Memory Cost of Quantum Contextuality

Patrik Harrysson
Master of Science Thesis in Physics

Memory Cost of Quantum Contextuality

Patrik Harrysson
LiTH-ISY-EX--16/4967--SE

Supervisor: Jan-Åke Larsson
isy, Linköpings universitet

Examiner: Jan-Åke Larsson
isy, Linköpings universitet

Division of Information Coding
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

Copyright © 2016 Patrik Harrysson
Abstract

This is a study taking an information theoretic approach toward quantum contextuality. The approach is that of using the memory complexity of finite-state machines to quantify quantum contextuality. These machines simulate the outcome behaviour of sequential measurements on systems of quantum bits as predicted by quantum mechanics. Of interest is the question of whether or not classical representations by finite-state machines are able to effectively represent the state-independent contextual outcome behaviour. Here we consider spatial efficiency, rather than temporal efficiency as considered by D. Gottesman\textsuperscript{a}, for the particular measurement dynamics in systems of quantum bits. Extensions of cases found in the adjacent study of Kleinmann et al.\textsuperscript{b} are established by which upper bounds on memory complexity for particular scenarios are found. Furthermore, an optimal machine structure for simulating any $n$-partite system of quantum bits is found, by which a lower bound for the memory complexity is found for each $n \in \mathbb{N}$. Within this finite-state machine approach questions of foundational concerns on quantum mechanics were sought to be addressed. Alas, nothing of novel thought on such concerns is here reported on.


## Contents

1 Introduction and purpose 1

2 Short on quantum mechanics 5
   2.1 An algebra for physics ................................. 6
   2.2 A Hilbert space formulation ............................ 8
      2.2.1 Observables, states and measurements ............ 9
      2.2.2 Commutativity as subspace invariance .......... 11
   2.3 The quantum bit ........................................ 12
   2.4 Quantum contextuality ................................ 14
      2.4.1 Prelude ............................................. 14
      2.4.2 Incompleteness and the Kochen-Specker argument .. 16
      2.4.3 The Peres-Mermin square ......................... 20
   2.5 The Pauli group and n-qubit observables ............ 22

3 Contextual machines 27
   3.1 Finite-state machines ................................ 27
      3.1.1 Definition ......................................... 28
      3.1.2 Usage in contextual scenarios ................... 29
   3.2 Deterministic machines ................................. 30
   3.3 Non-deterministic machines ............................ 33
      3.3.1 Model without realism ............................. 35
      3.3.2 Model with realism ................................ 37

4 Discussion and conclusion 43

Bibliography 49

A Appendicies 53
   A.1 Proof of Lemma 1 ..................................... 53
   A.2 Deterministic 27-state machine ....................... 55
   A.3 Proof of Theorem 1 ................................... 57
Introduction and purpose

Throughout the history of classical physics the usage of probability has had an unproblematic interpretation in that it signifies a compromise when considering physical systems of high complexity. A system may consist of parts which are significant to the overall behaviour but too numerous to keep track of. We do not have a language which captures the dynamics of all its parts mathematically exact while simultaneously retaining a feasible description. The compromise is that we relax the idea of bookkeeping the very detailed information of the system’s many parts in order to gain a description which is feasible. The result is that the state of the system is described by a probability distribution and, as such, probability is interpreted as nothing but our ignorance or uncertainty of its underlying parts. This is the case with classical statistical mechanics, through which we may explain thermodynamics.

The branch of quantum physics is famous in that its mechanics describes our relation to Nature as being in principle probabilistic. As the relation between physicists and probability is unavoidably conditioned on what came before in the history of physics and its philosophical content, it is not uncommon that attitudes of physicists toward quantum probability contain objections to the idea that the quantum mechanical description is the complete story. Not complete as in being the final story on physical reality, but complete in the sense of describing parts, or aspects, that we could unambiguously attribute to Nature as being the natural root causes for observed phenomenology. A complete description in such a sense would then not contain any compromise in terms of ignorance and uncertainty as any subjectiveness on our behalf is not included in the description.

As long as quantum mechanics has been around so too has the question of its completeness. Attempts have been made to supersede quantum mechanics by trying to embed it in some underlying realism; an extension of the quantum mechanical formalism. Such attempts, where quantum mechanics is taken as a
description of ignorance and uncertainty of some sharp underlying realism, have proven not to be unproblematic. Under such a premise of an underlying realism, the obstacle of interest in this thesis relates to the inability to straightforwardly talk about properties of physical systems in a non-contradictory way. What is deemed as “straightforward” is non-contextual reasoning. What non-contextual reasoning refers to here, is the ability to talk about the state of a system’s properties independently of the circumstances surrounding a measurement in which the actual knowledge of a state variable is retrieved. Such a circumstance is aptly called a measurement context and is loosely defined by what inferences an experimenter already have made, or simultaneously is making, by way of measurement. However, it is impossible to take quantum mechanics as relating to an underlying realism which is non-contextual. This is the content of the famous Kochen-Specker theorem [1], stating that realistic extensions of quantum mechanics are necessarily contextual. Hence quantum mechanics does not admit a separation from the act of measurement in the sense of that any realistic extension of it cannot be independent of the subject-object relationship. This is quantum contextuality and is a concept which comes about as a restrictive fact toward presupposing attitudes as to what quantum mechanics is about.

As realist extensions of quantum mechanics are necessarily contextual, there is the question of “how complex are the necessary contextual dynamics?” This thesis centres around this question and when asking about ‘how much of something’ there is the need for a measure. An information theoretic measure of complexity is herein used in order to attempt quantifying quantum contextuality. The approach uses the language of finite-state machines and these machines are to simulate quantum contextual behaviour. The machines use classical resources to produce the desired behaviour and as such their classical nature reflect the realist assumption in realist extensions of quantum mechanics. The measure of complexity is nothing but the memory complexity of these finite-state machines. This memory complexity quantifies the amount of classical resources required to produce some particular quantum contextual behaviour, hence a memory cost of quantum contextuality. The interest and impetus toward this study lies mainly in the investigation of the following:

- how computationally accessible quantum contextual behaviour is on classical machines,
- and whether we may learn something about our ideas of Nature.

In present-day considerations contextuality is seemingly of interest in the field of quantum computation. As have been popularized over recent years, quantum computation is to be a golden egg laid by the marriage of quantum theory, computer science and skilful engineering. The claim is supported by examples of computational tasks where the quantum computation model has been shown to outperform any classical attempt at a solution. Any general consensus as to what particular characteristics give quantum computation its apparent power, or advantage, has not yet been reached. But as speculated by some [2, 3], contextuality is a critical computational resource within one of the more promising models toward fault-tolerant quantum computation.
Even though solutions to computational tasks might originate as formulated within the quantum model of computation, it is not true that all such tasks are outside the reach of an efficient classical solution. An example being [4] in which solutions to tasks thought to be a prime display of the quantum model's power are given efficient classical formulations by means of using classical resources to imitate certain aspects of quantum mechanics. An interesting detail is that an aspect not imitated is contextuality, by which one may wonder if contextuality really is a defining computational resource within the quantum model. That is, defining in the sense that contextuality would offer a separation between classical and quantum models of computation by being efficiently accessible only from within the quantum model. If true, this would imply the fundamental superiority of quantum computation in comparison to any classical model of computation.

Models exhibiting contextuality can be differentiated from non-contextual ones in that they allow a behaviour with respect to measurement outcomes which violates certain inequalities any non-contextual model would abide. This is the case with quantum mechanics which in relatively simple settings produce distinct violations of such inequalities, thus implying its contextual nature [5, 6, 7]. More so, some inequalities are suitable for experimental investigations by which nature has ruled the quantum contextual predictions valid in that non-contextual inequalities were experimentally violated [8, 9]. Still, as contextual behaviour may or may not be dependent upon the actual quantum state, an interest in quantifying contextuality in different quantum scenarios by considering an interesting measure has been given efforts in [10, 11, 12]. Studies [11] and [12] are here especially relevant in that they make use of finite-state machines for the simulation of the contextual character in some simple quantum systems. Their classical machines are seen to saturate the quantum violations of relevant non-contextual inequalities.

The machine-equivalent to a measurement outcome is called an output, and a finite-state machine is said to be a description of some output behaviour if it can produce it. Producing some output behaviour requires a certain complexity associated the machine. Comparing the complexity between quantum machines and classical finite-machines provides an indicator of when any quantum scenarios are computationally approachable by classical means. As classical physics is a special (decoherent) case of quantum physics, the complexity in using classical resources is the same, or greater, than the complexity by using quantum resources. By using these finite-state machines the two theories can be put on an equal footing in that what the machines do is independent of either physical theory. The physics resides in how the machines do it, which places emphasis on the nature of the parts, that is, the information carriers, as they define the physical resources with which any machine operates in order to map input to output. Specifically, constructing a biased scenario by demanding that a set of input-output rules to mimic some quantum behaviour and the resulting complexity of the classical machine shows just how computationally accessible such behaviour is.

In this thesis a resource-view is taken by employing finite-state machines to classically simulate the outcome behaviour of measurements as performed on collections of quantum bits. These simple systems show a contextual behaviour in
measurement outcomes when already consisting of only two quantum bits. A central function of the machines is to act as bookkeepers of measurement history where this bookkeeping takes the form of an internal feedback function. This feedback defines the memory structure within the machines and is a necessary property if any contextual output behaviour is to be present. By this, one may associate a number to the simulated quantum contextual behaviour through the memory cost.

Pertaining to foundational importance is the question “can contextuality help us understand what quantum mechanics is about?” which is a question referring to the second point made above toward the questioning of our ideas and attitudes. An information theoretic measure such as a memory cost associated certain interpretational attitudes may enable ways of applying already established results of physics, some of which previously might have been seen rather disjunct in applicability, such as to restrict the range of what quantum mechanics may “be about.”

**Thesis outline.** There are two major parts to this thesis. The first part concerns the structure of quantum mechanics. Basics are covered in Sections 2.1 to 2.3 which is then followed by more foundational concerns surrounding quantum mechanics and its contextual character in Section 2.4. Section 2.5 marks the end the first part and contains some relevant facts about the Pauli group.

The second part is about the construction of contextual machines having a desired output behaviour. A definition of a finite-state machine is given in Section 3.1, along with why they are suitable for the investigation herein. In Sections 3.2 and 3.3 finite-state machines are given that simulates quantum contextual behaviour as seen in measurement outcomes from sequential measurements on systems of quantum bits.

Lastly there is Section 4 which contains discussions and conclusions gathered throughout the study.
Quantum mechanics is the name for the modern mathematical framework in which our relation to microscopic physical systems are described. Its concepts are not necessities for macroscopic creatures to function in the everyday-world, and perhaps as a consequence of this they might seem a lot more involved than some classical formulation. Thus, any good introduction on quantum mechanics should make an effort toward conveying how the quantum formalism encodes, in its mathematical objects, our experience of Nature. That is also the goal of this section.

The idea of a physical system entails considering some subset of the universe, whereas physics is the body of knowledge containing tools and concepts with which we can form an understanding of what goes on in these subsets. Some concepts might be very intuitive while others might be quite convoluted and hard to get across without resorting to esoteric mathematical statements. The latter describes the character of quantum mechanics.

As of now, physics consist of a patchwork where each patch contains concepts suitable in describing the behaviour in systems on some scale of energy and distance. Across the scales, concepts and/or principles about Nature’s phenomenology may not be wholly compatible in an obvious way, although their presence is called for by experiments in each respective domain. Indeed, significant departures from classical physics occurred by analysing new experimental data in the beginning of the twentieth century. This is when quantum theory and relativity entered the arena. This thesis concerns concepts brought by the former. The departure from classical physics into quantum physics is perhaps, in its origin, mathematically subtle. Still, any character of conceptual profoundness depends on which quantum church its user attends, as there is no consensus on what quantum mechanics is about, other than being a guide for making predictions about the results of measurements, that is.
The reference materials largely used when constructing Sections 2.1 to 2.3 are F. Strocchi [13], W. Heisenberg [14] and Nielsen & Chuang [15]. Other references are scattered about or stated explicitly at the top of a (sub)section.

2.1 An algebra for physics

In describing a physical system there are two central ideas being those of states and observables. The view is traditionally that Nature is (observable properties), about which agents inquire (state description). An observable corresponds to a measurable quantity of a physical system and performing measurements associated a system’s observables will generate information reflecting the state of the system. The mathematical objects associated this language can be seen to bring about different ways of constructing a suitable framework.

One intuitive approach can be to put emphasis on the geometry. In classical Hamiltonian mechanics the state is modelled by a point or distribution in some topological space $\Gamma$. This space is the so-called phase-space manifold where information about position $\vec{q}$ and (linear) momentum $\vec{p}$ is the nature of a point $(\vec{q}, \vec{p}) \in \Gamma$. In this classical instance observables are taken to be any element in either $\vec{q}$ or $\vec{p}$ and also suitable maps taking them as arguments. Hence observables are generally represented as mappings by real continuous functions $f(\vec{q}, \vec{p}) \in \mathbb{F}^{\Gamma}$ taking points in phase-space as arguments. As such, information about the physical system is encoded in the state on which the observables act as decoders, in the sense of unpacking information about a system’s observable quantities.

Even though it would take perfect measurements to discern, the intuitive classical assumption about the physical state of a system is that it is always seen to be mathematically sharp in its properties in the sense of being modelled by a phase-space point. What this means is that if one were to have a distribution over phase-space it only represents an agent’s ignorance about which point the system actually resides in. With the geometric construction of the phase-space in which the state is represented, the specifics of the system can be quite intuitive to read and interpret, which also is a strength of the approach.

