the probabilities of outcomes and the updating of the state after a reversible evolutions. The new formulation we provide affects the observables and the related measurements update rules. More precisely our issue, as already noticed, regards the updating-rule formula (2.38) for the shift vector $\mathbf{w}'$, which does not always hold when the dimension d is *non-prime*. In fact the vector $\boldsymbol{\gamma}$ such that $\boldsymbol{\Sigma}'^{T}\boldsymbol{\gamma} = 1$ does not always exist in that case. On the other hand, in prime dimensions, it always uniquely exists because $\boldsymbol{\gamma} = k^{-1}\boldsymbol{\Sigma}'$ and the inverse of the integer $k = \boldsymbol{\Sigma}'^{T}\boldsymbol{\Sigma}'$ always uniquely exists. Unlike the original formulation due to Spekkens, we will now characterise Spekkens' model in non-prime dimensions. In particular we characterise which are the observables that are problematic in the

above sense - the *coarse-graining* observables, like $3X$ in $d = 6$ - and we then find the update rules for a state subjected to the measurement of such observables by rewriting them in terms of non-problematic observables - the *fine-graining* observables.

In the next subsection we assume single-system observables (*i.e.* of the kind $\Sigma' = aX + bP$, $a, b \in \mathbb{Z}_d$) in order to soften the notation and facilitate the comprehension. This will bring more easily to the update rules even in the most general case of many systems (subsection 2.3.2). In this case we recall, without making any reference to the quantity k^{-1}, but just in terms of the vector $\boldsymbol{\gamma}$, the update rule for the shift vector $\mathbf{w}'$,

$$\mathbf{w}' = \mathbf{w} - x\boldsymbol{\gamma}, \tag{2.45}$$

where, as usual, $x = -\boldsymbol{\Sigma}'^{T}(\mathbf{r} - \mathbf{w})$.

2.3.1 Coarse-graining and fine-graining observables

We define a fine-graining observable as an observable that has *full spectrum*, *i.e.* it can assume all the values in $\mathbb{Z}_d$. On the contrary a coarse-graining observable has not full spectrum.

Lemma 1. *An observable O_{fg} has full spectrum,* i.e. *it is a fine-graining observable, if and only if it has the following form,*

$$O_{fg} = a'X + b'P, \tag{2.46}$$

where $a', b' \in \mathbb{Z}_d$ are such that they do not share *any integer factor or power factor of d.*

On the contrary a coarse-graining observable is written as

$$O_{cg} = aX + bP = D(a'X + b'P), \tag{2.47}$$

where $a', b' \in \mathbb{Z}_d$ are again such that they *do not share* any integer factor or power factor of d and D is a factor shared by $a, b \in \mathbb{Z}_d$. More precisely the factor D is called *degeneracy* and it is defined as

$$D = D_1^{n_1} \cdot D_2^{n_2} \cdot \ldots, \tag{2.48}$$

where $D_1, D_2, \ldots$ are different integer factors of d shared by a and b, and $n_1, n_2, \ldots$ are the

maximum powers of these factor such that they can still be grouped out from a and b. We take the maximum powers because we want the remaining part, $a'X + b'P$, to not share any common integer factor or power factor of d between a' and b'. In this way we can associate a fine-graining observable to a coarse graining one by simply dropping the degeneracy D from the latter.

Proof. Let us first prove that an observable of the kind (2.46), $O_{fg} = a'X + b'P$, is a full spectrum one. This can be proven by using Bezout's identity [99]: let a' and b' be nonzero integers and let D be their greatest common divisor. Then there exist integers X and P such that $aX + bP = D$. In our case the greatest common divisor D is equal to one, since a', b' are coprime. [6] Therefore we have proven that there exist values of the canonical variables $X, P \in \mathbb{Z}_d$ such that $O_{fg} = a'X + b'P = 1$. In order to reach all the other values of the spectrum we simply need to multiply both X and P in the previuos equation by $j \in \mathbb{Z}_d$.

We now prove the converse, *i.e.* that a full spectrum observable implies it to be written as (2.46). We prove this by seeing that an observable written as (2.47) has not full spectrum, *i.e.* we negate both terms of the reverse original implication. Proving the latter is straightforward, since the multiplication modulo d between an arbitrary quantity and a factor D, which is given by powers of integer factors of d, gives as a result a multiple of D. Since the multiples of D do not cover the whole $\mathbb{Z}_d$, then any observable of the form (2.47) has not full spectrum. [7] Since an observable of the form (2.46) is an observable that cannot be written as (2.47) by definition, we obtain that a full spectrum observable implies the observable to be written as (2.46).

□

Given lemma 1 we have got the expressions (2.47) and (2.46) for coarse-graining and fine-graining observables. We want now to prove the following lemma to ensure that fine-graining observables are characterised by precisely defined update rules.

Lemma 2. *The vector $\boldsymbol{\gamma}$ in the update rule (2.45) for the shift vector $\mathbf{w}'$ exists if and only if the observable is a fine-graining one.*

[6] It could be that a', b' share a factor which is not a factor of d. In this case the argument follows identically as if they were coprime.

[7] Multiples of D do not cover the whole spectrum of $\mathbb{Z}_d$ because D has not an inverse D^{-1} (it is not coprime with d) and so we cannot obtain the whole values σ of $\mathbb{Z}_d$ by simply finding X, P such that $a'X + b'P = D^{-1}\sigma$.

Proof. Let us prove that if we have a fine graining observable the vector γ exists. In our case $\Sigma' = (a', b')$ and, by definition of a', b' (as usual defined for fine-graining observables) and full spectrum, we can always find a vector $\boldsymbol{\gamma} = (\gamma_a, \gamma_b)$ such that $\boldsymbol{\Sigma'}^T\boldsymbol{\gamma} = a'\gamma_a + b'\gamma_b$ equals 1.

Let us prove the converse. We now have the vector $\boldsymbol{\gamma}$ such that $\boldsymbol{\Sigma'}^T\boldsymbol{\gamma} = a\gamma_a + b\gamma_b = 1$, where the coefficients $a, b \in \mathbb{Z}_d$ define our observable $aX + bP = \sigma$. We want to prove that σ can achieve all the values of $\mathbb{Z}_d$. Since $\boldsymbol{\Sigma'}^T\boldsymbol{\gamma} = 1$ we can set the values of (X, P) as equal to (γ_a, γ_b) in order to reach the value $\sigma = 1$. We can now achieve all the other values of the spectrum by simply redefining $\boldsymbol{\gamma}$ as $\tilde{\boldsymbol{\gamma}} = c\boldsymbol{\gamma}$, where c assumes all the values in $\mathbb{Z}_d$.

□

The above lemma 2 should convince us that in order to find the update rules in the presence of a coarse graining observable, it is appropriate to decompose it in terms of fine-graining observables. Let us assume that our coarse-graining observable is $O_{cg} = aX + bP = D(a'X + b'P) = \sigma$, and the associated isotropic subspace and representative ontic vector are $(V_{cg}, \mathbf{r}_{cg})$.[8] To this observable we can associate $\bar{D}$ different fine-graining observables $O_{fg} = a'X + b'P = \sigma_j$, where $j \in 0, \ldots, \bar{D} - 1$. The quantity $\bar{D}$ is the degeneracy D without the powers $n_1, n_2, \ldots$, *i.e.* $\bar{D} = D_1 \cdot D_2 \cdot \ldots$. Indeed the powers $n_1, n_2, \ldots$ simply represent multiplicities associated to each corresponding fine-graining observable. The associated isotropic subspaces and representative ontic vectors are $(V_{fg}, \mathbf{r}_{fg}^{(j)})$, where $V_{fg} = span\{(a', b')\}$ (see figure 2.10).

By definition the perpendicular isotropic subspaces are

$$V_{cg}^{\perp} = \{\mathbf{v} = (v_a, v_b) \in \Omega | v_a a + v_b b = D(v_a a' + v_b b') = 0 \mod d\} \tag{2.49}$$

$$V_{fg}^{\perp} = \{\mathbf{v}' = (v'_a, v'_b) \in \Omega | v'_a a' + v'_b b' = 0 \mod d\}. \tag{2.50}$$

It is clear that $V_{cg}^{\perp} \supset V_{fg}^{\perp}$ and we can therefore construct $V_{cg}^{\perp}$ as

$$V_{cg}^{\perp} = \bigcup_{j=0}^{\bar{D}-1} (V_{fg}^{\perp} + \mathbf{v}_j) = V_{fg}^{\perp} \oplus V_D, \tag{2.51}$$

where the subspace V_D provides all the vectors that we need to combine with the vectors of $V_{fg}^{\perp}$

[8] Notice that we use terms like "subspaces" and "vectors" even when we refer to modules and not proper vector spaces. The operations we will use, like the union and the direct sum, also hold for modules.

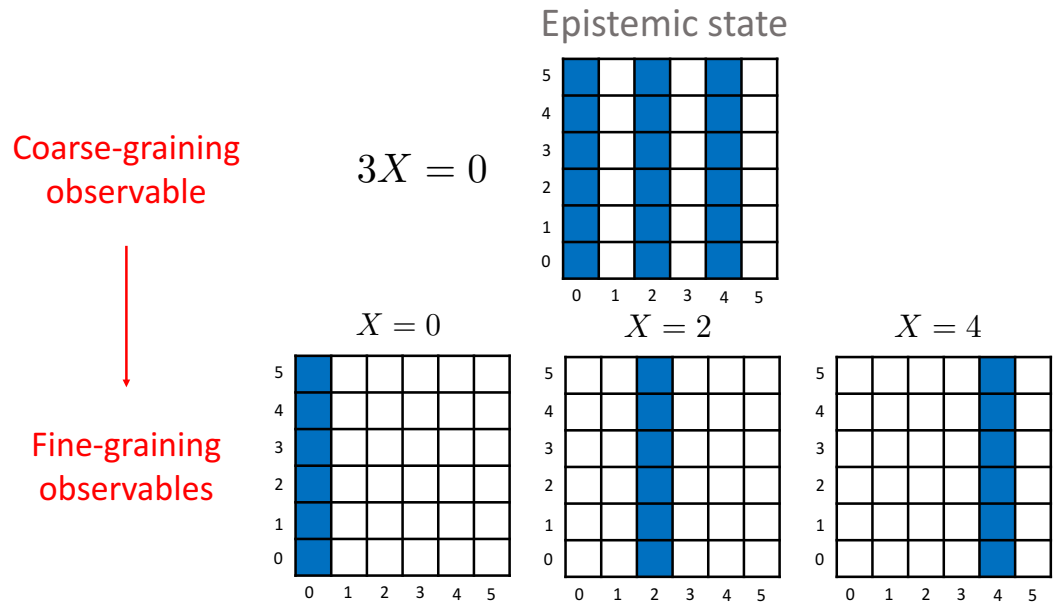


Figure 2.10: **Simple example of a coarse-graining observable and its decomposition in fine-graining observables in** $d = 6$**.** The coarse-graining observable $O_{cg} = 3X = 0$ in $d = 6$ shows degeneracy $D = 3$. The three fine-graining observables associated with O_{cg} are $O_{fg}^{(0)} = X = 0$, $O_{fg}^{(1)} = X = 2$ and $O_{fg}^{(2)} = X = 4$. The perpendicular subspaces of known variables are $V_{cg}^{\perp} = span\{(0,1),(2,0)\}$, $V_{fg}^{\perp} = span\{(0,1)\}$ and $V_D = span\{(2,0)\}$. A choice for the representative ontic vectors is $\mathbf{r}_{cg} = (0,0)$, $\mathbf{r}_{fg}^{(0)} = (0,0)$, $\mathbf{r}_{fg}^{(1)} = (2,0)$ and $\mathbf{r}_{fg}^{(2)} = (4,0)$. Notice that not all the values are possible for the coarse-graining observable $3X$ to be a valid observable. Only $3X = 0$ and $3X = 3$ are valid (indeed what would it be the epistemic state representation for *e.g* $3X = 2$?), as witnessed by the expression (2.54) for the associated fine-graining observables, that is valid only when the ratio $\frac{\sigma_{cg}}{D}$ exists.

to reach the whole $V_{cg}^{\perp}$. More precisely, it is defined as

$$V_D = \{\mathbf{v} \in \Omega | \alpha\mathbf{w} + \beta\mathbf{v} = \mathbf{t}, where\ \mathbf{w} \in V_{fg}^{\perp},\ \alpha, \beta \in \mathbb{Z}_d,\ \mathbf{t} \in V_{cg}^{\perp}\}. \tag{2.52}$$

We call the subspace V_D the *degeneracy subspace* because it encodes the degeneracy of V_{cg} with respect to V_{fg} . It has dimension 1 and size $\bar{D}$. This is consistent with the fact that the dimensions of $V_{cg}^{\perp}$ and $V_{fg}^{\perp}$ are respectively 2 and 1. The sizes are respectively $\bar{D} \cdot d$ and d. The size of $V_{fg}^{\perp}$ is d because it is always a maximally isotropic subspace and its dimension is 1 because from one generator we get all the other vectors of the subspace by multiplication with $j \in \mathbb{Z}_d$. The dimension $V_{cg}^{\perp}$ is 2 because it cannot be 1 (it would be the same subspace as $V_{fg}^{\perp}$)

and it cannot be greater than 2 since also the whole phase space $\Omega = \mathbb{Z}_d^2$ has dimension 2. In order to know the size of $V_{cg}^{\perp}$ we need to count all the $j\mathbf{v}$, where $j \in \{0, 1, \ldots, \bar{D} - 1\}$, that means $\bar{D} \cdot d$. Therefore it can be written as $V_D = span\{\mathbf{v}\}$, and all its $\bar{D}$ vectors are of the kind $\mathbf{v}_j = j\mathbf{v}$. The above reasoning easily extends to the case of n systems, where the dimensions are $\dim(V_{cg}^{\perp}) = 2n$, $\dim(V_{fg}^{\perp}) = n$, $\dim(V_D) = n$, and the sizes are $|V_{cg}^{\perp}| = \bar{D}^n d^n$, $|V_{fg}^{\perp}| = d^n$, $|V_D| = \bar{D}^n$.

We now define the shift vectors $\mathbf{r}_{fg}^{(j)}$ in terms of $\mathbf{r}_{cg}$ and see that we can encode the degeneracy expressed by V_D in there. The idea is schematically depicted in figure 2.11.

Given the shift vector associated to the coarse-graining observable $\mathbf{r}_{cg}$, the shift vectors $\mathbf{r}_{fg}^{(j)}$ associated to the corresponding fine-graining observables are of the kind

$$\mathbf{r}_{fg}^{(j)} = \mathbf{r}_{cg} + \mathbf{v}_j, \tag{2.53}$$

where $\mathbf{v}_j \in V_d$ and are therefore of the kind $j\mathbf{v}$, where $j \in \{0, \ldots, \bar{D} - 1\}$. This implies that if we assume the outcome associated to the coarse-graining observable to be σ_{cg}, *i.e.* $\mathbf{\Sigma}_{cg}^T \mathbf{r}_{cg} = \sigma_{cg}$, where $\Sigma_{cg} = (a, b)$, then the outcomes associated to the fine graining-observables are

$$\mathbf{\Sigma}_{fg}^T \mathbf{r}_{fg}^{(j)} = \mathbf{\Sigma}_{fg}^T (\mathbf{r}_{cg} + j\mathbf{v}) = \frac{\sigma_{cg}}{D} + jC, \tag{2.54}$$

where C is the *anti-degeneracy* and it is defined as a non-zero number belonging to $\mathbb{Z}_d$ such that $D \cdot C = 0 \mod d$. The idea is that the vector $\mathbf{v} \in V_D$ is such that $\mathbf{\Sigma}_{fg}^T \mathbf{v} = C \neq 0$, so it does not belong to $V_{fg}^{\perp}$, but it does belong to $V_{cg}^{\perp}$, since $D \cdot C = 0 \mod d$. An easy way to find one of the possible $\mathbf{v}$ is to calculate it as $C\mathbf{\Sigma}_{fg}$, where Σ_{fg} is the generator of V_{fg}. In this way we know that $D\mathbf{v} = 0$, but $\mathbf{v}$ does not belong to $V_{fg}^{\perp}$, *i.e.* $v_a a' + v_b b' \neq 0$ because Σ_{fg} is not in $V_{fg}^{\perp}$. It is important to notice that equation (2.54) implies that not all the outcomes are allowed for the fine-graining observables associated to the coarse-graining one; they are allowed only when the ratio $\frac{\sigma_{cg}}{D}$ exists. Figure 2.10 also explains this fact.

2.3.2 Measurement update rules

Let us assume to have n systems and to measure the coarse-graining observable $O_{cg} = a_1 X_1 + b_1 P_1 + \cdots + a_n X_n + b_n P_n = D(a_1' X_1 + b_1' P_1 + \cdots + a_n' X_n + b_n' P_n)$ with the outcome σ_{cg},

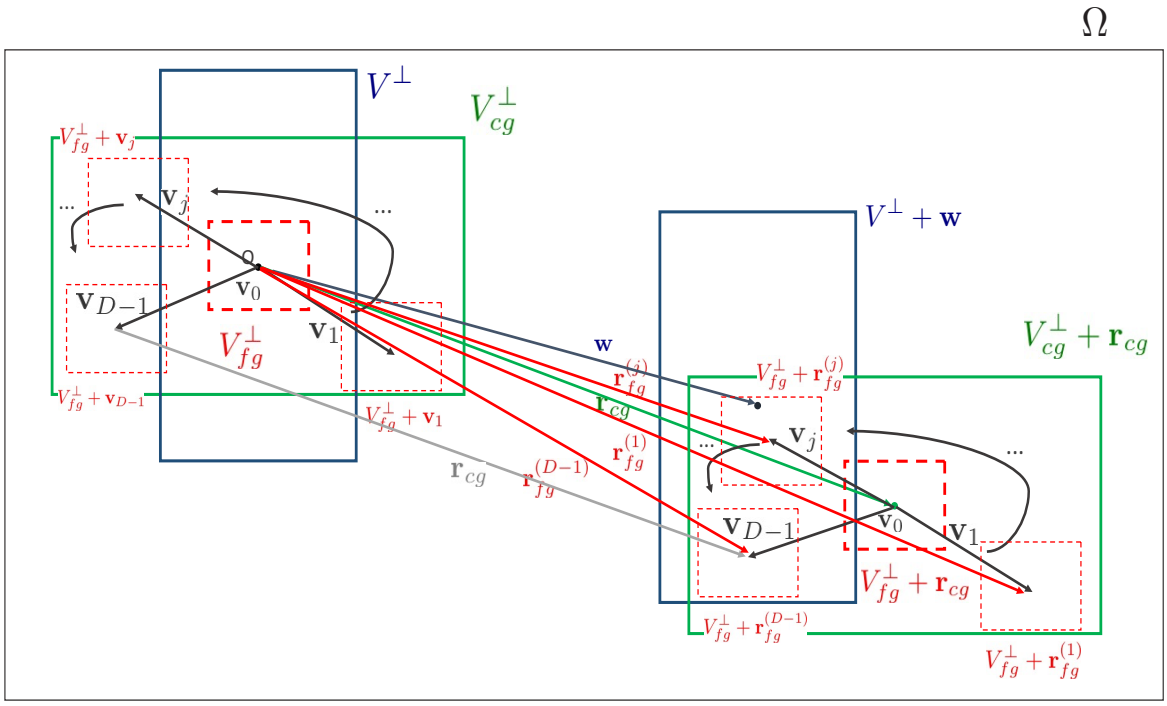


Figure 2.11: **Schematic representation of Coarse-graining decompositions into fine-graining observables.** The figure above schematically represents the relation between the subspaces $V^\perp, V_{cg}^\perp, V_{fg}^\perp$ and their corresponding shift vectors $\mathbf{w}, \mathbf{r}_{cg}, \mathbf{r}_{fg}^{(j)}$. The green rectangles represent the subspaces $V_{cg}^\perp$ and its translated $V_{cg}^\perp + \mathbf{r}_{cg}$. The latter can be seen to be equivalent either to the dashed red rectangle representing $V_{fg}^\perp$ shifted by the degeneracy vectors of V_D (light black arrows), this corresponding to $V_{cg}^\perp$, and then shifted by $\mathbf{r}_{cg}$ (green arrow or light grey arrow), or to the dashed rectangle representing $V_{fg}^\perp$ shifted by each $\mathbf{r}_{fg}^{(j)}$ (red arrows). Both are in accordance with the expressions of $V_{cg}^\perp + \mathbf{r}_{cg}$, $V_{cg}^\perp + \mathbf{r}_{cg} = V_{fg}^\perp \oplus V_D + \mathbf{r}_{cg} = V_{fg}^\perp + \sum_{j=0}^{\bar{D}-1}(\mathbf{r}_{cg} + j\mathbf{v})$, where $\mathbf{v}$ is the generator of V_D. Note that, as a consequence of the degeneracy characterising the coarse-graining observable, we could keep adding $\sum_{j=0}^{D-1} j\mathbf{v} = D\mathbf{v}$ without changing the validity of the expression of $V_{cg}^\perp + \mathbf{r}_{cg}$. We can see it just by noticing that $V_D + \sum_{j=0}^{D-1} j\mathbf{v} = V_D$ or, more simply, that $D\mathbf{v} = D \cdot C\boldsymbol{\Sigma}'_{\mathbf{fg}} = 0$, since $D \cdot C = 0 \bmod d$.

and corresponding isotropic subspace of known variables V_{cg} and shift vector $\mathbf{r}_{cg}$, on the state ρ defined by $\alpha_1 X_1 + \beta_1 P_1 + \cdots + \alpha_n X_n + \beta_n P_n = \sigma$, with corresponding isotropic subspace of known variables $V = span\{\Sigma_1, \ldots, \Sigma_n\}$ and shift vector $\mathbf{w}$. The idea in order to find the update rules for the state after measurement, the subspace of known variables V' and the representative ontic vector $\mathbf{w}'$ is to compute the update rule of the initial state ρ with the fine-graining observables that are associated to the coarse graining observable O_{cg}, *i.e.*

$O^{(j)}_{fg} = a'_1 X_1 + b'_1 P_1 + \cdots + a'_n X_n + b'_n P_n$ with outcome $\sigma^{(j)}_{fg}$ (indeed we know that the update rules are valid for them from lemma 2), and then combine them together. More precisely, the following theorem holds.

Theorem 3. *Given the epistemic state* $(V, \mathbf{w})$ *and the coarse-graining measurement* Π*, the epistemic state* $(V', \mathbf{w}')$ *after the outcome* σ_{cg} *associated to the measurement element* $(V_{cg}, \mathbf{r}_{cg})$ *has occurred, is described by*

$$V'^{\perp} = \bigcup_{j=0}^{\bar{D}-1} [(V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap (V^{\perp}_{fg} + \mathbf{r}^{(j)}_{fg} - \mathbf{w}')], \tag{2.55}$$

where the shift vector $\mathbf{w}'$ *is the shift vector deriving from the update rule of the state after the measurement of the fine-graining observable* $O^{(j)}_{fg}$*,*

$$\mathbf{w}' = \mathbf{w}'_j = \mathbf{w} + \sum_{i=0}^{n} \mathbf{\Sigma}'^{T}_i (\mathbf{r}^{(j)}_{fg} - \mathbf{w}) \boldsymbol{\gamma}_i, \tag{2.56}$$

where the vectors $\boldsymbol{\gamma}_j$ *are defined such that* $\mathbf{\Sigma}'^{T}_i \boldsymbol{\gamma}_j = \delta_{i,j}$*, and* Σ'_i *are the* n *generators of the subspace* V_{fg} *associated to the fine-graining observable* $O^{(j)}_{fg}$*. The subspace* $V^{\perp}_{commute}$ *is given by the original* V *after having removed the non-commuting part,* i.e. *equation* (2.43)*.*

The above theorem tells us that the way we combine the updating subspaces of the state with each individual fine-graining observables is through their union. This result is clear in terms of schematic diagrams (figure 2.11). The updated shift vector is just one of the updated shift vectors of the state with the fine-graining observables, because the information needed to update the shift vector of the state is encoded in just one of the fine-graining shift vectors. The degeneracy includes a meaningless multiplicity in the coarse-graining shift vector, and therefore every fine-graining observable can do the job of correctly updating the shift vector of the state. Actually every combination of the shift vectors $\mathbf{w}'_j$ can do the job, apart from the ones that sum to 0 mod d, like $\sum_{j=0}^{\bar{D}-1} \mathbf{w}'_j$.

Proof. We find the expression for the updated subspace $V'^{\perp}$ by simply reusing the already found formulas (2.37) and (2.43) of the prime-dimensional case and substituting $V^{\perp}_{\Pi}$ with $V^{\perp}_{cg}$

and $\mathbf{r}$ with $\mathbf{r}_{cg}$,

$$V'^{\perp} = (V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap (V^{\perp}_{cg} + \mathbf{r}_{cg} - \mathbf{w}').$$

If we now consider the decomposition of $V^{\perp}_{cg} + \mathbf{r}_{cg}$ as in (2.51) and (2.53), we obtain

$$V'^{\perp} = (V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap [\cup_{j=0}^{\bar{D}-1}(V^{\perp}_{fg} + \mathbf{r}^{(j)}_{fg}) - \mathbf{w}'].$$

Since the intersection of a union is the union of the intersections, we have proven the first part of the theorem,

$$V'^{\perp} = \bigcup_{j=0}^{\bar{D}-1} [(V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap (V^{\perp}_{fg} + \mathbf{r}^{(j)}_{fg} - \mathbf{w}')].$$

The second part of the proof regards $\mathbf{w}'$ being equal to any of the $\mathbf{w}'_j$. Because of the degeneracy, any $\mathbf{w}'_j$ is equivalent to the others (with different value of j) in order to provide us with $\mathbf{w}'$, indeed it is possible to find one from another just by adding a vector $\mathbf{v} \in V_D$. The latter can be proven as follows. For simplicity let us assume to be in the case $n = 1$ and that $\mathbf{v}$ is the generator of V_D. We know that, by the definition of state after measurement of a fine-graining observable, the updated shift vector $\mathbf{w}'_j$ is such that $\mathbf{\Sigma}^T_{fg}\mathbf{w}'_j = \frac{\sigma_{cg}}{D} + jC = \sigma^{(j)}_{fg}$, where $C = \mathbf{\Sigma}^T_{fg}\mathbf{v}$ is the antidegeneracy (equation (2.54)). It is straightforward to see that if we add $\mathbf{v}$ to $\mathbf{w}'_j$, we get $\mathbf{w}'_j + \mathbf{v} = \mathbf{w}'_{j+1}$, indeed $\mathbf{\Sigma}^T_{fg}(\mathbf{w}'_j + \mathbf{v}) = \frac{\sigma_{cg}}{D} + (j+1)C = \sigma^{(j+1)}_{fg}$. □

Figure 2.12 shows a basic example of theorem 3.

2.4 Equivalence of Spekkens' theory and stabilizer quantum mechanics in all odd dimensions

In [29] it has been shown that SQM and Spekkens' toy model are two operationally equivalent theories in odd prime dimensions via Gross' theory of discrete non-negative Wigner functions. We have generalised Spekkens' model to all discrete dimensions. The above equivalence does not hold in even dimensions, but we will now see that it holds in *all* odd dimensions. We will also state the equivalence in terms of the update rules, where all its elegance arises. We recall that SQM and Gross' theory of non-negative Wigner functions are equivalent in all odd dimensions [43].

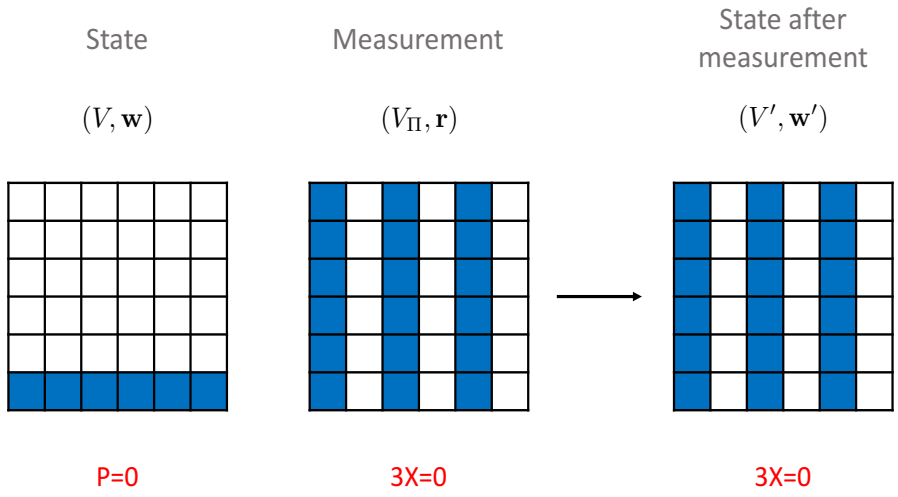


Figure 2.12: **Update rules in the non-prime non-commuting case.** The figure above shows a simple example (one system in $d = 6$) of theorem 3 regarding the update rule to predict the state after a sharp measurement that does not commute with the original state. The state after measurement is given by $V'^{\perp} + \mathbf{w}' = \bigcup_{j=0}^{\bar{D}-1}[(V^{\perp}_{commute} + \mathbf{w}) \cap (V^{\perp}_{fg} + \mathbf{r}^{(j)}_{fg})]$. In the above case the shift vectors are $\mathbf{w} = (0,0), \mathbf{r}^{(0)}_{fg} = (0,0), \mathbf{r}^{(1)}_{fg} = (2,0), \mathbf{r}^{(2)}_{fg} = (4,0), \mathbf{w}' = (0,0)$, the perpendicular subspaces are $V^{\perp}_{commute} = \Omega$, $V^{\perp}_{fg} = span\{(0,1)\}$, $V^{\perp}_{\Pi} = span\{(0,1),(2,0)\}$, and $V'^{\perp} = V^{\perp}_{\Pi}$.

