Imperial College London
Department of Physics

# A few connections between Quantum Computation and Quantum Non-Locality

Matthew Pusey

December 16, 2010

Supervised by Terry Rudolph and Jonathan Barrett

Submitted in part fulfilment of the requirements for the degree of Master of Research in Controlled Quantum Dynamics of Imperial College London and the Diploma of Imperial College London

# Abstract

Quantum mechanics enables some computational problems to be solved by quantum algorithms faster than any known classical algorithm. Is there a link between this and the other features of quantum mechanics that set it apart from classical theories?

A dramatic example of such a feature is quantum non-locality. This report begins by reviewing the stabilizer formalism, an efficient way of doing some quantum mechanical calculations. The formalism is then applied in Chapter 2 to review quantum non-locality. A brief review of quantum computation follows in Chapter 3, which compares the circuit and measurement-based approaches and provides a simple argument for a link with quantum non-locality.

The novel content in this report begins in Chapter 4 where Spekkens' example of a theory that shares some features with quantum mechanics is reviewed and a stabilizer-like notation for it is introduced. The theory is local by construction and so does not exhibit anything similar to quantum non-locality. With this in mind, the limited computational power of the theory is noted.

Finally Chapter 5 defines a non-standard notion of locality for quantum circuits, and shows that circuits that can be described within the stabilizer formalism fit this definition. It is known that this type of quantum circuit is not useful for quantum computation because it can be efficiently simulated on a simple classical computer. Therefore this chapter provides further support for a link between quantum computation and non-locality.

Chapter 6 presents a summary and some plans for extending this work.

With thanks to my supervisors Terry and Jon for providing excellent insights and suggestions and then putting up with me ignoring many of them!

# Contents

**1 Stabilizer formalism for qubits** **6**

1.1 States . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

1.1.1 Check vectors . . . . . . . . . . . . . . . . . . . . . . 8

1.1.2 Pure states . . . . . . . . . . . . . . . . . . . . . . . 9

1.2 Transformations . . . . . . . . . . . . . . . . . . . . . . . 9

1.3 Measurements . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.4 Effective measurements . . . . . . . . . . . . . . . . . . . . 11

**2 Quantum non-locality** **13**

2.1 Can stabilizer-formalism description of physical reality be considered complete? . . . . . . . . . . . . . . . . . . . . . . . 13

2.2 The GHZ paradox . . . . . . . . . . . . . . . . . . . . . . . 14

2.3 Local Hidden Variable models . . . . . . . . . . . . . . . . . 15

**3 Quantum computation** **16**

3.1 Quantum circuits . . . . . . . . . . . . . . . . . . . . . . . 16

3.2 Measurement-based computation . . . . . . . . . . . . . . . . 19

3.3 Heuristic argument for a link with non-locality . . . . . . . . 20

**4 Spekkens' toy theory** **22**

4.1 States . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22

4.2 Transformations . . . . . . . . . . . . . . . . . . . . . . . 29

4.3 Measurements . . . . . . . . . . . . . . . . . . . . . . . . . 30

4.4 Computation in the toy theory . . . . . . . . . . . . . . . . 33

4.4.1 Circuit computation . . . . . . . . . . . . . . . . . . . 33

4.4.2 Measurement-based computation . . . . . . . . . . . . 34

**5 LHV models for quantum circuits** **35**

5.1 Definition . . . . . . . . . . . . . . . . . . . . . . . . . . . 36

5.2 "Almost classical" circuits . . . . . . . . . . . . . . . . . . . 37

5.3 Clifford circuits . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

5.3.1 Example . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42

**6 Conclusions and outlook 44**

**A Appendices 46**

A.1 Linking the stabilizer formalism to the Hilbert space formalism 46

A.1.1 States . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

A.1.2 Transformations . . . . . . . . . . . . . . . . . . . . . . 48

A.1.3 Measurements . . . . . . . . . . . . . . . . . . . . . . . 48

A.2 Stabilizer generators versus pseudo-stabilizer generators . . . 49

A.3 Joint probabilities as expectation values of products . . . . . 50

**Bibliography 51**

# 1 Stabilizer formalism for qubits

I begin by reviewing the stabilizer formalism for qubits (two-level quantum systems). This formalism was introduced by Gottesman [17] to analyse certain quantum error-correcting codes, but as a compact and efficient way of analysing an interesting collection of quantum operations it has many applications beyond that.

The purpose of this review is to define notation I will use later on, and to establish a variant of the formalism in which certain mixed states are included along with the usual pure states. The link with the standard formalism of quantum mechanics (known as the Hilbert space formalism) is made in Appendix A.1. However, it is helpful for what follows to put that to one side and view the stabilizer formalism simply as a self-contained theory of certain quantum preparations, transformations and measurements.

For another introduction to the formalism see [28, Section 10.5.1]. For a thorough development of the mathematics see [9] and [12].

## 1.1 States

The stabilizer formalism makes extensive use of the "Pauli group". For our purposes a group is simply a non-empty set of $d \times d$ matrices satisfying

- *Closure under multiplication* — if $A$ and $B$ are in the group, then so is their product $AB$;
- *Inverses* — if $A$ is in the group, then its inverse $A^{-1}$ exists and is in the group.

Note that since the group is non-empty it contains some matrix $A$, and so it contains $A^{-1}$, and so it contains $A^{-1}A$ which is the $d \times d$ identity matrix.

The elements of the Pauli group on $n$ qubits, denoted $P_n$, are the $2^n \times 2^n$ matrices of the form

$$\alpha p_1 \otimes \cdots \otimes p_n \tag{1.1}$$

where $\otimes$ indicates the tensor product [28, Section 2.1.7], $\alpha$ is a "phase factor" 1, $-1$, $i$, or $-i$, and the $p_k$ are chosen from

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (1.2)$$

Therefore the Pauli group for a single qubit, $P_1$, consists of the 16 matrices $I$, $-I$, $iI$, $-iI$, $X$, $-X$, $iX$, $-iX$, $Y$, $-Y$, $iY$, $-iY$, $Z$, $-Z$, $iZ$ and $-iZ$.