The observables follow the logic of an algebra, meaning that there is a certain structure in the mathematical space $\mathbb{F}^\Gamma$. The algebra for these observables is called an abelian C*-algebra. The property of the space being abelian means that its elements (observables) commute, e.g., let $f_1, f_2 \in \mathbb{F}^\Gamma$ then for some point $\gamma \in \Gamma$ the product is the pointwise composition $(f_1 f_2)(\gamma) = f_1(\gamma) f_2(\gamma) = f_2(\gamma) f_1(\gamma) = (f_2 f_1)(\gamma)$. Since the geometry naturally points to observables being represented as real continuous functions, one can argue that the algebraic structure of the observables follows from the geometry of the state space.

This classical set up is to be contrasted with the quantum one, which is here to be given a similar, although partly historical, account.

It is said that the old quantum theory stretched from about 1900 to 1925,
being a body of mathematical constructions which was incomplete in its ability to describe all of the encountered quantum phenomena in a unified way. Here the idea of the quanta, as originated by A. Einstein’s considerations of previous work by M. Planck, was central and did make possible explanations as to why an atom was stable, say, “why does the atom not radiate all its energy at once?” which also connects to the so-called ultraviolet catastrophe, as in “why does the sun not cast off all its energy at once?” Induced currents by light-matter interactions, i.e., the photo-electric effect, was also a phenomena that could be explained. Although this old quantum theory carried the idea of the quanta it was insufficient as to be a general framework and should rather be seen as the first quantum corrections imposed on the classical description.

While working on the, at the time peculiar, spectra of hydrogen, Werner Heisenberg, in collaboration with Max Born and Born’s (other) student Pascual Jordan, held a central position in setting up the mathematical structure of modern quantum theory. Heisenberg recognized that an algebraic structure of observables was not of secondary importance relative to the geometry of the state space. His insight led him to an understanding that the act of performing a measurement on microscopic systems should leave an unavoidable disturbance on their state, in the sense that the representation of physical states can not be, in principle, seen as mathematically sharp; being something which is very different from the classical case. The old quantum theory is said to have ended with the publication of matrix mechanics by Heisenberg and his collaborators in 1925 which incorporates these new ideas. To quote Heisenberg on his discovery [16]:

*It was about three o’ clock at night when the final result of the calculation lay before me. At first I was deeply shaken. I was so excited that I could not think of sleep. So I left the house and awaited the sunrise on the top of a rock. – W. Heisenberg*

**Heuristic argument on measurement disturbance.** The archetypical observables discussed in this context are those of position $q$ and momentum $p$. Consider attempting to prepare the state of a system where the system is some small particle-like object, whereby the quantities of interest to prepare are position and momentum. Realistically, the aim of the preparation procedure is not to make the position $q$ and momentum $p$ mathematically sharp, but rather to specify the mean square deviations $(\Delta q)^2$ and $(\Delta p)^2$. Let the localization of the object be inferred from interactions with light, meaning that photons are sent in to resolve the object’s whereabouts. The resolving power of the photons depends upon their wavelength, where the smaller the wavelength the better the resolution. A good resolution correspond to a small $(\Delta q)^2$. Now, since photons carry momentum they will impart momentum on the particle upon scattering, hence altering $p$ and increasing $(\Delta p)^2$ of the object. The connection here is the relation between the momentum and wavelength of photons: the smaller the wavelength $\lambda$, the larger the momentum $p$. This is expressed in Einstein’s equation $p = h/\lambda$. Hence, an experimenter is unable to make both deviations arbitrarily small at the same
time. The measured response in measuring $q$ and $p$ is then sensitive to their relative ordering. Interestingly, an assumption of classical physics is the possibility of making the mentioned deviations arbitrarily small given an arbitrarily precise measuring apparatus.

The picture outlined above is certainly not counter to intuition. For instance, take a system and consider some intricate assembly of measurement apparatuses whose purpose is to measure a number of the system’s observable quantities. A simple observation is the fact that measurements are interactions with the system. Then, it is not inconceivable that the measured response in any quantity might be influenced by the relative order and timing of measurements of other quantities. The descriptive value of measurement outcomes, pertaining to the state of a system, should certainly be allowed to, at least, diminish while considering some subsequent interactions under the guise of measurement. Such considerations motivate an interest in describing the relations between measurements of observable quantities, thus leading directly to the algebraic structure of observables.

Now, if the relative ordering is to be of importance the property which must be dropped is the abelian property of the C*-algebra, which was implied by Heisenberg’s analysis. This is a step in a more general direction as commutation is a special case within a non-commutative (i.e. non-abelian) algebra. From this non-commutative structure it follows that observables are not to be represented by something as simple as real continuous functions since their structure is too restricting. As such, objects accommodating a more general algebraic structure is needed. And, consequentially, what is a suitable mathematical environment for the state to live in?

This approach is clearly one which places the algebra of observables in the spotlight, while the geometry associated the idea of a state is somewhat secondary in the sense of being reconstructed from the representation of observables. This non-commutative character of observables is what separates the classical from the quantum and is somewhat of a hallmark of quantum mechanics.

### 2.2 A Hilbert space formulation

The non-commutative algebra of observables is suitably given as an algebra of operators on a Hilbert space. Naturally, the ideas of observables and states are still present, although now the mathematical objects to which they correspond are different. This is because the subject-object relationship is in need of a revamp as was emphasised in the previous section.

The correspondence is such that a state is an element in an abstract Hilbert space $\mathcal{H}$ on which operators associated observables map the space onto itself, i.e., $A$ being an observable’s operator acts as $A : \mathcal{H} \to \mathcal{H}$. This formulation is due to J. von Neumann in his 1932 publication [17] which established a rigorous mathematical framework for quantum mechanics. In order to give some insight as to how this formulation captures physics one must dwell on the mathematics and is what this subsection is about. As the theory of matrices and vector spaces
2.2 A Hilbert space formulation

(linear algebra) is believed familiar to most, a representation in terms of such language seems appropriate. Only finite-dimensional Hilbert spaces are considered in order to gain simplicity and all indexes appearing are enumerable in \( \mathbb{N} \). The observables and their associated operators are taken as to have a one-to-one relation, whereby the wordings may be used interchangeably. Almost exclusively, the usage of the word ‘observable’ is taken to refer to its operator.

2.2.1 Observables, states and measurements

Observables are subject to spectral theory wherein their spectra, i.e., collection of eigenvalues, denotes the possible values observable in experiment. In this instance the observables are not just any matrix, but are taken to be Hermitian matrices which guarantees that the spectra is real, which is a highly sensible property if they are to make sense of outcomes. All spectra are here taken to be discrete and will be discrete throughout the thesis.

Any observable \( A \) can be given in the form of its spectral decomposition, written as

\[
A = \sum_j \lambda_j(A) P_j ,
\]

where \( \{\lambda_j(A)\} \) is the collection of eigenvalues associated the eigenspaces \( \{E_j(A)\} \) onto which the projectors \( \{P_j\} \) project. The eigenspaces span the complex vector space on which \( A \) acts in the sense of \( \mathbb{C}^n = E_1(A) \oplus \cdots \oplus E_m(A) \) where \( m \leq n \). Equivalently, there is a unique eigenbasis of vectors \( \{|a_k\rangle\} \) spanning the vector space. As an observable relates to a degree-of-freedom of a physical system, the vector space it generates by its eigenbasis represents the space in which its corresponding state is expressed. As such, an observable’s associated space \( \mathbb{C}^n \) is called its state space of dimension \( n \).

Quantum states are represented as unit vectors \( |\psi\rangle \) in state spaces associated the observables, and as such they may be expressed as any linear combination of some arbitrary basis \( \{|e_j\rangle\} \),

\[
|\psi\rangle = \sum_j c_j |e_j\rangle = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{C}^n .
\]

The notation used is due to Dirac, giving a compact way of writing.

That a state is represented as a vector of unit length derives from that its magnitude is suitably seen as the sum of a probability mass. This association is due to Born’s rule, being a postulate which, among other things, tells us how we ought to derive applicable information from the quantum state. Let the state \( |\psi\rangle \) be expressed in the eigenbasis of \( A \), that is

\[
|\psi\rangle = \sum_j c_j |a_j\rangle ,
\]

then Born’s rule states that the mapping of the complex coefficients as \( c_k \to |c_k|^2 \) gives the probability of finding the outcome \( \lambda_k(A) \), associated the basis state \( |a_k\rangle \),
upon measurement. This mapping is by way of the inner-product, which is a linear functional \( F_\psi : \mathbb{C}^n \rightarrow \mathbb{C} \) defined by any \( |\psi\rangle \in \mathbb{C}^n \) such that

\[
F_\psi(|\phi\rangle) = \langle \psi | \phi \rangle = [c_1^* \ldots c_n^*] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_i \sum_j c_i^* b_j \in \mathbb{C} 
\]

(2.4)

for any \( |\phi\rangle \in \mathbb{C}^n \). The space of linear functionals on \( \mathbb{C}^n \) is called its dual space, denoted by \( (\mathbb{C}^n)^* \), where the relationship between elements of these spaces is one-to-one. Then, the unity in any quantum state \( |\psi\rangle \) is expressed as

\[
\langle \psi | \psi \rangle = \sum_i \sum_j c_i^* c_j \langle a_i | a_j \rangle = \sum_i |c_i|^2 = 1 ,
\]

(2.5)

which is also called the normalization condition. With this condition we always have probability weights over the basis states modelling mutually exclusive measurement outcomes.

Discernible events by measurement are defined by the eigenvalues \( \{ \lambda_j \} \) in the sense that, if a measurement is performed and a value \( \lambda_k \) is observed, we update our quantum state to be an element of the eigenspace \( E_k \subset \mathbb{C}^n \). Being cautious in talking about the act of measurement, we can only claim as much that we describe how our knowledge changes upon new information from our interaction with the physical system. That is, a prior state \( |\psi\rangle \) is conditioned upon the measurement result by which we update it to a posterior state as

\[
|\psi\rangle \xrightarrow{\text{"measurement"}} \frac{P_k |\psi\rangle}{\sqrt{p_k}},
\]

(2.6)

where \( P_k \) is the projector onto the subspace identified with the eigenvalue \( \lambda_k \). This subspace has the probability weight \( p_k \) and is defined as

\[
\text{Probability of outcome } \lambda_k \equiv p_k = \langle \psi | P_k |\psi\rangle .
\]

(2.7)

Note that a projector is not a measurement, as it only aids in describing the aftermath of a projective measurement. The act of applying a projector on the state has a perfectly deterministic result. Thus, having an initial state, the theory tells us how it changes as conditioned on subsequent information gained by measurement.

Talking about measurements in the context of an observable \( A \) and an arbitrary state \( |\psi\rangle \), we may write an expectation value for the associated measurement as

\[
\langle A \rangle _\psi = \langle \psi | A |\psi\rangle = \langle \psi | \left( \sum_j \lambda_j(A) P_j(A) \right) |\psi\rangle = \sum_j \lambda_j(A) \langle a_j | P_j(A) |\psi\rangle = \sum_j \lambda_j(A) p_j(A) .
\]

(2.8)
2.2 A Hilbert space formulation

What we have, then, is that the quantum state contains information of the probable response as per interaction with a system. This information is stored in what is often called amplitudes, which is the set \( \{ c_j \} \) of complex coefficients.

Summarizing, the way quantum states captures physics is by encoding weights of probability \( |c_j|^2 \) associated the outcomes \( \{ \lambda_j(A) \} \) on the frame supplied by the relevant observable \( A \)'s eigenspaces \( \{ E_j(A) \} \).

2.2.2 Commutativity as subspace invariance

Recall that the impetus described above for altering the algebraic structure was to capture how observables relate to each other under measurement. Measurements associated an observable “takes place” in the eigenbasis of the observable, i.e., the probabilistic predictions are made with respect to its eigenbasis. Such a basis might at first seem sufficiently distinct among different observables given that there is an entire continuum available, but in fact two (or more) observables may have a shared eigenbasis. A shared eigenbasis forms a non-trivial measurement context where measurements results of the associated observables can simultaneously be represented sharply. Observables which share eigenbases are called compatible and if not they are said to be incompatible. Among incompatible observables, a state which is identical an element of one of the eigenbases will in the contexts of all others be seen distributed, or say, spread out.

The mathematical statement concerning simultaneous measurability (compatibility) between properties takes form in commutation relations among the associated observables. If observables \( A \) and \( B \) commute, i.e., \( AB = BA \), they are said to be compatible, and if they do not commute they are incompatible. The concept of subspace invariance of an operator is helpful in emphasising this mathematical structure. Simply put, an operator \( X : V \rightarrow V \) being invariant on a subspace \( W \subseteq V \) is one which preserves \( W \) in the sense that the restriction \( X : W \rightarrow W \) is true. Then, take observables \( A \) and \( B \) and suppose they commute. These observables act on some state space \( \mathbb{C}^n \) which, as a consequence of the spectral theorem, can be decomposed as a direct sum of eigenspaces. Let \( E_j(A) \) be the eigenspace spanned by \( A \)'s eigenvectors having eigenvalue \( \lambda_k(A) \). Decompose the state space as \( \mathbb{C}^n = E_1(A) \oplus \ldots \oplus E_m(A) \), where \( m \leq n \), and take some state \( |\psi\rangle \in E_k(A) \). Then, as the observables commute,

\[
A\left( B|\psi\rangle \right) = B\left( A|\psi\rangle \right) = B\left( \lambda_k(A)|\psi\rangle \right) = \lambda_k(A)\left( B|\psi\rangle \right),
\]

(2.9)

showing that \( B|\psi\rangle \) must also be an element of \( E_k(A) \). Seemingly it does not matter whether \( B \) operates on \( |\psi\rangle \) or not, the image is always within the eigenspace. As the eigenspace \( E_k(A) \) was chosen arbitrarily, it follows that \( B : E_k(A) \rightarrow E_k(A) \) for \( 1 \leq k \leq m \). Hence \( B \) shows subspace invariance for all of \( A \)'s eigenspaces. A common basis can be constructed by finding a restricted eigenbasis for \( B \) within each \( E_j(A) \) and then patching all such restricted bases together via direct product in order to span \( \mathbb{C}^n \). The same argument could of course be done by considering \( A \) on \( B \)'s eigenspaces. As such, commuting observables respect each others eigenspaces in the sense of subspace invariance. On the other hand, if \( A \) and \( B \)
do not commute, there is no common eigenbasis for the observables. Then, as the associated measurements are to be seen in each respective eigenbasis, we cannot simultaneously have sharp values about non-commuting observables at the same time.