2.4.1 Stabilizer quantum mechanics - update rules

We now analyse the update rules for the stabilizer state ρ under the stabilizer measurement Π, where, as already described in subsection 2.1.2, equation (2.16), we denote ρ and Π in terms of their stabilizer generators g_j and p_k respectively, where $j \in \{1, \ldots, N \leq n\}$, and $k \in \{1, \ldots, M \leq n\}$,

$$\rho \to \langle g_1, \ldots, g_N \rangle ,$$

$$\Pi \to \langle p_1, \ldots, p_M \rangle .$$

where p_k is a stabilizer generator of Π and $k \in \{1, \ldots, M \leq n\}$. We analyse the update rules first in the commuting case ($[\rho, \Pi] = 0$) and then in the general case.

1. For *non-disturbing* (commuting) measurements, the state after measurement ρ' is given by adding the stabilizer generators of the measurement Π and the state ρ. The generating

set is then given by removing the linear dependencies between the two set of generators. For simplicity, we here assume the generators of the state and the generators of the measurement to be linearly independent.

$$\rho' \to \langle g_1, g_2, \ldots, g_N, p_1, p_2, \ldots, p_M \rangle \,. \tag{2.57}$$

This formula means that the state ρ' is now

$$\rho' = \frac{1}{\mathcal{N}} \prod_j^{N^*} \sum_i^{d-1} r_j^i,$$

where $N^* = N + M$ and r_j is a stabilizer generator of ρ', *i.e.* it is either a valid (commuting) generator g_j or p_j.

2. For *disturbing* (non-commuting) measurements (the most general case) the idea is that if we remove the non-commuting factors ρ_j from the state ρ, *i.e.* $[\rho_j, \Pi] \neq 0$, this case reduces to the previous commuting one. We assume the state ρ to have only one non-commuting factor, say ρ_N, which corresponds to the stabilizer generator g_N. The state after measurement ρ' is given by removing the non-commuting generator and adding the remaining ones of the state and measurement.

$$\rho' \to \langle g_1, g_2, \ldots, g_{N-1}, p_1, p_2, \ldots, p_M \rangle \,, \tag{2.58}$$

where we have here considered the case in which no generators coincide. This formula means that the state ρ' is now

$$\rho' = \frac{1}{\mathcal{N}} \prod_j^{N^*} \sum_i^{d-1} r_j^i,$$

where $N^* = N + M - 1$ and r_j is a stabilizer generator of ρ', *i.e.* it is either a valid (commuting) generator g_j or p_j.

To sum up, in the commuting case we add generators of state and measurement to obtain the state after measurement. In the non-commuting case we remove the non-commuting generator of the state and add all the others as in the commuting case. This structure is perfectly analogue

to Spekkens' update rules, which are just motivated by the classical complementarity principle.

2.4.2 Gross' Wigner functions - update rules

Let us consider a stabilizer state $\rho = \rho_1 \cdot \rho_2 \cdot \cdots \cdot \rho_n$, where n is the number of qudits (odd *prime* dimensions), and a measurement Π on the stabilizer state $\Pi = \Pi_1 \cdot \Pi_2 \cdots \Pi_m$, where, in general, $m \leq n$. Let us assume $m = n$ in order to consider "total" measurements (not only to a part of the state).

Theorem 4. Commuting case. *Let us assume the state and measurement to commute,* i.e. $[\rho, \Pi] = 0$. *The Wigner function of the state after the outcome σ' of the measurement has occurred is*

$$W_{\rho'}(\lambda) = \frac{1}{N} W_\rho(\lambda) R_\Pi(\lambda), \tag{2.59}$$

where $\lambda \in \Omega$ and R_Π denotes the Wigner function (also called response function) associated with the measurement element of Π associated with the outcome σ'. The normalisation factor N is

$$N = \sum_{\lambda \in \Omega} W_\rho(\lambda) R_\Pi(\lambda).$$

Proof. We rewrite the formula (2.59) by replacing the Wigner functions with their definition in terms of Spekkens' subspaces,

$$\delta_{\lambda, V'^\perp + \mathbf{w}'} = \delta_{\lambda, V^\perp + \mathbf{w}} \cdot \delta_{\lambda, V_\Pi^\perp + \mathbf{r}}. \tag{2.60}$$

The proof is straightforward. The RHS is one if and only if both the deltas are one; this means that λ has to belong simultaneously to $V^\perp + \mathbf{w}$ and $V_\Pi^\perp + \mathbf{r}$, *i.e.* $\lambda \in (V^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r})$. If we recall equation (2.37) (and figure 2.8), we see that

$$(V^\perp + \mathbf{w}) \cap (V_\Pi^\perp + \mathbf{r}) = (V'^\perp) + \mathbf{w}',$$

and we can conclude that the RHS of equation (2.60) is one if and only if the LHS is one. At this point we can insert the normalisation factors on the RHS and the LHS. These guarantee that $\sum_{\lambda \in \Omega} W_{\rho'}(\lambda) = 1$ and the uniformity as expected. □

In the commuting case the update rule in SQM consists of the *addition* of the stabilizer generators of state and measurement (equation (2.57)). In ST the update rule consists of the *intersection* of the perpendicular isotropic subspaces (equation (2.37)). In GT addition and intersection translate into the *product* of the Wigner functions (equation (2.59)). In particular this stage consists of introducing zeros to the Wigner function in correspondence of the addition of generators to the subspace of known variables V (and so removing generators from the subspace $V^{\perp}$). We call this process - where we *learn* information about the state - the *localization stage*.

Theorem 5. Non-commuting case. *Let us assume the measurement, in general, not to commute with the state,* i.e. $[\rho, \Pi] \neq 0$. *The Wigner function of the state after the outcome σ' of the measurement has occurred is*

$$W_{\rho'}(\lambda) = \frac{1}{N} \sum_{t \in V_{other}} W_{\rho}(\lambda - t) R_{\Pi}(\lambda), \tag{2.61}$$

where $\lambda \in \Omega$, V_{other} is the set spanned by the non-commuting generators of Spekkens' subspace V associated to the state ρ. The normalisation factor N is

$$N = \sum_{\lambda \in \Omega} \sum_{t \in V_{other}} W_{\rho}(\lambda - t) R_{\Pi}(\lambda).$$

Note that we could have stated the theorem in terms of stabilizer generators instead of Spekkens' generators. The former being related to the latter as follows,

$$g_j = \hat{W}(J^{-1} \mathbf{\Sigma}_j), \tag{2.62}$$

where J is the usual invertible matrix used in symplectic geometry, Σ_j are Spekkens' generators and g_j the corresponding stabilizer generators. The relation (2.62) follows from the relation between ST and GT previously described, where the bridge between the two formulations is given by the matrix J.

Proof. In general the state after measurement in quantum mechanics (up to a normalization) is $\rho' = \Pi \rho \Pi$. If $[\rho, \Pi] = 0$ then $\rho' = \rho \Pi$.

In order to simplify the proof, let us assume the case of only one non-commuting generator, say ρ_n. In the present case we know, from the structure of SQM and Spekkens' update rules (adding the commuting factors between state and measurement and removing the non-commuting ones), that the state after measurement is $\rho' = \rho^*\Pi$, where $\rho^* = \rho_1 \cdot \ldots \rho_{n-1}$. This means that we can write the state after measurement as a product of two commuting terms: ρ^* and Π. Therefore we can write the Wigner function of ρ' according to the product rule for the commuting case (equation (2.59)):

$$W_{\rho'}(\lambda) = \frac{1}{N} W_{\rho^*}(\lambda) R_\Pi(\lambda),$$

where $N = \sum_\lambda W_{\rho^*}(\lambda) R_\Pi(\lambda)$. We want now to prove that equation (2.61) is equal to the latter. This means we want to prove the following:

$$W_{\rho'}(\lambda) = \sum_{t \in V_{other}} W_\rho(\lambda - t) R_\Pi(\lambda) = W_{\rho^*}(\lambda) R_\Pi(\lambda).$$

We can simplify the terms $R_\Pi(\lambda)$, thus getting

$$\sum_{t \in V_{other}} W_\rho(\lambda - t) = W_{\rho^*}(\lambda). \tag{2.63}$$

At this point, in order to prove the above theorem, we rewrite the formula (2.61) by replacing the Wigner functions with their definition, *i.e.* Kronecker deltas,

$$\sum_{t \in V_{other}} \delta_{\lambda - t, V^\perp + \mathbf{w}} = \delta_{\lambda, V^\perp_{commute} + \mathbf{w}}, \tag{2.64}$$

where $V^\perp_{commute} = V^\perp \oplus V_{other}$. Note that we have removed the response function of the measurement. We now want to see that the LHS of equation (2.64) is different from zero exactly when the RHS is. The LHS is different from zero when at least one $t \in V_{other}$ is such that $\lambda - t \in V^\perp + \mathbf{w}$. The latter corresponds to $\lambda \in V^\perp + \mathbf{w} + \mathbf{t}$. This means that $\lambda \in V^\perp \oplus V_{other} + \mathbf{w}$, *i.e.* $\lambda \in V^\perp_{commute} + \mathbf{w}$, which is precisely what makes the RHS different from zero.

□

In the most general non-commuting case, in addition to the localization stage, in SQM we also have to *remove* the non-commuting generators from the state (equation (2.58)). In ST this consists of the *union and shifts* in the perpendicular subspace (equation (**??**)). In GT removal and union translate into the *averaging out* of the Wigner function (equation (2.61)). In particular this stage consists of introducing ones to the Wigner function in correspondence of the removal of generators from the subspace of known variables V (and so adding generators to the subspace $V^{\perp}$). We can think of this process as the one where, after having learned some information in the localization stage, we need to forget something, otherwise we would get too much information about the ontic state, which is forbidden by the classical complementarity principle. This also explains why non-commuting measurements are also called *disturbing* measurements. We call this forgetting-part of the process the *randomization stage*. Finally note that the general-case formula (2.61) reduce to the product rule (2.59) in the commuting case. Figure 2.13 summarises the update rules in the three theories in prime dimensions.

In the non-prime dimensional case, we can rephrase all the reasonings already done in ST in terms of Wigner functions.

Lemma 3. *The Wigner function $W_{cg}(\lambda)$ of the coarse-graining observable $O_{cg} = a_1X_1 + b_1P_1 + \cdots + a_nX_n + b_nP_n = D(a'_1X_1 + b'_1P_1 + \cdots + a'_nX_n + b'_nP_n)$ with outcome σ_{cg}, can be written in terms of the Wigner functions $W^{(j)}_{fg}(\lambda)$ of the associated fine graining observables $O^{(j)}_{fg} = a'_1X_1 + b'_1P_1 + \cdots + a'_nX_n + b'_nP_n$ with outcomes $\sigma^{(j)}_{fg}$ as*

$$W_{cg}(\lambda) = \frac{1}{\bar{D}} \sum_{j=0}^{\bar{D}-1} W^{(j)}_{fg}(\lambda). \tag{2.65}$$

Proof. First of all the normalisation factor $\frac{1}{\bar{D}}$ is due to the fact that we are adding $\bar{D}$ Wigner functions, each of them having a normalisation factor of $\frac{1}{d}$, since they are Wigner functions of maximally isotropic subspaces (of dimension d). The proof of the rest of the formula is straightforward. According to the definition of Wigner functions, we need to prove that

$$\delta_{V^{\perp}_{cg}+\mathbf{r}_{cg}} \propto \sum_j \delta_{V^{\perp}_{fg}+\mathbf{r}^{(j)}_{fg}}. \tag{2.66}$$

From the decomposition of the isotropic subspaces and shift vectors in Spekkens' model, equa-

	Non-disturbing Measurements (Localization stage) $[\rho, \Pi] = 0$	Disturbing Measurements (Localization + randomization stage) $[\rho, \Pi] \neq 0$
Stabilizer Quantum Mechanics	$\rho \to \langle g_1, \ldots, g_N \rangle$ $\Pi \to \langle p_1, \ldots, p_M \rangle$ Add generators $\rho' \to \langle g_1, g_2, \ldots, g_N, p_1, p_2, \ldots, p_M \rangle$	$\rho \to \langle g_1, \ldots, g_N \rangle$ $\Pi \to \langle p_1, \ldots, p_M \rangle$ Add generators / Remove g_N $\rho' \to \langle g_1, g_2, \ldots, g_{N-1}, p_1, p_2, \ldots, p_M \rangle$
Spekkens Theory	$V' = V \oplus V_\Pi$ $V'^{\perp} = V^{\perp} \cap V_\Pi^{\perp}$ $\mathbf{w}' = \mathbf{w} + \sum_i^n \boldsymbol{\Sigma}_i'^T (\mathbf{r} - \mathbf{w}) \gamma_i$	$V' = V_{commute} \oplus V_\Pi$ $V'^{\perp} = (V^{\perp} \oplus V_{other}) \cap V_\Pi^{\perp}$ $\mathbf{w}' = \mathbf{w} + \sum_i^n \boldsymbol{\Sigma}_i'^T (\mathbf{r} - \mathbf{w}) \gamma_i$
Wigner Functions	$W_{\rho'}(\lambda) = \frac{1}{N} W_\rho(\lambda) R_\Pi(\lambda)$	$W_{\rho'}(\lambda) = \frac{1}{N} \sum_{\mathbf{t} \in V_{other}} W_\rho(\lambda - \mathbf{t}) R_\Pi(\lambda)$

Figure 2.13: **Equivalence of three theories in odd dimensions in terms of measurement update rules: Spekkens' toy model, stabilizer quantum mechanics and Gross' theory.** The table above shows the update rules in the three mentioned theories in odd prime dimensions both for the commuting and the more general non-commuting case. In SQM the update rules were already known: if state and measurement commute then the final state ρ' is given by the stabilizer generators of both ρ and Π. If, more generally, they do not commute, we also need to remove the non-commuting generators (g_N in the table above) of the original state. In Spekkens' model the update rules for the epistemic state $(V, \mathbf{w})$ and the measurement $(V_\Pi, \mathbf{r})$ have the same structure of the ones in SQM. At the level of the perpendicular subspaces, the update rules involve the intersection and also the direct sum (union and shifts) of the state perpendicular subspace $V^{\perp}$ with the non-commuting subspace V_{other}. The update rules for the representative ontic vector $\mathbf{w}$ are written in terms of the measurement generators Σ_i' and the vector $\boldsymbol{\gamma}_i$ such that $\boldsymbol{\Sigma}_i'^T \boldsymbol{\gamma}_i = 1$. The table above does not show, for aesthetics reasons, the influence of the shift vectors $\mathbf{w}, \mathbf{r}, \mathbf{w}'$ on the perpendicular subspaces. The actual update rule would be $V'^{\perp} = (V_{commute}^{\perp} + \mathbf{w} - \mathbf{w}') \cap (V_\Pi^{\perp} + \mathbf{r} - \mathbf{w}')$. In Gross' theory the update rule for Wigner functions of stabilizer states are given by a simple product of the Wigner functions associated to the state, W_ρ, and measurement, R_Π, in the commuting case, and an averaging over the non-commuting subspace V_{other} in the general case. It is easy to see that the latter formula reduces to the previous in the commuting case (*i.e.* $V_{other} = \{(0,0)\}$).

tions (2.51) and (2.53), we already know that $V_{cg}^{\perp}+\mathbf{r}_{cg} = V_{fg}^{\perp}\oplus V_D+\mathbf{r}_{cg} = V_{fg}^{\perp}+\sum_{j=0}^{\bar{D}-1}(\mathbf{r}_{cg}+j\mathbf{v})$, which exactly proves that the RHS of (2.66) is one if and only if the LHS is one. □

From the above construction and theorem 3 we can immediately write the Wigner function of a stabilizer state after a coarse-graining measurement, thus generalising theorem 5.

Theorem 6. *Given the state ρ of n-qudit systems, where the dimension d is a non-prime integer, and the (non-commuting) measurement Π, the Wigner function of the state ρ' after the outcome σ_{cg} of the measurement has occurred is*

$$W_{\rho'}(\lambda) = \frac{1}{N}\frac{1}{\bar{D}}\sum_{t\in V_{other}}\sum_{j=0}^{\bar{D}-1} W_{\rho}(\lambda - t)R_{fg}^{(j)}(\lambda), \tag{2.67}$$

where $\lambda \in \Omega$, V_{other} is the set spanned by the non-commuting generators of Spekkens' subspace V associated to the state ρ. The response function of the $j-th$ fine-graining measurement element associated with the outcome $\sigma_{fg}^{(j)}$ is denoted by $R_{fg}^{(j)}$. The normalisation factor N is

$$N = \sum_{\lambda\in\Omega}\sum_{t\in V_{other}} W_{\rho}(\lambda - t)R_{\Pi}(\lambda),$$

where $R_{\Pi}(\lambda) = \frac{1}{D}\sum_{j=0}^{\bar{D}-1} R_{fg}^{(j)}(\lambda)$.

Proof. We just need to apply lemma 3 to the response function of the coarse graining measurement of theorem 5. □

Figure 2.14 summarises the update rules in ST and Gross' theory in prime and non-prime dimensions.

2.5 Discussion

The importance of completing Spekkens' theory with update rules to determine the state after a sharp measurement depends upon their application in future works. In particular we think that it would be interesting to explore how quantum computational schemes can be represented by ST. In order to do this it is appropriate to first characterise ST in terms of its computational power.

	Prime dimensional systems	Non-prime dimensional systems
Spekkens Theory	$V'^{\perp} = (V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap (V^{\perp}_{\Pi} + \mathbf{r} - \mathbf{w}')$ $\mathbf{w}' = \mathbf{w} + \sum_{i}^{n} \Sigma'^{T}_{i} (\mathbf{r} - \mathbf{w}) \gamma_i$	$V'^{\perp} = \bigcup_{j=0}^{\bar{D}-1} [(V^{\perp}_{commute} + \mathbf{w} - \mathbf{w}') \cap (V^{\perp}_{fg} + \mathbf{r}^{(j)}_{fg} - \mathbf{w}')]$ $\mathbf{w}' = \mathbf{w}'_j = \mathbf{w} + \sum_{i=0}^{n} \Sigma'^{T}_{i} (\mathbf{r}^{(j)}_{fg} - \mathbf{w}) \gamma_i$
Wigner Functions	$W_{\rho'}(\lambda) = \frac{1}{N} \sum_{\mathbf{t} \in V_{other}} W_{\rho}(\lambda - \mathbf{t}) R_{\Pi}(\lambda)$	$W_{\rho'}(\lambda) = \frac{1}{N} \frac{1}{\bar{D}} \sum_{\mathbf{t} \in V_{other}} \sum_{j=0}^{\bar{D}-1} W_{\rho}(\lambda - \mathbf{t}) R^{(j)}_{fg}(\lambda)$

Figure 2.14: **Measurement update rules in Spekkens' toy model and Gross' theory in prime and non-prime dimensions.** The table above shows the update rules (for the general non-commuting case) of ST and Gross' theory in prime dimensions, first column, and non-prime dimensions, second column. The former have been already depicted in table 2.13. The latter regard the case of a coarse-graining measurement observable O_{cg}. In terms of perpendicular subspaces the update rules consist of the union of the updating subspaces of the original state $(V, \mathbf{w})$ with each of the $\bar{D}$ individual fine-graining observables $(V_{fg}, \mathbf{r}^{(j)}_{fg})$. The updated shift vector $\mathbf{w}'$ is just one of the updated shift vectors $\mathbf{w}'_j$ of the state with the fine-graining observables. In terms of Wigner functions, the union translates into a sum of $\bar{D}$ terms, and the response functions of the fine-graining observables are denoted as $R^{(j)}_{fg}$.

We can perform a simple analysis of the computational complexity of simulating Spekkens' theory on a classical computer following the same approach as Aaronson and Gottesman's analysis of the simulation of stabilizer circuits [100]. In ST, each epistemic state is described by the isotropic subspace of known variables V, which is defined by n generators, and the shift vector $\mathbf{w}$, which attributes n values to the n variables. Each generator is specified by $2n$ components. Therefore we need $2n^2 + n$ digits to specify an epistemic state $(V, \mathbf{w})$. To find the perpendicular subspace $V^{\perp}$, we need a further n^2 operations to check all the inner products

between the generators.

Simulating dynamics requires computing symplectic affine transformations involving about $(2n)^2+n \simeq n(n+1)$ digits for each generator of the epistemic state that has $2n \simeq n$ components, since the product between a matrix and a vector involves $\mathrm{O}(n)^2$ modular arithmetic operations and the affine translation $2n$ operations. Therefore the total is $n(n+1)\cdot n \simeq n^3$.

It should be possible to find more efficient algorithm using some of the ideas in [100]. However we are not aiming to optimise this simulation complexity in the current work, just show that it is classically efficient. The update rules for the measurements in the prime case (2) involve first adding the generators of the subspaces V and V_Π and then removing the non-commuting ones (this involves to check their symplectic inner product, which means about n^3 operations). In total we would have $n^3 + 2n = n(n^2+2) \simeq n^3$ operations for finding V', the isotropic subspaces of known variables after the measurement. The updated shift vector involves the sum of two inner products between vectors of $2n$ components, which roughly means n^3 operations. In the non-prime case (3) further operations are needed, namely the ones to recover the degeneracy factor $\bar{D}$, which consist of dividing the $2n$ components of the generators by each of the d possible integer factors and then do the division again for surely less than d times, which means no more than $2n \cdot d^2$ operations. A final operation of checking whether the results of the divisions of the $2n$ components give the same value must be considered. It implies another factor of $2nd$. This allows us to compute the operations to perform the union of the perpendicular subspaces in (3), *i.e.* $\bar{D}n \simeq (nd)^3$ operations. This approximate analysis wants just to show that, even with basic simulation schemes, the computational complexity to perform a classical simulation of ST is polynomial in the number of systems. This is in line, as expected, with the computational power of SQM.

The next chapter is dedicated to an application of ST in the above direction, where it is used as a non-contextual hidden variable model that represents the classically simulatable part of some state-injection schemes of quantum computation. ST and its subtheories which are operationally equivalent to subtheories of QM play the role of witnesses of non-negativity of the Wigner functions and non-contextuality, and are therefore used as a unifying framework for state injection schemes where negativity and contextuality are resources for universal quantum computation, similarly to [54–57, 79–81]. Moreover, our framework in non-prime dimensions

could be used to extend some of the cited results, such as [54].

The result about the equivalence between ST and SQM and the associated update rules in prime and *non-prime* odd dimensions can provide a powerful new way to use and analyse SQM in non-prime dimensions, about which almost nothing is known. For example we are now facilitated to state, given a set of commuting Pauli operators, whether the joint eigenstate that they represent is pure. In non-prime dimensions the latter issue is not trivial because for coarse-graining observables the number of independent generators is not equal to the number of observables. However, from our construction to decompose coarse-graining into fine-graining observables, we know that the number of independent generators is equal to the number of fine-graining observables. Therefore if the set of commuting Pauli operators has the number of independent generators that equals the number of fine-graining observables, then the state is pure. Indeed fine-graining observables are associated to pure states. In addition, in the field of quantum error correction it could be interesting to study if the coarse-graining observables have any usefulness. The course-grained observables considered here are an example of degenerate observables. Degenerate observables, such as a parity measurement, play a central role in quantum error correction theory. The degeneracy means that errors can be detected without collapsing the logical state. It would be interesting to investigate whether the course-grained observables in compound dimension SQM have any utility for novel forms of quantum error correction.

Finally, the enforced equivalence of SQM, ST and Gross' theory in odd dimensions can be exploited to address a given problem from different perspectives, where, depending on the cases, one theory can be more appropriate than another. An example is the already mentioned one of addressing protocols based on SQM with ST instead of SQM or Wigner functions.

2.6 Conclusion

Spekkens' toy model is a very powerful model which has led to meaningful insights in the field of quantum foundations and that seems to have interesting applications in the field of quantum computation. We have extended it from prime to arbitrary dimensional systems and we have derived measurement update rules for systems of prime dimensions when the state

and measurement commute, equations (2.37)(2.38), when they do not, equations (2.42)(2.38), and for systems of non-prime dimensions (theorem 3). These results directly derive from the basic axiom of the theory: the classical complementarity principle. The latter characterises a structure for the update rules which is the same as in stabilizer quantum mechanics: the state after measurement is composed by the generators of the measurement and the compatible (*i.e.* commuting) generators of the original state.

Spekkens showed the equivalence between SQM and ST in odd prime dimensions via Gross' Wigner functions. We have extended this result to all odd dimensions and we have translated the update rules of ST in terms of Wigner functions (theorems 4, 5, 6). We stress again that Spekkens' model and our measurement update rules hold in all dimensions, in even dimensions too. However the equivalence between ST and SQM only holds in odd dimensions. The main reason is that SQM in even dimensions shows contextuality, while ST does not. One of the main future challenges is to find a hidden variable toy model which is also equivalent to qubit SQM.

We treat the problem with systems of non-prime dimensions, which arises from the problem of defining an inverse in $\mathbb{Z}_d$, by decomposing the problematic (coarse-graining) observables in terms of the non-problematic (fine-graining) ones. This approach naturally suggests the form of the update rules. By comparing the update rules in the three mentioned theories we highlight the beauty and the elegance of this equivalence, where addition and removal of generators in SQM correspond to intersection and union in ST and product and randomization in GT. This correspondence is schematically depicted, for the prime-dimensional case, in table 2.13. The non-prime case correspondence is represented in table 2.14. We believe that the fresh perspective gained by moving from one theory to another can give powerful new tools for new insights in the field of quantum computation.

Chapter 3

State-injection schemes of quantum computation in Spekkens' toy theory

In the previous chapter we showed the operational equivalence between ST and qudit SQM by representing the latter through Gross' Wigner functions [43]. These turn out to be exactly equivalent to Spekkens' epistemic states and measurements. The measurement update rules are consistent and positiveness-preserving. Clifford gates are mapped into consistent symplectic affine transformations [40].

In the case of qubits the above equivalence does not hold. The toy theory is non-contextual by construction, while qubit SQM shows state-independent contextuality, as witnessed by the Peres-Mermin square argument [46–48]. This is reflected into the impossibility of finding a non-negative Wigner function that maps non-negatively qubit SQM into ST [44, 45]. Even if these two theories are not operationally equivalent, some restricted versions of them do show the same statistics. Our aim is to identify the subtheories of ST that are operationally equivalent to subtheories of qubit SQM. We define subtheories of ST compatible with subtheories of SQM as closed subtheories of quantum mechanics whose states and measurements are non-negatively represented by covariant Wigner functions.

In this chapter we use Spekkens' subtheories to provide an application in the field of quantum computation. More precisely, we relate them with state-injection schemes of quantum computation [51]. The latter constitute nowadays one of the leading models to implement fault

tolerant universal quantum computation (UQC). These schemes are composed by a "free" part, which consists of quantum circuits that are efficiently simulatable by a classical computer (usually stabilizer circuits), and by magic resources (that are usually distilled from many copies of noisy states through magic state distillation [53]) that boost the computation to universal. In 2014 Howard *et al.* [54] proved that in a state-injection scheme of qudits (odd prime dimensions), with the free part composed by stabilizer circuits, the contextuality possessed *solely* by the magic resource is a necessary resource for universal quantum computation. In terms of systems of dimensions 2 a similar result due to Delfosse *et al.* [55] holds for rebits, where the classical non-contextual free part is composed by Calderbank-Steane-Shor (CSS) circuits [87]. However, as already pointed out, an analogue version of Howard's result for qubit cannot be found, since qubit SQM is already contextual. Nevertheless it has been proven [56] that in any state-injection scheme of qubits where we get rid of the state-independent contextuality (*e.g.* Peres-Mermin square), the contextuality possessed *solely* by the magic resource is necessary for universal quantum computation. A more complete version of this result is also treated in [57], where a general framework for state-injection schemes of qubits with contextuality as a resource is provided.