It is easy to check that

- If an element of the Pauli group $g$ is written in the form (1.1) then $g^2 = \alpha^2 I^{\otimes n}$ ($I^{\otimes n}$ means the tensor product of $I$ with itself $n$ times, which is the $2n \times 2n$ identity matrix); and
- If $g$ and $h$ are in $P_n$, then either $gh = hg$ (in which case we say $g$ and $h$ commute) or $gh = -hg$ (in which case we say they anticommute).

Any element of a group can be written as some product of what are known as the group's generators. A useful set of generators for $P_n$ is $iI^{\otimes n}$ along with $X_1, X_2, \ldots, X_n$ and $Z_1, Z_2, \ldots, Z_n$. The shorthand $X_k$ denotes $X$ acting on the $k$-th qubit with the identity $I$ acting on the rest, i.e. $I^{\otimes k-1} \otimes X \otimes I^{\otimes n-k}$. Similarly $Y_k$ denotes $Y$ on the $k$-th qubit, and $Z_k$ denotes $Z$ on the $k$-th qubit. We say a list of generators is independent if none of the generators can be written as a product of the rest. The list of generators for $P_n$ given above is independent.

A subgroup is a subset of the elements of a group that itself forms a group. Let $S$ be a subgroup of $P_n$ that does not contain $-I^{\otimes n}$. None of the elements with imaginary phase factors $\alpha = i$ or $\alpha = -i$ are in $S$, since they square to $-I^{\otimes n}$. Therefore all the elements are Hermitian matrices, which correspond to quantum mechanical observables [28, Section 2.2.5]. Since they square to the identity their eigenvalues are all 1 or $-1$. In fact they are commuting observables, since if two elements anti-commute (which is the only other option in $P_n$) it is easy to check that their product has an imaginary phase factor. We define the quantum state of $n$ qubits $\rho_S$ by requiring the expectation values of these observables to be 1, whilst the expectation values of the other Pauli observables (i.e. the Hermitian $g$ in $P_n$ with neither $g$ nor $-g$ in $S$) are required to be zero. Put another way, we are certain that we will get the +1 outcome for some Pauli measurements $S$ and have no knowledge about the rest.

The subgroup $S$ is often specified by writing a set of independent generators in angle brackets, for example the two-qubit "singlet state" is given by $S = \langle -X_1X_2, -Z_1Z_2\rangle = \{I^{\otimes 2}, -X_1X_2, -Y_1Y_2, -Z_1Z_2\}$. Since the elements of $S$ commute and square to the identity, it is easy to check that if $S$ has $l$ independent generators then it has exactly $2^l$ elements.

### 1.1.1 Check vectors

Let $g = \alpha p_1 \otimes \cdots \otimes p_k$ be an element of $P_n$. Then we define a "check vector" for $g$ as $r(g) = (x_1, \ldots, x_n, z_1, \ldots, z_n)$ where

- If $p_k = I$ then $x_k = 0$, $z_k = 0$;
- If $p_k = X$ then $x_k = 1$, $z_k = 0$;
- If $p_k = Y$ then $x_k = 1$, $z_k = 1$; and
- If $p_k = Z$ then $x_k = 0$, $z_k = 1$.

Note that the check vector completely specifies $g$ except that it gives no information about the phase factor $\alpha$.

Up to some phase factor we have $p_k = X^{x_k}Z^{z_k}$. Since $X$ and $Z$ square to identity and commute up to some phase factor we see that if that if $g$ and $h$ are elements of $P_n$ then $r(gh) = r(g) \oplus r(h)$ where $\oplus$ denotes component-by-component addition modulo 2. For example $XY = iZ$ and $r(X) \oplus r(Y) = (1, 0) \oplus (1, 1) = (1 \oplus 1, 0 \oplus 1) = (0, 1) = r(iZ)$. This means that if some elements of $P_n$ are independent, then their check vectors must be linearly independent.

Check vectors are a useful way of checking if two elements of the Pauli group commute. Let $g = \alpha p_1 \otimes \cdots \otimes p_n$ and $h = \alpha' q'_1 \otimes \cdots \otimes q'_n$ be two elements of $P_n$. For each $k$, $p_k$ commutes or anticommutes with $p'_k$. Let $a$ be the number of $k$ for which $p_k$ and $p'_k$ anticommute. Then $gh = (-1)^a hg$ since each anticommuting pair gives a minus sign. This shows that $g$ and $h$ commute for $a$ even and anticommute for $a$ odd.

Consider the check vectors $r(g) = (x_1, \ldots, x_n, z_1, \ldots, z_n)$ and $r(h) = (x'_1, \ldots, x'_n, z'_1, \ldots, z'_n)$. By checking every combination we see that $p_k$ and

$p_k'$ anticommute if and only if $x_k z_k' \oplus z_k x_k' = 1$. Therefore the condition that $a$ is even gives that

$$x_1 z_1' \oplus z_1 x_1' \oplus x_2 z_2' \oplus z_2 x_2' \oplus \cdots \oplus x_n z_n' \oplus z_n x_n' = 0 \tag{1.3}$$

if and only if $g$ and $h$ commute. This can also be written $r(g)^T J_n r(h) = 0$ where addition modulo 2 is implied and $J_n$ is a $2n \times 2n$ matrix that can be written in terms of the $n \times n$ identity matrix $I_n$ as

$$J_n = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}. \tag{1.4}$$

### 1.1.2 Pure states

Suppose we have a subgroup $S$ of $P_n$ containing $2^n$ elements but not $-I^{\otimes n}$. Suppose we try to add another element $h$. We cannot add $\alpha g$ for any phase factor $\alpha \neq 1$ and element $g$ of $S$ because that would mean $\alpha g g = \alpha I^{\otimes n}$ would have to be added as well. Therefore the new element $h$ must have a check vector $r(h)$ different to all those of the current elements of $S$. Furthermore $h$ must commute with the existing elements of $S$, in particular it must commute with some set of $n$ independent generators of $S$. This means $r(h)$ must satisfy be the linear equations defined by (1.3) for each of these independent generators. Since the check vectors of these generators are linearly independent the linear equations are independent and so the solutions are a subspace of dimension $2n - n = n$, which has $2^n$ elements. But these must be exactly the existing check vectors of $S$. Therefore we cannot add a new element $h$.