As a continuation on what was said in Section 2.1 to be the scope of Heisenberg’s heuristic argument on how some observable properties are disturbed when measuring others, an answer by the quantum formalism comes in the form of inequalities. Take again observables $P$ and $Q$, representing momentum and position operators respectively. These operators (or matrices) do not commute, hence being incompatible to some degree. Let $\Delta P$ and $\Delta Q$ be the standard deviations of each observable property, defined as $\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$ for some observable $A$, quantum mechanics holds that

$$\Delta P \Delta Q \geq \frac{\hbar}{2},$$

showing an explicit trade-off in sharpness between the two. The $\hbar$ is the so-called reduced Planck’s constant, which is on the order $10^{-34}$ Js.

Recalling that an object treated classically is one having, in principle, sharp values for all observable quantities regardless of an agent’s interaction with it. This means that for all states described by a distribution over the classical phase-space there is a true underlying state which is a single point. Hence distributed classical states model ignorance in the sense that our description is only due to a lack of knowledge about the true state, whereby such a distribution could be reduced to the one true state by increased precision in measurements. In contrast, states as described by quantum mechanics inherently lack this feature of preciseness, in that its indeterminacy is not something which could be eliminated given measurement apparatuses of arbitrary precision.

2.3 The quantum bit

Quantum mechanics gives us a model for ascribing fleeting degrees-of-certainty to propositions concerning measurement outcomes from interactions with quantum systems. With the projectors associated any quantum observable we can, given some quantum state, learn about the probabilities, i.e., degrees-of-certainty about whether a proposition is true (or false) upon an actual test of the proposition; a test being nothing but the performance of a suitable measurement.

The simplest quantum systems to make use of in settling propositions are those which only have two mutually exclusive outcome events as distinguished by measurement. From an informational perspective, observing such an event reveals one bit of information. The title of this section—the quantum bit—is what such a simple system is called in the context of quantum information, where it occupies the same role as the classical bit does in classical information theory. For a single qubit (brief for quantum bit), propositions settled by measurements upon it relate to the two-dimensional eigenbases given by a set of observables called
the Pauli observables. In matrix form the observables are written

\[
X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.
\] (2.11)

That \(Z\) is the only one in its diagonal form is because these observables do not commute, followed by that viewing the state space from \(Z\)’s eigenbasis \(\{|0\rangle, |1\rangle\}\) is the conventional choice and is called the computational basis. In this basis, the eigenvectors of each Pauli observable are

\[
Z : \quad \{|0\rangle, |1\rangle\}
\]

\[
Y : \quad \{+i \equiv \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), -i \equiv \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\}
\] (2.12)

\[
X : \quad \{+ \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), - \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}
\]

where each pair constitute an orthonormal basis spanning the two-dimensional state space of a qubit. The Pauli observables are called dichotomic in that their discrete spectra consist of eigenvalues \(\{\pm 1\}\).

As the importance of algebraic relations between quantum observables was emphasised in the previous section, then, for the three above it is such that they are not only mutually non-commuting, but in fact are mutually anti-commuting. An anti-commutator of two observables \(A\) and \(B\) is written \(\{A, B\} = AB + BA\) and if the anti-commutator is zero \(A\) and \(B\) are said to anti-commute.

Among quantum observables, the significance in the anti-commutator complements what is told by the commutator in the sense that, all the while commuting observables are compatible, anti-commuting observables are maximally incompatible. What is meant by ‘maximally’ is often expressed as a certain unbiasedness between the eigenbases of anti-commuting observables. A case can be illustrated by considering the eigenbases of, say, \(Z\) and \(X\) as it holds that

\[
|\langle 0|\pm \rangle|^2 = \frac{1}{2}, \quad |\langle 1|\pm \rangle|^2 = \frac{1}{2}.
\] (2.13)

What this shows is that the vectors \(\{|0\rangle, |1\rangle\}\) are both evenly distributed on the eigenbasis of \(X\). This means is that if the observable property \(Z\) is sharply manifest by some preparation procedure (i.e., the state is represented by either of its eigenvectors), the state of the quantum system pertaining to the observable property of \(X\) is completely random as each element of \(X\)’s eigenbasis have equal probabilistic weights. This property holds among all three bases of the Pauli observables by which any pair of them are said to be mutually unbiased. A more general statement of mutually unbiased bases can be made by considering some observable \(A\), with eigenbasis \(\{|a_j\rangle\}\), which anti-commutes with an observable \(B\) having an eigenbasis \(\{|b_j\rangle\}\), and if it follows that

\[
|\langle a_i|b_j \rangle|^2 = \frac{1}{d} \quad \forall i, j \in \{1, \ldots, d\}
\] (2.14)
where $d$ is the dimension of the state space in question, then these bases are said to be mutually unbiased.

Generally, the quantum state of a qubit is written

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$  \hspace{1cm} (2.15)

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. In the literature, a neat graphical representation of a qubit is often shown and the same will be done here. With the coefficients being complex numbers, rewriting them in polar-representation is suitable. First, see that

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = r_\alpha e^{i\phi_\alpha}|0\rangle + r_\beta e^{i\phi_\beta}|1\rangle = e^{i\phi_\alpha} \left( r_\alpha|0\rangle + e^{i(\phi_\beta-\phi_\alpha)} r_\beta|1\rangle \right).$$  \hspace{1cm} (2.16)

Next, rewriting the above under the following: (i) let $\phi_\beta - \phi_\alpha \equiv \phi$, (ii) $|\alpha|^2 + |\beta|^2 = 1$ translates into $r_\alpha^2 + r_\beta^2 = 1$, as such let $r_\alpha \equiv \cos(\frac{\theta}{2})$ and $r_\beta \equiv \sin(\frac{\theta}{2})$, and (iii) the global phase $e^{i\phi_\alpha}$ holds no observable effect and can hence be omitted. The result is

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$  \hspace{1cm} (2.17)

where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$. This representation is called the Bloch-sphere and is depicted in figure 2.1.

### 2.4 Quantum contextuality

This section contains the concept at heart in this thesis. As mentioned in the introduction: quantum contextuality comes about as a concept which articulates a restrictive fact toward presupposing attitudes as to what quantum mechanics is about. The source of such attitudes and the predicament that is the understanding of quantum mechanics is briefly considered in Section 2.4.1 and is somewhat of a continuation to what was conveyed in the introduction. The famous Kochen-Specker theorem, which makes explicit the above mentioned restriction, is considered in Section 2.4.2 along with some historical details. This is followed up by the Peres-Mermin square in Section 2.4.3, being a simple construction as to give insight into the proof structure of the Kochen-Specker theorem.

#### 2.4.1 Prelude

Our direct experiences of Nature on the macroscopic scale consist of a succession of impressions as perceived by our minds. We may, say, open our eyes to see, or place our hands to touch, for such impressions to arise as the cause of what we deem to be the world outside of ourselves; the complement of our minds. Nature has an apparent divisibility in the sense that impressions are caused by distinct physical objects which exist in such an external world. The experience of objects partake in the macroscopic phenomenology about which we communicate in what may be called a conventional language. With such language we can
express propositions with reference to things external to mind, thereby using a rational basis of argumentation in the sense that truth-values of such propositions are mind independent. This mind independence refers to what is called the objective reality, manifesting itself as Nature. What the word reality, or realism, in the context of physics refers to is the idea of objects in Nature having, at all times, an independent physical reality in the sense of being the way they are, whether or not interacted with. That is, a reality with intrinsic physical properties existing, albeit independent, of an external world; an autonomous existence.

It is satisfying that we have the classical theories which capture these intuitive aspects of our macroscopic experience. Microscopic phenomena, we have found, is not as accommodating. The conventional language is found inadequate in the microscopic regime as the physics of objects require, say, both particle and wave aspects as complementary concepts in order to attempt explaining observation. Our empirical distillate in describing microscopic phenomena, i.e., quantum mechanics, can’t seem to capture how things are in the sense of realism. This is because we have this irreversible process called ‘measurement’ denoting the procedure that transforms a scenario of possibilities into a single factuality, thus also being a procedure to which the formalism does not provide a description as to any root cause of such a single factual outcome. This is commonly referred to as the ‘measurement problem’ in that we have no consensus as to why this is or
what it means. Although being in conflict, there exist several claims of an understand- ing to this problem.

An understanding is expressed by the relation between the knowing subject and body of interest. For instance, a physicist’s relation to a subset of natural phenomena is expressed by Newton’s laws, being a set of principles with which one may, to some extent, understand Nature. Another example is Einstein’s principles of relativity. These theories of Nature have been endowed a bottom-up structure in that their mathematical formalisms are seen to express the logical consequences of the underlying physical ideas. It is in this light that quantum mechanics is seen problematic; quantum mechanics is not expressing any physical ideas. Its abstract mathematical axioms reside on a heuristic level, emphasising no explicit depth of understanding as would be articulated by a set of principles. In this sense quantum mechanics is still a top-down model of Nature. Although admittedly, quantum mechanics is a very successful tool toward which the attitude “shut up and calculate” may be healthy if the formalism is to be used in a practical sense. But as the quantum description is not a unifying framework with respect to Nature’s phenomenology, an understanding of what it is about would surely prove useful for further developments in physics. What does it mean to say that the world is quantum mechanical?

The top-down character of quantum physics is troublesome in that it leaves room for guesswork about the nature of its probabilistic statements. In this room physicists have been—and still are—busy attempting to endow the quantum formalism with meaning by interpreting ‘quantum probability’ [18]. For a bystander the situation may seem similar to “the God of the gaps,” in that there is this pocket of scientific ignorance in which space is given to invoke one’s own personal belief as to what quantum physics is about. The central mathematical object is the quantum state and how such a state of affairs is related to reality is what interpretative accounts aim to settle, thus also resolving the measurement problem. It is in these lines of thought, i.e., interpretations of quantum probability, that the concept of quantum contextuality arise.

2.4.2 Incompleteness and the Kochen-Specker argument

The common attitude is that the mind independent world is what physical theories should, on some level, be describing. Such a world is often believed to be rather rigid in the sense of dealing in absolutes at its core, not unbecoming to our sense of realism. Hence the history of quantum mechanics contains objections to it being a complete description, that is, complete in the sense of describing the mind independent objective features of physical systems. “Do quantum probabilities exist in Nature?” If taking the probabilities of quantum mechanics as a description of ignorance and uncertainty about an underlying objective reality then, no, they do not, whereby one thinks of the quantum description as being incomplete. Such a stance naturally suggests itself from our experience with the relation between classical mechanics and statistical mechanics. A complete description of quantum phenomena is then sought to be constructed by underpinning the mathematical elements of the quantum formalism with relations to
2.4 Quantum contextuality

some elements of reality. Such an extension anchors the quantum description to an objective basis which rational agents must, at all times, be coherent with. Thus, quantum probabilities are taken to not exist in Nature, although they are thought to pertain toward Nature’s objective features. Hence realism is restored.

The formalism of a realistic theory should allow one to predict, with certainty, measurement results of unperformed experiments, i.e., counter-factual measurements. According to the 1935 paper of Einstein, Podolsky and Rosen (EPR) [19], such physical quantities should correspond to what they call ‘elements of physical reality’ in their reality criterion:

*If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity. – EPR*

This trio of physicists deemed quantum theory incomplete in the sense that its formalism do not include simultaneous reality for incompatible observables. The gist of their argument was that correlations between two arbitrarily separated quantum systems, as described by the quantum formalism, allows for predicting measurement results for one of them with certainty, by performing measurements only on the other. A crucial point is that the state of the unmeasured system could be in either of two *incompatible* eigenstates depending on how the other system is measured. Then, by their reality criterion, both observable properties of the unmeasured system must correspond to elements of reality, which the quantum formalism does not capture, i.e., uncertainty inequalities as per Heisenberg between non-commuting (incompatible) observables. Hence they conclude under their notion of realism that the quantum description is incomplete.

The EPR approach toward reality in physical theories relates to what is called *counter-factual definiteness* (CFD), where results of measurements performed are factual and results of unperformed measurements are counter-factual, i.e., “counter to the facts.” In assuming CFD one is allowed to speak meaningfully about the definiteness of counter-factual measurement results, where answers to “what-if” questions hold viable descriptive qualities. This means that the observables associated a physical object by the theory are assumed to denote sharp definite values whether or not the corresponding property is measured, being values with which we may argue meaningfully.