More precisely, in [57], Raussendorf *et al.* develop a framework for building non-negative and non-contextual subtheories of qubit SQM from the choice of the phase function defining the Weyl operators and consequently the Wigner functions (see equation (3.10)). Furthermore, in [83], Wallman and Bartlett address the issue of finding the subtheories of qubit quantum mechanics that are non-negative in certain quasi-probability representation (and so are classically simulatable and correspond to non-contextual ontological models). They construct the so-called $8-$state model, which can be seen as a generalisation of ST with an enlarged ontic space. The non-negativity for states and measurements is guaranteed by considering both the possible Wigner representations of a qubit [45].

It is important to point out that in the mentioned frameworks of [56] and [57], the definition of state-injection schemes is broader than the one we consider here and it also includes schemes based on measurement-based quantum computation with cluster states [15]. Measurement-based quantum computation is a model of quantum computation, alternative to the standard circuit model, based on local one-qubit measurements on some particular resource entangled

states [15]. We here consider only state-injection schemes as developed by Zhou *et al.* in [51] – defined in 3.1.1 – like the ones in [54] and [55], and we show that Spekkens' subtheories are an intuitive and effective tool to treat these cases. We first use Spekkens' subtheories to represent the non-contextual free part of the known examples of state-injection schemes, both for qubits and qudits, where contextuality arises as a resource [54, 55]. These can be unified in the following framework (figure 3.1): *Spekkens' subtheory + Magic state(s)* $\rightarrow$ *UQC*. Secondly, we prove in theorem 11 that qubit SQM can be obtained from a Spekkens' subtheory via state-injection since all its objects that do not belong to the Spekkens' subtheory, namely non-covariant Clifford gates, can be injected, where the circuit needed for the injection is always made of objects belonging to the Spekkens' subtheory. This means that Spekkens' subtheories contain all the tools for performing state-injection schemes of quantum computation. There is no need to consider bigger non-covariant subtheories in the free part.

The proof of theorem 11 suggests a novel state-injection scheme, where contextuality is a resource, based on injection of CCZ states. State-injection schemes with the related Toffoli ($CCNOT$) gates are already known [101–107], but our scheme differs from them as our non-contextual free part of the computation – a strict subset of the CSS rebit subtheory considered in [55] – is such that it is not possible to remove any object from it without denying the possibility of obtaining universal quantum computation via state-injection. The price to pay for this minimality is the injection of the control-Z state, $CZ\left|++\right\rangle$, too (which also provides the Hadamard gate). By analysing this example, we can associate different proofs of contextuality to different state-injections of non-covariant gates. More precisely, we show how the Clifford non-covariant CZ gate (as well as the phase gate S) can provide proofs of the Peres-Mermin square contextuality and the GHZ paradox. Moreover, the injection of the $T\left|+\right\rangle$ magic state, where T is the popular $\frac{\pi}{8}$ non-Clifford gate, allows, in addition to the previous proofs of contextuality (as $T^2 = S$), also to obtain the maximum quantum violation of the CHSH inequality [59].

In the reminder of the chapter we start, in section 3.1, by covering some background material on state-injection schemes, contextuality and the known results on contextuality as a resource for quantum computation. We then provide the definition of a Spekkens' subtheory in section 3.2 and we prove that Howard's and Delfosse's cases for qudits [54] and rebits [55], respectively, fit in our framework where the free parts of the computation, qudit SQM and CSS rebits,

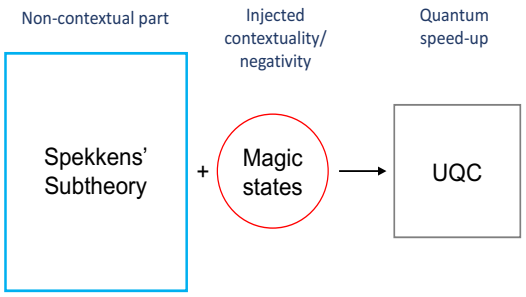


Figure 3.1: **Computational scheme.** Schematic representation of the computational scheme treated in this chapter.

respectively, are Spekkens' subtheories. We then set the instructions to construct a Spekkens' subtheory from the choice of a non-negative Wigner function in section 3.3. We do so in line with [57] and we find that the main difference from Raussendorf *et al*'s formulation (and also from the 8-state model) consists of demanding for the covariance of the Wigner function with respect to the allowed gates. By exploiting this comparison we then prove theorem 11. It basically shows that any state-injection scheme can be obtained from our framework, since any object not present in the considered Spekkens' subtheory can be injected by using an injection scheme made of objects in the Spekkens' subtheory. In section 3.4 we provide a novel example of state-injection for qubits based on CCZ magic states. We analyse the presence of different proofs of contextuality in correspondence of different state injected gates in section 3.5 and we recap all the results and the future directions in the conclusion section.

3.1 Background

This section does not contain original material. The references that have been used will be specified in the corresponding subsections.

3.1.1 State-injection schemes of quantum computation

Building a large-scale quantum computer is a challenging task, mostly due to the fragile coherence of quantum systems. The field of quantum error correction provides techniques and methods to face this problem and protect the information encoded in the quantum system [38]. In particular, it is important to protect the quantum information not only when it is stored

or transmitted, but also when it dynamically undergoes computation. This consideration leads to the concept of fault-tolerant quantum computing, that we here introduce. Fault tolerance is the property that enables a system to continue performing the operations it is meant to perform, even if some parts of the system fail, *i.e.* involve errors. The idea is to prevent single errors on each component, *e.g.* input systems prepared in certain states, unitary gates and measurements, from spreading. In particular, it is required that the error probability p is below a certain constant threshold p_{th}, for the quantum computation to be efficient and reliable, as stated by the threshold theorem (*verbatim* from [38]),

Theorem 7. *A quantum circuit containing $p(n)$ gates may be simulated with probability of error at most ε using*

$$O(poly(\log p(n)/\varepsilon)p(n))$$

gates on hardware whose components fail with probability at most p, provided p is below some constant threshold, $p < p_{th}$, and given reasonable assumptions about the noise in the underlying hardware.

There are several approaches to compute error thresholds, which depend upon the error model, the choice of error correcting code and the choice of the fault tolerant construction for gates and error-detecting measurements.[1] Moreover, once it is guaranteed that the errors do not spread in accordance with the threshold theorem, they can be suppressed through the concatenation technique [38], which is a recursive method to encode each qubit, starting from the original quantum circuit, in quantum codes. This technique exponentially decreases the probability of failure.

We have just seen that settings where the components are fault tolerantly implemented ensure the computation to work. With this respect, SQM, already extensively defined in subsection 2.1.2, comes in again, as it involves components that can all be fault-tolerantly realised. We recall that it is the subtheory of quantum mechanics composed by common eigenstates of Pauli operators, Clifford unitaries and Pauli observables, and it describes many of the mostly used error correcting codes [85–87]. We now describe the central role that SQM plays in quan-

[1] A couple of examples for well-studied codes: analytic thresholds for the Steane code [86] are of the order of $p_{th} = 10^{-5}$, while numerical thresholds are of the order of $p_{th} = 10^{-3}$; the Bacon-Shor code [108] leads to an analytic threshold of the order of $p = 10^{-4}$.

tum computation, by first stating the theorem that shows its computational power. [2]

Theorem 8. *Any quantum circuit on n qudits acting on an initial stabilizer state consisting of a polynomial in n Clifford group gates, and Pauli observable measurements, can be effectively simulated in poly(n) gates on a classical computer.*

The Gottesman-Knill theorem states that stabilizer circuits, even if they contain systems in maximally entangled states, such as Bell states, are easy to simulate classically and thus not enough for universal quantum computing. Nevertheless we can still use stabilizer circuits to perform universal quantum computation by adding a single non-Clifford unitary, as proven in the Nebe-Rains-Sloane theorem [109] (here paraphrased):

Theorem 9. *Given a non-Clifford unitary U, the set*

$$\left\langle U, U^\dagger, CNOT, H, S \right\rangle \tag{3.1}$$

is approximately universal, i.e. *any unitary can be approximated to arbitrary accuracy by a gate sequence contained in the set.*

This means that, by using the set of gates in equation (3.1), we can simulate any other set, thus gaining approximate universality. The simulation is shown to be efficient by the Solvay-Kitaev theorem [110]. The single-qubit non-Clifford gate that is usually used to achieve UQC is called the T gate. It is defined as a rotation of $\frac{\pi}{4}$ around the z axis in the Bloch sphere,

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{pmatrix}. \tag{3.2}$$

Notice that the phase gate S defined in equation (2.19) is equal to T^2.

In this chapter we consider state-injection schemes of quantum computation as developed in [51]. They represent one of the leading models for fault tolerant universal quantum computation when combined with the magic state distillation procedure due to Bravyi and Kitaev in [53]. The latter allows one to distill non-stabilizer magic states from noisy copies of quantum states with a high threshold for the error rate (about 14.6%), given a setting where only Clifford

[2] We refer the reader to [84, 100] for a more rigorous statement and the proof of the theorem.

gates are fault tolerant. This is extremely important because most of the popular codes do not have fault tolerant non-Clifford gates [86, 111], and, even when they have, they show a poor threshold [112]. The key idea of state-injection is that a non-Clifford gate can be implemented with the combination of the magic state, a Clifford group circuit and Pauli measurement. More precisely, as was proved in [51], we can define state-injection schemes for implementing any diagonal unitary gate U as follows.

Definition 1 (Zhou-Leung-Chuang state-injection [51])**.** *Given a n-qubit unitary gate U, we say that it is achieved via **state injection**, if it is implemented via a circuit of the form in figure 3.2 comprising the elements*

- *The injected state $U|+\rangle^{\otimes n}$.*
- *n CNOT gates, applied transversally.*
- *n Pauli Z measurements, with the output of the jth measurement denoted as $s_j = (-1)^{m_j}$, where $m_j \in \{0, 1\}$.*
- *The correction gate $UX^mU^\dagger$, where $m = m_1 \dots m_n$ is the bitstring of meausurement outcomes and $X^m = X^{m_1} \otimes \cdots \otimes X^{m_n}$.*

Figure 3.2 depicts the state-injection scheme just defined.

3.1.2 Contextuality

In the previous subsection we introduced state-injection schemes of quantum computation, that achieve quantum universality with the injection of magic states. It is natural to ask which quantum ingredients must be present in these states to allow for the quantum computational speed-up. We here provide the definition of contextuality, which is one of the main features studied for explaining the source of quantum power and one of the main characters of this work.

The original notion of contextuality arose in 1967 due to Kochen-Specker's no-go theorem [33]. We here follow [113] for the statement.

Theorem 10. *In a Hilbert space of dimension $d \geq 3$, it is impossible to associate definite numerical values, 1 or 0, with every projection operator P_k, in such a way that, if a set of*

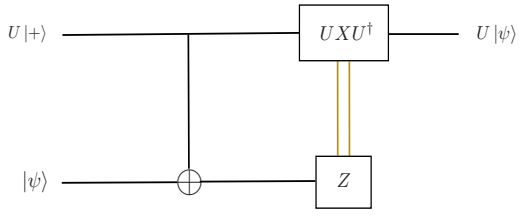


Figure 3.2: **State-injection schemes of quantum computation.** The state-injection schemes that we consider in this work are the ones developed by Zhou, Leung and Chuang in [51]. The diagonal gate U can be injected in the circuit by using objects that are allowed in the free part of the computation. The injected state is $U|+\rangle$, which is subjected to a controlled not with the input state $|\psi\rangle$. Conditioned on the outcome of the measurement of the Pauli Z on the state $|\psi\rangle$ after the $CNOT$, the correction $UXU^\dagger$ is applied to the state $U|+\rangle$. At the end we obtain the gate U applied to the input state $|\psi\rangle$.

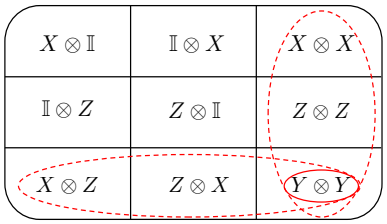
commuting P_k *satisfies* $\sum_k P_k = \mathbb{I}$, *the corresponding values, namely* $v(P_k) = 0$ *or* 1, *also satisfy* $\sum_k v(P_k) = 1$.

This theorem can be rephrased by saying that a hidden variable theory – a theory with variables that unambiguously determine the outcome of each sharp measurement[3] – which reproduces the statistical properties of quantum theory must be *contextual*, *i.e.* the outcome of a projective measurement depends on the other commuting measurements that we perform with it. In the present case this means that, if three operators P_k, P_j and P_l have commutators $[P_k, P_j] = [P_k, P_l] = 0$ and $[P_j, P_l] \neq 0$, the result of a measurement of P_k cannot be independent of whether P_k is measured alone, or together with P_j, or together with P_s. The original proof of the theorem involves 117 vectors in $\mathbb{R}_3$. It was then improved by Peres in [113], by using 33 unit vectors in $\mathbb{R}^3$ grouped in sixteen orthogonal triads (with each unit vector belonging to several triads). However, the simplest proof is due to Cabello in [114], which only involves 18 four-dimensional vectors.

Probably the most simple and intuitive way of expressing the above notion of contextuality is through the Peres-Mermin square argument [47], depicted below.

The square is composed by nine Pauli observables on a two-qubit system. Each row and

[3]When talking about general theories (possibly different or larger than quantum mechanics) sharp measurements correspond to projective measurements in quantum mechanics.



each column is composed by commuting (simultaneously measurable) observables. With the assumption that the functional relation between commuting observables is preserved in terms of their outcomes (*e.g.* if an observable C is the product of two observables A, B, also its outcome c is the product of the outcomes a, b of A, B) and the outcome of each observable does not depend on which other commuting observables are performed with it (non-contextuality), the square shows that it is impossible to assign the outcome of each observable among all the rows and columns without falling into contradiction. For example, if we start by assigning values, say ± 1, to the observables starting from the first (top left) row on, the contradiction can be easily seen when we arrive at the last column and last row (red circles), that bring different results to the same observable YY, as witnessed by the following simple calculation, $(XZ) \cdot (ZX) = YY$, and $(XX) \cdot (ZZ) = -(YY)$. Kochen-Specker contextuality refers to the fact that the outcome of a sharp measurement does depend on the other compatible measurements that we perform with it (*i.e.* on the contexts). The Peres-Mermin square is an example of *state-independent contextuality*, as the argument does not depend on the state the observables are measured on.

Another popular manifestation of contextuality is the Greenberger-Horne-Zeilinger (GHZ) paradox [46]. In this case the key ingredient for the argument is the GHZ state, $\frac{|000\rangle+|111\rangle}{\sqrt{2}}$, and the mutually commuting observables XXX, XYY, YXY, YYX. The GHZ state is the common eigenstate of these four operators, with the eigenvalues being $+1, -1, -1, -1$ respectively. By considering these observables, the quantum predictions are in conflict with any non-contextual hidden variable model that assigns definite pre-existing values, $+1$ and -1, to the local Pauli observables X, Y. Let us denote these definite values as $\lambda_{x1}, \lambda_{x2}, \lambda_{x3}, \lambda_{y1}, \lambda_{y2}, \lambda_{y3}$ in correspondence of each local Pauli X and Y. The product of the three observables XYY, YXY, YYX, that must yield the outcome -1, in the hidden variable model means the following expression $\lambda_{x1}\lambda_{x2}\lambda_{x3}\lambda_{y1}^2\lambda_{y2}^2\lambda_{y3}^2 = -1$. However this is in neat contradiction with the outcome of XXX which is $+1$, and corresponds to $\lambda_{x1}\lambda_{x2}\lambda_{x3} = +1$. Notice that, unlike the Peres-Mermin square

example, the GHZ paradox is an example of *state-dependent contextuality*, as the argument crucially relies on considering the GHZ state.

We mention that there are ways to quantify contextuality based on the graph theoretic approach [115–118]. The idea is to associate contextual experiments to exclusivity graphs $\mathcal{G}$, where events e_j, that are denoted with conditional probability distributions of outcomes given tests, are associated to the vertices of the graph and mutually exclusive events are connected by adjacent edges. Two events e_i, e_j are mutually exclusive if there exist two joint observables O_i, O_j, each associated to one of the events, that distinguish between them. From the graph it is possible to calculate linear combination of probabilities of a subset of events, $S = \sum_i w_i p(e_i)$, with some weight $w_i \geq 0$, that we can assume to be $w_i = 1$ for every i to simplify the argument. It results that, when considering non-contextual hidden variable models, these expressions are bounded by the independence number of the graph $\alpha(\mathcal{G})$, which is the maximum number of non-connected vertices in the graph $\mathcal{G}$. This result is known as Cabello-Severini-Winter inequality,

$$S \leq \alpha(\mathcal{G}). \tag{3.3}$$

Upper bounds for quantum theories and more general theories exist, but we do not formulate them as we will not study post-quantum theories in the current work. We do not treat the sheaf theoretic approach to contextuality, as developed by Abramsky and Brandenburger [119], that also provides a resource theory for contextuality [120]. In our research that involves state-injection schemes based on SQM, this approach does not look ideal as their notion of *strong* contextuality is present already in qubit SQM, which, as already illustrated, is efficiently simulatable by a classical computer by the Gottesman-Knill theorem 8. Nevertheless we mention a result, obtained from the framework of Abramsky and Brandenburger, showing that contextuality is a necessary resource when computing non-linear Boolean functions in measurement-based quantum computation [78].

We now introduce the generalised notion of contextuality introduced by Spekkens in 2005 [34]. It is defined in the so-called operational framework of a physical theory. According to this framework the role of the operational theory is just to provide the rules to compute the statistics that we observe in the laboratory, without any reference to some physical reality. The

primitive elements of the operational approach are the experimental procedures, *i.e.* lists of instructions to be implemented in the laboratory, usually classified in preparation procedures, transformation procedures and measurement procedures. More precisely, a physical theory has to provide the rules to compute the probability $p(k|P,T,M)$, of some outcome k, given the preparation P, transformation T and measurement M on the system considered. According to this approach, the only thing that matters about quantum mechanics is the Born rule, *i.e.* $p(k|\rho,\varepsilon,O) = \mathrm{Tr}(\varepsilon(\rho)\Pi_k)$, where the preparation procedure is associated with the quantum state ρ, the transformation procedure with the completely-positive trace-preserving map ε and the measurement procedure with the POVM $\{\Pi_k\}$. Two experimental procedures, *e.g.* preparations, P, P' are said to be equivalent if they provide the same statistics:

$$p(k|P,T,M) = p(k|P',T,M) \;\; \forall T, M. \tag{3.4}$$

We denote with $e(P)$ the equivalence class of equivalent preparation procedures. Similarly, we denote with $e(M)$ the equivalence class for measurement procedures and with $e(T)$ the equivalence class for transformation procedures.

A natural way of justifying why an operational theory works is to assume that there exist physical systems that are the subjects of the experiment. These systems have properties independent on the fact that an experiment is performed and on the existence of an observer that knows about them. According to the *ontological model framework* [34] the physical properties of the system are specified, at a given time, in the ontic state of the system, which is represented by a point λ in a measurable set Λ. Ontological models are usually taken as synonyms of *hidden variable models*. However, we adopt the former terminology because, in principle, we could build an ontological model where none of the variables are hidden (see for example [121]). The ontological model of an operational theory associates the experimental procedures to probability distributions on the ontic space Λ. A preparation procedure P prepares an ontic state of the system and is represented by a probability distribution $\mu_P(\lambda)$ over the ontic space, $\mu_P : \Lambda \to [0,1]$ such that $\int \mu_P(\lambda)d\lambda = 1$. A transformation procedure T is represented by transition matrix $\Gamma_T(\lambda',\lambda)$ over the ontic space, $\Gamma_T : \Lambda \times \Lambda \to [0,1]$ such that $\int \Gamma_T(\lambda',\lambda)d\lambda' = 1$. A measurement procedure M with associated outcomes labeled by k is

represented by a set of indicator functions $\{\xi_{M,k}(\lambda)\}_k$ over the ontic space, $\xi_{M,k} : \Lambda \to [0,1]$ such that $\sum_k \xi_{M,k}(\lambda) = 1$. The ontological model framework reproduces the predictions of the operational theory according to the law of classical total probability,

$$p(k|P,T,M) = \int d\lambda' d\lambda \xi_{M,k}(\lambda') \Gamma_T(\lambda', \lambda) \mu_P(\lambda) \tag{3.5}$$

for all P, T and M. As already mentioned, Spekkens' theory is an example of ontological model, where the ontic space is the phase space.

An ontological model of an operational theory is preparation non-contextual if $\mu_P(\lambda) = \mu_{e(P)}(\lambda) \ \ \forall P$. An ontological model of an operational theory is transformation non-contextual if $\Gamma_T(\lambda', \lambda) = \Gamma_{e(T)}(\lambda', \lambda) \ \ \forall T$. An ontological model of an operational theory is measurement non-contextual if $\xi_{M,k}(\lambda) = \xi_{e(M),k}(\lambda) \ \ \forall M, k$. An ontological model of an operational theory is universally non-contextual if it is non-contextual for all the experimental procedures. When considering quantum mechanics, the equivalence classes of preparation procedures are the density operators ρ describing the quantum states. The equivalence classes of transformation procedures are the completely-positive trace-preserving maps ε describing the quantum transformations. The equivalence classes of measurement procedures are the POVM $\{\Pi_k\}$ describing the quantum measurements. It can be shown that a universally non-contextual ontological model of quantum mechanics is impossible (even for just one qubit, unlike Kochen-Specker no-go theorem 10) [34].

Examples of contexts for preparation procedures in quantum mechanics are the convex decompositions of a non-rank-1 quantum state. Examples of contexts for transformation procedures are the convex sums of unitaries that define certain completely-positive trace-preserving maps. Examples of contexts for measurement procedures are the convex decomposition of non-maximal measurements (*i.e.* POVM where at least one element is not rank 1). The contexts considered in the Kochen-Specker definition of contextuality, that only deals with sharp measurements, are a particular kind of the ones considered here. More precisely, if we consider projectors P_k, P_j and P_l with commutators $[P_k, P_j] = [P_k, P_l] = 0$ and $[P_j, P_l] \neq 0$, the non-rank-1 projector P_k has a common diagonalising basis with P_j and another one with P_l. These two different bases provide two different convex decompositions of P_k. In light of this

new definition of contextuality, the original definition by Kochen and Specker manifests a hidden assumption: the indicator functions $\xi_{M,k}$ are always assumed to take values either 0 or 1. This assumption is called *outcome determinism*. Kochen-Specker non-contextuality is Spekkens non-contextuality for sharp measurements plus outcome determinism.

In conclusion, Spekkens' contextuality generalizes the original notion of contextuality by Kochen-Specker, separating it from the notion of outcome determinism and extending it to unsharp measurements, preparations and transformations. Nevertheless, in the next subsection the notion due to Kochen-Specker will be assumed when treating the literature regarding the role of contextuality in state-injection schemes of quantum computation. One of the reasons is that with the generalised notion, also simple classically simulatable theories, like *one* qubit SQM, are contextual [122]. We will provide few examples of quantum computational advantages related to the generalised notion of contextuality in the next chapter, subsection 4.1.2. Figure 3.3 depicts the relations between the notions of non-contextuality defined in this subsection.

3.1.3 Contextuality as a computational resource

In 2014, Howard *et al.* showed that the contextuality possessed *solely* by the magic state is a necessary resource for UQC in state-injection schemes on qudits of odd prime dimensions [54]. More precisely, they proved that a noisy input state must violate Cabello-Severini-Winter inequality (equation (3.3)) in order to be distilled into a magic state. Therefore UQC via magic states implies that the magic state manifests contextuality. The free part of the computation is composed by stabilizer circuits and the inequality is built considering a scenario where the vertices of the graph represent stabilizer rank-1 projectors.

This result triggered a whole field of research focused on studying the role of contextuality in quantum computation. The main question that remains open is to find a result analogue to Howard's for qubits. As already discussed, multi-qubit SQM shows state-independent contextuality, which means that the result of Howard *et al.* cannot be extended to the qubit case. However, it is possible to restrict the free part of the computation to subtheories of SQM and obtain analogous results to the qudit case. In 2015, Delfosse *et al* considered state-injection schemes on rebits, where the free part is composed by the CSS circuits, presenting no state-independent contextuality. More precisely, we define the CSS subtheory, that we denote with

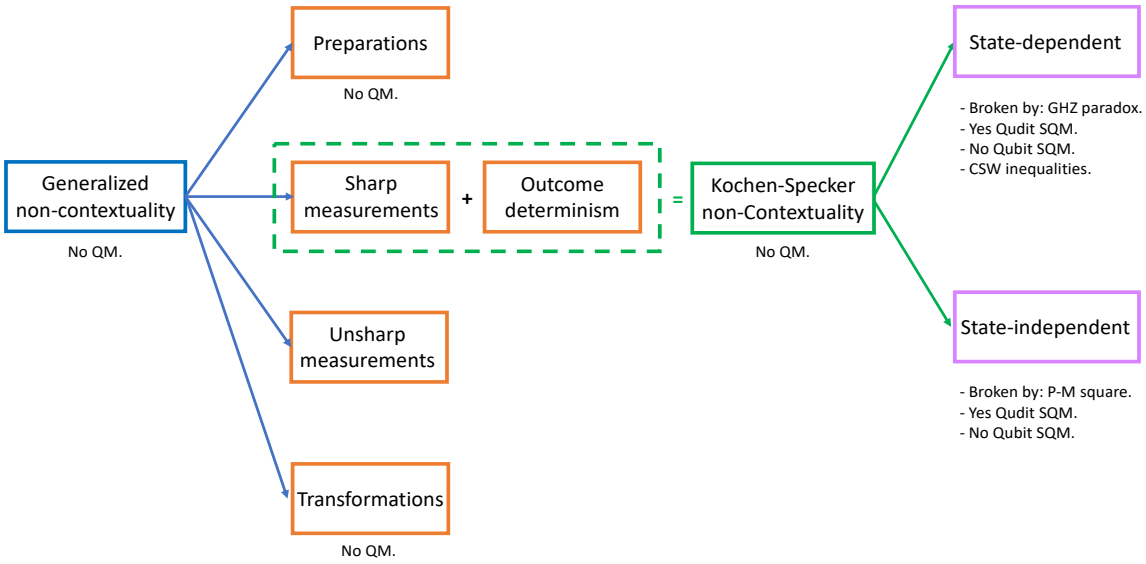


Figure 3.3: **Schematic representation of the relations between different notions of non-contextuality.** The notion of non-contextuality due to Spekkens [34] generalises the notion due to Kochen-Specker [33], separating it from the notion of outcome determinism and extending it to unsharp measurements, preparations and transformations. Non-contextual ontological models of quantum mechanics are impossible for both notions (and also for just preparation and transformation non-contextuality, but not for measurement non-contextuality, as proven in [34]). There exist examples of Kochen-Specker contextuality that, given a set of projective measurements, arise only for certain states. In this case we talk about state-dependent contextuality (the usual example of this is the GHZ paradox). Violations of Cabello-Severini-Winter (CSW) inequality quantify this contextuality. When the contextuality arguments hold for any quantum state (*e.g.* in the Peres-Mermin (P-M) square) we talk of state-independent contextuality. It results that qubit SQM is Kochen-Specker contextual, while odd dimensional qudit SQM is not, due to the intrinsic difference in the structure of the Pauli groups in the two cases.

$(\mathcal{S}_r, \mathcal{T}_r, \mathcal{M}_r)$, where $\mathcal{S}_r, \mathcal{T}_r, \mathcal{M}_r$ are the sets of allowed quantum states, transformations and measurements respectively, as follows. The set $\mathcal{S}_r$, a subset of the stabilizer states, is composed by CSS states [87], *i.e.* stabilizer states $|\psi\rangle$, whose corresponding stabilizer group $S(|\psi\rangle)$ decomposes into an X and a Z part; *i.e.* $S(|\psi\rangle) = S_X(|\psi\rangle) \cup S_Z(|\psi\rangle)$, where all elements of $S_X(|\psi\rangle)$ and $S_Z(|\psi\rangle)$ are of the form $X(x)$ and $Z(p)$, respectively, where $x, p \in \mathbb{Z}_2^n$. CSS states

are the eigenstates of the allowed observables belonging to the set M_r,

$$\mathcal{M}_r = \{X(x), Z(p) | x, p \in \mathbb{Z}_2^n\}. \tag{3.6}$$

The set of allowed transformations is composed by the CSS preserving gates, subset of the Clifford group $C_{n,2}$,

$$\begin{aligned} \mathcal{T}_r &= \{g \in C_{n,2} | g\,|\psi\rangle \in \mathcal{S}_r, \forall\, |\psi\rangle \in \mathcal{S}_r\} \\ &= \left\langle \bigotimes_{i=1}^{n} H_i, CNOT(i,j), X_i, Z_i \right\rangle, \end{aligned} \tag{3.7}$$

where $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. The universal quantum computation is reached by injecting two particular magic states to the free subtheory of CSS rebits just described [55].