In conclusion, $n$-qubit stabilizer states with a subgroup $S$ of size $2^n$ are states of maximal knowledge (in that we cannot be certain about any more measurement outcomes), which we will refer to as "pure states".

## 1.2 Transformations

The set of reversible quantum transformations that always take stabilizer states to stabilizer states is known as the Clifford group. Within the stabilizer formalism such a transformation is represented by a reversible function $f$ from the Pauli group $P_n$ to itself. To update a stabilizer subgroup $S$ we just apply $f$ to each element. If we specify $S$ by a list of generators than we can simply apply $f$ to each generator.

The linearity of quantum mechanics means that $f$ will be

- *Homogeneous* — for any $g$ in the Pauli group and any phase factor $\alpha$, $f(\alpha g) = \alpha f(g)$; and
- *A group homomorphism* — for any $g$ and $h$ in $P_n$, $f(gh) = f(g)f(h)$. This implies that $f(I^{\otimes n}) = I^{\otimes n}$.

Therefore to specify $f$ we need only specify its action on $X_1, \ldots, X_n$ and $Z_1, \ldots, Z_n$, as any element of $P_n$ can be written as a product of some of those (possibly with a phase factor) and we can use the properties above to determine the action of $f$. For example the single-qubit Hadamard transformation, which I will denote $f_H$, is completely specified by stating that $f_H(X) = Z$ and $f_H(Z) = X$. We can then calculate that

$$f_H(Y) = f_H(iXZ) = if_H(XZ) = if_H(X)f_H(Z) = iZX = -Y. \tag{1.5}$$

Another single-qubit transformation is the "phase" transformation $f_S$ defined by $f_S(X) = Y$ and $f_S(Z) = Z$. Any single-qubit transformation in the Clifford group can be decomposed into sequence of $f_H$ and $f_S$ transformations. The controlled-NOT transformation on two qubits, with control qubit 1 and target qubit 2, is defined by

$$f_{CNOT}(X_1) = X_1X_2, \qquad f_{CNOT}(Z_1) = Z_1 \tag{1.6}$$

$$f_{CNOT}(X_2) = X_2, \qquad f_{CNOT}(Z_2) = Z_1Z_2. \tag{1.7}$$

Any transformation in the Clifford group can be decomposed into a sequence of $f_H$, $f_S$ and $f_{CNOT}$ transformations acting on the various qubits [9].

## 1.3 Measurements

Measurements in the formalism are restricted to "Pauli observables", which is to say the Hermitian elements of the Pauli group. Their expectation values were given in the definition of the states above. A useful shortcut is that the expectation value of an observable is zero if it anticommutes with at least one of the generators of $S$. In the special case of pure state, the converse is also true: if the expectation value of an observable is zero then it anticommutes with at least one of the generators [28, Section 10.5.3].

If the expectation value of an observable $O$ is 1 or $-1$ then that value is returned by the measurement and the state is unchanged.

If the expectation value is zero then there is a set of generators for $S$ with at most one element that anticommutes with $S$ (if starting with a list where more than one does, just multiply the additional such elements by the first such element). The measurement returns 1 or $-1$ with equal probability and to find the new state $O$ or $-O$ respectively is added to the list of generators, and the anti-commuting element (if present) is removed.

## 1.4 Effective measurements

Two types of "effective measurement" will be useful in what follows. Suppose we are interested in implementing some Pauli measurement $O$. By effective measurement I mean a procedure that has the same probabilities of giving each outcome as $O$. The procedure may well leave the system in a different state to the actual measurement. But if we don't care about the state of the system after the measurement then the effective measurement is completely interchangeable with the actual one.

The first type is straightforward. Consider applying some Clifford group transformation $f$ with $f(O) = O'$, followed by a measurement of $O'$. Denote the initial state of the system $S$, and the state application of $f$ as $S'$. Certainly $O'$ is in $S'$ if and only if $O$ is in $S$ since $f$ is reversible. Also $-O'$ is in $S'$ if and only if $-O$ is in $S$ since $f$ is homogeneous (and, as before, reversible). Therefore the application of $f$ followed by the measurement of $O'$ is an effective measurement of $O$.

The second type is a bit more subtle, and is best illustrated through an example. Suppose $O = X_1X_2$. Then an effective measurement of $O$ can be implemented by measuring $X_1$ (which only requires access to the first qubit) and $X_2$ (only requiring the second qubit) and multiplying the results (i.e. the numbers $+1$ or $-1$). This will work regardless of which order $X_1$ and $X_2$ are measured. This can be seen by checking every case. Denote the state before the measurement as $S$. The cases for $X_1$ being measured first (just interchange the labels 1 and 2 for the cases where $X_2$ is measured first) are

1. $O$ is in $S$, but $X_1$ is not,
2. $O$ and $X_1$ are in $S$,
3. $-O$ is in $S$, but $X_1$ is not,
4. $-O$ and $X_1$ are in $S$,
5. Neither $O$, $-O$ nor $X_1$ is in $S$; and
6. Neither $O$ nor $-O$ is in $S$, but $X_1$ is.