The EPR argument is an early case for what Mermin described as *the dream of hidden variables* [20]. Hidden variables (HV) models are attempts to regain control in the sense that these variables give a model with enough resolution on Nature such that the probabilistic aspect of quantum theory is lifted. A hidden variable is synonymous to an element of reality in the EPR sense and reflects a means to talk about physics in a conventional language. Still, if considering hidden variables models attractive as viable extensions, any such model must of course be compatible with the predictions of quantum mechanics (wherever quantum mechanics is correct, that is).

In the 1960’s came significant papers by J. Bell, S. Kochen and E. Specker
addressing the very question of how compatible the idea of hidden variables is with quantum mechanics. Bell put out two papers on this matter, one which addressed the EPR argument [21] and the other considering the wrongness of an analysis von Neumann had done years earlier [22]. The essence of Bell’s analysis was proving that something called local realism was not compatible with quantum mechanics. Again, realism refers to the idea of hidden variables themselves, meaning that objects in nature all have sharp properties which predetermine the results of measurements. Locality refers to that of physical influences, in that physical systems are only influenced by their immediate environment. Influences between systems must be mediated by a field having disturbances propagating with finite speed as limited by the speed of light. Hence arbitrarily separated systems cannot influence each other instantaneously if locality is to hold. In a nutshell, Bell’s result proves that if a hidden variables model is local it cannot agree with quantum mechanics, and if it does agree with quantum mechanics it is necessarily non-local. As such, either locality has to be thrown out the window, or the idea of hidden variables is to be thrown out, whereby the latter effectively eliminates the locality issue.

Kochen and Specker made no explicit reference to locality in their paper [1] as the case was made with the broader notion of non-contextuality. Thus also, incidentally, aligning with some critique Bell had aimed at the assumptions in his own previous analysis. Realistic models, i.e., hidden variable extensions of quantum mechanics, brings about a sharpness in a quantum system’s properties in that they are modelled with definite values. Non-contextuality refers here to the assumption that the possession of quantum properties, as made definite in all observables by the hidden variables extension’s realism, is independent of the circumstances of their subjective actualization through measurement. Hence, in communicating about the definite values describing a quantum system, one does not need supply any context as to how any such values are known in order to—in a logically consistent way—convey the quantum physical state of affairs in its entirety. The “how”-part refers explicitly to an interacting agent being part of the description. Thus a non-contextual hidden variable description of quantum phenomena is without any reference to an inquiring subject, something which is certainly not unbecoming for a realistic description as one may see no reason to include details surrounding the act of measurement when talking about pre-existing descriptive values.

However, Kochen and Specker showed that non-contextual hidden variable models are in conflict with the predictions of quantum mechanics. Their result is the Kochen-Specker theorem and can be cast into a rather simple statement with few technicalities as will be shown below.

Again, the set of possible events associated a projective measurement is expressed by an orthonormal basis in a Hilbert space. The one-dimensional subspace generated by any such basis element is often referred to as a ray and may be represented by the projector mapping all vectors to it. Each ray then corresponds to an event about which we can associate a proposition which may be thought of as written ‘observable A has value a’ and the orthogonality among rays refers to the mutual exclusiveness of their associated events. As quantum mechanics fa-
mously do not deal in absolutes, the truth-value of any such proposition is given a mere probability of being true upon measurement. This probability is given by projecting the quantum state onto a ray. Now, if considering features of realism added through a non-contextual hidden variable extension, one does not need to discuss the act of measurement in relation to such propositions. Contrary to the quantum description, one does deal in absolutes as the possession of physical properties is made definite by realism, whereby only one proposition among the rays in an orthonormal basis may be assigned the truth-value ‘true’.

Kochen-Specker: In Hilbert spaces of dimension $\geq 3$ it is impossible to associate the truth-value ‘true’ to exactly one ray within each orthonormal basis.

Kochen and Specker proved this in three-dimensional Hilbert space using a finite set of orthonormal bases where a total of 117 rays over the set of bases were associated either ‘true’ or ‘false’ by which a logical contradiction could be established; the contradiction being that of reaching a single base where either no proposition was, or more than one proposition were, necessarily true.

As an orthonormal basis of rays is just a joint eigenbasis of compatible quantum observables, associating a ray with ‘true’ makes the state of those observables definite. Non-contextuality makes the particular value in any single observable persist over all joint eigenbases in which the observable partakes. This means that the value itself is decoupled from any such basis in the sense that its assignment is to be independent of which other observables are simultaneously considered. Here the “other” observables in any such joint eigenbasis comprise what is called the measurement context.

By invoking the measurement context as an additional degree of freedom for the assignment of definite values, i.e., dropping the assumption of non-contextuality, the logical contradiction as per Kochen-Specker may be avoided. This contextual retreat articulates a central implication of the Kochen-Specker theorem: non-contextual hidden variable extensions of quantum mechanics are impossible. What follows is that in order to speak in a logically consistent way about the state of a quantum system—now definite in all properties by realism—one must provide the (measurement) context in which any particular values of observables are registered. These circumstances describe how an observable was measured in the sense of what context its value was retrieved in or is considered, meaning whether or not any other compatible measurements were made just prior, or simultaneously, to the measurement in question. Hence these kinds of realistic extensions cannot stand on their ‘objective own’, whereas instead the description is necessarily voiced through the measurement context, that is, the how of an agents interaction with a quantum system of interest.
2.4.3 The Peres-Mermin square

As mentioned, the original proof by Kochen and Specker involved no less than 117 rays in a three-dimensional Hilbert space. The sheer number of rays makes that proof a rather tedious exposition whereas instead a structurally similar, although less complicated, scenario of Kochen-Specker type is here considered. The construction is referred to as the Peres-Mermin square [20, 23], employing only 24 rays over six different measurement contexts in a four-dimensional Hilbert space. A measurement context will further be referred to only as a context, with the suitable definition of being a maximal set of compatible quantum observables. Thus when referring to the context of any single observable, one is talking about which set of compatible observables it is considered as being part of. Furthermore, an observable may not be indifferent to what values the other observables in one of its contexts have been associated with as per measurement. This is because for a single context \( \{U_1, U_2, U_3, \ldots\} \) where each element shown represents an observable, any functional identity \( f(U_1, U_2, U_3, \ldots) = \hat{0} \) translates into a corresponding polynomial relation \( f(v(U_1), v(U_2), v(U_3), \ldots) = 0 \) between each observable’s possible values as given by their spectra. The symbol \( \hat{0} \) represents the zero matrix. Here a single value for an observable \( U_j \) is denoted \( v(U_j) \). Then, if a context contains \( n \) observables, with such a relation one may infer with a probability of unity what any observable will put out under an ideal measurement if the state values of the other \( n - 1 \) observables are known.

To set the stage, consider figure 2.2 where there are nine non-trivial observables constructed from the Pauli observables for which the measurement results are taken to be in \( \{\pm 1\} \).

The physical scenario is that of two qubits and the experiments readily available are the corresponding ones for the given observables. The structure here is that each row and each column form a single context. Any two of these contexts overlap in a single observable or, conversely, each observable partakes in two contexts. The product of the associated operators within each context form an identity, an example being the second column \((IX)(YI)(YX) = +I\). Writing this equality as

\[
f(IX, YI, YX) - I = \hat{0}
\]

and having the commuting matrices simultaneously diagonalized

\[
D_{IX}D_{XI}D_{YX} - I = \begin{bmatrix}
\ddots & v_k(IX)v_k(YI)v_k(YX) - 1 \\
v_k(IX)v_k(YI)v_k(YX) - 1 & \ddots 
\end{bmatrix} = \hat{0},
\]

results in a set of equations having the form \( v(IX)v(YI)v(YX) - 1 = 0 \). Hence two outcomes of these observables will unambiguously specify the third. Up to the sign of the identity, this applies to the other five contexts as well.

Adopting a realist-view for the scenario wherein granting each observable a value as pre-existing independently of measurement is where the assumption of non-contextuality is made. Hence each entry in the square is assigned a value
2.4 Quantum contextuality

Figure 2.2: A Peres-Mermin construction with two-qubit Pauli observables which may be used in a Kochen-Specker proof of four dimensions. The notation is such that, e.g., $X \otimes Z \equiv XZ$, and so on. The observables within each row and column all mutually commute and as such there are six measurement contexts. Each context associates four rays denoting states spanning the Hilbert space, by which there is a total of 24 rays present. The elements in the left square denotes the actual Pauli operators. Operator products for each row and column is shown to the left and below the square. All contexts have a product resulting in positive identities with the exception of the third column’s negative identity. The rightmost square contain values associated the Pauli observables. The assignment of values is done non-contextually and the expected product of values are shown to the right of, and below, the square.

\[
\begin{array}{ccc}
+I &=& \begin{array}{ccc}
XI & IX & XX \\
II & YI & YY \\
XY & YX & ZZ \\
\end{array} \\
\to NC & & \begin{array}{ccc}
A & B & C \\
a & b & c \\
\alpha & \beta & \gamma \\
\end{array} = +1 \\
& & = +1 = +1 = -1
\end{array}
\]

An observation is then that the product of all rows and columns should equal to minus one because of the odd number of negative identities. But, using the notation as shown in figure 2.2, the equality in the product

\[
(ABC) \times (abc) \times (\alpha \beta \gamma) \times (Aaa) \times (Bb\beta) \times (Cc\gamma) = -1
\]

(2.20)
can never be satisfied for any assignment of values. Combining the contexts in an additive fashion wherein each term is a measurable quantity due to compatibility, an inequality may be formed as

\[
\langle ABC \rangle + \langle abc \rangle + \langle \alpha \beta \gamma \rangle + \langle Aaa \rangle + \langle Bb\beta \rangle - \langle Cc\gamma \rangle \leq 4 \quad (\text{NC})
\]

= 6 \quad (\text{QM})

(2.21)

where the brackets in the sum denote the product of expectation values for each trio of measurements. This inequality is an example of a non-contextual inequality of the likes mentioned in the introduction. The bound denoted NC is the so-called non-contextual bound, and the bound denoted QM is what quantum mechanics prescribes. The existence of such an inequality is proof that any non-contextual description of this scenario is insufficient. Obviously, the problem with any non-contextual realist assignment in the Peres-Mermin square owes to the identities and overlap of contexts, i.e., the algebraic structure in the selected set of observables.
2.5 The Pauli group and n-qubit observables

The main reference material for this section is M. Waegell [24], which is a good read if one is interested in the structure of the Pauli group. Also, an accessible introduction on groups can be found in the appendices of Nielsen & Chuang [15].

Propositions about an assembly of $n$ qubits is the subset of quantum theory which is central to this thesis. The observables associated $n$ qubits are the $n$-fold tensor products of Pauli operators and these are contained in what is called the Pauli group $P_n$. As have been previously emphasised, relations among quantum observables hold physical significance and as such the purpose of this section is to become somewhat familiar with the physics of qubit observables by way of their group structure. Importantly, what is considered will be of use in subsequent sections.

When referring to a group in the mathematical sense one talks about a collection of objects $G$ which is equipped with a group-operation $\xi$ combining two elements to form a third. The operation is such that it satisfies closure ($\xi : G \times G \to G$, i.e., the image is always within the group), associativity, and the group contains an identity element $I$ such that $\xi(g,I) = g$ for all $g \in G$. Also, for all $g \in G$ inverses exist. As elements of the Pauli group are represented as matrices, the group-operation is nothing but the ordinary matrix multiplication. An example could be the simplest level of the group

$$P_1 \equiv \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}, \tag{2.22}$$

where then, e.g., $\xi(X, -iY) = -iXY = -i(iZ) = Z$. Notably, this group does not only contain elements eligible as quantum observables because complex coefficients render some of them non-Hermitian. The reason why all these coefficients $\{\pm 1, \pm i\}$, and the identity, are included is in order for the collection to be a proper group. As such, from the view of quantum theory, it is not the whole of a group $P_n$ which is interesting, but rather some of its substructures. These substructures are foremost separated among on the basis of commutation, by which maximal sets of mutually commuting observables are considered. Such a maximal set is nothing but what have previously been called a context. Each of these contexts relate to certain subgroups within $P_n$ and the ones of interest are the maximal real Abelian subgroups. Abelian is just another word for that all elements are mutually commuting and being real refers to that no complex coefficients are attached to any element. An important fact about being a proper group is that the identity is included. This inclusion means that by combining group elements under the group-operation (matrix multiplication) one may produce the identity, and as we have seen such relations lead to a dependence between eigenvalues of observables (cf. Eq.(19)).

When considering an arbitrary large collection of $n$ qubits the observables are, for brevity, not written explicitly with the tensor product. Say, $X \otimes Y \otimes X$ is instead written as $XYX$, and so on, and when writing a stand-alone $I$ what is really referred to is $I_1 \otimes I_2 \otimes \ldots \otimes I_n$ depending on the current choice of $n$. Then, in order
to expand upon the above, subgroups of $P_2$ make a suitable illustration because they allow for just enough complexity. A select few of them are

$$\{I, XI, IX, XX\}, \{I, XI, -IX, -XX\},$$
$$\{I, -XI, IX, -XX\}, \{I, -XI, -IX, XX\},$$

with an obvious commonality in that they are all constructed from the non-trivial observables $\{XI, IX, XX\}$ having suitable coefficients. In fact, each of these groups define a state which is a joint eigenstate for all their elements (observables) since the associated measurements are all compatible, as seen by that the state is invariant under each element (i.e., having eigenvalue one). For example, let a state $|\psi\rangle$ be defined by $\{I, XI, -IX, -XX\}$ through

$$I|\psi\rangle = |\psi\rangle, \quad XI|\psi\rangle = |\psi\rangle,$$
$$-IX|\psi\rangle = |\psi\rangle, \quad -XX|\psi\rangle = |\psi\rangle,$$

from which it is clear that $|\psi\rangle$ is an eigenstate of the non-trivial collection $\{XI, IX, XX\}$ with eigenvalues $\{+1, -1, -1\}$ ordered respectively. Similarly do the remaining three groups in Eq.(2.23) define other, mutually exclusive, states.