The proof that contextuality is a necessary resource for UQC injected with the magic states is based on the construction of a contextuality witness similar to the Cabello-Severini-Winter inequality, in the sense that they both consist of linear operators for which the range of expectation values allowed by quantum mechanics is strictly greater than the one allowed for non-contextual ontological models. Moreover, they also developed a Wigner function that, analogously to Gross Wigner function for odd qudit stabilizer states, is non-negative if and only if it represents CSS states. It is defined as follows:

$$A_r(\lambda) = \frac{1}{2^n} \sum_{T(\lambda') \in \mathcal{A}} (-1)^{[\lambda, \lambda']} \hat{W}(\lambda'), \tag{3.8}$$

where $\hat{W}(\lambda) = Z(p)X(x)$, $\lambda = (x, p)$, and $\mathcal{A} = \{\hat{W}(\lambda) | \mathbf{x} \cdot \mathbf{p} = 0 \mod 2\}$. The set $\mathcal{A}$ is the set of inferred observables. "Inferred" means that these observables may not be directly measurable, but they can be inferred by multiple measurements. For example, in the case of two qubits, the set M_r and $\mathcal{A}$ are $M_r = \{\mathbb{II}, \mathbb{I}X, \mathbb{I}Z, X\mathbb{I}, Z\mathbb{I}, XX, ZZ\}$, and $\mathcal{A} = \{\mathbb{II}, \mathbb{I}X, \mathbb{I}Z, X\mathbb{I}, Z\mathbb{I}, XX, ZZ, XZ, ZX, YY\}$, *i.e.* the set of all rebits observables. Delfosse's Wigner function shares most of the properties that Gross' Wigner function possesses, as defined in the previous chapter. However, it is not factorizable (equation (2.26)), *i.e.* it is composed by phase-point operators of n qubits that are not given by the tensor products of the ones for the single qubit, *e.g.*

$A_r((0,0),(0,0)) \neq A_r(0,0) \otimes A_r(0,0)$.

There is a strict relation between the notions of contextuality and negativity of the Wigner functions (and more general quasi-probability representations). More precisely, as proven in [123], the existence of a non-negative quasi-probability distribution representing a subtheory is equivalent to the existence of a non-contextual ontological model for it. Therefore, finding a non-negative Wigner function for a subtheory guarantees that it is non-contextual. The converse is not true, as it may be that quasi-probability representations different from the Wigner functions are non-negative even if the Wigner functions unavoidably show some negativity. When restricted to Pauli measurements, in the case of qudits of odd dimensions, Delfosse *et al.* proved that the negativity of the Wigner functions and Kochen-Specker contextuality are equivalent notions [124]. This result does not hold for a single qudit, for which there are states that fit in a non-contextual ontological model, even if no non-negative Wigner functions exist.

When considering more generic subtheories of multi-qubit SQM that represent the free part of the computation, the obstacle to be avoided in order to find results showing contextuality as a resource possessed solely by the magic state is again the presence of state-independent contextuality. In 2017, Bermejo-Vega *et al.* showed that, with the condition that the free part of the computation does not show state-independent contextuality, the contextuality of the magic state is a necessary resource for universal quantum computation. The proof they developed is quite different from the previous ones based on the construction of a contextuality witness. It relies on a characterization of non-contextual ontological models of state-injection schemes on qubits, which shows a contradiction if universality of the scheme and the non-contextuality of the magic state are assumed. This result is also treated in the framework of Raussendorf *et al.*, which is based on Wigner functions [57]. We now describe this framework, that will be crucial to state the results of this chapter.

3.1.4 Raussendorf *et al* framework and the 8-state model

As already mentioned and witnessed by the Wigner function of equation (3.8) used for the state-injection scheme on rebits, in the case of qubits a Wigner function that shares all the properties (in particular the non-negativity of the free part of the computation) of the Wigner function defined in equation (2.22) by Gross does not exist. We now introduce a framework

for Wigner functions that includes all the Wigner functions defined so far and that will provide the bridge between subtheories of Spekkens' toy model and subtheories of quantum mechanics. This framework describes a class of Wigner functions parametrized by a function $\gamma : \Omega \to \mathbb{Z}_q$, for some integer q. We start by redefining the Weyl operators of equation (2.11), that we now denote as $\hat{W}^\gamma(\lambda)$, taking into account the parameter γ,

$$\hat{W}^\gamma(\lambda) = w^{\gamma(\lambda)} Z(p) X(x), \tag{3.9}$$

where, as usual, the phase-space point is $\lambda = (x, p) \in \Omega$, and $w^{\gamma(\lambda)} = i^{\gamma(\lambda)}$ in the case of qubits and $w^{\gamma(\lambda)} = e^{-2^{-1}\frac{2\pi i}{d}\gamma(\lambda)}$ in the case of qudits of odd dimensions. The Wigner function of a quantum state ρ is defined as[4]

$$W^\gamma_\rho(\lambda) = \mathrm{Tr}(A^\gamma(\lambda)\rho), \tag{3.10}$$

where the phase-point operator is

$$A^\gamma(\lambda) = \frac{1}{N_\Omega} \sum_{\lambda' \in \Omega} \chi([\lambda, \lambda'])\hat{W}^\gamma(\lambda'). \tag{3.11}$$

The function χ is defined as $\chi(a) = (-1)^a$ for qubits and $\chi(a) = e^{\frac{2\pi i}{d}a}$ for qudits of odd dimensions. We will omit the superscript γ in the future in order to soften the notation. The normalisation N_Ω is such that $\mathrm{Tr}(A(\lambda)) = 1$. The operators $X(x), Z(p)$ represent the (generalised) Pauli operators of equations (2.12) and (2.13). The function γ, which is only constrained by the requirement that the Weyl operators are Hermitian, will be specified in the next section according to the subtheory the Wigner function non-negatively represents. The above Wigner functions are quasi-probability distributions satisfying the properties of equations (2.23), (2.24), (2.25) , but in general they do not satisfy the factorizability property of equation (2.26). Moreover we now state again the most important property of the Wigner functions for this chapter: the property of covariance for a unitary U (see equation (2.27)). It means that, for all the allowed states ρ in the theory,

$$W_{U\rho U^\dagger}(\lambda) = W_\rho(S\boldsymbol{\lambda} + \mathbf{a}), \tag{3.12}$$

[4]The Wigner function associated to a measurement element Π_k is defined analogously.

where S is a symplectic transformation and $\mathbf{a}$ is a translation vector. This property guarantees that the transformations in quantum mechanics correspond to symplectic affine transformations in the phase space and it is not necessarily implied by the definition of the Wigner functions above.

In [57] Raussendorf *et al.* considered only the case of state-injection schemes on *qubits* and characterised the subtheories of qubit SQM for the free part of the computation that have non-negative Wigner functions for states and measurements. They also imposed two further requirements: the free measurements preserve the non-negativity of the Wigner function and are tomographically complete, *i.e.* they allow to fully measure the density matrix ρ of any $n-$qubit quantum state. [5] The first requirement ensures the "classicality" of the free part of the computation, meaning that a classical simulation procedure of it can be implemented [57, Section IV]. This also guarantees the absence of state-independent contextuality. In principle, the requirement of positivity preservation under free measurements tends to restrict the number of possible allowed observables in the subtheory. Therefore they also require tomographic completeness, *i.e.* any state can be fully measured by the observables allowed in the free part of the scheme, to be sure that the set of free observables is still large enough for the state-injection scheme to function. Notice that, unlike the case of measurements, they allow gates that introduce negativity in the Wigner functions, *i.e.* non-covariant gates, the reason being that the gates can always be absorbed in the measurements without altering the outcome distribution of the computation (so not affecting the classical simulation of the subtheory).

Given these assumptions, the choice of γ in the Wigner function of equation (3.10) uniquely defines a non-contextual subtheory of qubit SQM for the free part of the state-injection scheme. Before seeing how, let us recall, from [57], that the function γ uniquely specifies the function $\beta(\lambda, \lambda')$, defined such that $\hat{W}(\lambda)\hat{W}(\lambda') = w^{\beta(\lambda,\lambda')}\hat{W}(\lambda + \lambda')$. More precisely,

$$\beta(\lambda, \lambda') = \gamma(\lambda) + \gamma(\lambda') - \gamma(\lambda + \lambda') - xp', \tag{3.13}$$

where $\lambda = (x, p)$ and $\lambda' = (x', p')$. It results that the observable $\hat{W}(\lambda)$ preserves non-negativity if and only if $\beta(\lambda, \lambda') = 0 \; \forall \; \lambda' \; s.t. \; [\lambda, \lambda'] = 0$, as proven in [57, lemma 1]. Here below is the list of instructions to obtain the non-contextual subtheory of qubit SQM, described by the sets of

[5] Tomographic completeness can also be checked via the condition stated in [57, Equation 17].

quantum states $\mathcal{S}$, transformations $\mathcal{T}$ and observables $\mathcal{M}$, from the choice of γ and given the two requirements described above.

1. The function γ uniquely defines the set of allowed observables, $\mathcal{M} = \{\hat{W}(\lambda) \mid \beta(\lambda, \lambda') = 0 \; \forall \; \lambda' \; s.t. \; [\lambda, \lambda'] = 0\}$.

2. The set of allowed states $\mathcal{S}$ is given by the states corresponding to common eigenstates of commuting observables in $\mathcal{M}$.

3. The set of allowed gates is $\mathcal{T} = \{U \in C_{n,d} | \; U\rho U^\dagger = \rho' \in \mathcal{S} \; \forall \; \rho \in \mathcal{S}\}$. This is a subset of the Clifford unitaries.

With this framework, Raussendorf *et al.* proved that negativity and contextuality of the magic states are necessary resources for UQC.

Another important work on non-negative representations of qubit SQM is due to Wallman and Bartlett [83]. They exploited the broader formalism of the quasi-probability distributions to find a non-negative representation and a non-contextual ontological model (in terms of measurements and preparations) of one qubit SQM. They built the so-called *8-state model*, where the one-qubit quantum states (and measurement elements) are represented as uniform probability distributions over an ontic space of dimension 8 (doubling the dimension of the standard phase space) and the Clifford transformations – generated by the Hadamard H and phase gate S – are represented by permutations over the ontic space. In the previous chapter, section 2.1.3, we introduced in equation (2.31) the standard Wigner function for one qubit SQM firstly developed by Wootters, Gibbons and then extensively studied by Galvao and collaborators [44, 45, 96, 97]. In [97] it is shown that non-negative Wigner functions for one qubit stabilizer with the desired properties (equations (2.23), (2.24), (2.25), (2.26) and (3.12)) can only be of two kinds: either with phase point operators as in equation (2.31), with an even number of minuses, or with phase-point operators obtained from $A_-(0,0) = 1 + X + Y - Z$ by applying the X, Y, Z Pauli operators by conjugation, thus having an odd number of minuses. We denote the two Wigner functions with W_+ and W_- respectively. Both these Wigner functions fit in the framework of Raussendorf *et al.* defined above, with the choice of γ as $\gamma(\lambda) = \mathbf{p} \cdot \mathbf{x} \mod 4$, and up to the swapping of $Z(p)$ and $X(x)$ in the definition of the Weyl operators (equation (3.9)). The

states that are non-negatively represented according to W_+ and W_- can be represented in the quantum states space, *i.e.* the Bloch sphere (figure 3.4).

The idea of the 8-state model is to use a quasi-probability distribution that takes into account both the Wigner functions W_+ and W_-. It is clear, as shown in the figure 3.4, that the states that are non-negative in both the Wigner functions constitute the octahedron that describes the stabilizer states in the Bloch sphere. The 8-state model has been recently proven to be transformation contextual (as well as the one qubit SQM), since there are completely-positive trace-preserving maps that are operationally equivalent, but unavoidably different at the ontological level. For example, the completely depolarizing channel ε, defined such that $\varepsilon : \rho \to \mathbb{I}/2 \ \ \forall\rho$, can be given by two decompositions, $\varepsilon_1 = \sum_P P\rho P$, where P denotes a Pauli operator, and $\varepsilon_2 = \sum_P (HP)\rho(HP)$, where H is the usual Hadamard gate, that can always be distinguished at the ontological level (they correspond to two different permutations) [122]. Transformation contextuality is here caused by the non-covariant gates (*e.g* the Hadamard gate), thus suggesting a strict relation between the notions of transformation contextuality and non-covariance.

The proposed and straightforward generalisation of the 8-state model to more than one qubit consists of considering the distributions built from the tensor products of the phase-point operators of the single qubit. The resulting subtheory of qubit SQM described by the model is the one composed by all the product states of tensors of Pauli X, Y, Z observables and all the *local* Clifford unitary gates (generated by local H and S). No entanglement is present. Nevertheless it is possible to reach universal quantum computation from this subtheory by performing measurement-based quantum computation [15] with a particular entangled cluster state, as shown in [56, 57].

The remainder of the chapter mainly reports the material contained in [49], which is a joint work with Nadish De Silva and Dan Browne, and [50], which is a joint work with Dan Browne.

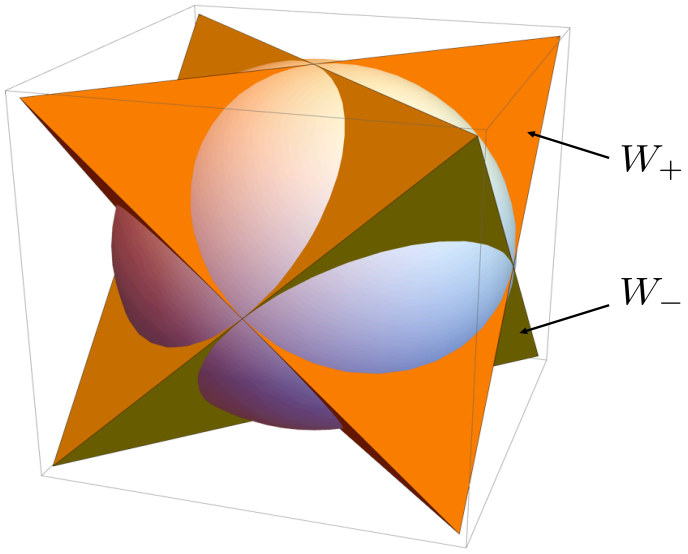


Figure 3.4: **Non-negative Wigner functions of one qubit SQM.** The figure above shows the states that are non-negative according to the two possible Wigner functions W_+ and W_- for one qubit defined in [97]. The intersection of the two non-negative sectors in the Bloch sphere forms the stabilizer octahedron. The idea behind the 8-state model of Wallman and Bartlett [83] is to consider a quasi-probability distribution over an ontic space with twice the dimension of the standard phase space and consider both the two Wigner functions.

3.2 Characterisation of Spekkens' subtheories

3.2.1 Definition

A Spekkens' subtheory is defined as a set of quantum states, transformations and measurements, $(\mathcal{S}, \mathcal{T}, \mathcal{M})$, which satisfies the following conditions.

1. *Subtheory.* The set must be *closed*, which means that any allowed gate cannot bring from one allowed state to a non-allowed one.

$$\forall\, U \in \mathcal{T}, U\rho U^\dagger \in \mathcal{S}\ \forall \rho \in \mathcal{S}. \tag{3.14}$$

2. *Spekkens representability.* There must be an operational equivalence between the subtheory of quantum mechanics $(\mathcal{S}, \mathcal{T}, \mathcal{M})$, defined by sets of quantum states, transformations and measurements, and a subtheory of Spekkens' toy theory $(\mathcal{S}_s, \mathcal{T}_s, \mathcal{M}_s)$, defined by sets of epistemic states, symplectic affine transformations and measurements as defined in the

previous chapter (subsection 2.1.1).

The operational equivalence means that the statistics of the two subtheories $(\mathcal{S}, \mathcal{T}, \mathcal{M})$ and $(\mathcal{S}_s, \mathcal{T}_s, \mathcal{M}_s)$ are the same. We state this equivalence by finding a *non-negative* Wigner function that maps the states ρ and the measurement elements Π_k in $(\mathcal{S}, \mathcal{M})$ to epistemic states and measurement elements in $(\mathcal{S}_s, \mathcal{M}_s)$, *i.e.*

$$W_\rho(\lambda) = \frac{1}{N}\mathrm{Tr}(\rho A(\lambda)) = \frac{1}{N}\delta_{(V^\perp + \mathbf{w})}(\lambda); \tag{3.15}$$

$$W_{\Pi_k}(\lambda) = \frac{1}{N'}\mathrm{Tr}(\Pi_k A(\lambda)) = \frac{1}{N'}\delta_{(V_\Pi^\perp + \mathbf{r}_k)}(\lambda). \tag{3.16}$$

The N, N' are the normalisation factors so that $\sum_{\lambda\in\Omega} W(\lambda) = 1$ for all the above Wigner functions. Notice at this point that the measurement update rules defined in the previous chapter (theorem 6) guarantee that also a state after a measurement is non-negatively represented if the measurement and the original state have non-negative Wigner functions, as they involve only sums and products of Wigner functions. Moreover we are considering subtheories with duality between states and measurement elements, therefore it is enough to check only the properties of the Wigner functions of states (or measurements) to guarantee Spekkens representability.

The operational equivalence in terms of transformations is implied if the Wigner function (3.15) satisfies the property of covariance (3.12) for the allowed unitaries $U \in \mathcal{T}$, which guarantees that the transformations in quantum mechanics correspond to transformations that preserve the epistemic restriction in ST. Notice that the property of covariance is defined in terms of Wigner functions and not directly in terms of the phase point operators.[6] Therefore we do not necessarily need to demand for the standard Wigner function of the transformation, defined as $W_U(\lambda/\lambda') = \frac{1}{N'}\mathrm{Tr}(A(\lambda)UA(\lambda')U^\dagger)$, to be non-negative in all the elements, once the previous requirements, non-negativity and covariance of W_ρ, are satisfied. The transition matrix corresponding to the allowed permutation of the phase points can be always found, as shown by the following lemma.

Lemma 4. *Given a non-negative Wigner function representation, $W_\rho, W_{\rho'}$, of any two*

[6] This will make a difference only in the next subsection, where we consider factorisable Wigner functions for the CSS rebit case.

allowed states $\rho, \rho' \in \mathcal{S}$ *such that* $\rho' = U\rho U^\dagger$, *where* $U \in \mathcal{T}$, *and covariance holds,* i.e. $W_{\rho'}(\lambda) = W_\rho(S\boldsymbol{\lambda}+\mathbf{a})$, *there always exists a (non-negative) transition matrix* $P_U : \Omega\times\Omega \to [0,1]$ *representing the transformation* $U \in \mathcal{T}$,

$$P_U(\lambda/\lambda') = \frac{1}{N''}\delta_{\lambda, S\boldsymbol{\lambda}'+\mathbf{a}}, \tag{3.17}$$

where N'' *is the normalisation factor, such that*

$$W_{\rho'}(\lambda) = \sum_{\lambda'\in\Omega} P_U(\lambda/\lambda')W_\rho(\lambda'). \tag{3.18}$$

Proof. A matrix made of non-negative elements $P_U(\lambda/\lambda')$ proportional to Kronecker deltas always exists because it corresponds to the transition matrix representing the permutation that brings W_ρ to $W_{\rho'}$. More precisely, non-negative solutions $P_U(\lambda/\lambda')$ to the equations (3.18) for every λ, given the non-negative $W_\rho(\lambda'), W_{\rho'}(\lambda)$ defined in (3.10), always exist. For every fixed λ, the $P_U(\lambda/\lambda')$ are vectors with all zero components apart from one, *i.e.* they are proportional to Kronecker deltas. The covariance property (3.12) guarantees that this permutation corresponds to a symplectic affine transformation on the phase space points (independent on the state ρ that U is acting on). □

The non-negative functions (3.15), (3.16) and (3.17) can be interpreted as probability distributions and guarantee that the theories $(\mathcal{S}, \mathcal{T}, \mathcal{M})$ and $(\mathcal{S}_s, \mathcal{T}_s, \mathcal{M}_s)$ are operationally equivalent, *i.e.* they provide the same statistics:

$$\begin{aligned} p(k) &= \mathrm{Tr}(\Pi_k U\rho U^\dagger) \\ &= \sum_{\lambda\in\Omega} W_{\Pi_k}(\lambda) \sum_{\lambda'\in\Omega} P_U(\lambda/\lambda')W_\rho(\lambda'). \end{aligned} \tag{3.19}$$

To sum up, a Spekkens' subtheory is a (closed) subtheory of quantum mechanics whose states (and measurements) are represented by non-negative and covariant Wigner functions. We say that a Spekkens' subtheory is *maximal* if the set $(\mathcal{S}, \mathcal{T}, \mathcal{M})$ is such that by adding either another state, gate or observable to the set of allowed states, transformations and observables at least one of the conditions above is violated, *i.e.* it is no longer a subtheory or Spekkens-

representable. We will also talk about *minimal* non-contextual subtheories of SQM meaning those subtheories that can no longer be used for state-injection schemes after the removal of just one object from them.

3.2.2 Examples

We now show that the known examples by Howard *et al.* [54] and Delfosse *et al.* [55] of state-injection schemes with contextuality as a resource fit into the framework depicted in figure 3.1, *i.e.* that the free parts of those schemes are Spekkens' subtheories.

In the case of qudits of odd dimensions, as we already discussed in subsection 2.1.3, Gross' theorem [43] guarantees that there is a non-negative Wigner representation of all stabilizer states. This Wigner function is covariant and also Clifford transformations and Pauli measurements are non-negatively represented. Thus Gross' Wigner function proves the operational equivalence, in odd dimensions, between SQM and the whole ST, as shown in the previous chapter. Gross' Wigner function is defined in equation (2.22) and corresponds to equation (3.10) with the function γ given by $\gamma(\lambda) = \mathbf{x} \cdot \mathbf{z}$. In the scheme of Howard *et al.* the free part of the computation is given by SQM in odd prime dimensions, which, by Gross' Wigner functions, is a maximal Spekkens' subtheory.[7]

In the case of rebits studied by Delfosse *et al.* the Wigner function of equation (3.8) is always non-negative for CSS states and it is covariant for the CSS-preserving gates. In terms of the definition provided in equation (3.10), the function γ is $\gamma(\lambda) = 0$. This choice guarantees that the phase point operators are Hermitian. However the price to pay for the Hermitianity in this case is the non-factorisability of the Wigner function (equation (2.26)). One may wonder whether the non-factorisability of the Wigner function is necessary to treat the CSS case and preserve the non-negativity and covariance. Here we show that it is not. We define a Wigner function that, we argue, is more in line with the construction of ST, where the ontic space of n systems is made by the cartesian products of individual systems' subspaces. The non-negative, covariant and factorisable Wigner function for the CSS theory is built out from the single-qubit phase-point operators

$$A_f(0,0) = \mathbb{I} + X + Z + iY. \tag{3.20}$$

[7] Stabilizer quantum mechanics in odd dimensions is the unique maximal Spekkens' subtheory, since it coincides with the whole Spekkens' theory.

The phase point operators $A_f(0,1), A_f(1,0), A_f(1,1)$ are given by applying the Pauli X, Y, Z respectively by conjugation on $A_f(0,0)$. The phase point operators of many qubits are given by tensor products of the ones for single qubits $A_f(0,0), A_f(0,1), A_f(1,0), A_f(1,1)$. Notice that the phase point operators are not Hermitian; however the allowed observables are only present in their Hermitian part (*e.g.* for the single qubit in $\mathbb{I} + X + Z$). We now need to prove the following lemma.

Lemma 5. *The Wigner function of Delfosse et al.* $W_r(\lambda) = Tr(\rho A_r(\lambda))$, *given by* (3.8), *is equivalent to the factorisable Wigner function* $W_f(\lambda) = Tr(\rho A_f(\lambda))$, *given by* (3.20), *for any* $\rho \in \mathcal{S}_r$.

Proof. What we need to prove is actually that $A_r(\lambda) = \mathcal{H}(A_f(\lambda))$, where $\mathcal{H}(A_f(\lambda))$ indicates the Hermitian part of the phase point operator $A_f(\lambda)$. The non-Hermitian part of $A_f(\lambda)$ has zero contribution to the Wigner function. It is always composed by tensors of mixtures of Pauli operators with an odd number of Y's, that never form allowed observables and so are never in the stabilizer group of any $\rho \in \mathcal{S}_r$. This implies that the non-Hermitian part of $A_f(\lambda)$ has no contribution to the Wigner function as Pauli operators (apart from the identity) are traceless. However the non-Hermitian part of $A_f(\lambda)$ is important since when its operators compose into phase point operators for multiple qubits, they sometimes provide Hermitian operators that contribute to the Wigner function. We know that $A_r(\lambda)$ is defined as the sum of observables $\hat{W}(\lambda)$, where $\lambda = (x,p)$ such that $\mathbf{x} \cdot \mathbf{p} = 0 \mod 2$. We can now see that also $\mathcal{H}(A_f(\lambda))$ is given by the sum of observables subjected to the same condition of having zero inner product between the components. This condition indeed singles out all the rebit observables, which are the only ones we are interested in. Given an observable $\hat{W}(\lambda) = Z(p)X(x)$ in $A_f(\lambda)$, with $\lambda = (x,p)$ and $x, p \in \mathbb{Z}_d^n$, it is Hermitian if and only if $\hat{W}(\lambda) = \hat{W}(\lambda)^\dagger$. This means that

$$\hat{W}(\lambda)^\dagger = X(x)Z(p) = (-1)^{\mathbf{x}\cdot\mathbf{p}}\hat{W}(\lambda),$$

which holds if and only if $\mathbf{x} \cdot \mathbf{p} = 0 \mod 2$.

□

In conclusion, by using one of the above Wigner functions, (3.8) or (3.20) , given the duality between states and measurement elements, the covariance and lemma 4, we can conclude that

CSS rebits subtheory is Spekkens representable. Moreover, the definition of CSS-preserving transformations guarantees the closure property and the fact that CSS states are the ones and only ones with non-negative Wigner function [55] ensures that it is maximal. Therefore the CSS rebit subtheory of quantum mechanics is a maximal Spekkens' subtheory.

3.3 Spekkens' subtheories as toolboxes for state-injection

We now prove that qubit stabilizer quantum mechanics can be obtained from a Spekkens' subtheory, in the sense that within Spekkens' subtheories it is possible to build a state-injection scheme that injects all the objects of qubit stabilizer quantum mechanics that are not in the subtheories. We need to understand which objects do we actually need to inject to reach the full multi-qubit SQM from a Spekkens' subtheory. We here state the list of instructions to construct the *maximal* Spekkens' subtheory that corresponds to a given choice of γ, in analogy with the framework of [57] defined in subsection 3.1.4.[8]

1. The function γ uniquely defines the set of allowed observables, $\mathcal{M} = \{\hat{W}(\lambda) \mid \beta(\lambda, \lambda') = 0 \; \forall \; \lambda' \; s.t. \; [\lambda, \lambda'] = 0\}$.

2. The set of allowed states $\mathcal{S}$ is given by the states corresponding to common eigenstates of commuting observables in $\mathcal{M}$.

3. The set of allowed gates is $\mathcal{T} = \{U \in C_{n,d} | \; U\rho U^\dagger = \rho' \in \mathcal{S} \; \forall \; \rho \in \mathcal{S}, \; and \; W_{U\rho U^\dagger}(\lambda) = W_\rho(S\boldsymbol{\lambda} + \mathbf{a}) \; \forall \; \lambda \in \Omega\}$. This is a subset of the Clifford unitaries.

Let us point out that the above construction differs from [57] in that it does not require tomographic completeness and it does require covariance of the Wigner functions. With respect to the Wallman-Bartlett 8-state model [83] the difference holds for analogous reasons. We now show that the implementation of all the non-covariant Clifford unitaries would boost Spekkens' subtheories, composed by covariant Clifford gates, to the full qubit SQM.

[8] Notice that with the following construction a given γ provides the *maximal* Spekkens' subtheory, but the Wigner function from the same γ can, obviously, be used to represent any smaller subtheory of the maximal Spekkens' subtheory.