Take the first case. Since $X_1$ is not in $S$, either outcome, $\pm 1$, will be returned with equal probability and then $\pm X_1$ will be added to $S$. Any elements of $S$ that anticommute with $X_1$ will be removed, but this certainly does not include $O$. Since the subgroup is closed $\pm X_1 O = \pm X_2$ will end up in the subgroup. Therefore an $X_2$ measurement will return $\pm 1$ with certainty. Hence the product of the two outcomes is $(\pm 1)(\pm 1) = 1$ as required since $O$ was in $S$. The remaining five cases are similar.

This argument can easily be generalized to the statement that measuring $p_1$ on the first qubit, $p_2$ on the second qubit etc, and multiplying the results, is an effective measurement of $O = p_1 \otimes p_2 \otimes \cdots \otimes p_n$.

# 2 Quantum non-locality

In this chapter I make use of the stabilizer formalism to briefly review the concept of quantum non-locality. Mermin [24, 26] has written several engaging and accessible introductions to this topic. More detailed discussions can be found in [5] and [22].

## 2.1 Can stabilizer-formalism description of physical reality be considered complete?

The title of this section is based on that of the seminal paper “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” [16] by Einstein, Podolsky and Rosen (EPR). Their argument was recast in terms of qubits by Bohm [7], and the version here is based on that.

Consider two qubits (for example, the spins of two electrons) prepared in the “singlet state”, $S = \{I^{\otimes 2}, -X_1X_2, -Y_1Y_2, -Z_1Z_2\}$ and then taken to separate locations. The state tells us that if we measure $X$, $Y$ or $Z$ on one qubit, we will get the opposite result for the same measurement on the other qubit since their product will be $-1$ with certainty (recall the second type of effective measurement in §1.4). But the state does not predict the result of the first measurement. Since $S$ has $2^2$ elements it is a pure state, and so no additional predictions, for example of the result $X_1$ alone, can be made by any valid quantum state.

EPR posit the hypothesis that “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”. They also assume that a measurement on the first qubit cannot disturb the second, since they can be arbitrarily far apart. Since a measurement on one qubit allows the prediction with certainty of the same measurement on the other qubit, it follows that there must be an “element of physical reality” that determines all three measurement outcomes ($X$, $Y$

and $Z$) on each qubit. Since the quantum mechanical state does not and cannot describe these elements of reality, it is not a complete description.

EPR concluded that "While we have thus shown that the [quantum state] does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible."

## 2.2 The GHZ paradox

This belief was definitively shown to be in conflict with the physical predictions of quantum mechanics by Greenberger, Horne and Zeilinger [18] (GHZ). They presented an argument that was a particularly compelling version of what is now known as Bell's theorem [4]. For an excellent discussion see [25].

Consider three qubits prepared in the "GHZ state"

$$S = \{I^{\otimes 3}, X_1X_2X_3, Z_1Z_2, Z_2Z_3, Z_1Z_3, -Y_1Y_2X_3, -X_1Y_2Y_3, -Y_1X_2Y_3\} \tag{2.1}$$

and taken to three separate locations. Since by carrying out $X$ measurements on the second and third qubit we can predict the outcome of an $X$ measurement on the first (we know their product will be 1), according to EPR there is an element of physical reality corresponding to $X_1$. We can make similar arguments for $X_2$, $X_3$, $Y_1$, $Y_2$ and $Y_3$. For some particular instance of the GHZ state, let $v(X_1)$ be the outcome of an $X_1$ measurement which, according to EPR, must be pre-determined (even if the measurement is not made). Similarly $v(X_2)$ denotes the outcome of an $X_2$ measurement, and so on. Note that the $v(\cdot)$ are just ordinary numbers equal to $+1$ or $-1$.

To agree with the quantum mechanical predictions the values must satisfy

$$v(X_1)v(X_2)v(X_3) = 1, \qquad v(Y_1)v(Y_2)v(X_3) = -1, \tag{2.2}$$

$$v(X_1)v(Y_2)v(Y_3) = -1, \qquad v(Y_1)v(X_2)v(Y_3) = -1. \tag{2.3}$$

Since $v(X_1)v(X_1) = (\pm 1)^2 = 1$ and so on, multiplying all four equations together gives $1 = -1$. Therefore the predictions of quantum mechanics together with the requirement that all the measurement outcomes have pre-existing values independent of what was measured at the distant sites leads to a contradiction. (That requirement was in turn derived from EPR's assumptions.)

Bell's theorem uses similar arguments (in general of a more statistical nature than the particularly direct ones for the GHZ case) to derive restrictions on how measurement results for distant systems should be correlated, known as Bell inequalities. The fact that the predictions of quantum mechanics do not obey these inequalities is known as "quantum non-locality".

Note that since, for example $X_1$ is not in $S$, either outcome for an $X_1$ measurement is equally likely, and this is true regardless of what measurements are made on qubits 2 and 3. Hence the "non-locality" cannot be used to send messages from one site to another. This is always the case in quantum mechanics, and is known as the no-signalling principle.

## 2.3 Local Hidden Variable models

Exactly what set of assumptions are strictly necessary to derive Bell's inequalities is still a subject of some debate (see [6] and [23] for a recent example), and I do not seek to take a position on that debate here. What is not controversial is that Bell's inequalities must be obeyed by local hidden variable (LHV) models.

Suppose two systems $a$ and $b$ are prepared in some quantum state and then taken to distant locations where choices of measurements $s_a$ and $s_b$ respectively are made and implemented. In an LHV model the preparation would set some hidden variable (shared by both systems) $\lambda$ according to some probability distribution $\rho(\lambda)$, and then the measurement outcomes $o_a$ and $o_b$ would be sampled from some probability distributions $p_a(o_a; \lambda, s_a)$ and $p_b(o_b; \lambda, s_b)$. Note that, crucially, $p_a$ cannot depend on $s_b$ and vice versa by the assumption of locality. In such a model the joint probability for measurement outcomes $o_a$ and $o_b$ when measurements $s_a$ and $s_b$ are chosen would be found by integrating over the hidden variables

$$p(o_a, o_b; s_a, s_b) = \int d\lambda \rho(\lambda) p_a(o_a; \lambda, s_a) p_b(o_b; \lambda, s_b). \tag{2.4}$$

Not all conceivable $p(o_a, o_b; s_a, s_b)$ can be written in this form, and the restrictions on those that can are exactly the Bell inequalities for two systems.