Another commonality, perhaps subtle but important, is that none of these groups include the negative identity, which is a necessary condition for them to define non-trivial states. Indeed, if $-I$ were to be included it would nullify what other elements within the group would bring about toward defining a state because

$$-I|\psi\rangle = |\psi\rangle \implies |\psi\rangle = 0 .$$

As such, the structure of the groups in Eq.(2.23)—mutually commuting observables with real coefficients and no negative identity—is the only one admitted from a physical point of view as it define viable states. These four groups can be seen abbreviated by the de facto context $\{XI, IX, XX\}$ in that the joint eigenstates of the context, and those defined by the groups of Eq.(2.23), are the same.

Again considering a collection of $n$ qubits, the number of distinct $n$-fold tensor products of the Pauli observables (including the identity) is $4^n$ because of the four ways of selecting a single tensor product argument. Then, with $P_n$ growing exponentially in size as $n$ increases, the number of physically admissible subgroups, and the contexts to which they are associated, is most certain to explode by the combinational possibilities. In fact, they do, and certain facts about these structures for arbitrary $n$ are here very useful in order to study contextuality in qubit systems.

**Lemma 1.** For the Pauli group $P_n$ with $n \in \mathbb{N}_+$, (i) let $C$ denote the number of contexts, then

$$C(n) = \prod_{k=0}^{n-1} \left(2^{n-k} + 1\right),$$
where each context contains $2^n - 1$ non-trivial Pauli observables and is associated $2^n$ maximal quantum states. (ii) For $n = 2$, let $A$ and $B$ denote two contexts with eigenbases $\{a_j\}$ and $\{b_j\}$ respectively, where $j \in \{1, 2, 3, 4\}$. Also, let $W_{B\perp A}(k)$ be the set of indices for which $\{b_j\}$ is non-orthogonal to some $|a_k\rangle$. Then, for any $k$,

$$|\langle a_k | b_j \rangle|^2 = \frac{1}{|W_{B\perp A}(k)|}$$

and

$$\sum_j |\langle a_k | b_j \rangle|^2 = 1,$$

for all $j \in W_{B\perp A}(k)$.

**Proof:** See Appendix A.1.

The second point made in Lemma 1 means that, in any context, the associated quantum states are unbiased over the non-orthogonal remainder of any other context’s eigenbasis.

Considering the bipartite case of $n = 2$ it follows by Lemma 1 that there are 15 contexts. Explicitly, these are

\[
\begin{align*}
\{XI,IX,XX\}_+ & \quad \{XI,IY,XY\}_+ & \quad \{XI,IZ,XZ\}_+ \\
\{YI,IX,YX\}_+ & \quad \{YI,IY,YY\}_+ & \quad \{YI,IZ,YZ\}_+ \\
\{ZI,IX,ZX\}_+ & \quad \{ZI,IY,ZY\}_+ & \quad \{ZI,IZ,ZZ\}_+ \\
\{XX,YZ,ZY\}_+ & \quad \{XZ,YY,ZX\}_+ & \quad \{YX,XY,ZZ\}_+ \\
\{XX,YY,ZZ\}_- & \quad \{XY,YZ,ZX\}_- & \quad \{XZ,YX,ZY\}_-
\end{align*}
\]

where each subscript denotes the sign of the identity which results from the product of all elements. These observables can be arranged in an array as shown in figure 2.3.

The array is similar to the PM square which, in fact, contain it as a substructure. This larger collection of contexts also constitute a non-classical structure in the sense that its observables supply logical contradictions under the assumption of non-contextuality. The separation due to the assumption of non-contextuality over the observables shown in figure 2.3 may be expressed by the non-contextual
2.5 The Pauli group and n-qubit observables

\[
\begin{array}{cccc}
IZ & XZ & YZ & ZZ \\
IY & XY & YY & ZY \\
IX & XX & YX & ZX \\
& XI & YI & ZI \\
\end{array}
\]

**Figure 2.3:** The array of quantum observables for a two-qubit system; showing a compact representation of the quantum observables. This square contain no less then 15 contexts, whereby there are many which overlap.

inequality

\[
\langle (XI)(IX)(XX) \rangle + \langle (YI)(IY)(YY) \rangle + \langle (ZI)(IZ)(ZZ) \rangle + \\
+ \langle (XI)(IY)(XY) \rangle + \langle (XI)(IZ)(XZ) \rangle + \langle (YI)(IX)(YX) \rangle + \\
+ \langle (YI)(IZ)(YZ) \rangle + \langle (ZI)(IX)(ZX) \rangle + \langle (ZI)(IY)(ZY) \rangle + \\
+ \langle (XX)(YZ)(ZY) \rangle + \langle (XY)(YX)(ZZ) \rangle + \langle (XZ)(YY)(ZX) \rangle - \\
- \langle (XX)(YY)(ZZ) \rangle - \langle (XY)(YX)(ZX) \rangle - \langle (XZ)(YX)(ZY) \rangle \leq 9 \quad \text{(NC)},
\]  

(2.26)

where each term in the sum corresponds to one of the 15 contexts. The brackets denote the average value of the product of outcomes which can be obtained either by measuring simultaneously or sequentially. Note that the terms in the sum with a positive coefficient are those for which the product of all three observables result in +I, and those with a negative coefficient result in −I. As such, the bound predicted by quantum mechanics is exactly 15. Contrasting the PM square and its difference in upper bounds between the NC and QM cases being 2, this larger set of observables holds a difference of 6.

One might suspect that as the number of qubits \(n\) increase the difference between the NC and QM bound may increase when considering what other non-classical structures \(P_n\) may contain. How this separation grows is in fact an unsolved problem, as finding the NC bound for arbitrary \(n\) is quite a difficult optimization-problem.
This marks the second part of the thesis in which abstract machines defining input-output processes are considered. These machines simulate quantum contextual behaviour, whereby each then earns the name ‘contextual machine’. The following subsections go through the descriptions of such machines having different premises in their simulation models. Whether or not these machines give any insight on quantum contextuality from an information theoretic viewpoint is investigated.

### 3.1 Finite-state machines

That a physical system produce a response when excited by an interacting agent is a natural situation of an input being mapped to an output; a computation. Knowing the responsive behaviour of physical systems allow us to use them—as far as our control of them go—in building ever more intricate responsive systems to our liking. Using suitable systems a physical circuit may be realised which comprise a collection of ordered logical operations, allowing for a useful input-output process. Such a thing is called a computer and its response to excitation is programmable or, at least, is manufactured to specification.

When analysing the capabilities and limitations of a computer one is often not interested in a computer’s specific physical realization, whereas the main components with which it operates are then preferably given rather abstract as elements in a set. Such a set defines a model of computation which is also called a *machine*.

Nevertheless, computation by any machine is still something that process information by manipulation of the information carriers, whereby the capabilities and limitations of machines is ultimately tied to the physical nature of its information carriers. That is, how the carriers themselves interact and may relate to each other. The classical information carrier—the bit—is the carrier in a useful
type of machine called a finite-state machine. Such a machine defines a most basic loop and may be realised by a logical circuit with access to memory. The usefulness of memory is that the machine can store intermediate values for computational tasks, amounting to the ability of generating an output which is conditional on previous values.

### 3.1.1 Definition

Reference materials used in the construction of this section were J. Savage [25] and Bengtsson & Danielsson [26].

For the purposes here, a deterministic finite-state machine (FSM) $M$ is defined as a 5-tuple $M = (X, Y, S, \delta, \lambda)$, where $X = \{x_1, \ldots, x_p\}$ is the input alphabet, i.e., ways of exciting the machine, and $Y = \{y_1, \ldots, y_q\}$ is the output alphabet, i.e., ways the machine can respond to excitation. The set $S = \{s_1, \ldots, s_n\}$ is the finite set of states. The machine is always initialized in some state of $S$ and at any given time the machine is said to occupy a single state $s_i \in S$ which, along with being given an input $x_j \in X$, the machine uses for the mapping of an output $y_k \in Y$. This is made explicit by the response function $\lambda : X \times S \rightarrow Y$ for which at each time step $t$ it holds that

$$y_k(t) = \lambda(x_j(t), s_i(t)).$$

By this, any state the machine currently occupies functions as a container which captures all past events relevant on which the determination of an output to a given input is conditioned. This is the function of the machine's memory, enabling the response to any excitation to be situational through the occupation of a state. The details lodged in a state is only limited by how much memory the machine has access to. The state space dynamics is governed by the next-state function $\delta : X \times S \rightarrow S$ which evaluates at each time step, that is

$$s_i(t + 1) = \delta(x_j(t), s_i(t)),$$

whereby possibly changing the state of the machine. This memory structure, serving the act of remembrance, is realised through a feedback-loop which couples to any subsequent input under $\lambda$ and $\delta$. A machine such as this is said to operate with a sequential logic.

A more general consideration is that of a non-deterministic finite-state machine (NFSM), which refers to a machine not evolving deterministically in its state space. Rather than being a next-state function, $\delta$ is a next-set function acting as $\delta : X \times S \rightarrow 2^S$, where $2^S$ is the power set of $S$, i.e., the set of all subsets of $S$. Note that $\delta$ itself is deterministic in that it deterministically selects a subset of states. The non-determinism comes about in the selection of a successor state within the subset picked out by $\delta$, a selection which is based on some stochastic source defining a suitable probability distribution.

A nice way to think about a NFSM is as a deterministic machine but that there is an additional hidden, or uncontrollable, input supplied by what may be called
3.1 Finite-state machines

A machine is said to be contextual if its output behaviour violates any non-contextual inequality associated its input alphabet. Such behaviour entails that there are correlations in output behaviour between symbols of the input alphabet; correlations being such that, a single listing of each input symbol being associated an output symbol fulfilling the correlations does not exist. The non-existence of such a list satisfying correlations associated the input alphabet is the defining character of contextuality and implies that the machine necessarily attains different states over the course of sequential inputs, as otherwise the correlations in output would fail to be represented. Hence the machine’s response to input is situational in the sense that ‘a situational instance’ must define the relevant interaction-history between user and machine, upon which any subsequent response is conditional. As such, a contextual machine needs the ability to act as a bookkeeper of a user’s interaction-history, which is nothing but the usage of memory as represented by states.

The simulation of quantum systems with contextual responsive behaviour us-
ing finite-state machines will require some number of machine states where, if the number of states used is $n$, the amount of memory $b$ used is $b(n) = \log_2(n)$ bits. Hence a bit-address of size $\lceil b \rceil$ is what is put through the feedback-loop in order to point out a single current state.

The idea with the approach of using finite-state machines is to investigate the memory needed for simulation of the contextual response behaviour in certain quantum scenarios. A “quantum scenario” is just short for a set of quantum observables associated a physical system. Then, such scenarios are sought to be associated a number which in turn is related to the memory used by the performing machine.

Although, there may be more than one finite-state machine producing a certain output behaviour, whereby the interesting machine is one using the least number of states for a given scenario. Such a machine is said to be minimal (or optimal) in that it uses a minimal amount of memory for producing the desired response behaviour to sequential input. Finding representations of machines minimal in simulating certain contextual quantum scenarios is thought of as an important step toward talking about any memory cost associated quantum contextuality.

### 3.2 Deterministic machines

This section centres around the approach and content of Kleinmann et al. [11]; a study wherein an information theoretic measure of quantum contextuality is proposed. The measure there defined is associated the state space memory usage by finite-state machines simulating a contextual quantum scenario. Again, the measure of memory $b$ is simply $b(n) = \log_2(n)$, where $n$ is the number of states used by the machine. As this measure of memory applies to any finite-state machine, the meaningfulness of it toward contextuality is made clear from the premises of the simulation model. Here, it is such that each state is realistic and that the state space dynamics are deterministic. What is meant here by realism pertains to the simulation model, meaning that each state contains a pre-determined output to all input, whereas the determinism refers to that each next state is selected deterministically. In fact, the determinism implies that the simulation cannot accommodate the statistical character of quantum mechanics, i.e., there is no “quantum uncertainty” in the simulation model. Under these settings, any non-contextual scenario would require only a single state, thus $b(n) = 0$ for $n = 1$, whereas any contextual scenario would require strictly more than one state, thus $b(n) > 0$ for $n > 1$. Hence the deviation of $b(n)$ from zero is the number associated the contextual correlations among a set of observables, a number then seen to quantify the memory complexity required of a finite-state machine in order to display the particular contextual behaviour under the model’s premises.