Theorem 11. *Qubit stabilizer quantum mechanics can be obtained from a Spekkens' subtheory via state-injection: all the possible non-covariant Clifford gates can be state-injected via a circuit made of objects in the Spekkens' subtheory.*

Proof. In order to prove that qubit stabilizer quantum mechanics can be obtained from a Spekkens' subtheory via state-injection we need to show that all the objects needed for injecting any non-covariant Clifford gate are present in at least one Spekkens' subtheory. We recall that in order to generate the whole Clifford group we need, in addition to the $CNOT$, also the generators of the local single gates, *e.g.* the usual phase and Hadamard gates, S, H. Let us consider the following subtheory, which corresponds to the CSS rebit subtheory, equations (3.6) and (3.7), with no global Hadamard gates:

- The allowed observables are, analogously to equation (3.6), non-mixing tensors of X and Z Pauli operators, $\mathcal{M} = \{X(x), Z(p) | x, p \in \mathbb{Z}_2^n\}$.
- The allowed gates are the ones generated by the $CNOT$ and the Pauli rotations X, Z, *i.e.*

$$\mathcal{T} = \langle CNOT(i,j), X_i, Z_i \rangle \,. \tag{3.21}$$

- The allowed states are, as usual, the eigenstates of the allowed observables.

This is a smaller subtheory than CSS rebit (the difference being the absence of the global Hadamard gate). It possesses all the objects needed for state-injection of non-covariant gates. The Z observables and the $CNOT$ gate are present. The correction gates are always Pauli gates, as for any injected Clifford unitary U, even when U is non-covariant, $UX^{\otimes n}U^\dagger$, by definition of a Clifford gate, gives back a Pauli gate. All the objects of this subtheory can be non-negatively represented by the Wigner functions for the CSS rebit theory of the previous section and also in ST, as shown in figure 3.5. Therefore this subtheory is closed and Spekkens representable, *i.e.* a Spekkens' subtheory, and it is possible for it to reach universal quantum computation via state-injection, as proven in details in the next section.

We are here interested in showing that we can obtain the whole Clifford group via state-injection. Once we have it, we can map any of the allowed states and observables to any other in qubit SQM. The whole Clifford group can be achieved by first injecting the CZ gate, as shown

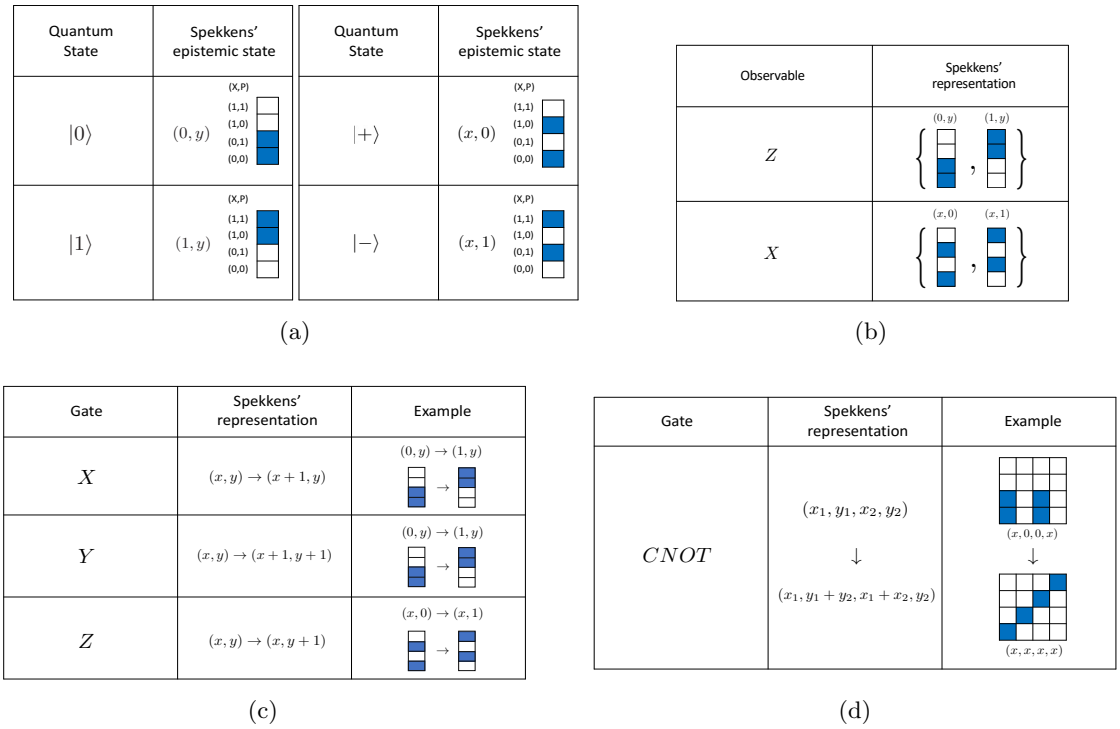


Figure 3.5: **Representation of a minimal non-contextual subtheory of qubit stabilizer quantum mechanics in Spekkens' theory.** In the figure above the allowed pure states 3.5a, observables 3.5b and gates 3.5c,3.5d of the non-contextual subtheory of qubit SQM considered in the proof of theorem 11 are represented in ST. In figure 3.5a, 3.5c and 3.5d we have also indicated the probability distributions associated to the epistemic states and how the gates act on them. In the examples of figures 3.5c,3.5d we have considered the scenarios corresponding to acting with X, Y on $|0\rangle$, with Z on $|+\rangle$ and with $CNOT$ on $|+0\rangle$ (thus obtaining the Bell state $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$).

in figure 3.6, that also provides a construction for the Hadamard gate (figure 3.7) and then the phase gate S. The correction gate for the state-injection scheme of the S gate is given by the Pauli Y gate, which is present, up to a global phase, in our Spekkens' subtheory as a composition of X and Z Pauli rotations. Notice that, even if the state-injection, as defined in 3.1.1, allows to only inject diagonal unitary gates, we can here obtain also all the non-diagonal ones since we have obtained a generating set of gates for all the Clifford unitaries. The other peculiarity of the Spekkens' subtheory we are considering is that it is minimal, as we cannot remove any object from it without denying the possibility of achieving universal quantum computation. One could think of removing the $X(x)$ observables, but these play a crucial role in obtaining the Hadamard gate. Moreover Spekkens' subtheories with observables given by tensors of a

single Pauli operator and basis preserving gates do not allow any state-injection scheme to inject objects outside of them.

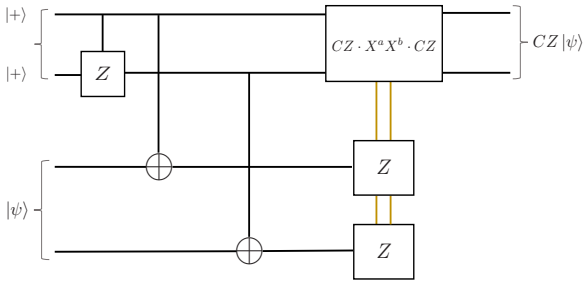


Figure 3.6: **CZ injection.** The first injection of our scheme is the injection of the $CZ\,|{+}{+}\rangle$ state, needed in order to perform the correction in the injection scheme for the CCZ gate and to produce the Hadamard gate. In the figure above the correction is $CZ \cdot X^a X^b \cdot CZ$ conditioned on obtaining x, y outcomes from the Z measurements, where $(-1)^a = x$ and $(-1)^b = y$. For example if the outcomes are $x = 1, y = -1$, the correction is $CZ \cdot \mathbb{I}X \cdot CZ = XZ$, which is an allowed gate in our subtheory.

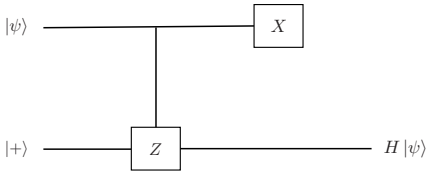


Figure 3.7: **Hadamard gate via CZ.** The Hadamard gate can be obtained once the CZ gate is available. Obtaining H from CZ and X measurements was originally shown in [125].

□

The above theorem guarantees that a state-injection scheme can always be recast in the following structure: *Spekkens' subtheory + Magic state(s)* $\rightarrow$ *UQC*. ST contains all the tools to do state-injection of non-covariant Clifford gates that appear in qubit SQM and also in its subtheories bigger than Spekkens' subtheories as formalised in [57]. Notice that the current approach looks at the minimal non-contextual subtheories of SQM such that we can still reach universal quantum computation via state-injection. All the times these minimal subtheories are Spekkens' subtheories. In the next section we describe how the minimal subtheory of rebit

SQM used in the proof of theorem 11, equations (3.6) and (3.21), can reach universal quantum computation through the injection of CCZ magic states.

3.4 State-injection schemes with CCZ states

The idea of reaching fault-tolerant universal quantum computation by exploiting Toffoli gates, or $CCNOT$ - control control X, goes back to Peter Shor in 1997 [101]. It is known that the Toffoli gate (enough for universal classical computation) and Hadamard gate allow one to reach universal quantum computation [102,103], and, in a sense, this is the most natural universal set of gates, since the Toffoli allows one to achieve all the classical operations, and just by adding the Hadamard gate, which creates superposition, they lead to universal quantum computation. The same result holds if we use the related CCZ, control control Z, gate instead of the Toffoli gate. Other examples of fault tolerant universal quantum computation involving Toffoli state distillation have been proposed [104–107]. We here propose a scheme of CCZ injection with the fewest possible objects in the free part of the computation. We reach that by also injecting the CZ state before the CCZ. The state-injection scheme is depicted in figure 3.8.

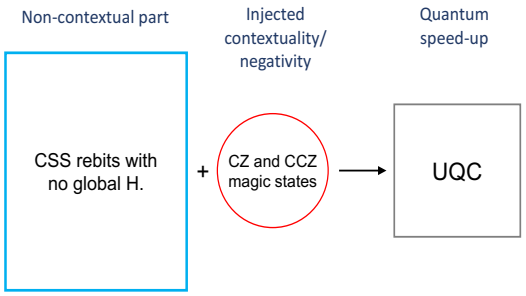


Figure 3.8: **Novel state-injection scheme based on CCZ injection.** By injecting first the CZ state and then the CCZ state we can boost the subtheory of rebit SQM made of observables that are tensors of non-mixing Pauli $\mathbb{I}, X, Z$, and the gates generated by $CNOT$ and the Pauli rotations X, Z to universal quantum computation.

The free part is the one described in the proof of theorem 11 and defined by the equations (3.6) and (3.21). This is the subtheory of CSS rebits with no global Hadamard, thus it is a Spekkens' subtheory. The state-injection scheme works with two state-injections: first the state $CZ\,|{+}{+}\rangle$ (figure 3.6), where the correction is given by the $CZ \cdot X^a X^b \cdot CZ$ conditioned

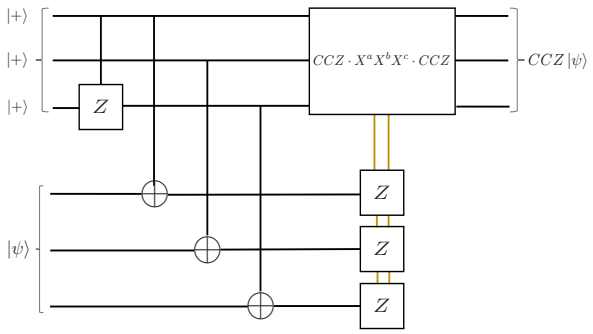


Figure 3.9: **CCZ injection.** The second injection of our scheme is the injection of the $CCZ\,|+++\rangle$ state. In the figure above the correction is $CCZ \cdot X^a X^b X^c \cdot CCZ$ conditioned on obtaining x, y, z outcomes from the Z measurements, where $(-1)^a = x, (-1)^b = y$ and $(-1)^c = z$. For example if the outcomes are $x = -1, y = 1, z = 1$ the correction is $CCZ \cdot X\mathbb{I}\mathbb{I} \cdot CCZ = X \cdot CZ$, which is an allowed gate in our subtheory.

on obtaining x, y outcomes from the measurements of Z's, where $(-1)^a = x$ and $(-1)^b = y$. Just to give an example, for outcomes $x = 1, y = -1$ the correction is $CZ \cdot \mathbb{I}X \cdot CZ = XZ$. Secondly, the injection of the state $CCZ\,|+++\rangle$ (figure 3.9), where the correction is given by $CCZ \cdot X^a X^b X^c \cdot CCZ$, with outcomes x, y, z of the measurments of Z's such that $(-1)^a = x$, $(-1)^b = y$ and $(-1)^c = z$, *e.g.* $CCZ \cdot X\mathbb{I}\mathbb{I} \cdot CCZ = X \cdot CZ$ if the outcomes are $-1, 1, 1$. Notice that the injection of the CZ allows also to obtain the Hadamard gate, as shown in figure 3.7. With Hadamard and CCZ gates we then have a universal set for quantum computation. The contextuality, which is not present in the subtheory of CSS rebits, is clearly present after the two injections that lead to universal quantum computation.

A few comments on the free part of the scheme are needed. As already said, it is *minimal*, in the sense that it is not possible to remove any object from the free part of the computation without denying the possibility of obtaining universal quantum computation via state-injection. Also it is a strict subtheory of the CSS rebit, where we allow all the same objects apart from the global Hadamard. We argue that this is desirable, since in principle the Hadamard gate is a local gate; we want to keep only the entangling gates to have a global nature.

3.5 Proofs of contextuality and state-injections

The Spekkens' subtheory used for the CCZ state-injection scheme of the previous subsection allows us to establish a relation between the different resources injected and different proofs of contextuality. Proofs of contextuality like the Peres-Mermin square argument and the GHZ paradox [46–48] defined in subsection 3.1.2 are not present within the Spekkens' subtheory, which, as we know, always represents the absence of any form of contextuality. We now explicitly show how these cases are obtained after the injections that bring in either the CZ gate or the S gate. We also show that the popular non-Clifford gate T (equation (3.2)), in addition to the Peres-Mermin square argument and GHZ paradox (with the condition that we can apply the T gate at least two times, as $T^2 = S$), also allows one to formulate the argument that brings to the CHSH inequality maximum violation [59]. These examples demonstrate that specific states injected to the minimal Spekkens' subtheory treated here can be considered resources for specific manifestations of contextuality.

- *Peres-Mermin square [47].* Let us consider the free Spekkens' subtheory of the CCZ injection scheme supplemented with the injection of the CZ state. It allows us to construct a circuit to perform the Peres-Mermin square argument (subsection 3.1.2) . While in our original Spekkens' subtheory we are only allowed to perform the observables in the first two rows of the square, with the presence of the CZ we can obtain the last row too, since $CZ \cdot X\mathbb{I} \cdot CZ = XZ$, $CZ \cdot \mathbb{I}X \cdot CZ = ZX$ and $CZ \cdot XX \cdot CZ = YY$. Figure 3.10 shows a circuit where we can perform all the contexts of the Peres-Mermin square on an arbitrary input state $|\psi\rangle$ by just using objects belonging to the Spekkens' subtheory and CZ injections.

  We can obtain the Peres-Mermin square argument also with the injection of the S gate. This time the observables considered in the square are $\mathbb{I}X, X\mathbb{I}X, XX, Y\mathbb{I}, \mathbb{I}Y, \mathbb{I}Y, YX, XY, ZZ$. The ones containing Y can be obtained by applying S to the X observable, while the others are already present in our Spekkens' subtheory.

- *GHZ paradox [46].* In order to obtain the GHZ paradox we need to be able to implement the GHZ state $\frac{|000\rangle+|111\rangle}{\sqrt{2}}$, already present in our Spekkens' subtheory, and the mutually commuting observables XXX, XYY, YXY, YYX. While the first observable XXX is

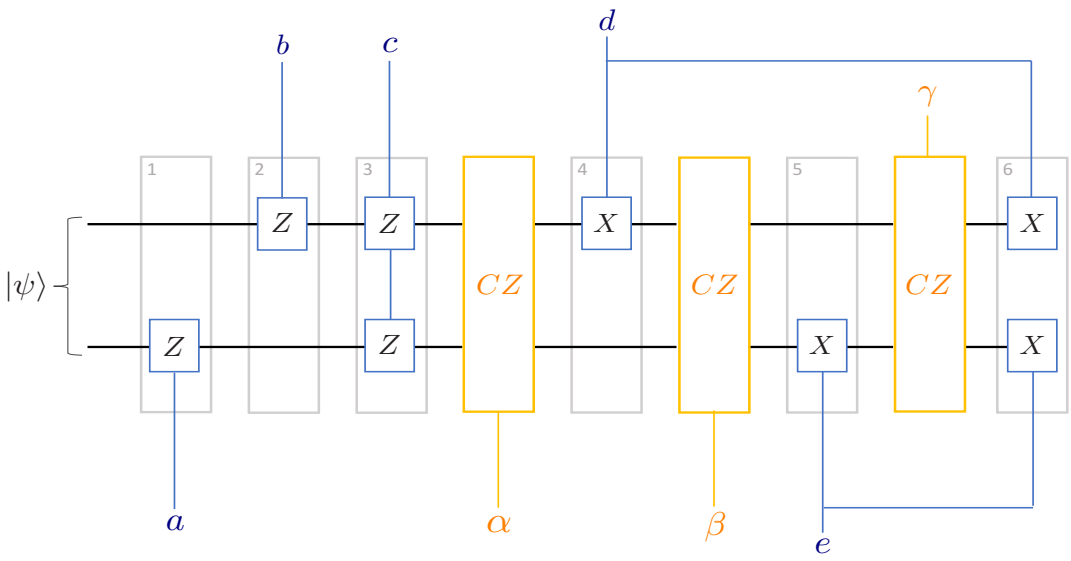


Figure 3.10: **Peres-Mermin square via Spekkens' subtheories and CZ gates.** The above circuit provides a way of implementing all the contexts of the Peres-Mermin square. Each block denoted by CZ corresponds to the injection scheme of figure 3.6 endowed also with a swap gate (which is present in our Spekkens' subtheory as it can be made of a series of three alternated CNOT gates) in order to set the output state $CZ\,|\psi\rangle$ as a precise modification of the input state $|\psi\rangle$ (and not of the ancillary resource state $CZ\,|{+}{+}\rangle$). Each context can be selected according to some combinations of the classical control bits $a, b, c, d, e, \alpha, \beta, \gamma$ that can take values in $\{0, 1\}$. The value 0 indicates that the corresponding gate is not performed. The value 1 indicates that it is performed. At the end of every grey block (labelled by numbers) we assume that we can read the output outcome. The three row contexts of the Peres Mermin square are identified by the variables (d, e), (a, b, c) and $(\alpha, \beta, \gamma, d, e)$ assuming value one, respectively. The three column contexts are identified by (a, d, γ) (b, e, γ) and (c, d, e, γ). Notice that in the last case where we implement the context XX, ZZ, YY, the measurement of XX is implemented by performing $\mathbb{I}X$ first and then $X\mathbb{I}$, and in this case we consider the outputs related to the blocks labelled by $3, 5, 6$.

already present in our Spekkens' subtheory, the others can be obtained either by local S gate or CZ gate on two of the three single Pauli operators composing each observable.

- *CHSH argument [59].* If we consider our Spekkens' subtheory with the addition of the T gate we can obtain the maximum quantum violation of the CHSH inequality (equation (4.1)). The CHSH game is an example that shows neat computational advantages of certain quantum strategies over classical ones – expressed by the violation of the CHSH inequality – and it will be precisely defined in the next chapter (subsection 4.1.1). For

the purpose of this section, we just need to know that by using the Bell state

$$\frac{|00\rangle - |11\rangle}{\sqrt{2}}, \tag{3.22}$$

which is a -1 eigenstate of XX and $+1$ eigenstate of ZZ, and the observables

$$A_0 = Y,\ A_1 = X,\ B_0 = TYT^\dagger = \frac{Y-X}{\sqrt{2}},\ B_1 = TXT^\dagger = \frac{X+Y}{\sqrt{2}}, \tag{3.23}$$

the maximum quantum violation of the CHSH inequality in equation (4.1) is obtained, as $\langle A_0B_0\rangle = \langle A_0B_1\rangle = \langle A_1B_0\rangle = -\langle A_1B_1\rangle = \frac{1}{\sqrt{2}}$. Notice that the Bell state that we consider is present in our Spekkens' subtheory, and that the observables are provided by the presence of the T gate.

3.6 Conclusion

In this chapter we have defined the subtheories of ST that are compatible with quantum mechanics. They are the closed subtheories of ST that have non-negative and covariant Wigner function representations. SQM is the maximal Spekkens' subtheory in odd dimensions, as it corresponds to the full ST. This is not true for qubits, as SQM is contextual and the toy theory does not reproduce its statistics. We have used Spekkens' subtheories as a unifying framework for known examples of state-injection schemes of quantum computation with contextuality as an injected resource [54,55], in the sense that they fit into the scheme of figure 3.1, *i.e. Spekkens' subtheory + Magic states* $\rightarrow$ *UQC*, where Spekkens' subtheories represent the non contextual part of the computation and the contextuality arises in the injection of the magic states. Even more, we have proven, in theorem 11, that multi-qubit SQM can be obtained from a Spekkens' subtheory via state-injection, as all the objects which do not belong to the subtheory but are present in SQM (namely non-covariant gates) can be state-injected with a circuit made of objects in the Spekkens' subtheory. This means that for quantum computation via state-injection we only need to study Spekkens' subtheories, *i.e.* the part of ST that coincides with quantum mechanics, because we can always generate the whole Clifford group by injection and therefore all the other state-injection schemes of quantum computation can be mapped to our framework

by injection.

In order to prove theorem 11 we have constructed a Spekkens' subtheory which is a strict subtheory of the CSS rebit theory (used in [55]) and provides a novel state-injection scheme that allows to reach UQC by injections of CZ and CCZ states. This subtheory is minimal, meaning that it is not possible to remove any object from it without denying the possibility of reaching UQC via state-injection. By analysing the different injection processes in the above scheme we have also associated different proofs of contextuality to specific state-injections of non-covariant gates. In particular we have explicitly shown how the CZ and S gates are resources for the Peres-Mermin proof of contextuality and the GHZ paradox, and how the T gate, used in the most popular state-injection schemes [53], is a resource also for the CHSH argument.

With respect to previous related works, we have often referred to Raussendorf *et al*'s framework [57] as the main reference, since it is very general and it includes, for example, also the subtheory of multi-qubit SQM that arise from the 8-state model formulation of Wallman and Bartlett [83]. Raussendorf *et al*'s framework differs from our Spekkens' subtheories as it requires tomographic completeness and it does not demand for covariant Wigner functions. In this work we have preferred the tools provided by Spekkens' toy theory, which has a less abstract structure, it is an intuitive and fully non-contextual ontological model and it is intrinsically related to non-negative and covariant Wigner functions, because, as we prove in theorem 11, they are enough to treat state-injection schemes of quantum computation. In particular, the CSS rebit subtheory with no global Hadamard considered in section 3.4 and 3.5 is useful for quantum computation via state-injection despite the fact that it is covariant and not tomographically complete, as shown in section 3.5.

Notice that the state-injection schemes we have considered are the ones developed in [51], which means that we are not considering cluster state computation, unlike in [57]. An open question is how to extend theorem 11 to these more general schemes. A suggestion in this direction comes from the example of cluster state computation provided in [56] and [57]. It consists of a non-contextual free subtheory made of tensors of X, Y, Z Pauli observables, their product eigenstates and all the local Clifford gates, and the resource is a specific entangled cluster state. The free subtheory in this case is not a Spekkens' subtheory, as the S and H gates are not covariant if we allow all the product eigenstates of tensors of X, Y, Z Pauli

observables. However, if we remove these local gates we can still implement the computational scheme (which never needs to use those gates indeed) and obtain UQC with the same resource state. In the latter case the free part is now a Spekkens' subtheory. Therefore this example can be actually recast in our framework.

Finally, we point out that all previous works [56, 57, 83] look at the biggest non-contextual subtheories of SQM that allow to perform state-injection schemes of computation with contextuality as a resource. Here instead, we have focused also on the smallest free subtheories such that it is still possible to reach UQC via state-injection.

We believe that the results presented here suggest some related future projects. The importance of the property of covariance highlighted by our result could inspire one to study further its relationship with non-contextuality. A recent work on contextuality in the cohomological framework could give the right tools to address this question [126]. In particular, covariance seems to be strictly related to Spekkens' transformation non-contextuality [34]. As an example of this, the single qubit SQM, already argued to be not covariant, shows transformation contextuality (even if a preparation and measurement non-contextual model for it - *e.g.* the 8-state model - exists) [122]. A big open question regards which notion of contextuality is actually the proper resource for UQC as, for example, it is known that qubit SQM, despite being contextual, is efficiently classically simulatable [84]. It would be desirable to match the notion of non-classicality in quantum foundations, namely contextuality, with the notion of non-classicality in quantum computation, *e.g.* non-efficient classical simulatability. Finally, we think that it would be interesting to extend ST to obtain a psi-epistemic ontological model of the multi-qubit SQM. Possibly some extensions of the 8-state model can achieve this. In this case it would be interesting also to know which epistemic restrictions these models would imply. This would also help us understand which notion of contextuality is more appropriate to be considered as a resource for UQC.

Chapter 4

Tsirelson's bound and Landauer's principle in a single-system game

So far we have focused on Spekkens' theory, which has been a powerful tool for applications also in quantum computation. However, we now take a different approach. In the previous chapter we studied the role of contextuality in state-injection schemes of quantum computation, where it acts as a resource to achieve UQC. In this chapter we consider a more restricted scenario that manifests quantum computational advantages, where contextuality (in its standard definitions of section 3.1.2, [33, 34]) is not present.

Computational protocols where strategies based on quantum mechanics perform better than classical strategies have long been an important focus of study. A well-known example is the CHSH game [127], a way of recasting the Clauser-Horne-Shimony-Holt (CHSH) formulation of Bell's celebrated theorem [32, 59] into a game for which quantum strategies can provide an advantage. The CHSH game is a game between two players, Alice and Bob, who are separated and unable to communicate with each other, and a referee who asks them binary questions. They win if the sum of their answers is equal to the product of the questions (arithmetic modulo 2). The bound on the performances of classical strategies is known as the Bell bound, while in the quantum case the maximum winning probability is bounded by the so-called Tsirelson bound [60]. The CHSH game can be generalised to mod$q$ arithmetic in the CHSH$_q$ game, which has been studied in [64–67]. Naturally, a key focus of these studies has been to find

the Bell bound and Tsirelson bound for these games. However, success has been limited. Upper bounds on the Tsirelson bound given by a precise mathematical expression have been provided in [67] when q is a prime or prime power, but these are not known to be tight. Moreover, numerical analysis on lower and upper bounds suggest different values [66]. The CHSH game is of great importance because the sensitivity of its optimal success probability depending on the underlying physical model gives us a tool to distinguish different types of theories experimentally, and allows us to test nature. It also reveals insights into a non-classical feature of quantum mechanics (known colloquially as "non-locality"), which has proven to be a resource for quantum technologies, such as device independent cryptography [61, 63].

Other protocols showing similar features to the CHSH game exist [68, 69]. In particular, in quantum random access codes (QRACs),[1] where Alice encodes m bits in $n < m$ information carriers to communicate to Bob the value of one of the bits (randomly chosen), the optimal classical and quantum strategies are closely related to the ones used in the CHSH protocol and provide the same bounds.

Inspired by these works, we here propose and investigate a single-system protocol, which is a simple single-player variant of the CHSH game. To play the game, the player has a system in a fixed initial state, two gates controlled by classical input bits and a measurement at the end (figure 4.4). The task is to output a non-linear function – the product – of the input bits in $\mathrm{mod}\,2$ arithmetic. Due to its similarity with the CHSH game we call it the *CHSH* game*. However, unlike the CHSH game that involves two space-like separated parties, the CHSH* game cannot involve any non-locality argument to explain the computational advantages. Similarly, it does not show any contextuality (at least in its usual formulations [33, 34]), as there are no contexts as usually defined (the projective measurement is fixed, the system is, in principle, in a fixed pure state and transformations are unitaries that do not form operationally equivalent decompositions of a completely-positive trace-preserving map).

We study the probability of success of the CHSH* game in different settings. We first show that, when the player applies unitary dynamics and projective measurements on a qubit system, the maximum probability of success of the game is equal to Tsirelson's bound; this is proven via an explicit mapping from the strategies in the CHSH* game to the strategies in CHSH

[1] We will sometimes use the acronym RACs, instead of QRACs, to address the cases that, in principle, may not use quantum resources.

game (lemma 6). We then illustrate that the game is sensitive to a broad range of properties of the system used, specifically whether the system is quantum or classical, what is the set of operations allowed to the player (namely reversible versus irreversible and Clifford versus non-Clifford) and what is the dimension of the system. We demonstrate that the Bell bound holds for classical reversible strategies and quantum strategies involving only Clifford computation, while the possibility of performing irreversible computation allows one to win the game with certainty. Moreover, following Landauer's assertion that only reversible operations are truly fundamental, we show that bit erasure is a powerful tool for increasing the winning probability, shedding light on the source of quantum advantage in this game. We finally conjecture that our results also apply to the CHSH_q^* game for any dimension q, by considering the case of $q = 3$.