A LHV model for three or more systems would be similar.

# 3 Quantum computation

Here I briefly review two forms of quantum computation. Although different in many respects, they are equivalent in computational power.

## 3.1 Quantum circuits

The quantum circuit model of computation was proposed by Deutsch [14] in 1989. For a historical review see [15], for comprehensive coverage see [28].

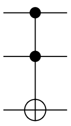
Any classical computation can be viewed as a boolean circuit, which is composed of wires and gates. A wire carries a single bit (0 or 1) of information between gates. A gate receives some bits on its input wires, computes a function of them and then outputs the resulting bits to its output wires.

Any classical computation can be written as a circuit involving just a single type of gate, the Toffili gate [28, Section 3.2.5], which has three input wires and three output wires. The function computed by this gate is shown

| Input | Output |
|---|---|
| 000 | 000 |
| 001 | 001 |
| 010 | 010 |
| 011 | 011 |
| 100 | 100 |
| 101 | 101 |
| 110 | 111 |
| 111 | 110 |

Table 3.1: Truth table for the Toffili gate.

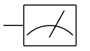
in Table 3.1. This gate is represented in a circuit diagram as



(3.1)

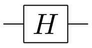
Note that the computation flows from left to right, and wires are indicated by horizontal lines. Since any computation can be carried out using this gate it is known as a "universal" gate. The Toffili gate is also reversible, unlike the Negated-AND gate (which is also universal).

The quantum circuit model is a generalization of the boolean circuit. Instead of carrying one bit of classical information, wires now carry a single qubit. Gates now represent reversible quantum transformations (unitary matrices in the Hilbert space formalism). An important difference from a classical circuit is that in order to provide useful information to the user, the output of the computation must be converted into classical bits by carrying out a $Z$ measurement on each qubit at the end. This process is represented using a meter, as



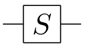
(3.2)

The Hadamard transformation $f_H$ discussed in §1.2 is written



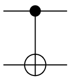
(3.3)

similarly the phase transformation $f_S$ is written



(3.4)

However the controlled-NOT transformation $f_{CNOT}$ has the special notation



(3.5)

where the small black circle indicates the control qubit.

The qubit states represented in the stabilizer formalism by $S = \langle Z \rangle$ and $S = \langle -Z \rangle$ are written as $|0\rangle$ and $|1\rangle$ respectively, and described as the "computational basis". Any classical reversible gate has an equivalent quantum gate that acts on $|0\rangle$ and $|1\rangle$ in the same was as the classical gate acts on

0 and 1. Therefore quantum circuits are at least as powerful as classical circuits.

Since quantum mechanics can be simulated on a classical computer, it is not possible for a quantum computer to solve a problem that a classical computer cannot. However, this simulation is typically very inefficient and so it is feasible that quantum computers may be able to solve problems faster than classical ones. In computer science an algorithm is usually described as efficient if the amount of resources (e.g. time) used by it is always less than some polynomial in the size of the input [28, Section 3.2.2].

Shor [35] has devised a quantum circuit for the efficient decomposition of an integer into its prime factors. No efficient classical algorithm for this problem is known, and it is suspected that such an algorithm may be impossible. This problem is of practical interest since the security of the popular public-key cyptography scheme RSA [33] relies on the assumption that factoring integers is difficult.

Any quantum circuit can be approximated using the quantum version of the Toffili gate and the Hadamard gate $H$ (defined in §1.2) [34]. These are therefore an example of a universal set of quantum gates. Another important example of a universal set is controlled-NOT together with arbitrary single qubit gates [28, Section 4.5.2]. (Single qubit gates can in turn be approximated using a finite set of gates [28, Section 4.5.3].)

The Toffili gate is not in the Clifford group (i.e. it cannot be described in the stabilizer formalism). Indeed a circuit of only Clifford group gates is no more powerful than a classical computer, since such a computation can be simulated by a classical computer using the stabilizer formalism. In fact their simulation is in a classical complexity class $\oplus\mathbf{L}$, which is related to classical circuits consisting of NOT and controlled-NOT gates only [1]. $\oplus\mathbf{L}$ is assumed to be weaker than full classical computation.

Finally we note that a circuit diagram may seem to suggest qubits travelling along wires, for example photons travelling along optical fibre with a qubits encoded in their polarization states, between gates that are fixed in place, for example optical elements. However, in other implementations the qubits may be stationery, for example they may be encoded in the energy levels of trapped ions, and the gates will be applied to them in place, for example using various laser pulses. The quantum circuit model is not concerned with these details, it is abstract enough that both of these physical

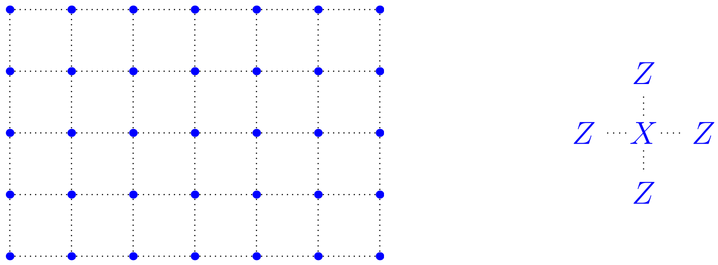


Figure 3.1: A cluster state of $n = 7 \times 5$ qubits (each represented by a blue dot) is depicted on the left. The dashed lines indicates which qubits are considered neighbours. A typical stabilizer generator is shown on the right.

implementations are equally valid.