The particular scenario considered is that of the Peres-Mermin square, which is here already introduced in Section 2.4; the reader might want to review. As such, the physical picture is that of measurements on a bipartite system of qubits.
Adopting the common shorthand notation, the Peres-Mermin square is written as in figure 3.2,

\[
\begin{array}{ccc}
A & B & C \\
a & b & c \\
\alpha & \beta & \gamma \\
\end{array}
\]

\[
AB \alpha = + \\
abc = + \\
\alpha \beta \gamma = + \\
\]

\[
ABC = + \\
Aaa = + \\
Bb\beta = + \\
Cc\gamma = - \\
\]

**Figure 3.2:** The Peres-Mermin square and the correlations among its elements. For brevity the 1’s are omitted, whereby only the signs ± are appearing.

where the correlations among output values are such that the entries in each row and column have a product of +1, with the exception of column three having a product of −1. Then, letting \( R_i \) and \( C_j \) denote any permutation of the elements of row \( i \) and column \( j \) of the Peres-Mermin square respectively, the correlations require the machine to display a response as

\[
\lambda'(R_i) = +, \quad \forall i \in \{1, 2, 3\} \\
\lambda'(C_1) = + \\
\lambda'(C_2) = + \\
\lambda'(C_3) = -
\]

(3.3)

where \( \lambda' \) denotes a map taking the input sequence in argument into a product of the corresponding output sequence. As this scenario is contextual, the non-existence of a single list of assignments fulfilling the correlations carries over to a non-existence of a single state doing the same. Hence the responsive behaviour experienced by a user of the machine necessarily derives from dynamics over multiple states. As such, for a simulation of quantum contextual responsive behaviour under sequential input, Kleinmann et al. propose a machine using only four states. These states are given in figure 3.3 along with the associated update-tables showing any next state. They note the fact that this four-state machine does not accommodate all input sequences in the sense of the associated output values not fulfilling all certain quantum predictions. In other words, this means that a user may obtain an output value which do not adhere to the correlations prescribed by quantum mechanics. Such an instance is referred to as a contradiction and in the case of the Peres-Mermin square there is a minimum of one contradictory row or column under any assignment of values. With such contradictions being possible in their machine there is a possibility that the simulation would not reach the quantum violation of the Peres-Mermin square’s associated non-contextual inequality.

Finding a contradiction is straightforward by using the fact that the correlations among input symbols in any of the contexts (i.e. any row or column) pro-
Figure 3.3: The four states of Kleinmann et al. Each state contains a single contradiction, whereby a change of state is needed in order to accommodate the correlations as described by quantum mechanics. States $s_1$ and $s_4$ have their contradictions in column three, whereas states $s_2$ and $s_3$ have contradictions in row three. An encirclement denotes an instance of deterministic update by $\delta$, which points to a next state as given by the corresponding element in the associated update-table $u_k$.

vides equations, each with three unknowns, whereby knowing two will specify the third. E.g., say the machine is in state $s_1$ (cf. figure 5) and the user feeds it the input $(A, B)$, by which the response is $(A = +, B = +)$. As $ABC = +$ it then follows that an input of $C$ must with certainty result in $C = +$; this much can be asserted while in $s_1$ for a machine fulfilling the contextual correlations. Further, with the machine still occupying $s_1$, the user may input $c$ whereby the response $c = +$ is followed by a subsequent change of state to $s_3$. The input $C$ is compatible with all elements in row one and column three, reflecting the fact that quantum mechanics would have the value observed in the observable corresponding to $C$ reproducible as long as only measurements of observables corresponding to elements in row one or column three are performed. As such, the user may now in $s_3$ reach a contradiction through output $C = -$ when it is expected that $C = +$ as it should be reproducible from earlier inferences. The problem here is that with such implicit inferences allowed by the correlations, any value in either state of their four-state machine is potentially known by the user. By this it is impossible for the machine to guard—by way of the assigned update-instances—a user against reaching a contradiction, as none of the values intended to change is given room to do so.

The remedy to this problem is having a sufficient set of values being impossible for the user to know in each state; which presumably also was the intent with the assignment of update instances resulting in the four-state machine. Thus in order to restrict a user’s possibility to know about a state there has to be more instances of update in each state. This can be thought of as the need for each state to have a “buffer-of-ignorance” in which values may change under state space dynamics. In fact, Kleinmann et al. mentions in one of their appendices a 10-state machine which fulfils the correlations for all sequences of input. It is however not made explicit in the sense of being presented in the same form as their four-state machine, nor is any claim of minimality associated with it; rather, it sets an
3.3 Non-deterministic machines

Using non-deterministic machines enables a construction showing the statistical character of quantum mechanics. Such machines are explored in this section.

Before viewing any explicit machine it is instructive to consider the physical scenario which they aim to simulate: a system of \(n\) qubits subjected to sequential measurements for which the outcomes are predicted by quantum mechanics. As stated in previous sections, these systems of qubits have correlations such that the upper bound of the memory needed in a machine accommodating all sequences of excitations.

Considering the premises as per Kleinmann et al., the search of a minimal machine led here to a 10-state machine as well, being similar in structure to what Kleinmann et al. proposed. This machine is here given an explicit form in figure 3.4.

Furthermore, by the same token as the 10-state machine was constructed, a machine accommodating the full set of bipartite correlations was found constructable. This contextual scenario is referred to as the ‘extended Peres-Mermin square’ in Kleinmann et al. and is shown in Section 2.5 to contain no less than 15 non-trivial observables, a set in which there are 15 contexts. The machine constructed for this scenario make use of 27 states whose explicit form can be found in Appendix A.2. No claim of minimality is here made for either machine, i.e., the 10-state one or the 27-state one, as any formal proof is yet to be completed. As such, what is said here about this kind of simulation model, being one of realism and determinism, is that the contextual correlations in the full bipartite scenario is upper bounded by \(\log_2(27) \approx 4.75\) bits of memory.

### Figure 3.4: The 10-states machine accommodating the contextual quantum response of all input sequences.

\[
\begin{array}{cccccc}
\text{s}_1: & + & + & + & + & + \\
\text{s}_2: & + & + & + & + & + \\
\text{s}_3: & + & + & + & + & + \\
\text{s}_4: & + & + & + & + & + \\
\text{s}_5: & + & + & + & + & + \\
\text{s}_6: & + & + & + & + & + \\
\text{s}_7: & + & + & + & + & + \\
\text{s}_8: & + & + & + & + & + \\
\text{s}_9: & + & + & + & + & + \\
\text{s}_{10}: & + & + & + & + & + \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{u}_1: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_2: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_3: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_4: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_5: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_6: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_7: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_8: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_9: & \_ & \_ & \_ & \_ & \_ \\
\text{u}_{10}: & \_ & \_ & \_ & \_ & \_ \\
\end{array}
\]
a non-contextual realist description is impossible at already \( n = 2 \) qubits. Of course, this depends on if one considers a rich enough set of measurements to perform, i.e., a structurally rich enough set of observables. Such scenarios are the ones considered throughout.

Consider two machines as in figure 3.5, where everything about them is identical except the nature of each respective state space. One machine draws its output from a system of \( n \) qubits being kept internally, while the other one is a proper finite-state machine using a number \( b \) of bits. The idea of simulation is to have these two machines indistinguishable to any user having access to both. As such, with the quantum machine being the target of simulation, the size of the classical state space is seen related to the number \( n \) of qubits, that is, \( b = b(n) \).

The quantum machine shows reproducibility in outcomes of previously per-

![Figure 3.5: Two machines with differing nature in their state spaces. Machine (i) contains a quantum system of \( n \) qubits as its state space, whereas machine (ii) is a classical one whose state space consists of \( b(n) \) bits, i.e., classical pointers. A user of any machine can give it an input by pushing any button on the front, here designated by the letters of the Peres-Mermin square. A registered input is shown in the input-field \( I \) along with an associated output in the output-field \( O \).](image)

formed measurements given that any subsequent measurements performed were compatible. Again, this notion of compatibility refers only to that the observables associated a measurement commutes. E.g., consider measurements over two different sets of mutually commuting observables \( \{A, B, C\}, \{A, D, E\} \) where \( A \) is the only common observable; let \( A \) be measured (in any of these contexts) thus associating it with an outcome, then the very same outcome is expected for \( A \) in all subsequent measurement sequences over both sets. That is, \( \{B, C, D, E\} \) are all compatible with \( A \), meaning that the outcome associated \( A \) is never “disturbed” by the subsequent measurements, hence remaining a descriptive value throughout all compatible measurement sequences.

The statistical character refers to the frequency of particular outcomes among sequences of incompatible measurements. As expressed by Lemma 1 the situa-
tion is one of simplicity; joint eigenstates of contexts are, if not orthogonal, unbiased over the joint eigenstates of any other context. This means that whenever a subsequent incompatible measurement is to be performed there is always an even chance of obtaining either +1 or −1 as outcome. Hence the probability assignment over outcomes among incompatible measurements is that of a simple coin toss.

3.3.1 Model without realism

By dropping the premise of realism in the states, that is, that any state is associated pre-existing outputs for all inputs, the freedom gained in machine construction makes way for a simple machine which is straightforwardly generalizable. As such, the case of $n = 2$ qubits is here made explicit, which is then followed by a theorem concerning its optimality in memory usage for any $n$.

Again, the specification of a machine entails specifying its parts, hence elements of a 5-tuple $(X, Y, S, \delta, \lambda)$ are to be made precise. The input alphabet $X$ is to reflect the full set of bipartite observables and is given as the array of 15 symbols in figure 3.6. As the observables are dichotomic the output alphabet is set as the usual $Y \equiv \{+,−\}$. With the demands from quantum mechanics, the states comprising the state space $S$ are suitably constructed as shown in figure 3.7 and is later supported by Theorem 1.

$$X \equiv \begin{array}{cccc}
IZ & XZ & YZ & ZZ \\
IZ & XY & YY & ZY \\
IX & XX & YX & ZX \\
XI & YI & ZI \\
\end{array}$$

*Figure 3.6: The input alphabet of the bipartite machine where each entry symbolises the tensor product of two Pauli observables.*

As this is a non-deterministic machine $\delta$ is a next-set function which if not performing the trivial update (i.e. returning the same state) will, in fact, always point to a subset of $S$ containing two states. This is because of a symmetry in the observables associated the input alphabet; any selection of an observable not in some context $Q$ will be part of some other context $W$ having exactly one observable in common with $Q$. This is seen directly by inspection of figure 3.7. Due to reproducibility, this translates directly into that there is always some informational baggage necessarily carried into a next state, here in the form of a single output symbol. As this single output symbol is necessarily matched in the subsequent state, there is a unique next state associated the outcome of the current input being (+) and another unique state associated the outcome being (−). This is because any third symbol in a trio $(i, j, k)$ is specified through a relation $ijk = \pm$, hence fully determining a state. An example is supplied in figure 3.8.
Figure 3.7: Compact view of the state space for the case of $n = 2$ qubits. The 15 arrays on display show the 15 maximal commuting sets, or measurement contexts, here adapted to the machine. The states in $S$ are specifications of these, with each trio of $(i, j, k)$ coming in four configurations only, as a consequence of the correlations present in the structure of the Pauli group. For simplicity, the products are written as $ijk = \pm$ and are shown below the collection of arrays. The shaded boxes denotes which outputs the state does not have stored, whereas an output is instead acquired via the choice agent upon query.

Figure 3.8: A deterministic selection of a next set of states by $\delta$ as the user has given the input of $x = YY$. The states $s_j$ and $s_k$ both accommodate that $x = XX$ is to be preserved. The output as to be given by the choice agent for the query $x = YY$ is not shown here, as it would select either of the states $s_j$ and $s_k$. 
The next-set function $\delta$ can be turned into a next-state function $\delta'$ given that it takes the argument $\xi$ of the choice agent. As the choice agent only need to produce the statistics of a simple coin toss, its specification at any time can be represented by $\xi \in \{+, -\} \equiv CA$. Then, the next-state function works as $\delta' : X \times CA \times S \rightarrow S$. Lastly, also making use of the specification of the choice agent is the output function $\lambda$, which simply assigns the output as $y = \xi$ for any query with an output not specified in any currently occupied state. This completes the specification of the machine.

The structure of the state space is such that the states are one-to-one with the joint eigenstates of the contexts associated the chosen set of observables. Or, in other words, each state is associated exactly one ray in an ON-basis of rays, where each such ON-basis denotes a context, i.e., a common ON-basis. Hence the response of exciting an arbitrary number $n$ of qubits is simulated with a state space being one-to-one with the set of rays associated the chosen observables. This construction is in fact memory optimal for all $n$ under the present premises and is the content of Theorem 1.

**Theorem 1** ($n$-partite optimality). A finite-state machine simulating the quantum mechanical predictions in measurement outcomes associated a set $Q_n$ of Pauli observables on an $n$-partite system of qubits is memory optimal if there is a one-to-one correspondence between the set of machine states and the set of rays associated $Q_n$. The number of states in the optimal machine is then

$$2^n \prod_{k=0}^{n-1} (2^{n-k} + 1)$$

for $n \in \mathbb{N}_+$, from which it follows that the memory complexity is polynomially bounded as $O(n^2)$.

**Proof**: See Appendix A.3.

The fact that the machine functions as intended is trivial in that its states directly model the information of necessary quantum states associated the input alphabet.

### 3.3.2 Model with realism

Setting the context for this section is a study emphasising the memory usage of finite-state machines in order to quantify quantum contextuality [12]. The idea behind the study was to consider a non-contextual realist toy model, called Spekkens’ toy model [27], for which an extension was to be constructed such that the extended model had a desired contextual behaviour. The difference in memory usage between the initial toy model and its extension was proposed to be a
suitable measure—a memory cost—of contextuality. The premise for each state is realism, in the same sense as that of Kleinmann et al. of section 3.2, implying that to all input symbols there is a pre-existing output symbol in each state. Non-deterministic machines with such realism in the model are to be considered here.