In the reminder of the chapter we start by covering some background material on non-locality and CHSH games, other protocols related to the CHSH game and Landauer's principle (section 4.1.1). In section 4.2 we then introduce the CHSH* game, its relation with the CHSH game and its characterisation in terms of the system and gates used. Given the crucial role of irreversible versus reversible computation for the performances of the protocol, we draw a connection with Landauer's principle in section 4.3. In the same section we also discuss the presence of a new notion of contextuality in certain quantum strategies for the CHSH* game. We briefly treat the case of the CHSH_q^* game in section 4.4 and we sum-up and propose possible future projects in the conclusion section.

4.1 Background

This section does not contain original material. The references that have been used will be specified in the corresponding subsections.

4.1.1 Non-Locality and CHSH game

The philosophical consequences of quantum mechanics troubled many prominent physicists since the early stages, as witnessed by the Einstein-Podolsky-Rosen (EPR) paper [36] in 1935, that aimed at demonstrating the incompleteness of quantum mechanics. The EPR argument is based on the assumptions of locality and realism (also known as local realism), that, in the

famous words by Einstein, can be summed up as "no spooky action at distance" and "the moon is there even if no one is looking at it". In 1964 John Stewart Bell, inspired by this work, provided a mathematical precise formulation of these concepts and proved the inconsistency of quantum mechanics with local realism [128]. This inconsistence can be stated as the violation of the so-called Bell's inequality. We now treat the Bell scenario as formulated in 1969 by Clauser-Horne-Shimony-Holt [59].

Let us consider two parties, Alice and Bob, initially together and then taken apart to space-like separated regions. Each party can measure one of two observables labelled by $a, b \in 0, 1$, $A_{a=0}, A_{a=1}$ and $B_{b=0}, B_{b=1}$ respectively, with outcomes q for Alice's observables and p for Bob's observables, where $q, p \in \{-1, +1\}$ (figure 4.1). Freedom of choice of the measurements of Alice and Bob is assumed. [2] We further assume locality, which means that an event can only be affected by events happening within its past light-cone, *i.e.* there is no instantaneous (faster than light) influence between the two parties that are space-like separated. Therefore the outcomes of Alice's observables do not depend on which measurement Bob chooses to perform on his system and vice versa. We also assume realism, which here means that we assume the ontological model framework, *i.e.* the outcomes of the measurement observables depend on some set of pre-existing properties of each party's system and they exist even if the measurements are not performed. With the assumptions of local realism the values in $\{-1, +1\}$ can be assigned simultaneously to all four observables and this implies the following inequality, known as CHSH inequality:[3]

$$| \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle | \leq 2, \tag{4.1}$$

where the angular brackets denote the expectation values of the outcomes associated to the observables. This inequality can be used to test any operational theory. Let us consider quantum mechanics. If Alice and Bob share a Bell state like the one of equation (3.22) and the four observables are particular Hermitian operators like the ones in (3.23), then the above inequality is violated (see section 3.5). Tsirelson [60] proved that the maximum violation, when considering quantum mechanics, is $2\sqrt{2}$. We stress that, despite entanglement being necessary for providing the violation of CHSH inequality, it is not true that every entangled state can

[2] See [129] and [130] for the consequences of dropping it.
[3] This inequality actually belongs to a class of CHSH inequalities for this scenario [131].

provide such a violation [132]. However, it has been proven that for every entangled state ρ_1 there exists another state ρ_2 not violating the CHSH inequality, such that $\rho_1 \otimes \rho_2$ violates it [133]. In this sense any entangled state actually encodes an amount of non-locality.

Bell's theorem states that no local realistic theory is compatible with quantum mechanics. They cannot provide the same statistics for outcomes of measurements distributed in space. Moreover, in 2015 a loophole-free Bell experiment that confirms the validity of Bell's theorem was performed [134].

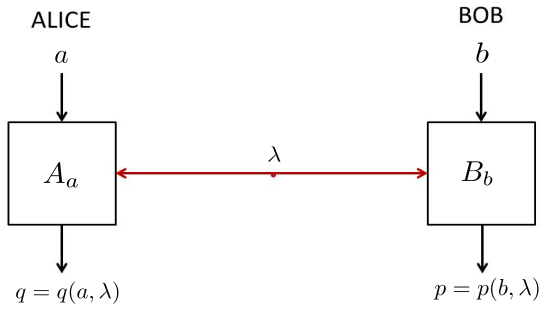


Figure 4.1: **Bell's scenario.** The figure above schematically depicts Bell's scenario, where two space-like separated parties, Alice and Bob, each perform a measurement observable that they can freely choose between two possible ones labelled by bits a for Alice and b for Bob. The outcomes of the observables are denoted with q for Alice and p for Bob. In a local realistic model these outcomes depend on some ontic states (hidden variables) here denoted with λ and the choice of the observable.

Bell's theorem is often colloquially restated as implying that quantum mechanics is non-local. Furthermore, non-local correlations have been treated as an information-theoretic resource [135–137]. With (bipartite) non-local correlations we mean correlations that cannot be given by the set of probabilities of the form

$$p(q, p|a, b) = \sum_{\lambda} p(\lambda) p(q|a, \lambda) p(p|b, \lambda), \tag{4.2}$$

where λ is the ontic state that, together with the choice of observables, influences the outcomes q, p of Alice and Bob in a local realistic model. The main application in this direction concerns the field of device independent quantum information processing, where the security of the

protocols does not rely on any assumption about the properties of the device. In this setting all the details of the devices are ignored and only the statistics of measurements matters. Protocols like quantum key distribution [63, 138] and randomness generation [139, 140] can be carried out with this approach. The former protocol shows that the cryptographic security against general attacks by a postquantum eavesdropper (limited only by the impossibility of superluminal signaling) is guaranteed by the violation of a Bell inequality. The latter protocol exploits the non-locality of quantum systems to certify the presence of genuine randomness.

The Bell scenario just described can be recast into a game called the *CHSH game* (figure 4.2). In the CHSH game a referee uniformly asks questions $a, b \in \{0, 1\}$ to Alice and Bob, respectively, who agree on a strategy beforehand to then answer, when separated and unable to communicate, with bits $x, y \in \{0, 1\}$, respectively. They win the game if $x \oplus y = a \cdot b \bmod 2$. Notice, with respect to the Bell scenario above, the notation here is such that $q = (-1)^x$ and $p = (-1)^y$.

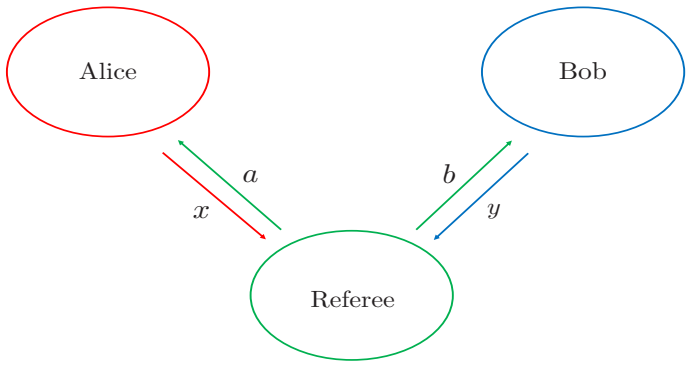


Figure 4.2: **CHSH game.** In the CHSH game a referee asks binary questions, a, b, to Alice and Bob, who answer with bits x, y. After the questions have been asked, they can no longer communicate. They win the game if $x \oplus y = a \cdot b \bmod 2$.

In game theory, the optimal success probability for a game is called its *value*, which we denote by ω. The value of the CHSH game, $\omega(\text{CHSH})$, depends upon the physics of the systems exploited by Alice and Bob. Famously, if Alice and Bob employ only classical strategies, the value of the CHSH game is $\omega_C(\text{CHSH}) = 0.75$. An optimal classical strategy consists of always output $x = 0, y = 0$, thus winning 3 over 4 times (always except when the inputs are $a = 1, b = 1$). On the other hand, if they have access to quantum resources, $\omega_Q(\text{CHSH}) =$

$\cos^2(\frac{\pi}{8}) \approx 0.85$. More precisely, an optimal strategy, as already mentioned in section 3.5, involves the Bell state of equation (3.22) and the observables of equation (3.23). The value 0.85 can be found by noticing that the (weighted) expression in the CHSH inequality (4.1), $\frac{1}{4}(\langle A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1\rangle)$, represents the probability that Alice and Bob win minus the probability that they lose. Since $\langle A_0B_0\rangle = \langle A_0B_1\rangle = \langle A_1B_0\rangle = -\langle A_1B_1\rangle = \frac{1}{\sqrt{2}}$, the probability of success is $\frac{1}{2} + \frac{1}{2\sqrt{2}} \approx 0.85$. The limitation on the value of the game for classical systems is called a Bell inequality, and the value 0.75 is often called the Bell bound. The fact that the value of the game when using quantum resources violates the Bell inequality, but is nevertheless limited substantially below 1, was first noted by Tsirelson [60], and the value $\cos^2(\frac{\pi}{8})$ is known as Tsirelson's bound. Popescu and Rohrlich [141] noted that in more general theories than quantum mechanics, perfect strategies for the CHSH game that achieve a value of 1 could exist via a correlation now known as a Popescu-Rohrlich (PR) box, without violating the no-signaling assumption between Alice and Bob.

The CHSH game can be generalised to arithmetic modulo q, where $a, b, x, y \in \mathbb{Z}_q$. The so-called CHSH_q game was first introduced by Buhrman and Massar in 2005 [64]. It was defined for q being prime or prime power and studied for the case of $q = 3$. They found the Bell bound to be $\frac{2}{3} \approx 0.66$ and an upper bound on Tsirelson's bound which reads as $\frac{1}{3} + \frac{2}{3\sqrt{3}} \approx 0.71823$. In 2009, Liang *et al.* [66] developed a numerical analysis for lower and upper bounds (see [66, Table III]). In the case of $q = 3$ they agree on the value of 0.7124 for the Tsirelson bound. This is confirmed also by the lower bound analytically found by Ji *et al.* one year earlier [65], $\frac{1}{3} + \frac{2}{3\sqrt{3}} \cdot \cos(\frac{\pi}{18}) \approx 0.7124$. However, an analytic proof of the actual value of Tsirelson's bound does not exist and, in 2015, Bavarian and Shor [67], exploiting new tools from incidence geometry and arithmetic combinatorics, provided analytic proofs of upper bounds on Tsirelson's bounds for any prime or prime power q that agree with the upper bound provided by Buhrman and Massar for $q = 3$ (and in line with the $q = 2$ case known to be tight [60]):

$$\omega_Q(\text{CHSH}) \leq \frac{1}{q} + \frac{q-1}{q\sqrt{q}}. \tag{4.3}$$

The question on the tightness of this bound for $q > 2$ remains still open.

4.1.2 Other related games

There exist protocols related to the CHSH game, where the computational advantages that arise when exploiting quantum strategies can be associated to the presence of non-classical features different from non-locality. One of those protocols goes under the name of *quantum random access codes* (QRACs). It first appeared in a paper by Wiesner published in 1983 [142] and was then rediscovered by Ambainis *et al.* in [143] and studied by Galvao in his PhD thesis [68] in 2002. Let us imagine that Alice encodes m bits in $n < m$ information carriers that she sends to Bob, who wishes to learn the value of a single bit among the m ones (without Alice to know which one) with a probability at least p (figure 4.3). We denote this scheme with the notation $m \to n$. They have to agree on a particular efficient encoding to maximise the least probability of success. QRACs have been generalized and studied also considering qudits of arbitrary dimensions [144].

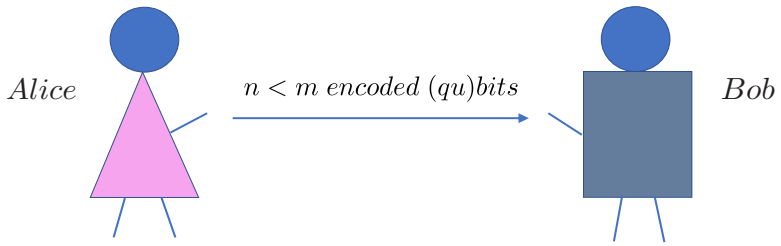


Figure 4.3: **Quantum random access codes.** QRACs consist of Alice encoding m bits in $n < m$ information carriers that she sends to Bob. He wants to know the value of one of the m bits, but Alice does not know which one. Their goal is to come up with an encoding strategy to maximise the least probability of success.

Let us focus, for simplicity, on the $2 \to 1$ protocol. It turns out that, analogously to the CHSH game, the optimal classical strategy succeeds with probability $\omega_C(\text{QRAC}) = 0.75$, while the optimal quantum strategy achieves $\omega_Q(\text{QRAC}) = cos^2(\frac{\pi}{8}) \approx 0.85$. A strategy for the former consists of Alice sending the bit 0 to encode the bits 00, the bit 1 to encode the bits 11 (thus succeeding with probability 1 in these two cases) and sending the bit 0 (or 1) for encoding 01 and 10 (thus succeeding with probability 0.5 in these two cases). Therefore, on average, the probability of success is $\frac{1}{4}(1 + 1 + 0.5 + 0.5) = 0.75$. An optimal strategy for the quantum case consists of Alice sending a qubit in the state $|\psi_{00}\rangle = R_z(\frac{\pi}{4})|+\rangle = T|+\rangle$ to

encode the bits 00, in the state $|\psi_{01}\rangle = R_z(\frac{7\pi}{4})|+\rangle = T^\dagger|+\rangle$ to encode the bits 01, in the state $|\psi_{10}\rangle = R_z(\frac{3\pi}{4})|+\rangle = ST|+\rangle$ to encode the bits 10, in the state $|\psi_{11}\rangle = R_z(\frac{5\pi}{4})|+\rangle = S^\dagger T^\dagger|+\rangle$ to encode the bits 11. Here $R_z(\theta)$ represents a rotation of angle θ around the z-axis in the usual Bloch sphere representation of the qubit and $S = R_z(\frac{\pi}{2})$ and $T = R_z(\frac{\pi}{4})$ are the gates already defined in equations (2.19) and (3.2). This means that the states above lie in the XY plane of the Bloch sphere. Bob then needs to measure on the X basis if he wants to know the first bit, and on the Y basis if he wants to know the second bit (the positive eigenvalues are associated to the bit 0 and the negative ones to the bit 1). The probability of obtaining the outcome corresponding to the correct bit is therefore $cos^2(\frac{\pi}{8}) \approx 0.85$ for each of the four cases above. This strategy is strictly related to the one described in details in section 4.2 and depicted in figure 4.9.

We have just seen that $2 \to 1$ QRACs turn out to be related to the CHSH game as they provide the same bounds and, as we will further advocate in the next section, the optimal classical and quantum strategies are strictly related to the ones used in the CHSH protocol. Indeed, these facts will also hold for the CHSH* game that we present in this chapter (section 4.2).[4] The source of non-classicality here derives from the fact that Bob uses non-commutative measurements and that the states sent by Alice "lie" in between these two measurements (figure 4.9). This possibility is not achievable when using only classical resources. Moreover, slight modifications of the protocol, where in the optimal quantum strategy Alice prepares her states with local projective measurements on an entangled state, show the presence of contextuality as a necessary resource for the quantum advantage [68].

We now describe a protocol similar to QRACs that again resembles the CHSH* game we treat in this chapter, even if, unlike the CHSH* game, it shows preparation contextuality as a necessary resource for the quantum computational advantage [69]. This protocol is called *parity oblivious multiplexing* (POM) and it was firstly introduced in 2009 by Spekkens *et al* [69]. Let us take the QRACs previously defined. Let us denote the $m-$bit string that Alice possesses with x and, instead of requiring the information carriers to be n systems (bits or qubits), let us impose a different constraint, called *parity obliviousness*: Alice cannot communicate to Bob the parity of the $m-$bit string x. More formally, let $s \in Par$, where $Par = \{r \in \{0,1\}^m \mid \sum_i r_i \geq 2\}$, *i.e.*

[4] Similar connections exist also with other protocols like quantum dense coding and remote state preparation, as shown in [68].

Par is the set of $m-$bit strings with at least two bits in the state 1; Alice cannot transmit to Bob any information about the *s-parity*, *i.e.* $s \cdot x = \bigoplus_i s_i x_i$, where $\oplus$ denotes the sum modulo 2. Let us denote the bit that Bob outputs as b. The integer y denotes which of the m bits b should correspond to, and x_y the actual bit in Alice's string.

The optimal classical probability of success satisfies $p(b = x_y) \leq \frac{m+1}{2m}$, as the only classical encoding that transfers some information to Bob without violating the parity obliviousness consists of encoding only a single bit x_i. Given that y is chosen at random, any bit x_i would perform the same. Therefore Alice and Bob can agree on Alice always sending x_1 and Bob outputting $b = x_1$. The probability of success is given by the probability that $y = 1$, which is $\frac{1}{m}$, and the probability that Bob outputs correctly (at random, with probability 0.5) in the other cases where $y \neq 1$, that occur with probability $\frac{(m-1)}{m}$. For this optimal classical strategy we obtain $p(b = x_y) = \frac{1}{m} + \frac{(m-1)}{2m} = \frac{m+1}{2m}$, as already stated. When $m = 2$, this amounts to $\omega_C(\text{POM}) = 0.75$, like the Bell bound of the CHSH game and QRACs. Spekkens *et al* proved the following theorem.

Theorem 12. *The optimal success probability in $m-$bit parity oblivious multiplexing of any operational theory that admits a preparation non-contextual ontological model satisfies $p(b = x_y) \leq \frac{m+1}{2m}$.*

This theorem means that preparation contextuality is a necessary resource for performing the $m-$bit parity oblivious multiplexing protocol with higher success probability than classical strategies. The proof is based on first showing that for preparation non-contextual ontological models, parity obliviousness at the operational level implies parity obliviousness at the ontological level. More formally, parity obliviousness at the operational level can be written as $\forall s \ \forall M \ \forall k \ \sum_{x|x\cdot s=0} p(P_x|k, M) = \sum_{x|x\cdot s=1} p(P_x|k, M)$, where P_x is a preparation procedure implemented by Alice and M is the measurement of Bob with outcome k performed to provide the output $b = k$. Parity obliviousness at the ontological level can be written as $\forall s \ \sum_{x|x\cdot s=0} p(P_x|\lambda) = \sum_{x|x\cdot s=1} p(P_x|\lambda)$, where λ is the hidden variable prepared by P_x. Then, it is enough to realise that λ provides a classical encoding of x without any information about the s-parity, which, as already shown, means that $P(b = x_y) \leq \frac{m+1}{2m}$. Thus, even if Bob could perfectly determine λ, the two parties could not achieve a better performance than $\omega_C(\text{POM})$.

Let us now consider, for simplicity, the case of $2-$bit parity oblivious multiplexing. It turns

out that, by using the same optimal quantum strategy of $2 \to 1$ QRACs, the probability of success is again $\omega_Q(\text{POM}) = \cos^2(\frac{\pi}{8}) \approx 0.85$. It can be shown that it is the maximal one [69]. Notice that, when Bob measures on X or Y basis, he cannot gain any information about the parity, as the parity 0 and parity 1 mixtures are represented by the same quantum state (and so the same probability), $\frac{1}{2}\rho_{00} + \frac{1}{2}\rho_{11} = \frac{\mathbb{I}}{2} = \frac{1}{2}\rho_{01} + \frac{1}{2}\rho_{10}$, where $\rho_{ij} = |\psi_{ij}\rangle \langle\psi_{ij}|$ and $|\psi_{ij}\rangle$ are the states defined previously.

4.1.3 Landauer's principle

We conclude the background section by briefly reviewing Landauer's principle. In 1961 Rolf Landauer [70] formulated his famous principle that we here state *verbatim* from [145].

Landauer's principle. *Any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information-bearing degrees of freedom of the information-processing apparatus or its environment.*

The principle arose because Landauer noticed that the logical states of the computation (*e.g.* logical bits) evolve sometimes irreversibly, with a single logical state resulting from several logical states. This irreversible operation of the information-bearing-degrees of freedom is associated to a decrease in entropy and, considering the reversibility of Hamiltonian/unitary dynamics (that preserves the entropy), it must therefore be compensated by a rise in the entropy of the non-information-bearing degrees of freedom (*e.g.* physical bits) and the environment. Landauer's principle assumes that irreversible operations are not fundamental. We can always imagine an irreversible operation as a reversible operation plus erasure of some information (or a copy of this information stored in another system), where the erasure consists of an increase in entropy. More precisely, we associate the erasure of a single bit with an increase in entropy of $kT \log_2 2$, where k is the Boltzmann constant and T the temperature of the system and environment.

The remainder of the chapter mainly treats the content of [58], which is a joint work with Luciana Henaut, Dan Browne, Shane Mansfield and Anna Pappa.

4.2 The CHSH* game

We now describe the CHSH* game (illustrated in Fig. 4.4). A single player has in her possession a single system of dimension d, that can be classical or quantum. She is given a specification of the state preparations, transformations and measurements that she is allowed to employ and in the course of the game, she is also provided with two uniformly random bits a and b. Choosing from the allowed operations, the player must specify in advance an initial state, controlled operations A_a and B_b and a final two-outcome measurement M. Once the player receives a and b, the corresponding operations are implemented in sequence and measurement M is performed, returning outcome c. The player wins the game when $c = a \cdot b \bmod 2$. We are interested in finding the value $\omega(\text{CHSH}^*)$ of this game, which corresponds to the average winning probability of the best possible strategies:

$$\omega(\text{CHSH}^*) = \max_{\text{all strategies}} \frac{1}{4} \sum_{a,b \in \mathbb{Z}_2} p(c = a \cdot b \mid a, b).$$

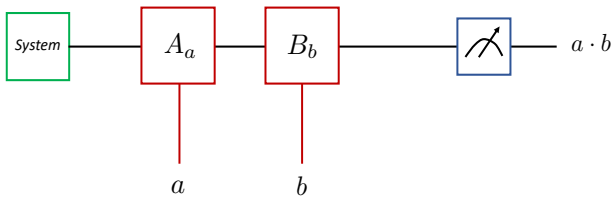


Figure 4.4: **CHSH* game.** An initial system is subjected to controlled transformations, with control bits a and b, respectively, and then measured. The goal is to maximise the probability that the value of the output is the product of the values of the input bits (arithmetic modulo 2).

4.2.1 Relationship with the CHSH game

In this chapter we will study the CHSH* game in a variety of *settings* (see Fig. 4.5), where we make different assumptions about the physics of the system available to the player. First, we consider the case where the player's system is a single qubit in the *unitary setting*, meaning that all transformations applied during the game are unitary. We further assume that the final measurement is a projective two-outcome measurement. In this setting, similarly to the

quantum cases in the CHSH game, $2 \to 1$ QRACs and $2-$bits parity oblivious multiplexing, the optimal success probability is bounded by the Tsirelson bound, as shown by the following proposition.

Name of setting	System Type	Initial states	Transformations	Measurements	ω(CHSH*)
Unitary	Quantum	Any	Any unitary gate	Any two-outcome PVM	$\cos^2(\frac{\pi}{8})$
Clifford	Quantum	Pauli eigenstates	Clifford group gates	Pauli measurements	0.75
Reversible Classical	Classical	Any	Reversible gates	n/a	0.75
Irreversible	Classical/Quantum	Any	Any	Any	1

Figure 4.5: **Several settings for the CHSH* game.** The four $d = 2$ settings we study with the CHSH* game. The value of the game is dependent on the setting and the dimension d of the system.

Proposition 2. *The value of the CHSH* game with a $d = 2$ quantum system in the unitary setting is* $\cos^2(\frac{\pi}{8})$.

This result follows directly from the following lemma.

Lemma 6. *For every strategy in the CHSH* game in the unitary setting with $d = 2$, we can derive an equivalent strategy for the two-player CHSH game such that both strategies lead to the same average success probability.*

Proof. We prove this explicitly. We first consider the CHSH* game and assume without loss of generality that the initial state is $|+\rangle$ and the measurement is the Pauli X observable (whose positive eigenvalue is associated to the output bit 0 and the negative to the output bit 1). A strategy thus consists of optimally choosing the gates A_0, A_1, B_0, B_1.

In Fig. 4.7, we show how, given a strategy for the CHSH* game, we can construct a strategy for the CHSH game. The key ingredient is a teleportation protocol that uses entanglement shared via the CNOT gate to teleport the effect of gate A_a from one site (Alice's) to another spatially separated site (Bob's). Since operations A_a are unitary, it holds that

$$A_a^T \otimes \mathbb{I}\Big(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\Big) = \mathbb{I} \otimes A_a\Big(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\Big).$$

The teleported state on Bob's side after Alice measures her qubit is $A_a Z^x |+\rangle$, where Z is the Pauli Z. The bits x and y are Alice's and Bob's outputs respectively. In order to prove the lemma, we will show that the success probabilities for obtaining $c = a \cdot b$ in the CHSH* game

and $x \oplus y = a \cdot b$ in the CHSH game are equal, i.e.:

$$\sum_{a,b} \Pr\left(c = a \cdot b | a, b\right) = \sum_{a,b} \Pr\left(x \oplus y = a \cdot b | a, b\right).$$

We proceed by showing that the terms in the above sums are pairwise equal, i.e. for every $a, b \in \{0, 1\}$,

$$\Pr\left(c = a \cdot b \mid a, b\right) = \Pr\left(x \oplus y = a \cdot b \mid a, b\right).$$

In the case that $x = 0$ this holds trivially; and when $x = 1$, this reduces to showing that

$$\begin{aligned} |\langle +| B_b A_a |+\rangle|^2 &= |\langle -| B_b A_a |-\rangle|^2 \\ |\langle -| B_b A_a |+\rangle|^2 &= |\langle +| B_b A_a |-\rangle|^2, \end{aligned}$$

which is necessarily true for any 2×2 unitary gate. □

To see that Lemma 6 implies Proposition 2 we recall that Tsirelson's bound upperbounds the CHSH game at probability $\cos^2(\frac{\pi}{8}) \approx 0.85$. A strategy which achieves this success probability involves the following gates: $A_0 = \mathbb{I}, A_1 = S, B_0 = T^\dagger, B_1 = T$. Indeed, the probability of success in this case is given by

$$\begin{aligned} p_{\text{suc}} &= \frac{1}{4} \sum_{a,b \in \mathbb{Z}_2} p(c = a \cdot b \mid a, b) \\ &= \frac{1}{4} \Big[|\langle +| B_0 A_0 |+\rangle|^2 + |\langle +| B_1 A_0 |+\rangle|^2 \\ &+ |\langle +| B_0 A_1 |+\rangle|^2 + (1 - |\langle +| B_1 A_1 |+\rangle|^2) \Big] \\ &= \frac{1}{4} \sum_{a,b \in \mathbb{Z}_2} \big[\frac{1}{2} + (-1)^{a \cdot b} \frac{\cos(\theta_{ab})}{2} \big], \end{aligned} \tag{4.4}$$

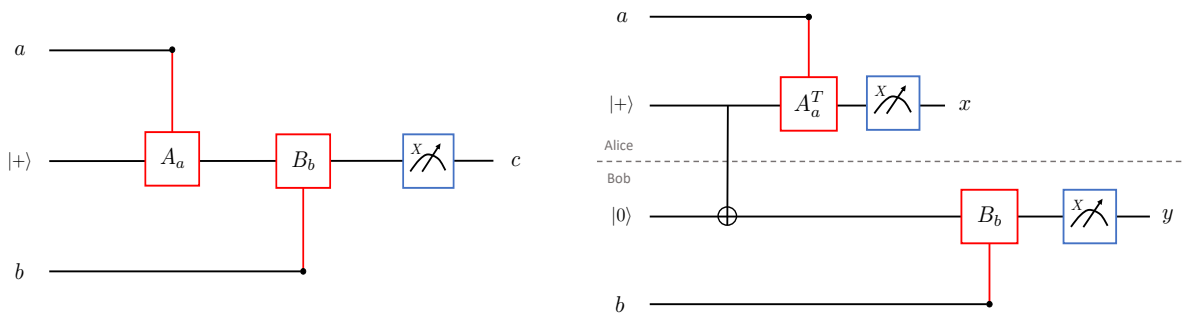
where the angle θ_{ab} is the overall phase resulting from the application of $B_b A_a = R_z(\theta_{ab})$ on the input state $|+\rangle$. With the choice of gates above we obtain $p_{\text{suc}} = \cos^2(\frac{\pi}{8}) \approx 0.85$. Figure 4.6 shows the states $B_b A_a |+\rangle$ and the values of the probabilities $p(c|a, b)$ for the four possible input bits a, b.