## 3.2 Measurement-based computation

The measurement-based model of quantum computation considered here is known as the one-way quantum computer or cluster state computer. It was proposed by Raussendorf and Briegel [31] in 2001. A more detailed account can be found in [32], and a very insightful review has been produced by Nielsen [27].

The model uses a class of qubit stabilizer states known as cluster states. $n$ qubits are arranged on a 2-dimensional rectangular grid. The stabilizer subgroup $S$ is generated by $n$ elements of the form $X_k \prod_k' Z_k'$, one for each qubit $k$. The $k'$ label qubit sites that neighbour $k$ in the grid, see Figure 3.1. The ability to perform arbitrary single-qubit adaptive measurements on a sufficiently large cluster state allows the efficient simulation of any quantum circuit [31]. On the other hand, the preparation and measurement of a cluster state can be efficiently carried out as a quantum circuit. Therefore the power of the two computational models is identical.

Note that the measurements must be adaptive, which is to say that the choice of measurement sometimes depends on the outcomes of previous measurement. A classical computer, known as the "control computer" must operate during the computation to keep track of the outcomes and make the appropriate measurement choices. Interestingly, this computer need not be a full classical computer, in fact since it only needs to keep track of the parity of various sequences of measurements a so-called "parity computer" (only capable of solving $\oplus\mathbf{L}$ problems) will suffice [2].

It was stated above that arbitrary single qubit measurements on certain stabilizer states permit universal quantum computation. What if measurements are restricted to Pauli observables? In that case the computation can be efficiently simulated by a classical computer using the stabilizer formalism. But universal classical computation is still possible by using the GHZ paradox to implement an AND gate [2]. To do this associate 0 with a $Y$ measurement and 1 with an $X$ measurement. For any two bits $a$ and $b$ carry out the $a$ measurement on the first qubit of a GHZ state, the $b$ measurement on the second and the $a \oplus b \oplus 1$ measurement on the third. From (2.1) we see that the product of the measurement outcomes will be 1 if and only if $a = b = 1$.

This result should be compared to the one mentioned in the previous section, that Clifford gates in the circuit model can only achieve $\oplus\mathbf{L}$ computations. Loosely speaking, the ability to choose measurements "invokes" the quantum non-locality and provides an improvement in computational power. I will later show that in a well-defined sense a Clifford circuit does not invoke non-locality.

The relationship between the computational power of measurement-based computation and quantum non-locality has been studied in [30]. This work makes use of the observation that any function of a single bit is linear, i.e. for any $f$ we can write $f(x) = f(0) + (f(1) - f(0))x$. In an LHV model each site being measured can only compute some function of the information it receives through the measurement choice. Therefore it can be shown that a function computed using a parity computer that can choose between two measurements at each site of some correlated resource can only be non-linear (i.e. not computable by the parity computer alone) if the measurement results violate a Bell inequality.

## 3.3 Heuristic argument for a link with non-locality

There would certainly not be a quantum computational speed-up over classical computers if quantum mechanics could be efficiently simulated by classical computers. Consider a physical theory compatible with an underlying classical hidden variables of the type that EPR sought for quantum mechanics. Then it is plausible to argue computation within such a theory can always be efficiently simulated on a classical computer.

The simulation would begin by sampling from an appropriate distribution of hidden variable states and would then proceed to update that hidden state according to the rules of the model. The exponential size of the quantum state vector (which uses $2^n$ complex numbers to describe $n$ qubits) would then be analogous to the exponential size of a classical probability distribution (which uses $2^n$ real numbers to specify a probability distribution over $n$ bits) which does not pose any difficulties to the simulation of classical probabilistic computation.

This argument is by no means watertight - for example there may be an infinite number of hidden variable states (see [19] for a proof that this would hold for quantum theory). That objection can perhaps be overcome by restricting attention to computations that are fault-tolerant [28, Section 10.6] in the theory, but others may remain.

On the other hand, this argument can be made rigorous for theories that live within in a certain operational framework, see [3, Theorem 15].

# 4 Spekkens' toy theory

A toy theory of restricted knowledge about local classical bits was introduced by Spekkens in [36] in order to argue that quantum theory is also a theory about restricted knowledge. I will refer to the definition provided in that paper as the "old" definition. The theory bears a striking similarity to the stabilizer formalism for qubits reviewed above. In this chapter I review the toy theory and define a notation for it that makes these similarities self-evident. I will refer to this as "pseudo-stabilizer" notation. This notation rests upon a more mathematically transparent (but presumed equivalent) definition of the theory that Spekkens has developed since the publication of [36]. This definition will be reviewed and compared with the original below, and can also be found in [37] (and partly in [38]). I will refer to it as the "new" definition.

## 4.1 States

The toy theory makes a distinction between ontic states, which are states of reality, and epistemic states, which are states of knowledge about that reality.

An elementary system in the theory is always in one of four possible ontic states. In the old definition the states are denoted $\{1, 2, 3, 4\}$. In the new definition an ontic state is written as two bits $(q, p)$, and so the four possible states are written $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$.

The ontic state of a composite system is simply a list of the ontic states of the elementary systems it is built from. For example the possible ontic states of two systems are $(a, b)$ where $a$ and $b$ are each one of the four possible ontic states for an elementary system. When the first system is dealt with in isolation, it behaves exactly as an elementary system in the state $a$, and similarly for the second system. This simple structure means that there is nothing resembling quantum entanglement at the ontic level.

In the new definition the ontic state space of a composite system composed of $n$ elementary systems, denoted $\Omega_n$, is identified with the vector space of length $2n$ bit-strings, denoted $(\mathbb{Z}_2)^{2n}$. Addition of two vectors in this vector space is component-by-component addition modulo 2. Hence the ontic state of $n$ elementary systems is written $(q_1, \ldots, q_n, p_1, \ldots, p_n)$. (In [37] it is written $(q_1, p_1, \ldots, q_n, p_n)$ but the convention adopted here makes comparison with stabilizer formalism slightly easier.)