What is here needed to know about Spekkens’ toy model is little, as it may be right-out cast as a non-deterministic finite-state machine. The paper itself is titled ‘in defense of the epistemic view of quantum states’, by which there are ontic states reflecting elements of reality, being nothing but the ordinary states of the machine, and there are epistemic states which each reflect a particular state of knowledge of a user of the machine. The relationship between the ontic and epistemic states is such that an epistemic state is a subset of the set of ontic states, i.e., a select region of the ontic state space, which expresses the knowledge of the knowing agent in the sense of “one of these ontic states is the one currently possessed by the system.” The region of possibly possessed ontic states varies with the agent probing the system (machine) through input, as knowledge is gained by the corresponding output. The non-deterministic (ontic) state space dynamics is governed by an information theoretic principle which Spekkens call the ‘knowledge-balance principle’, roughly stating that there is a persisting balance between knowledge and ignorance of the currently possessed ontic state for the inquiring agent. This safe-guards an ontic state from ever being fully revealed through an agent’s input sequences. Further, the non-determinism lies in that any next ontic state is randomized within some specific range, being the range also defining an agent’s epistemic state.

A problem seen in the extension of [12] is that the very same update structure as in [11] is used, being an update structure found inadequate as seen in the analysis of section 3.2. Hence the problem is, again, that a user of the machine may know too much about the current (ontic) state of the machine. Or, stated slightly different, the issue is that each ontic state can be associated with too many epistemic states; the number of epistemic states overlapping is too great. Still, this is what is seen to be the steps taken as per construction of the extended model, wherein each ontic state is endowed with special update instances not found in Spekkens’ model which effectively restricts the number of epistemic states compatible with any ontic state. The consequence is that the ontic state space grows, while the number of epistemic states remain the same. As such, there is an increased memory cost for running the extended machine as compared to Spekkens’ machine.

Continuing the same path by adding additional special update instances, machines accommodating all input sequences may be constructed. The more update instances associated an ontic state the less compatible it is with the epistemic states, simply as the machine then tends to dwell less in any same state when fed input sequentially. In fact, the preferred structure of update instances reflects a realistic extension of the minimal machine shown in (the previous) Section 3.3.1. This amounts to assigning pre-existing values in states, replacing those that previously where decided in situ by the choice agent. Then also, which ontic states that are compatible with which epistemic state is already set by the existing update instances.
The memory complexity that comes from assigning pre-existing values in states only reflects the statistical character of quantum mechanics. This is the reason why it is desirable to have a reference model, e.g., Spekkens’ toy model (or machine), operating under the same premises, such as to define a memory cost of contextuality by the difference in memory complexity between both machines having realism in their simulation model. If no reference machine would be available it would be rather unsatisfying in saying that the memory complexity reflects the presence of contextuality as there are more aspects the machine needs to fulfil, e.g., the statistical character.

The need is then to construct subsets in the state space reflecting a user’s maximal knowledge such that one of these states may be the one currently occupied by the machine. These subsets of states are precisely the subsets that the next-set function may point to, and are to be special in that they must be seen as statistical environments relative to the knowledge of the user. What this means is that a randomization over any such subset holds that any output not specified in any current state is, upon inquiry, given as either (+) or (−) with a chance as specified by quantum mechanics. The structural idea behind two slightly differing constructions, as devised from the minimal machine in Section 3.3.1, can be seen in figure 3.9.

Figure 3.9 shows an A scenario and a B scenario for a machine with an alphabet associated the full set of Pauli observables for a bipartite qubit system. The bipartite qubit case is simple in its statistical character, all quantum states are unbiased relative to all non-orthogonal states of ON-bases of other contexts (cf. Lemma 1). The A scenario is the simplest in that the statistically unbiased set of states only contain two states, in which one state has all unknowns specified as (+) and the other state has all unknowns specified as (−). The choice agent is here used for a randomization over the states in these kinds of subsets as to select a next state, where it is seen that the user is equally likely to encounter either (+) or (−) for any subsequent input. This addition of realism takes the states into twice as many, thus ending up with 120 states for this case. The B scenario has an additional constraint which might seem preferable in an aesthetic sense; the values known by the user reflects the states of the subsystems involved, i.e., each respective qubit. By this, one cannot just double-up the number of states as in the A scenario. Instead, all combinational possibilities are utilized such as to “lock” the specifications of states, whereby gaining an unbiased subset of states to randomize over should the need arise.

In counting the total number of states resulting from adding realism with regard to the subsystems, there are two cases to consider in the bipartite scenario. The first case is the combinational possibilities associated an ontic state compatible with an epistemic state reflecting the knowledge of any of the six non-local contexts, that is, a context which only contain observables pertaining to the correlations between subsystems, e.g., \{XX, YY, ZZ\}. That said, consider outcomes associated the context \{XX, YY, ZZ\} being known as (+,+,−)
Figure 3.9: Upon a user learning of the values shown in the upper box, cases A and B show different sets of states which are compatible with the user’s knowledge. These cases are taken to belong to two different machines. The next-set function $\delta$ in each machine would deterministically point to the respective case, whereas the choice agent would randomize over that particular set of states. This construction upholds the statistical character of quantum mechanics in both cases, as any next input would appear statistically unbiased. Entries with a diamond-square only highlights values which are necessarily fixed.

and have that the knowledge reflects the actual state in the subsystems by assigning values to either $\{XI,YI,ZI\}$ or $\{IX,IY,IZ\}$. Why the assignment to either of them is sufficient is because of relations as $(XI)(IX)(XX) = +$, hence if, say, $XI$ is assigned the value $(+)$ it follows that $IX$ is necessarily given the value $(+)$ because $XX$ is already assigned $(+)$. Thus, to have both subsystems reflecting the known information of the user, there are in this case two possibilities per observable of one of the subsystems.
3.3 Non-deterministic machines

Then as the assignment of values to the subsystems may also specify the other non-local observables, it completes the specification of the square. The number of combinations is $2^3 = 8$ and is nothing but the number of states constituting the statistical environment for the knowledge of $\{XX, YY, ZZ\}$. This set is what the next-set function picks out of the state space if a user attains such knowledge, as is seen in figure 3.9.

The second case is when a user has the knowledge of a local context, e.g., $\{XI, IX, XX\}$,

because here the combinational possibilities differ from the case mentioned above in that there are no $YY$ and $ZZ$ specified. This means that there is a complete freedom in selecting the values for $\{YI, ZI, IY, IZ\}$, amounting to $2^4 = 16$ possibilities.

These 16 combinations constitute a suitable statistical environment for the knowledge of $\{XI, IX, XX\}$.

As such, out of the 15 contexts there are six non-local contexts in which each of its four rays require an eight-state environment, and there are nine local contexts in which each of its four rays require a 16-state environment. The total number of states is then calculated as $4 \times (6 \times 8 + 9 \times 16) = 768$. This is a significant increase from the meagre 60 states when not forcing this aesthetic realism.
4

Discussion and conclusion

In the first part of this thesis quantum contextuality was discussed and was in the second part followed by finite-state machines simulating this contextual behaviour. The machines shown vary in their simulation model and as such they reach out to contextuality differently. The machine found optimal in its representation is one of non-determinism and non-realism. That a model of non-realism is present is in fact quite intuitive, because if the states would need to store answers in the sense that they pre-exist the time when a question asks about them, these pre-existing answers need to appear with a frequency as that of the statistical character of quantum mechanics. It follows that a state selected need to be picked from a subset of states in the state space which accommodates the statistics. Such a state space is then one which explicitly captures, not only contextual behaviour, but also the statistical character. Hence not forcing such realism entails eliminating the statistical character from adding to the state space complexity. It is in this way that the memory cost, i.e., memory complexity, of the state space is taken to measure contextuality, as the statistical character does not partake in its magnitude. The representation of the state space is found bounded polynomially as $O(n^2)$, meaning that the number of $b(n)$ bits required when linearly increasing the number of simulated qubits $n$ shows a growth less than exponential.

In fact, one may conclude that the optimal machine acts as an upper bound on the memory cost of the deterministic machines shown in Section 3.2. This is seen in that both types of construction have states which are in a one-to-one correspondence with the rays associated the set of observables, which in turn is associated the input alphabet. As the optimal machine makes use of all rays in any scenario, the deterministic machines make use of less. Hence the deterministic machines are also efficient in their representation.

As for the non-deterministic machines with realism in their model, any decisive result on their efficiency is not obtained here. It would require the knowledge
of all inner-products between the joint quantum states associated the contexts
of the Pauli group $\mathcal{P}_n$, because the number of states comprising a so-called statistical
environment associated a user’s maximal knowledge of a context of level $n$ is de-
pendent upon the statistical character. The question of efficiency is then whether
or not any factors which correspond to such environments grow exponentially in
the base two logarithm. It is however suspected that point (ii) of Lemma 1 may
hold for all $n$, by which all inner-products would be known (such a proof would
be an interesting result in and of itself).

Furthermore, on the topic of efficiency, a contextual computer was featured in
[28] which was deemed inefficient w.r.t. memory complexity. The memory struc-
ture of the computer is in fact of the same character as to what is here found to be
the optimal state space structure, i.e., “remember one context at a time”. In one
of its sections, namely ‘A classical contextual computer requires unlimited density
of memory’, the claim is that the memory complexity grows exponentially as the
number of simulated qubits grows linearly. Such a growth in memory complexity
makes the machine inefficient in the sense that a classical simulation is deemed
infeasible. Surely, inefficiency could be the case as it depends upon the particular
simulation model, but with a character the same as the one herein seen optimal,
and efficient, something must be amiss. Inspection reveals that Eq.(6) of [28] is
faulty and is the factor responsible for the exponential growth (the equation’s cor-
rectness would imply that $n$ qubits can ship $2^n - 2$ bits of information which is
a strong violation of the Holevo bound). With the contradiction resolved as to
what is presented in this thesis, the position that the optimal non-deterministic
contextual machine is efficient is kept.

Adding to—and killing—the discussion of efficiency is the study by D. Gottes-
man [29], where a theorem known as the Gottesman-Knill theorem is presented.
This theorem is a statement concerning the efficiency of simulating a subset of
quantum dynamics in an $n$ qubit system and is expressed by the stabilizer formal-
ism which makes use of states called stabilizer states. These states are precisely
the states associated the contexts of the Pauli group and as such the machines
considered here simulates the same measurement dynamics. What the theorem
states is that the simulation is time-efficient, meaning that the bits which rep-
resent the information about the quantum state may be manipulated in polyno-
mial time. But, the significance of the time-efficiency is by its support of an
efficient representation in memory. This representation is implicitly referred to
when talking about the Gottesman-Knill theorem and use a number of $2n(2n+1)$
bits for simulating $n$ qubit stabilizer dynamics. As such this representation is
also bounded in complexity as $O(n^2)$. Hence there is no asymptotic advantage
in memory between the optimal machine or the Gottesman-Knill representation.
Also, importantly, there is no problem of efficiency associated the simulation of
contextuality in measurement outcomes over systems of qubits. What follows is
that the state-independent contextuality found in systems of qubits does not offer
a computational separation between the quantum and the classical.

One may further argue, even though casting the Gottesman-Knill theorem
aside, that the question of efficiency in memory is straightforwardly resolved
by simply counting the number of joint quantum states associated the set of observables over which measurements are interesting. In the case of qubits one only needs knowledge of the Pauli group which—being a well explored set of operators—associates a number \(2^n\) (the number of orthogonal states associated any single context) times the number of contexts (cf. Lemma 1) of joint quantum states. Taking this number under the base two logarithm yields something polynomial in growth, as we have seen. Although that is simple and straightforward, one must recognize that part of the impetus for this study was to investigate the structure associated minimal representations in these contextual scenarios, as it might prove useful toward characterizing the interplay between an agent inquiring about an underlying hidden reality. That is, may we in some way characterize the epistemic view on the quantum mechanical formalism by way of the resulting memory complexity associated contextuality? Such a question has more dimensions to it than allowing itself answered by only knowing whether or not a simulation is efficient.

An adjacent study addressing the epistemic view toward quantum mechanics is Cabello et al. [30], being a view assuming that the quantum probabilities are determined by intrinsic properties of Nature. It is adjacent in that it also employs finite-state machines as the main tool to simulate outcome behaviour of simple qubit systems. Similar to this thesis, the memory complexity of a minimal machine is central to their investigation as it is used for arguing about the memory density associated a single qubit system should the epistemic view be true. What is found is that the memory density tends (linearly) to infinity as a proposed ideal experiment gains more ways of probing a single qubit. That the memory density would be infinite seem rather infeasible as such a spatially small system would then need to store an incomprehensible amount of detail beyond its information carrying capability. This kind of result may be seen to strengthen the position for quantum epistemologists.

Still, the result of an infinite memory density may be seen not too much of a surprise and can be paralleled with the minimal machine found here in Section 3.3.1. The machine in Cabello et al. simulates an ideal experiment consisting of sequential spin measurements in the \(\hat{x}\) and \(\hat{z}\) directions. The states of the machine gives the probabilities for the measurement outcomes, whereby the machine represents the objective aspect, i.e., reality, in the epistemic view. At first only measurements in the two directions of \(\hat{x}\) and \(\hat{z}\) are allowed, but is further followed by a relaxation in choosing the measurement direction by allowing measurements in an exponentially increasing set of directions in the \(xz\)-plane. As the probabilities of outcomes vary with varying directions there must also be a state associated each new direction. In the limit is the continuum of directions spanning the \(xz\)-plane and as such it is not so surprising to find the system of qubits being associated an infinite memory density. The parallel with the minimal machine herein presented is such that it would need to have its output sensitive to how it is rotated in a plane, i.e., for distinct angles of rotation the machine must be able to identify the machine’s rotation. Hence distinct angles of rotation must correspond to different states; a number of states tending to infinity for a single qubit. In this way the
Discussion and conclusion

representation becomes inefficient as it would take an infinite amount of memory to capture any continuum.