The unitaries of the optimal strategy just described are the gates mapping between the

a	b	$B_b A_a \|+\rangle$	$p(0\|a,b)$	$p(1\|a,b)$	$a \cdot b \bmod 2$
0	0	$T^\dagger \|+\rangle = R_Z(-\frac{\pi}{4}) \|+\rangle$	0.85	0.15	0
0	1	$T \|+\rangle = R_Z(\frac{\pi}{4}) \|+\rangle$	0.85	0.15	0
1	0	$ST^\dagger \|+\rangle = R_Z(\frac{\pi}{4}) \|+\rangle$	0.85	0.15	0
1	1	$ST \|+\rangle = R_Z(\frac{3\pi}{4}) \|+\rangle$	0.15	0.85	1

Figure 4.6: **Optimal quantum strategy for the CHSH* game.** The table above reports the state, $B_b A_a \|+\rangle$, before the measurement on the X basis and the probability $p(c|a,b)$ for each input bits a, b in the optimal quantum strategy, given by gates $A_0 = \mathbb{I}, A_1 = S, B_0 = T, B_1 = T^\dagger$. For every input bits a, b the probability of obtaining $a \cdot b \bmod 2$ is $\cos^2(\frac{\pi}{8}) \approx 0.85$.

observables typically used to attain the Tsirelson bound in the CHSH game when the parties share a Bell pair. This strategy is also strictly related to the optimal strategies used in other tasks involving one qubit, like $2 \to 1$ QRACs [68] and $2-$bits parity oblivious multiplexing [69] defined in subsection 4.1.2. Lemma 6 demonstrates a tight link between Tsirelson's bound for the CHSH game and the value of the CHSH* game in the above setting.



(a) (b)

Figure 4.7: **Mapping of the CHSH* game to the CHSH game.** Fig. 4.7a shows the single qubit scheme, with the initial qubit in state $|+\rangle$, controlled gates A_a, B_b, measurement on the X basis and output c. Figure 4.7b shows the corresponding CHSH game, where Alice and Bob share a Bell pair, and apply gates A_a^T, B_b to their systems to obtain measurement results x and y respectively.

4.2.2 Further settings

The proof of Lemma 6 relied on the fact that the transformations are unitary, and that the system in the CHSH* game had dimension 2. We will now study the game in other settings, and see that its value is strongly setting-dependent.

First, we relax the restriction that transformations must be unitary by considering the *irreversible setting*. We now allow irreversible transformations, such as the ERASE map, which maps any qubit state to the state $|0\rangle$. This may be achieved via a Z measurement and conditional X correction. Introducing irreversible transformations has a dramatic effect on the value of the CHSH* game.

Proposition 3. *The value of the CHSH* game with a $d = 2$ classical or quantum system in the irreversible setting is 1.*

Proof. Proof is via explicit example. Let the initial state be $|0\rangle$ and let $A_0 = \mathbb{I},\ A_1 = X,\ B_0 = \mathsf{ERASE}, B_1 = \mathbb{I}$. The final measurement is in the Z basis. Considering the 4 cases, we see that the output c will always be 0 unless both a and b are 1. Thus this strategy always wins the game. Every element of the strategy presented in this proof can be achieved in a classical system, hence we can conclude that this maximum value of 1 can be achieved even with no quantum dynamics at all. □

This increase in the value of the game depends crucially on the *irreversibility* of the ERASE map. As we see directly, if we restrict logic operations to be reversible, we find that the value of the game is reduced.

Proposition 4. *The value of the CHSH* game with a $d = 2$ classical system in the reversible setting is 0.75.*

Proof. To show that the value is at least 0.75, it suffices to describe a protocol which attains this success probability. This is given by the trivial protocol where the input bit is set to 0 and gates A_a and B_b are the identity, and thus the output is always 0. To see why this cannot be exceeded, we observe that all reversible one-bit functions are linear functions. The closest linear function to $a \cdot b$ is the constant function $f(a, b) = 0$. This result can also be found by just enumerating all the possible classical strategies. □

So far we have studied the CHSH* game with a variety of restrictions on the system and we have found values of the game of 0.75, $\cos^2(\frac{\pi}{8})$ and 1, depending on the setting. These precisely match the Bell bound, Tsirelson bound and PR-box value of the CHSH game. We now show that the CHSH* game is sensitive to further restrictions. We denote the *Clifford setting* as the setting where the initial system is a pure stabilizer state, all transformations are unitary Clifford and the measurement is a Pauli observable (definitions provided in subsection 2.1.2).

Proposition 5. *The value of the CHSH* game with a $d = 2$ quantum system in the Clifford setting is 0.75.*

Proof. The state $B_b A_a |+\rangle$ before the measurement is an eigenstate of Pauli operators, which, when measured on the Pauli X operator, will always yield one of the possible outcomes with probability 0, 0.5 or 1. Therefore the probability of success for any choices of input bits a and b will always take one of eight possible values in $\{0, \frac{1}{8}, \ldots, \frac{7}{8}, 1\}$. Since the maximum probability of success of our protocol is about 0.85 in the less restricted unitary setting, we conclude that the maximum attainable probability of CHSH* in the Clifford setting is 0.75. □

We see that restricting the CHSH* game to the Clifford setting gives a success probability equal to the reversible classical setting. This, again, resembles the CHSH game, where if states, operations and measurements are similarly limited, the Bell inequality value of 0.75 cannot be surpassed. We now show that when diagonal non-Clifford gates are available, one can always do better than this bound.

Proposition 6. *For a quantum system with $d = 2$, in the Clifford setting but with the addition of any pair of non-Clifford gates $R_z(\varepsilon)$ and $R_z(\varepsilon)^\dagger$, with $\varepsilon \in (0, \frac{\pi}{2})$, the value of the CHSH* game is greater than* 0.75.

Proof. The proof is via explicit construction. We adopt a strategy similar to the optimal quantum strategy in the unitary setting, where replacing T with $R_z(\varepsilon)$ and $T^\dagger$ with $R_z^\dagger(\varepsilon)$, achieves a probability of success p_{suc} greater than 0.75:

$$p_{\text{suc}} = \frac{1}{4}\left[\left(\frac{1}{2} + \frac{\cos(\varepsilon)}{2}\right) + \left(\frac{1}{2} + \frac{\cos(-\varepsilon)}{2}\right) + \left(\frac{1}{2} + \frac{\cos(\frac{\pi}{2} - \varepsilon)}{2}\right) + \left(1 - \frac{1}{2} - \frac{\cos(\frac{\pi}{2} + \varepsilon)}{2}\right)\right].$$

This probability is always greater than 0.75 when $\varepsilon \in (0, \frac{\pi}{2})$, and attains a maximum of $\cos^2(\frac{\pi}{8})$ when $\varepsilon = \frac{\pi}{4}$ as expected (see figure 4.8). □

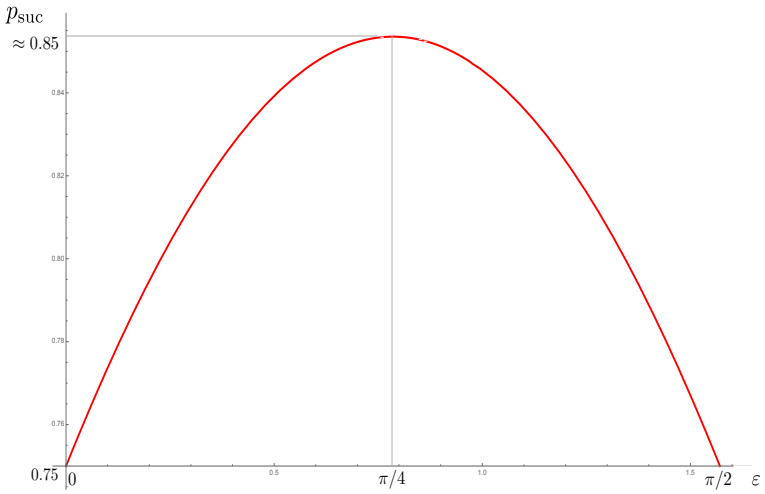


Figure 4.8: **Success probability varying** $\varepsilon \in (0, \frac{\pi}{2})$. Any pair of non-Clifford gates $R_z(\varepsilon)$ and $R_z(\varepsilon)^\dagger$, with $\varepsilon \in (0, \frac{\pi}{2})$, allow us to win the CHSH* game with probability greater than the classical value $\omega_C(CHSH^*) = 0.75$. Notice that the argument works the same for ε outside the interval $(0, \frac{\pi}{2})$ by rotating the controlled gates accordingly.

Figure 4.9 provides a geometrical comparison of optimal strategies in the three reversible settings we have considered.

Having seen that the value of the CHSH* game allows us to distinguish between various settings with systems of dimension 2, we will now consider systems of higher dimension, beginning with dimension 3.

Proposition 7. *For d-dimensional quantum or classical systems, in the reversible setting with* $d \geq 3$, *there always exists a perfect strategy (i.e. the value of the game is 1).*

Proof. We provide a qutrit strategy, and note that this can always be embedded into systems of dimension greater than 3. Without loss of generality we suppose that the system is prepared in the state $|0\rangle$, and the strategy consists of the gates $A_0 = \mathbb{I}, A_1 = X, B_0 = \mathbb{I}, B_1 = X$. The generalised Pauli X acts as $X|i\rangle = |i+1\rangle$, where $i \in \{0, 1, 2\}$ and the sum is mod3. The measurement is given by the PVM $\{|0\rangle\langle 0| + |1\rangle\langle 1|, |2\rangle\langle 2|\}$. If we associate the outcome 0 to the first element of the measurement and the outcome 1 to the second, we obtain $a \cdot b \bmod 2$

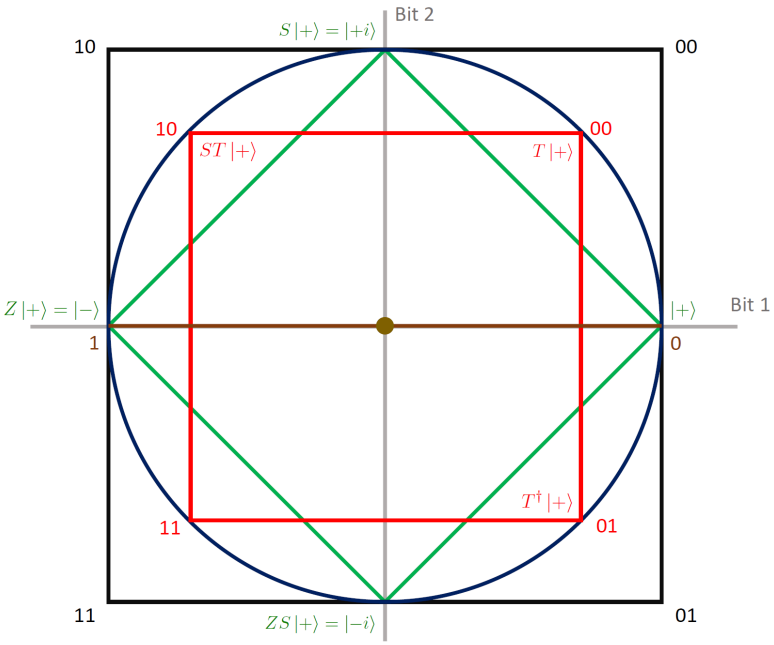


Figure 4.9: **Geometrical analysis of the protocol.** The figure shows the state space of two bits (vertices of the big black square), one qubit (XY plane of the Bloch sphere) both in the optimal winning strategy (the vertices of the red square) and restricted to Clifford computation (the vertices of the tilted green square), and one bit (the edges of the brown line). Notice that the measurement at the end of the protocol corresponds to the collapse of a state to the X axis. This geometrical representation provides an intuition of why the different settings give different values $\omega(CHSH^*)$. Two bits are needed to obtain the non-linear function with probability 1. This can be seen also as one of the two bits being erased in accordance with Landauer's principle (*i.e.* the irreversible setting). In the unitary setting, the single qubit in the optimal quantum strategy can be seen as two bits where the erasure is just partial (the red square can be seen as a smaller version of the black square). The Clifford setting does not allow more possibilities than the reversible classical setting – it indeed provides a value of 0.75 – even if the stabilizer qubit can reach more states than the single bit (as a curiosity, notice that the single stabilizer qubit corresponds to knowing one bit and being completely ignorant about the other, exactly like the epistemic states of ST in subsection 2.1.1). Outputting a random bit would correspond to the origin (that can be seen as an infinitesimally small square), which would always provide a success probability of 0.5.

with probability 1. Notice that this strategy can equally be applied in the case of a classical trit, using the obvious analogous state and reversible gates. □

This shows that, if the operations on the system are restricted to reversible gates, the CHSH* game is a *dimensional witness*, as it can witness when the dimension of the system is at least 3.

4.3 Sources of computational advantages

We now analyse what are the physical reasons for the different performances of the protocol in the settings that we considered. We first develop an analysis in terms of Landauer's principle and then we discuss the presence of contextuality for quantum strategies that achieve a probability of success higher than the Bell bound.

4.3.1 Connection to Landauer's principle

We have seen that under the assumption that only reversible gates are employed, the CHSH* game acts as a witness that distinguishes quantum and classical systems, and systems of different dimension. How reasonable is it to restrict the operations to reversible transformations? As described in subsection 4.1.3, it was first argued by Landauer that irreversible operations are not fundamental and that every irreversible classical operation on logical bits must be accompanied by a rise in the entropy of the non-information bearing degrees of the system or its environment [70]. This holds because in order to build an irreversible gate out of fundamentally reversible operations, we need to discard or erase information.

We have seen that erasure is a powerful tool that allows to win the CHSH* game with certainty. Reversible classical and quantum settings lead to distinct lower values for the game. This can be seen as a reflection of the non-classical nature of quantum information storage and measurement.

Moreover, following Landauer's approach, in the optimal strategy presented for the irreversible setting, winning the game with certainty requires the erasure of one bit for only one of the four input combinations of a and b. On average, the heat generated by the protocol is therefore $\frac{1}{4}kT\log_2 2$. With similar considerations, we can imagine an optimal quantum strategy in the unitary setting as implementing a *partial* erasure. We can quantify the heat generated by this partial erasure as, on average, $\frac{1}{4}0.41kT\log_2 2$ corresponding to the probability of success $\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}(\sqrt{2}-1)\approx 0.85$. In other words, to increase the winning probability for the game in the classical reversible setting to unity, $\frac{1}{4}kT\log_2 2$ information would need to be erased, whereas to do so for the unitary setting only $\frac{1}{4}0.59kT\log_2 2$ must be erased.

We can interpret the success probability as how much the chosen setting allows us to learn

about the irreversible function $a \cdot b$. The quantum resource in this protocol is the qubit's ability to simulate two classical bits (one of which is going to be erased). This is made even more explicit in Figure 4.9, which compares the state spaces of a pair of bits, a single qubit and a single bit. In particular, in the optimal quantum strategy the single-qubit state space (that mimics the two-bit state space) encodes the four possible input combinations as four quantum states. The measurement then extracts one bit of information. Since the four states are not all pairwise orthogonal, the system is not storing two independent bits prior to the measurement and can therefore perform better than the reversible classical and Clifford settings.

4.3.2 Connection to Contextuality

We have already argued that the CHSH* game never shows non-locality and contextuality in its standard definitions due to Kochen-Specker [33] and Spekkens [34] defined in subsection 3.1.2. We here report a notion of transformation non-contextuality, recently introduced by Mansfield and Kashefi in [71], where the contexts are sequences of transformations. They called it *sequential transformation non-contextuality* (STNC) and it refers to the fact that the same transformation in different sequences of transformations must have the same ontological representation. More precisely, if we consider a finite sequence $C = (U_i)_{i=1}^{t}$ of unitaries U_i, the ontological representation of the unitary Γ_{U_i}, is the same in any other sequence of unitaries C',

$$\Gamma_{U_i(C)} = \Gamma_{U_i(C')}. \tag{4.5}$$

We are here also assuming that the sequential composition is reflected at the ontological level, *i.e.* $\Gamma_{U_t \ldots U_1} = \Gamma_{U_t} \circ \cdots \circ \Gamma_{U_1}$. Despite being different from the other notions of non-contextuality, STNC still encodes the same counterfactual spirit of them, where pre-existing properties associated to each experimental procedure (here unitary transformations) must not depend on the contexts (here sequences) they belong to.

This notion of contextuality is useful as a resource for computational advantages in some particular computational models, called $l2-$TBQC.[5] $l2$-TBQC is a computational model consisting of a classical control computer that can only perform mod2-linear computation, and

[5] TBQC stands for transformation-based quantum computation, in analogy with MBQC [15].

it can interact with a resource (possibly quantum) to enhance its computational power. The CHSH* protocol is an example of $l2$-TBQC protocol. The natural ontological model to associate to $l2-$TBQC has ontic space $\mathbb{Z}_2^n$ for some $n \in \mathbb{N}$ and transformations that are mod2$-$ linear. Since unitaries must be represented by invertible functions at the ontological level, this implies that their action must correspond to addition of vectors in $\mathbb{Z}_2^n$, *e.g.*

$$\Gamma_U(\lambda) = \lambda + u$$

for some $u \in \mathbb{Z}_2^n$. We call this ontological model as $l2-$*ontological model*.

In [71] Mansfield and Kashefi proved that in a $l2$-TBQC protocol sequential transformation contextuality (with the assumption of $l2-$ontology) is necessary to enable quantum advantage over classical resources for the task of probabilistically computing any non-linear function. Of course, if we drop the assumption of $l2-$ontology the result does not hold, as already an ontological model for classical physics – which is intrinsically sequential transformation non-contextual – can reproduce a protocol that performs non-linear functions (which concerns problems belonging to the complexity class P). The point here is to impose natural restrictions on generic ontological models that reflect the artificial nature of the computation encoded by the protocol (in this case the restriction to $\bmod 2$ linear computation). The result is that either the natural assumption of STNC or the here natural assumption of $l2-$ontology is incompatible with quantum mechanics.

This work is relevant to the CHSH* game, since the result above applies to the CHSH* game too, which therefore shows sequential transformation contextuality. The result above can also be stated in our case as the following no-go theorem: a sequential transformation non-contextual $l2-$ontological model cannot in general reproduce the performance of the CHSH* game in the unitary setting.

Lastly, we report an interesting connection between contextuality and the entropic costs associated to the (partial) erasures considered in the previous subsection.[6] Let us define with *Landauer's Erasure* (LE) the fraction of erasure (in terms of the unit – $kT\log_2 2$ – of bit erased) associated to a given strategy that performs better than the optimal reversible classical

[6]This part refers to some preliminary results developed mainly with Shane Mansfield and still not contained in any published work.

strategy. More formally, let us define, in accordance with [120], the average distance between two boolean functions $f, g : \mathbb{Z}_2^n \to \mathbb{Z}_2$ as $d(f,g) = \frac{1}{2^n}|\{i \in 2^n | f(i) \neq g(i)\}|$, *i.e.* the fraction of the number of inputs for which the two functions differ. We also define the *non-linearity* $\nu(f)$ of a function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ as the distance between the function f and the closest $\mathbb{Z}_2$-linear function $g : \mathbb{Z}_2^n \to \mathbb{Z}_2$,

$$\nu(f) = \min_g \{d(f,g) | g : \mathbb{Z}_2^n \to \mathbb{Z}_2 \text{ is } \mathbb{Z}_2\text{-linear}\}. \tag{4.6}$$

The Landauer's erasure, $LE \in [0,1]$, associated to the probability of success of a strategy for the task of computing a non-linear function f, is defined as

$$LE = \frac{p_{\text{suc}} - p_{\text{suc}}^{rev}}{\nu(f)}, \tag{4.7}$$

where p_{suc}^{rev} denotes the probability of success of the optimal reversible classical strategy. For example, in the CHSH* game, when considering the optimal quantum strategy, the Landauer's erasure is $LE = \frac{\cos^2(\frac{\pi}{8}) - 0.75}{0.25} = 0.41$, which corresponds to the quantity already discussed in the previous subsection. Notice that this value represents an upper bound on the Landauer's erasures for the quantum strategies (unitary setting). Moreover, the definition (4.7) can be rearranged in the relation

$$p_{\text{fail}} \geq (1 - LE)\nu(f), \tag{4.8}$$

considering that $p_{\text{suc}}^{rev} = 1 - \nu(f)$ and that the probability of failure is $p_{\text{fail}} = 1 - p_{\text{suc}}$. The reason for doing this is that the relation (4.8) above is exactly the same relation for the contextual fraction (CF) – a way of quantifying contextuality in the sheaf-theoretic approach [119] – in [120, Theorem 3], $p_{\text{fail}} \geq (1 - CF)\nu(f)$. We leave the study of the relation between these two quantities for future reasearches. However, we suggest that the analogy of these two relations can lead to possible applications. In realistic scenarios involving reversible quantum computation and actual irreversible processes, we can consider a relation involving both the contributions due to the contextualilty and to the erasures: $p_{\text{fail}} \geq (1 - CF - LE)\nu(f)$. This consideration suggests a way of measuring the amount of contextuality (CF) in the computation, as, after n rounds of the experiment, the probability of failure is known and the average entropic

cost (encoded by LE) can be measured. Moreover, it seems that the notions of erasure and contextuality are interchangeable in these scenarios by rebalancing the amount of computation which is quantum and the one which is purely irreversible.

4.4 Generalisation to higher dimensions

We have introduced the CHSH* game as a modification of the CHSH game from two players to one player. It is natural to consider a similar one-player modification of the mod$q$ CHSH$_q$ game. We call such a game the CHSH$_q^*$ game. We leave the full investigation of the CHSH$_q^*$ for future work, but make some preliminary observations here.

An interesting question is whether Lemma 6 can be extended to a correspondence between strategies for the single qudit and CHSH$_q$ games. The current proof of the lemma does not directly generalise to systems of higher dimension since it utilises some special properties of 2x2 unitary matrices.

Nevertheless, we conjecture that the correspondence between the Tsirelson bound for the CHSH$_q$ game and the quantum value for the CHSH$_q^*$ game in the unitary setting holds for arbitrary dimensions. We here provide a support towards the validity of the conjecture, by focusing on the case of $q = 3$. The CHSH$_3^*$ game requires that the player's final measurement outputs $c = a{\cdot}b \pmod 3$, for inputs $a, b, c \in \{0, 1, 2\}$. For a classical trit with reversible gates, the value of the game (coinciding with the known Bell bound [64–67]) is $\omega_C(CHSH_3^*) = 2/3$. This can be found by listing all the possibilities for the different input values. One way to obtain it is to start with the trit in the state 0 and apply the gates $A_0 = A_1 = B_0 = B_2 = \mathbb{I}, A_2 = B_1 = X$.

Suppose now that we have a qutrit system prepared in state

$$T_3 \left|+\right\rangle = T_3 \frac{\left|0\right\rangle + \left|1\right\rangle + \left|2\right\rangle}{\sqrt{3}},$$

where the gate $T_3 = \mathrm{diag}(1, w^{-1/3}, w^{-2/3})$ is the dimension-3 equivalent of the non-Clifford gate T, and $w = \exp(\frac{2\pi i}{3})$. Let us choose the following control gates:

$$A_0 = B_0 = \mathbb{I}, A_1 = B_2 = V, A_2 = B_1 = W,$$

where $V = \text{diag}(1, w, w)$ and $W = \text{diag}(1, 1, w)$. Measuring the system in the X basis gives a success probability $p_{\text{suc}} \approx 0.71$. This strategy is inspired by the one used to obtain the Tsirelson bound for the CHSH_3 game in [65], thereby providing support for the conjecture that there exists a mapping from the CHSH_q^* game to CHSH_q game for $q \geq 2$.

4.5 Conclusion

In this chapter we have introduced the CHSH* game, a single player game inspired by the CHSH game. We have showed that the optimal success probability for the CHSH* game, called the value of the game, depends on many properties of the system available to the players. Defining these properties via settings, we have showed that the value of the game depends on the irreversibility, or otherwise, of the transformations available to the players, the quantum or classical nature of the system and the system dimension.

Furthermore, we have seen that the values obtained are equal to the Bell and Tsirelson bounds in the CHSH game (and the perfect strategies embodied by PR boxes). In particular, for the unitary quantum setting, Lemma 6 shows that any unitary strategy in CHSH* can be mapped to a quantum strategy in the CHSH game. This correspondence gives a new perspective on Tsirelson's bound, which arises due to the absence of irreversible transformations and the limited ability of quantum strategies with unitary gates and projective measurements to simulate erasure.

We have seen that in the more restricted Clifford setting, the value obtained is no better than the reversible classical setting, reflecting the crucial role of non-Clifford computation to obtain better than classical performance in quantum computation. We have shown that, under the assumption of reversible transformations, the CHSH* game acts as a dimensional witness, since any initial state of dimension $d > 2$ can in principle win the game with certainty. However, the restriction to reversible operations is not a limitation. In accordance with Landauer's principle, implementing irreversible transformations at the microscopic level requires ancillary bits which must then be erased. The presence of exactly these hidden ancilliary bits is detected by our protocol.

We have noted a similarity between the optimal unitary strategy for the CHSH* game and

$2 \to 1$ QRACs (and $2-$bits parity oblivious multiplexing). The latter have also been proposed as dimensional witnesses [146]. It is therefore important to emphasise the differences between RACs and the CHSH* game. The CHSH* game is able to detect the hidden information needed to implement irreversible gates. However, irreversible gates provide no advantage for the implementation of RACs. This means that a dimensional witness based on the RAC protocol will be blind to this kind of hidden information. Following Landauer's approach, we have asserted that the ability to detect irreversible dynamics should be an important desideratum for quantum dimensional witnesses. This has not been considered in prior work.

We have conjectured our results to hold also for the generalisation of the protocol to $\mathrm{mod}\, q$ arithmetics. We have supported this by examining the $q = 3$ case in the single system scenario, for which we have shown the validity of the Bell bound and we have further provided a strategy to achieve Tsirelson's bound. The validity of this conjecture may open the way to easier approaches for deriving Tsirelson's bounds in $\mathrm{mod} q$ arithmetics, by using our single-system protocol as a tool for proving tightness.

In light of Landauer's principle, we have further considered the entropic costs of the erasure associated with the CHSH* game. The lack of such an erasure operation in unitary quantum mechanics was a barrier to winning the game deterministically. Via the correspondence with Tsirelson's bound proven in Lemma 6, we have demonstrated a link between the reversibility in fundamental operations embodied by Landauer's principle, and the non-unity value of Tsirelson's bound. This work shows that Tsirelson's bound can be seen as arising from the restricted physics of a unitarily evolving single qubit system.

Finally, we have shown that theorem 1 in [71] applies to the CHSH* game, thus demonstrating that, by assuming only reversible computation, sequential transformation contextuality is necessary for our protocol to achieve a probability of success higher than the Bell bound. Other forms of contextuality have been studied from the single-particle perspective [69], but they do not apply here. Our work shows that assumptions of reversibility in transformations can have a dramatic effect on the capabilities of the system, motivating further study of the relationship between non-classicality and irreversible dynamics, as suggested by some considerations on the connection between the entropic costs associated to the erasures and contextuality.

Chapter 5

Summary and outlook

The heresy of one age becomes the orthodoxy of the next.

H. Keller, *Optimism*

Quantum technologies are already available in the market and universal quantum computers are a dream believed to become true in few decades. Nevertheless, it is still unknown which physical principles are responsible for the quantum computational speed-up. In this thesis we have tried to take a step closer to the understanding of this crucial matter, by studying notions of non-classicality that act as resources for computational advantages. We have mainly focused on contextuality, which emerges as an inherent non-classical feature from studies in quantum foundations, like Spekkens' toy theory. The latter indeed shows that almost all the other phenomena usually associated to quantum mechanics can be reproduced in the phase-space formalism of classical mechanics with a restriction on what can be known about the reality – a sort of uncertainty principle built in the symplectic structure of classical physics.

In chapter 2 we have endowed the toy theory with measurement update rules and we have generalised it to systems of arbitrary finite dimensions, non-prime too. A complete formulation of the model has allowed us to fully study its operational equivalence with subtheories of quantum mechanics. In the case of odd dimensional systems, we have proven that ST and SQM share analogous structural properties and, more importantly, they reproduce the same statistics of outcomes. This is proven by using the tool of Gross' Wigner functions, which non-negatively

represent odd qudit SQM. An elegant manifestation of the operational equivalence of the three theories arises when comparing their measurement update rules, which have been provided for Gross' Wigner functions too. Given the importance of SQM in quantum computation, the hope is that its representation in ST can open the way to the generalization of results based on SQM to the non-prime case. A possible example is the result due to Howard *et al* in [54], demonstrating contextuality – in its original definition due to Kochen-Specker [33] – to be necessary for state-injection schemes of computation on qudits of odd prime dimensions. Moreover, the way we have treated the coarse-graining observables in ST may give suggestions on how to characterise SQM in non-prime dimensions, which still lacks of an unambiguous mathematical formulation [77].