The allowed epistemic states are slightly more complicated. In the old definition they are constrained by a "knowledge balance principle", which I quote from [36] in full rather than attempting to paraphrase.

> "If one has maximal knowledge, then for every system, at every time, the amount of knowledge one possesses about the ontic state of the system at that time must equal the amount of knowledge one lacks."

For an elementary system this results in the allowed epistemic states consisting of pure states where one knows that the ontic state is one of two possibilities (for example $1 \vee 2$, where $\vee$ means "or"), and a completely mixed state where one knows nothing about the ontic state. A simple characterization of the allowed epistemic states of composite systems is not provided by the old definition.

The new definition speaks of "canonical variables", which are elements of the dual space to the space of ontic states $\Omega_n$, denoted $\Omega_n^*$. The dual space is the space of functions from ontic states, $\Omega_n$, to a single bit, $\{0, 1\}$. We denote the dual basis $Q_1, Q_2, \ldots, Q_n, P_1, P_2, \ldots, P_n$. Being the dual basis simply means if we have an ontic state $m = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ then $Q_i(m) = q_i$ and $P_i(m) = p_i$. Any canonical variable $F$ in $\Omega_n^*$ may be written $F = a_1Q_1 + a_2Q_2 + \cdots + a_nQ_n + b_1Q_1 + b_2Q_2 + \cdots + b_nP_n$ for some bit-string $(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n)$ in $(\mathbb{Z}_2)^{2n}$.

To enforce the knowledge balance principle we define a "Poisson bracket" of two canonical variables $F = a_1q_1 + \cdots + a_nq_n + b_1p_1 + \cdots + b_np_n$ and

$G = a'_1 Q_1 + \cdots + a'_n Q_n + b'_1 P_1 + \cdots + b'_n P_n$ as

$$\{F, G\} = \begin{pmatrix} a_1 & \dots & a_n & b_1 & \dots & b_n \end{pmatrix} J_n \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \\ b'_1 \\ \vdots \\ b'_n \end{pmatrix}, \tag{4.1}$$

where addition modulo 2 is implied (so $\{F, G\}$ is always 0 or 1) and we recall from (1.4) that

$$J_n = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}. \tag{4.2}$$

If two canonical variables $F$ and $G$ satisfy $\{F, G\} = 0$ then we say $F$ and $G$ are jointly-knowable, just as in quantum mechanics we say two Hermitian observables $A, B$ satisfying the commutation relation $[A, B] = 0$ are jointly-measurable. If we know the value of two jointly-knowable variables $F$ and $G$ then adding the values gives us knowledge of $F + G$. This never causes an inconsistency because $\{F + G, F\} = \{F, F\} + \{G, F\} = 0$, and so $F + G$ is jointly-knowable with $F$ (and similarly with $G$).

Notice that for any ontic state we certainly know the value of the trivial variable 0 (it is always the bit 0). This is analogous to knowing that a measurement of $I$ will always return 1 for any quantum state.

The allowed epistemic states in the new definition are then knowledge of the exact value of some subspace of pairwise jointly-knowable variables, and complete ignorance of (i.e. a uniform probability distribution for) the remaining variables. For an elementary system this means we can know the value of at most one of the non-trivial variables, which are $Q$, $P$ and $Q + P$.

The "pseudo-stabilizer" notation combines a variable with it's known value. We begin by considering an elementary system. We write $\mathcal{X}$ to denote the knowledge that the ontic state $m$ satisfies $Q(m) = 0$, and we write $-\mathcal{X}$ to denote the knowledge $Q(m) = 1$. Similarly $\mathcal{Z}$ denotes $P(m) = 0$, $-\mathcal{Z}$ denotes $P(m) = 1$, $\mathcal{Y}$ denotes $(P + Q)(m) = 0$ and finally $-\mathcal{Y}$ denotes $(P + Q)(m) = 1$. We will also write $\mathcal{I}$ to denote the trivial knowledge that $0(m) = 0$, and it will also be useful to write $-\mathcal{I}$ to denote the clearly incorrect "knowledge" that $0(m) = 1$.

We can turn this into a group, denoted $G_1$, by associating[1]

$$\mathcal{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ \mathcal{X} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{4.3}$$

$$\mathcal{Y} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ \mathcal{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{4.4}$$

Minus signs negate the matrices in the ordinary way, so for example

$$-\mathcal{X} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

We can read which of the ontic states $\{(0,0),(1,0),(0,1),(1,1)\}$ are consistent with the represented knowledge by the location of 1s on the diagonal.

The multiplication table for $G_1$ is shown in Table 4.1. But what does multiplication of two elements of $G_1$ correspond to? It is easy to check that if $g$ represents knowledge that $F(m) = a$ and $h$ represents knowledge that $G(m) = b$, then $gh$ represents knowledge that $(F+G)(m) = a \oplus b$. Notice that this knowledge can be "derived" from the knowledge $g$ and $h$, so multiplication represents the deriving of new knowledge from existing knowledge. If we attempt to combine the knowledge $\mathcal{X}$ and $-\mathcal{X}$ then we get $-\mathcal{X}\mathcal{X} = -\mathcal{I}$ (knowledge that $0(m) = 1$) showing that these two pieces of knowledge are inconsistent.