Other aspects to the idea of an epistemic view of quantum mechanics is whether any principle governing the dynamics between some reality and an inquiring agent exist. Such a discussion may perhaps best be taken in the language of ontic and epistemic states, reflecting machine states and user states of knowledge respectively. It is the relation between these two concepts of state which is expressed by the state space dynamics in the finite-state machines and this relation is interesting from a foundational point of view. For instance, above we considered the efficiency of such a relation.

To set the context, consider again the study of R. Spekkens [27] which explores the relation between an inquiring agent’s knowing (epistemic states) of an underlying hidden reality (ontic states). What Spekkens seemed to aim at, was to show that simple information theoretic principles could be used to articulate an interesting interplay between an inquiring agent’s interaction with some fine-grained hidden reality, in that the generated abstract systems may show a host of behaviours typical of the quantum mechanical description. The fact that contextuality is not present is raised in the paper, wherein Spekkens sets the challenge of finding a principle which similarly may generate an abstract system showing contextual behaviour, at least close to, that of quantum contextuality.

The resulting structure of an optimal machine was thought to perhaps give some hints as to the formulation of an information theoretic principle with which one could, at least, generate an abstract system showing some contextual behaviour. It is believed that this was part of the impetus for seeking a contextual extension of Spekkens’ toy model [12], a study in which the ‘knowledge-balance principle’ as stated by Spekkens was given a relaxation in the form of a ‘knowledge-imbalance principle’. The relaxation is necessary in that it was noted that “a balance” could not be upheld as the ontic state space grew while the epistemic possibilities of state remained fixed. This is a consequence of associating less epistemic states to any single ontic state, whereby the epistemic states—pictured as uniform distributions over select portions of the ontic state space—cease to overlap as much. A larger underlying ontic state space is then needed as the epistemic states separate. This separation of epistemic states is due to the required fulfillment of contextual correlations, as it is such correlations that demand a certain amount of ignorance associated the ontic states in order to be upheld.

As is seen in this thesis, the amount of ignorance in the optimal machine, and the non-deterministic realistic machine, is maximal in the sense that only the knowledge of single context is associated an ontic (machine) state. What this means is that the optimal scenario is when no two epistemic states overlap which, one could say, reflects the limit of the mentioned attempt at constructing a contextual extension to Spekkens’ toy model, as the separation of the epistemic states are now total.

Does this hint at some information theoretic principle capturing some contextual behaviour? As the epistemic states seem to enjoy not overlapping, they kind of lose their character as states of mere information. Because what is truly defin-
ing of epistemic states as pertaining to an underlying reality is the fact that they may overlap, contrasting states of realism which either ‘are’ or ‘are not’. Hence the simple state space structures shown herein are not seen to help the cause of epistemology as applied to quantum mechanics.

Also, perhaps being a curious relation, a somewhat controversial theorem in the debate of “quantum states as epistemic states over an underlying reality” is the PBR theorem [31], which roughly states that if epistemic states overlap the predictions of quantum mechanics cannot be upheld; a so-called no-go theorem toward these kinds of realist presuppositions. The author is however still sceptical about the content of the theorem as any caveats in its proof has not yet been considered personally, but at a distance it does corroborate what is seen here with the separation of epistemic states.

Using the language of finite-state machines we have here taken an information theoretic approach to quantum contextuality. The classical resources associated quantum contextuality was investigated in order to see if any argument, positive or negative, toward an epistemic view of the quantum formalism could be formulated. Although being an interesting approach, no novel grip on the epistemic view was found, nor does the type of contextuality considered here show any computational separation between the quantum and the classical w.r.t. memory complexity.


A.1 Proof of Lemma 1

For arbitrary $n \in \mathbb{N}_+$, the fact that each context contain $2^n - 1$ non-trivial Pauli observables is shown in [32], whereas that the number of pure stabilizer states is

$$2^n \prod_{k=0}^{n-1} \left( 2^{n-k} + 1 \right)$$  \hspace{1cm} \text{(A.1)}

is shown in Proposition 2 of [33].

The Holevo bound (see e.g. [15]) certifies that $n$ qubits may (at most) ship $n$ bits. Then, as each context defines what is maximally simultaneously measurable, it follows that values associated its observables may come in $2^n$ different configurations; each configuration being nothing but a pure stabilizer state associated the context. Hence by dividing the total number of pure stabilizer states Eq.(A.1) with the number of pure stabilizer states associated any single context, i.e., $2^n$, one retrieves the total number of contexts.

Point (ii) of the lemma makes use of Theorem 8 in [34]:

Let $|\psi\rangle$ and $|\phi\rangle$ be two non-orthogonal stabilizer states. Over all sets of generators $\{P_1, \ldots, P_n\}$ and $\{Q_1, \ldots, Q_n\}$ for the associated stabilizer groups $S(|\psi\rangle)$ and $S(|\phi\rangle)$ respectively, let $s$ denote the minimum number of generating elements not in common. It holds that

$$|\langle \psi | \phi \rangle|^2 = \frac{1}{2s} \ .$$  \hspace{1cm} \text{(A.2)}

In the bipartite case of $n = 2$ each context is associated four joint quantum states. Such a joint quantum state is uniquely defined by a stabilizer group (cf. Section
2.5 Eq.(2.24)) which may be generated from any pair of its non-trivial elements. There are only two cases to consider here as any pair of contexts have either no common Pauli observables or one common Pauli observable.

Let $A$ and $B$ be two contexts having no Pauli observable in common, by which it follows that any generating set for its (four) stabilizer groups cannot have any common elements. Thus $s = 2$, giving that for any joint quantum state associated $A$, its inner-product with all (four) joint quantum states of $B$ is $1/4$. This also holds for all states of $B$, whereby the ON-bases of $A$ and $B$ are mutually unbiased (cf. Section 2.3 Eq.(2.14)).

Next, let $A$ and $B$ be two contexts having exactly one Pauli observable in common. This means that for any stabilizer group $S(|\psi\rangle)$ associated $A$ and any stabilizer group $S(|\phi\rangle)$ associated $B$, the generators of each group may maximally have a single same element. When this is the case, there is only one element not in common, hence $s = 1$. It follows that the inner-product between joint quantum states defined by such stabilizer groups is $1/2$.

The case where contexts overlap, i.e., they have a common Pauli observable, is such that a state of either basis set is orthogonal to two states of the other basis set. This is because states associating different values to the same Pauli observable represent mutually exclusive events. Then, because of the Pauli observables being dichotomic, the single common element among these contexts makes two of the four orthogonal. It is then seen that the states of a context $A$ is unbiased over non-orthogonal remainder of a context $B$.

Both these cases have in common that any joint quantum state associated a context is evenly distributed over any of the other context’s non-orthogonal joint quantum states. This is precisely what is expressed by the lemma.
Shown here is the 27-state machine's state space for when the machine is deterministic over the complete set of two-qubit observables. An encircled entry in a state $s_j$ denotes an instance of update. The next state associated such an update is found in the corresponding entry of the associated update-table $u_j$. 

<table>
<thead>
<tr>
<th>S1:</th>
<th>S2:</th>
<th>S3:</th>
<th>S4:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="state1.png" alt="State Diagram" /></td>
<td><img src="state2.png" alt="State Diagram" /></td>
<td><img src="state3.png" alt="State Diagram" /></td>
<td><img src="state4.png" alt="State Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>U1:</th>
<th>U2:</th>
<th>U3:</th>
<th>U4:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="update1.png" alt="Update Diagram" /></td>
<td><img src="update2.png" alt="Update Diagram" /></td>
<td><img src="update3.png" alt="Update Diagram" /></td>
<td><img src="update4.png" alt="Update Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S5:</th>
<th>S6:</th>
<th>S7:</th>
<th>S8:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="state5.png" alt="State Diagram" /></td>
<td><img src="state6.png" alt="State Diagram" /></td>
<td><img src="state7.png" alt="State Diagram" /></td>
<td><img src="state8.png" alt="State Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>U5:</th>
<th>U6:</th>
<th>U7:</th>
<th>U8:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="update5.png" alt="Update Diagram" /></td>
<td><img src="update6.png" alt="Update Diagram" /></td>
<td><img src="update7.png" alt="Update Diagram" /></td>
<td><img src="update8.png" alt="Update Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S9:</th>
<th>S10:</th>
<th>S11:</th>
<th>S12:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="state9.png" alt="State Diagram" /></td>
<td><img src="state10.png" alt="State Diagram" /></td>
<td><img src="state11.png" alt="State Diagram" /></td>
<td><img src="state12.png" alt="State Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>U9:</th>
<th>U10:</th>
<th>U11:</th>
<th>U12:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="update9.png" alt="Update Diagram" /></td>
<td><img src="update10.png" alt="Update Diagram" /></td>
<td><img src="update11.png" alt="Update Diagram" /></td>
<td><img src="update12.png" alt="Update Diagram" /></td>
</tr>
</tbody>
</table>
A finite-state machine which is optimal is one using the least number of states to perform its task.

**Proof of Theorem 1**: For each \( n \in \mathbb{N}_+ \) let a finite-state machine \( M_n \) have a state space \( S_n \) and let its input alphabet \( X_n \) be symbols having a one-to-one relationship with a set \( Q_n \) of Pauli observables associated an \( n \)-partite system of qubits. For the two statements

A: \( M_n \) can simulate the measurement outcome behaviour of an \( n \)-partite system of qubits as predicted by quantum mechanics.

B: There is a set \( R_n \subseteq S_n \) of states in which each state is associated *at least* the informational content equivalent to that of a ray in an ON-basis of rays (context), such that all rays associated \( Q_n \) are represented.

It holds that \( A \Rightarrow B \) (necessity) and \( B \Rightarrow A \) (sufficiency). The necessity part is due to the fact that a user of \( M_n \) is able to give a sequence of inputs corresponding to a maximal set of compatible measurements associated \( Q_n \), where the output of such a measurement sequence is the informational equivalent of a ray in an ON-basis associated a subset of \( Q_n \). By a user’s freedom of choice in selecting input symbols there must exist a collection \( R_n \) of states which supports instances where a user’s knowledge of any single ray is required reproducible. The sufficiency part is trivial in that a user of \( M_n \) can always be satisfied by at least one state in \( R_n \) as they cover all cases where some previous output-history is required reproducible.

By \( A \iff B \) follows that \( S_n = R_n \) may be taken, whereby minimizing the number of states in \( R_n \) implies the minimization of \( M_n \).

As each state in \( R_n \) is at least associated the informational content of a single ray, one may minimize the number of states by a single requirement; output values associated a single ray are the only values specified in a state and are all accessible (cf. figure 3.7). This requirement puts \( R_n \) in a one-to-one relation with the set of rays associated \( Q_n \). The content of the following is just an assertion of the intuitive: forcing more work onto the choice agent reduces the complexity of the state space.

To see that the above requirement is minimizing, assume that each state in \( R_n \) do in fact abide by the above requirement. Take an arbitrary state

\[
r_k = \{v_1^{(k)}, \ldots, v_p^{(k)}\} \in R_n
\]  

(A.3)

where the entries are the \( p \) reproducible values associated a single ray of \( Q_n \). That values \( \{v_1^{(k)}, \ldots, v_p^{(k)}\} \) of a state is accessible refers to that there are no instances of
update associated any query about them. If there were, an update could only point to a next set of states in which all states are associated the same ray, hence a clear redundancy. Further, attempt extending the information contained in any $r_k$ by appending to it a value $w \in \{+,-\}$, that is

$$r_k \rightarrow r_k'(w) = \{v_1^{(k)}, \ldots, v_p^{(k)}, w\},$$

which is the most simple case of extension. For the value $w$ to appear with a frequency as that of quantum mechanic's statistical character to a user having knowledge of $\{v_1^{(k)}, \ldots, v_p^{(k)}\}$, the next-set function $\delta$ demands the inclusion of both possibilities for $w$ so that it can select a suitable subset of states. That is, if $r_k'(w = +1)$ is included, then in order to have a viable state space at least some $r_k'(w = -1)$ need also be included. As such, any state space $R'_n$ containing extended states in the sense of $r_k'$ necessarily contains more states that $R_n$.

Then by Lemma 1, the one-to-one relationship between states of $M_n$ and rays associated the set of Pauli observables $Q_n$ holds that

$$|S_n| = 2^n \prod_{j=0}^{n-1} (2^{n-j} + 1),$$

which is then the number of states used by a minimal machine $M_n$.

The memory required to address each state in the state space is then

$$\log_2(|S_n|) = \log_2\left(2^n \prod_{j=0}^{n-1} (2^{n-j} + 1)\right) = n + \sum_{j=0}^{n-1} \log_2(2^{n-j} + 1),$$

where

$$\log_2(2^{(n-j)} + 1) < \log_2(2^{(n-j)} + 2) = n + \log_2\left(\frac{2^n}{2^j} + \frac{1}{2^j}\right) \leq \left|\frac{2}{2^n} \leq 1 \forall n \in \mathbb{N}_+, \right. \left. \frac{1}{2^j} \leq 1 \forall j \in \{0, \ldots, n-1\}\right| \leq n + \log_2(1 + 1) = n + 1,$$

by which one may see

$$\log_2(|S_n|) = n + \sum_{j=0}^{n-1} \log_2(2^{(n-j)} + 1) < n + n(n + 1) = n^2 + 2n.$$

Hence the memory complexity is bounded polynomially as $O(n^2)$.