In the important case of qubits, the operational equivalence between ST and SQM does not hold, due to the contextual character of qubit SQM. We therefore dedicate chapter 3 to identifying which subtheories of qubit SQM are operationally equivalent to subtheories of ST. These are the subtheories that can be represented by non-negative and covariant Wigner functions. We have used this definition to group in the same scheme the known results on state-injection schemes of universal quantum computation with contextuality as a resource for odd prime qudits [54] and rebits [55], where the non-contextual free part of the computation is represented by a Spekkens' subtheory, and the contextuality comes in with the magic states. Furthermore, we have proven that the multi-qubit SQM can be obtained from a Spekkens' subtheory by circuits of state-injections made of objects only belonging to the Spekkens' subtheory. This result has also shown us a Spekkens' subtheory for the non-contextual free part of a novel state-injection scheme with CZ and CCZ magic states. In this scheme different manifestations of contextuality can be associated to different state-injections. The property of covariance of the gates composing the free part of the computation is the main difference between our framework and other similar works [57, 83]. This property is necessary because it guarantees that the epistemic restriction is satisfied when the gates are applied. However, a classical simulation protocol based on non-negative Wigner functions for some non-covariant subtheories of qubit SQM exists [57]. This fact questions the role of covariance and its relation to other notions of classicality referring to transformations, like the positivity preservation of the Wigner functions and transformation non-contextuality (recently shown to be violated in one-qubit SQM [122]).

Stepping back from the classical simulations based on Wigner functions – that can always be associated to psi-epistemic ontological models – Gottesman-Knill theorem [84] shows that n-qubit SQM, despite manifesting contextuality [46, 47], can be efficiently simulated by a classical computer. This highlights the need to distill which contextuality and, more in general, which notions of non-classicality in quantum foundations match the notion of non-classicality in quantum computation, *i.e.* non-efficient classical simulatability.

The results above show that contextuality provides a justification of the quantum computational power only in particular schemes of computation and the question of whether similar results hold in other models is still open. In chapter 4 we have focused on a restricted computational scenario that shows quantum advantages and it is free of contextuality (in its standard versions, [33] and [34]) as well as non-locality, the other feature that is usually considered as inherently non-classical and proven to be a resource in several information processing tasks [61, 63, 135–140]. We have considered a single-system protocol subjected to controlled gates and a fixed measurement – the CHSH* game – that computes a non-linear function in arithmetic modulo 2 with different success probabilities depending on the settings considered. In particular, in the classical reversible setting it achieves the Bell bound and in the quantum unitary setting it achieves Tsirelson's bound, via a direct mapping to the popular CHSH game. If restricted to Clifford computation, it cannot perform better than the classical reversible case, while the possibility of using non-Clifford gates provides strategies that overcome the Bell bound, since they allow to exploit the more powerful non-classical storage of quantum systems. Moreover, by allowing irreversible gates, the non-linear function can be obtained with certainty. The crucial role of irreversibilty has suggested an analysis of the performances in terms of Landauer's principle, that associates entropic costs to erasures of information. A new notion of contextuality – sequential transformation contextuality [71] – has been proven to be necessary for the quantum computational speed-up and we have depicted an interesting connection between Landauer's erasures and the contextual fraction, that can possibly trigger further studies. An open question regards the tightness of the Tsirelson bound in the CHSH_q game for $q > 2$. We have conjectured that the mapping with our CHSH_q^* game holds also for $q > 2$ by analysing the $q = 3$ case, where the alleged optimal quantum strategy inherited from the $CHSH_3$ game provides the same upper bound. Preliminary analyses seem to show that

the optimal quantum strategies in the CHSH_q^* game manifest similar patterns in terms of the gates when varying q, thus suggesting a possible direction to study the tightness of Tsirelson's bounds in the CHSH_q game for arbitrary q using the CHSH_q^* game.

Coming back to the original question of where does the quantum computational speed-up originate, the present work suggests that the answer is not to be expected to come from a single feature (*e.g.* a given notion of contextuality), but it depends on the scenario considered. For a more satisfactory answer, a possible solution would be to develop a novel and inclusive notion of non-classicality that manifests itself in different forms depending on the computational scenario. Most, if not all, of the features considered as inherently non-classical in quantum foundations – the ones deriving from no-go theorems [32, 33, 147, 148] – share the characteristic to require fine-tunings, *i.e.* they require that properties that are always valid at the operational level – like locality, non-contextuality and no-retrocausality – do not hold at the ontological level (thus also providing nature with a conspiratorial connotation). It would be desirable to formally define a notion of fine-tuning in order to subsume all the current notions of non-classicality. Hopefully, with this notion we will be able to match the notions of non-classicality in quantum foundations and quantum computation, thus possibly find results for universal quantum computation also concerning the sufficiency (in correspondence of a certain amount of fine-tuning), and not just the necessity. We believe that understanding what is really peculiar about quantum theory and what explains the quantum computational power is crucial to unveil the mysterious nature of quantum reality, build new quantum technologies and solve open problems in theoretical physics, like developing the quantum version of gravity.

Bibliography

[1] D. Schroeder and M. Peskin, *An Introduction To Quantum Field Theory*. Reading, Mass: Westview Press, first edition ed., 1995.

[2] M. Planck *Verh. Deut. Phys. Ges*, vol. 2, pp. 237–245, 1900.

[3] A. Einstein, "Uber einen die erzeugung und verwandlung des lichtes betreenden heuristischen gesichtspunkt," *Annalen der Physik*, vol. 17, pp. 132–148, 1905.

[4] N. Bohr *The London Edinburgh and Dublin Philosophical Magazine and Journal of Science*, vol. 26, pp. 1–25, 1913.

[5] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.

[6] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, pp. 467–488, 1982.

[7] D. Deutsch, "Quantum theory, the church-turing principle and the universal quantum computer," *Proceedings of the Royal Society of London A*, vol. 400, pp. 97–117, 1985.

[8] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM J. Comput.*, vol. 26 (5), pp. 1484–1509, 1997.

[9] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, 2, pp. 120–126, 1978.

[10] D. Deutsch and R. Jozsa, "Rapid solutions of problems by quantum computation," *Proceedings of the Royal Society of London A*, pp. 439–553, 1992.

[11] D. R. Simon, "On the power of quantum computation," *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pp. 116–123, 1996.

[12] L. K. Grover, "A fast quantum mechanical algorithm for database search," *Proc. 28th Annual ACM Symp. Theory of Comp.*, 1996.

[13] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum algorithm for solving linear systems of equations," *Phys. Rev. Lett.*, vol. 103, p. 150502, 2009.

[14] S. Aaronson and L. Chen, "Complexity-Theoretic Foundations of Quantum Supremacy Experiments," *arXiv:1612.05903 [quant-ph]*, 2016.

[15] R. Raussendorf, D. Browne, and H. Briegel, "Measurement-based quantum computation on cluster states," *Phys. Rev. A*, vol. 68, p. 022312, 2003.

[16] A. M. Childs, D. Gosset, and Z. Webb, "Universal Computation by Multiparticle Quantum Walk," *Science*, vol. 339, no. 6121, pp. 791–794, 2013.

[17] E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, T. Magerlein, E. Solomonik, and R. Wisnieff, "Breaking the 49-Qubit Barrier in the Simulation of Quantum Circuits," *arXiv:1710.05867 [quant-ph]*, 2017.

[18] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices," *Nature Physics*, vol. 14, no. 6, pp. 595–600, 2018.

[19] N. Walenta, A. Burg, D. Caselunghe, J. Constantin, N. Gisin, O. Guinnard, R. Houlmann, P. Junod, B. Korzh, N. Kulesza, M Legr, C. W. Lim, T. Lunghi, L. Monat, C. Portmann, M. Soucarros, R. T. Thew, P. Trinkler, G. Trolliet, F. Vannel, and H Zbinden, "A fast and versatile quantum key distribution system with hardware key distillation and wavelength multiplexing," *New J. Phys.*, vol. 16, p. 013047, 2014.

[20] W. Heisenberg *Zeitschrift fr Physik*, vol. 33, pp. 879–893, 1925.

[21] E. Schroedinger *Ann. d. Physik*, vol. 28, pp. 1049–1070, 1926.

[22] P. A. M. Dirac, "The physical interpretation of the quantum dynamics," *Proc. R. Soc. Lond. A*, vol. 113, pp. 621–641, Jan. 1927.

[23] D. Hilbert, J. von Neumann, and L. Nordheim, "Uber die grundlagen der quantenmechanik," *Mathematische Annalen*, vol. 98, no. 1, pp. 1–30, 1928.

[24] J. von Neumann, "Mathematische grundlagen der quantenmechanik," *Springer-Verlag*, 1955.

[25] H. Weyl, "The theory of groups and quantum mechanics," *Dover Publications*, 1950.

[26] A. Cabello, "Interpretations of quantum theory: A map of madness," *What is Quantum Information? Cambridge University Press*, pp. 138–144, 2017.

[27] A. Einstein, "Zur elektrodynamik bewegter krper," *Annalen der Physik*, vol. 17, pp. 891–921, 1905.

[28] R. W. Spekkens, "Evidence for the epistemic view of quantum states: A toy theory," *Phys. Rev. A , Published*, vol. 75, no. 032110, 19 March 2007.

[29] R. W. Spekkens, "Quasi-quantization: classical statistical theories with an epistemic restriction," *Fund. Theor. Phys.*, vol. 181, pp. 83–135, 2016.

[30] E. Schroedinger *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, pp. 555–563, 1935.

[31] C. H. Bennett, G. Brassard, C. Crepeau, R. Josza, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, 1993.

[32] J. S. Bell, "On the problem of hidden variables in quantum mechanics," *Rev. Mod. Phys.*, vol. 38, no. 447452, 1966.

[33] S. Kochen and E. Specker, "The problem of hidden variables in quantum mechanics," *J. Math. Mech.*, vol. 17, pp. 59–87, 1967.

[34] R. W. Spekkens, "Contextuality for preparations, transformations, and unsharp measurements," *Phys. Rev. A*, vol. 71(5), no. 052108, 2005.

[35] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*. Addison Wesley, third edition ed., 2002.

[36] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Phys. Rev.*, vol. 47, pp. 777–780, 1935.

[37] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, second edition ed., 1964.

[38] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, tenth anniversary edition ed., 2000.

[39] V. Guillemin and S. Sternberg, *Symplectic techniques in physics*. Cambridge University Press, 1984.

[40] L. Catani and D. E. Browne, "Spekkens' toy model and its relationship with stabiliser quantum mechanics," *New journal of physics*, vol. 96, no. 5, p. 052112, 2017.

[41] D. Gottesman, "Stabilizer codes and quantum error correction," *PhD thesis, California Institute of Technology*, 1997.

[42] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, 1932.

[43] D.Gross, "Hudson's theorem for finite-dimensional quantum systems," *J. Math. Phys.*, vol. 47, p. 122107, 2006.

[44] W. K. Wootters, "A wigner-function formulation of finite-state quantum mechanics," *Annals of Physics*, vol. 176, pp. 1–21, 1987.

[45] E. F. Galvao, "Discrete wigner functions and quantum computational speedup," *Phys. Rev. A*, vol. 71, no. 042302, 2005.

[46] D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, "Bell's theorem without inequalities," *Am. J. Phys.*, vol. 58, no. 1131, 1990.

[47] N. D. Mermin, "Simple unified form for the major no-hidden-variables theorems," *Phys. Rev. Lett.*, vol. 65, pp. 3373–3376, 1990.

[48] A.Peres, "Incompatible results of quantum measurements," *Phys. Lett. A*, vol. 151, pp. 107–108, 1990.

[49] L. Catani, N. D. Silva, and D. E. Browne, *Spekkens' toy model in quantum computation and contextuality as a resource.* IOS Press Amsterdam, SIF Bologna: accepted for publication in Proceedings of the International School of Physics Enrico Fermi, course 197, foundations of quantum theory, edited by e. m. rasel, w. p. schleich and s. woelk ed., 2017.

[50] L. Catani and D. E. Browne, "State-injection schemes of quantum computation in spekkens' toy theory," *arXiv:1711.08676 [quant-ph]*, 2017.

[51] X. Zhou, D. W. Leung, and I. L. Chuang, "Methodology for quantum logic gate construction," *Phys. Rev. A*, vol. 62, no. 052316, 2000.

[52] F. G. Brando and G. Gour, "Reversible Framework for Quantum Resource Theories," *Phys. Rev. Lett.*, vol. 115, no. 7, p. 070503, 2015.

[53] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal clifford gates and noisy ancillas," *Phys. Rev. A*, vol. 71, no. 022316, 2005.

[54] M. Howard, J. Wallman, V. Veitch, and J. Emerson, "Contextuality supplies the 'magic' for quantum computation," *Nature*, vol. 510, pp. 351–355, 2014.

[55] N. Delfosse, P. A. Guerin, J. Bian, and R. Raussendorf, "Wigner function negativity and contextuality in quantum computation on rebits," *Phys. Rev. X*, vol. 5, no. 021003, 2015.

[56] J. Bermejo-Vega, N. Delfosse, D. Browne, and R. R. C. Okay, "Contextuality as a resource for models of quantum computation with qubits," *Phys. Rev. Lett.*, vol. 119, no. 120505, 2017.

[57] R. Raussendorf, D. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, "Contextuality and wigner function negativity in qubit quantum computation," *Phys. Rev. A*, vol. 95, no. 052334, 2017.

[58] L. Henaut, L. Catani, D. E. Browne, S. Mansfield, and A. Pappa, "Tsirelson's bound and Landauer's principle in a single-system game," *arXiv:1806.05624 [quant-ph]*, 2018.

[59] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, p. 880, 1969.

[60] B. S. Tsirelson, "Quantum generalizations of bell's inequality," *Letters in Mathematical Physics*, vol. 4, p. 93, 1980.

[61] S. Pironio, A. Acn, N. Brunner, N. Gisin, S. Massar, and V. Scarani, "Device-independent quantum key distribution secure against collective attacks," *New J. Phys.*, vol. 11, no. 4, p. 045021, 2009.

[62] R. Colbeck, "Quantum And Relativistic Protocols For Secure Multi-Party Computation," *arXiv:0911.3814 [quant-ph]*, 2009. arXiv: 0911.3814.

[63] J. Barrett, L. Hardy, and A. Kent, "No Signaling and Quantum Key Distribution," *Phys. Rev. Lett.*, vol. 95, no. 1, p. 010503, 2005.

[64] H. Buhrman and S. Massar, "Causality and Tsirelson's bounds," *Phys. Rev. A*, vol. 72, no. 5, p. 052103, 2005.

[65] S.-W. Ji, J. Lee, J. Lim, K. Nagata, and H.-W. Lee, "Multisetting Bell inequality for qudits," *Phys. Rev. A*, vol. 78, no. 5, p. 052103, 2008.

[66] Y.-C. Liang, C.-W. Lim, and D.-L. Deng, "Reexamination of a multisetting Bell inequality for qudits," *Phys. Rev. A*, vol. 80, no. 5, p. 052116, 2009.

[67] M. Bavarian and P. W. Shor, "Information Causality, Szemerdi-Trotter and Algebraic Variants of CHSH," in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, ITCS '15, (New York, NY, USA), pp. 123–132, ACM, 2015.

[68] E. F. Galvao, "Foundations of quantum theory and quantum information applications," *arXiv:quant-ph/0212124*, 2002. arXiv: quant-ph/0212124.

[69] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, "Preparation Contextuality Powers Parity-Oblivious Multiplexing," *Phys. Rev. Lett.*, vol. 102, no. 1, p. 010401, 2009.

[70] R. Landauer, "Irreversibility and Heat Generation in the Computing Process," *IBM Journal of Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.

[71] S. Mansfield and E. Kashefi, "Quantum advantage from sequential transformation contextuality," *arXiv:1801.08150 [quant-ph]*, 2018.

[72] M. F. Pusey, "Stabilizer notation for spekkens' toy theory," *Found Phys*, vol. 42, pp. 688–708, 2012.

[73] B. Coecke, B. Edwards, and R. Spekkens, "Phase groups and the origin of non-locality for qubits," *Electron. Notes Theor. Comput.*, vol. 270, no. 29, 2011.

[74] L. Disilvestro and D. Markham, "Quantum protocols within Spekkens' toy model," *Phys. Rev. A*, vol. 95, no. 5, p. 052324, 2017.

[75] P. Blasiak, "Quantum cube: A toy model of a qubit," *Phys. Lett. A*, vol. 377, pp. 847–850, 2013.

[76] J. Larsson, "A contextual extension of spekkens' toy model," *AIP Conf. Proc.*, vol. 1424, no. 211, 2012.

[77] D. Gottesman, "Stabilizer codes with prime power qudits," *invited talk at Caltech IQIM seminar (Pasadena, California)*, 2014.

[78] R. Raussendorf, "Contextuality in measurement-based quantum computation," *Phys. Rev. A*, vol. 88, no. 2, p. 022322, 2013.

[79] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation," *New J. Phys.*, vol. 16, no. 013009, 2014.

[80] V. Veitch, N. Wiebe, C. Ferrie, and J. Emerson, "Efficient simulation scheme for a class of quantum optics experiments with non-negative Wigner representation," *New J. Phys.*, vol. 15, no. 013037, 2013.

[81] C. Ferrie, R. Morris, and J. Emerson, "Necessity of negativity in quantum theory," *Phys. Rev. A*, vol. 82, no. 044103, 2010.

[82] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, "Negative quasi-probability as a resource for quantum computation," *New Journal of Physics*, vol. 14, no. 113011, 2012.

[83] J. J. Wallman and S. D. Bartlett, "Non-negative subtheories and quasiprobability representations of qubits," *Phys. Rev. A*, vol. 85, no. 062121, 2012.

[84] D. Gottesman, "The heisenberg representation of quantum computers," *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physic*, pp. 32–43, 1999.

[85] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A*, vol. 52, no. 4, pp. R2493–R2496, 1995.

[86] A. Steane, "Multiple-particle interference and quantum error correction," *Proc. R. Soc. Lond. A*, vol. 452, no. 1954, pp. 2551–2577, 1996.

[87] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum Error Correction and Orthogonal Geometry," *Phys. Rev. Lett.*, vol. 78, no. 3, pp. 405–408, 1997.

[88] R. W. Spekkens, "Negativity and contextuality are equivalent notions of nonclassicality," *Phys. Rev. Lett.*, vol. 101, no. 020401, 7 July 2008.

[89] C. Ferrie and J. Emerson, "Framed hilbert space: hanging the quasi-probability pictures of quantum theory," *New Journal of Physics*, vol. 11, no. 063040, p. 33, 2009.

[90] *Wigner Function*. Wiley-Blackwell, 2005.

[91] F. A. Buot, "Method for calculating $\mathrm{Tr}{\mathcal{H}}^{n}$ in solid-state theory," *Phys. Rev. B*, vol. 10, no. 8, pp. 3700–3705, 1974.

[92] J. H. Hannay and M. V. Berry, "Quantization of linear maps on a torus-fresnel diffraction by a periodic grating," *Physica D Nonlinear Phenomena*, vol. 1, pp. 267–290, 1980.

[93] L. Cohen and M. O. Scully, "Joint Wigner distribution for spin-1/2 particles," *Found Phys*, vol. 16, no. 4, pp. 295–310, 1986.

[94] R. P. Feynman, *Negative Probability. Quantum Implications: Essays in Honour of David Bohm*. Methuen, 1987.

[95] D. Galetti and A. F. R. de Toledo Piza, "An extended Weyl-Wigner transformation for special finite spaces," *Physica A: Statistical Mechanics and its Applications*, vol. 149, no. 1, pp. 267–282, 1988.

[96] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, "Discrete phase space based on finite fields," *Phys. Rev. A*, vol. 70, no. 6, p. 062101, 2004.

[97] C. Cormick, E. F. Galvao, D. Gottesman, J. P. Paz, and A. O. Pittenger, "Classicality in discrete wigner functions," *Physical Review Letter*, vol. 73, no. 012301, 2006.

[98] N. Harrigan and R. W. Spekkens, "Einstein, Incompleteness, and the Epistemic View of Quantum States," *Foundations of Physics*, vol. 40, no. 2, pp. 125–157, 2010.

[99] E. Bezout, "Theorie generale des equations algebriques," *Paris, France: Ph.D. Pierres*, 1779.

[100] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," *Phys. Rev. A*, vol. 70, no. 052328, 2004.

[101] P. Shor, "Fault-tolerant quantum computation," *arXiv:9605011*, 1997.

[102] D. Aharonov, "A simple proof that toffoli and hadamard are quantum universal," *arXiv:0301040*, 2003.

[103] Y. Shi, "Both toffoli and controlled-not need little help to do universal quantum computation," *Quant. Inf. Comput.*, vol. 3, no. 84, 2003.

[104] B. Eastin, "Distilling one-qubit magic states into toffoli states," *Phys. Rev. A*, vol. 87, no. 032321, 2013.

[105] C. Jones, "Novel constructions for the fault-tolerant toffoli gate," *Phys. Rev. A*, vol. 87, no. 022328, 2013.

[106] C. Jones, "Composite toffoli gate with two-round error detection," *Phys. Rev. A*, vol. 87, no. 052334, 2013.

[107] A. Paetznick and B. W. Reichardt, "Universal fault-tolerant quantum computation with only transversal gates and error correction," *Phys. Rev. Let.*, vol. 111, no. 090505, 2013.

[108] D. Bacon, "Operator quantum error-correcting subsystems for self-correcting quantum memories," *Phys. Rev. A*, vol. 73, no. 1, p. 012340, 2006.

[109] G. Nebe, E. M. Rains, and N. J. A. Sloane, "The invariants of the Clifford groups," *arXiv:math/0001038*, 2000.

[110] A. Y. Kitaev, "Quantum computations: algorithms and error correction," *Russ. Math. Surv.*, vol. 52, no. 6, 1997.

[111] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation," *Phys. Rev. A*, vol. 86, no. 3, p. 032324, 2012.

[112] D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," *Transactions of the I.R.E. Professional Group on Electronic Computers*, vol. EC-3, no. 3, pp. 6–12, 1954.

[113] A. Peres, *Quantum Theory: Concepts and Methods.* Dordrecht: Springer, new edition edition ed., May 2008.

[114] A. Cabello, J. Estebaranz, and G. Garca-Alcaine, "Bell-Kochen-Specker theorem: A proof with 18 vectors," *Physics Letters A*, vol. 212, no. 4, pp. 183–187, 1996.

[115] A. Cabello, S. Severini, and A. Winter, "Graph-Theoretic Approach to Quantum Correlations," *Phys. Rev. Lett.*, vol. 112, no. 4, p. 040401, 2014.

[116] A. Acn, R. Duan, D. E. Roberson, A. B. Sainz, and A. Winter, "A new property of the Lovsz number and duality relations between graph parameters," *Discrete Applied Mathematics*, vol. 216, pp. 489–501, 2017.

[117] R. Kunjwal and R. W. Spekkens, "From the Kochen-Specker Theorem to Noncontextuality Inequalities without Assuming Determinism," *Phys. Rev. Lett.*, vol. 115, no. 11, p. 110403, 2015.

[118] R. Kunjwal and R. W. Spekkens, "From statistical proofs of the Kochen-Specker theorem to noise-robust noncontextuality inequalities," *Phys. Rev. A*, vol. 97, no. 5, p. 052110, 2018.

[119] S. Abramsky and A. Brandenburger, "The sheaf-theoretic structure of non-locality and contextuality," *Phys. Rev. A*, vol. 13, no. 11, p. 113036, 2011.

[120] S. Abramsky, R. S. Barbosa, and S. Mansfield, "Contextual Fraction as a Measure of Contextuality," *Phys. Rev. Lett.*, vol. 119, no. 5, p. 050504, 2017.

[121] E. G. Beltrametti and S. Bugajski, "A classical extension of quantum mechanics," *J. Phys. A: Math. Gen.*, vol. 28, p. 3329, 1995.

[122] J. W. P. Lillystone and J. Emerson, "Contextuality and the single-qubit stabilizer subtheory," *arXiv:1802.0612*, 2018.

[123] R. W. Spekkens, "Negativity and Contextuality are Equivalent Notions of Nonclassicality," *Phys. Rev. Lett.*, vol. 101, no. 2, p. 020401, 2008.

[124] N. Delfosse, C. Okay, J. Bermejo-Vega, D. E. Browne, and R. Raussendorf, "Equivalence between contextuality and negativity of the Wigner function for qudits," *New J. Phys.*, vol. 19, no. 12, p. 123024, 2017.

[125] M. J. Bremner, R. Jozsa, and D. J. Shepherd, "Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy," *Proc. Royal Soc.*, vol. A, p. 0301, 2010.

[126] C. Okay, S. Roberts, S. D. Bartlett, and R. Raussendorf, "Topological proofs of contextuality in quantum mechanics," *Quant. Inf. and Comp.*, vol. 17, pp. 1135–1166, 2017.

[127] B. van Dam, "Nonlocality & communication complexity," *PhD thesis, University of Oxford, Department of Physics*, 2000.

[128] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, pp. 195–200, 1964.

[129] J. Barrett and N. Gisin, "How Much Measurement Independence Is Needed to Demonstrate Nonlocality?," *Phys. Rev. Lett.*, vol. 106, no. 10, p. 100406, 2011.

[130] M. J. W. Hall, "Relaxed Bell inequalities and Kochen-Specker theorems," *Phys. Rev. A*, vol. 84, no. 2, p. 022102, 2011.

[131] A. Fine, "Hidden Variables, Joint Probability, and the Bell Inequalities," *Phys. Rev. Lett.*, vol. 48, no. 5, pp. 291–295, 1982.

[132] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Phys. Rev. A*, vol. 40, no. 8, pp. 4277–4281, 1989.

[133] L. Masanes, Y.-C. Liang, and A. C. Doherty, "All Bipartite Entangled States Display Some Hidden Nonlocality," *Phys. Rev. Lett.*, vol. 100, no. 9, p. 090403, 2008.

[134] B. Hensen, H. Bernien, A. E. Drau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abelln, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, no. 7575, pp. 682–686, 2015.

[135] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, "Nonlocal correlations as an information-theoretic resource," *Phys. Rev. A*, vol. 71, no. 2, p. 022101, 2005.

[136] J. I. d. Vicente, "On nonlocality as a resource theory and nonlocality measures," *J. Phys. A: Math. Theor.*, vol. 47, no. 42, p. 424017, 2014.

[137] S. Pironio, L. Masanes, A. Leverrier, and A. Acn, "Security of Device-Independent Quantum Key Distribution in the Bounded-Quantum-Storage Model," *Phys. Rev. X*, vol. 3, no. 3, p. 031007, 2013.

[138] L. Masanes, S. Pironio, and A. Acn, "Secure device-independent quantum key distribution with causally independent measurement devices," *Nature Communications*, vol. 2, p. 238, 2011.

[139] S. Pironio, A. Acn, S. Massar, A. B. d. l. Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, "Random numbers certified by Bells theorem," *Nature*, vol. 464, no. 7291, pp. 1021–1024, 2010.

[140] R. Gallego, L. Masanes, G. D. L. Torre, C. Dhara, L. Aolita, and A. Acn, "Full randomness from arbitrarily deterministic events," *Nature Communications*, vol. 4, p. 2654, 2013.

[141] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Found Phys*, vol. 24, no. 3, pp. 379–385, 1994.

[142] S. Wiesner, "Conjugate Coding," *SIGACT News*, vol. 15, no. 1, pp. 78–88, 1983.

[143] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, "Dense Quantum Coding and a Lower Bound for 1-way Quantum Automata," *arXiv:quant-ph/9804043*, Apr. 1998.

[144] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, "Quantum Random Access Codes Using Single d-Level Systems," *Phys. Rev. Lett.*, vol. 114, no. 17, p. 170502, 2015.

[145] C. H. Bennett, "Notes on Landauer's principle, reversible computation, and Maxwell's Demon," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 34, no. 3, pp. 501–510, 2003.

[146] S. Wehner, "Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities," *Phys. Rev. A*, vol. 73, no. 2, p. 022110, 2006.

[147] M. F. Pusey, J. Barrett, and T. Rudolph, "On the reality of the quantum state," *Nature Physics*, vol. 8, no. 6, pp. 475–478, 2012.

[148] M. S. Leifer and M. F. Pusey, "Is a time symmetric interpretation of quantum theory possible without retrocausality?," *Proc. R. Soc. A*, vol. 473, no. 2202, p. 20160607, 2017.