For composite systems we use the group $G_n$. The elements of $G_n$ are

[1] This representation of the group was devised by Terry Rudolph.

|  | $\mathcal{I}$ | $\mathcal{X}$ | $\mathcal{Y}$ | $\mathcal{Z}$ | $-\mathcal{I}$ | $-\mathcal{X}$ | $-\mathcal{Y}$ | $-\mathcal{Z}$ |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{I}$ | $\mathcal{I}$ | $\mathcal{X}$ | $\mathcal{Y}$ | $\mathcal{Z}$ | $-\mathcal{I}$ | $-\mathcal{X}$ | $-\mathcal{Y}$ | $-\mathcal{Z}$ |
| $\mathcal{X}$ | $\mathcal{X}$ | $\mathcal{I}$ | $\mathcal{Z}$ | $\mathcal{Y}$ | $-\mathcal{X}$ | $-\mathcal{I}$ | $-\mathcal{Z}$ | $-\mathcal{Y}$ |
| $\mathcal{Y}$ | $\mathcal{Y}$ | $\mathcal{Z}$ | $\mathcal{I}$ | $\mathcal{X}$ | $-\mathcal{Y}$ | $-\mathcal{Z}$ | $-\mathcal{I}$ | $-\mathcal{X}$ |
| $\mathcal{Z}$ | $\mathcal{Z}$ | $\mathcal{Y}$ | $\mathcal{X}$ | $\mathcal{I}$ | $-\mathcal{Z}$ | $-\mathcal{Y}$ | $-\mathcal{X}$ | $-\mathcal{I}$ |
| $-\mathcal{I}$ | $-\mathcal{I}$ | $-\mathcal{X}$ | $-\mathcal{Y}$ | $\mathcal{Z}$ | $\mathcal{I}$ | $\mathcal{X}$ | $\mathcal{Y}$ | $\mathcal{Z}$ |
| $-\mathcal{X}$ | $-\mathcal{X}$ | $-\mathcal{I}$ | $-\mathcal{Z}$ | $\mathcal{Y}$ | $\mathcal{X}$ | $\mathcal{I}$ | $\mathcal{Z}$ | $\mathcal{Y}$ |
| $-\mathcal{Y}$ | $-\mathcal{Y}$ | $-\mathcal{Z}$ | $-\mathcal{I}$ | $\mathcal{X}$ | $\mathcal{Y}$ | $\mathcal{Z}$ | $\mathcal{I}$ | $\mathcal{X}$ |
| $-\mathcal{Z}$ | $-\mathcal{Z}$ | $-\mathcal{Y}$ | $-\mathcal{X}$ | $\mathcal{I}$ | $\mathcal{Z}$ | $\mathcal{Y}$ | $\mathcal{X}$ | $\mathcal{I}$ |

Table 4.1: Multiplication table for $G_1$.

$4^n \times 4^n$ diagonal matrices of the form

$$\pm\, g_1 \otimes g_2 \otimes \cdots \otimes g_n \tag{4.6}$$

where the $g_k$ are chosen from $\{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$. The group is generated by $-\mathcal{I}^{\otimes n}$, $\mathcal{X}_1, \ldots, \mathcal{X}_k$ and $\mathcal{Z}_1, \ldots, \mathcal{Z}_k$. The shorthand $\mathcal{X}_k$ means $\mathcal{I}^{\otimes k-1} \otimes \mathcal{X} \otimes \mathcal{I}^{\otimes n-k}$, i.e. knowledge that $Q_k(m) = 0$. Similarly $\mathcal{Z}_k$ means $\mathcal{I}^{\otimes k-1} \otimes \mathcal{Z} \otimes \mathcal{I}^{\otimes n-k}$.

Recall that multiplication of two elements gives another piece of knowledge that can be derived from them. Therefore if we have the knowledge represented by some list of elements of $G_n$, then we have the knowledge of the entire subgroup generated by those elements. Therefore we can associate epistemic states with subgroups of $G_n$. Which subgroups represent allowed knowledge?

Certainly we must not have $-\mathcal{I}^{\otimes n}$ in our subgroup, since that represents knowledge that $0(m) = 1$. This requirement ensures that the represented knowledge is consistent, but we must also respect the knowledge balance principle.

Our notation suggests an natural function $m$ from $G_n$ to the Pauli group $P_n$. For example $m(\mathcal{ZIZ}) = ZIZ$, $m(-\mathcal{XY}) = -XY$ etc. For $g$ in $G_n$ we can define a check vector using $m$, as $r(g) = r(m(g))$. If $g$ represents knowledge about the canonical variable $F$ then $r(g)$ is exactly $F$ written as a vector $(a_1, \ldots, a_n, b_1, \ldots, b_n)$.

Recall from §1.1.1 that two elements of the Pauli group $g$ and $h$ commute if their check vectors $r(g)$ and $r(h)$ satisfy $r(g)^T J_n r(h) = 0$. Comparing this

| Picture from [36] | Old | New | Pseudo-stabilizer |
|---|---|---|---|
| 🟦⬜🟦⬜ | $1 \vee 3$ | $Q = 0$ | $\langle \mathcal{X} \rangle$ |
| ⬜🟦⬜🟦 | $2 \vee 4$ | $Q = 1$ | $\langle -\mathcal{X} \rangle$ |
| 🟦🟦⬜⬜ | $1 \vee 2$ | $P = 0$ | $\langle \mathcal{Z} \rangle$ |
| ⬜⬜🟦🟦 | $3 \vee 4$ | $P = 1$ | $\langle -\mathcal{Z} \rangle$ |
| 🟦⬜⬜🟦 | $1 \vee 4$ | $Q + P = 0$ | $\langle \mathcal{Y} \rangle$ |
| ⬜🟦🟦⬜ | $2 \vee 3$ | $Q + P = 1$ | $\langle -\mathcal{Y} \rangle$ |
| 🟦🟦🟦🟦 | $1 \vee 2 \vee 3 \vee 4$ | No knowledge | $\langle \rangle$ |

Table 4.2: The allowed epistemic states for an elementary system in three notations. The pictures in the first column show the allowed ontic states in blue. In the third column $F = a$ is shorthand for "knowledge that the ontic state $m$ satisfies $F(m) = a$".

with (4.1) we see that two elements of $g$ and $h$ of $G_n$ are jointly-knowable if and only if $m(g)$ and $m(h)$ commute. Indeed it is shown in Appendix A.2 that the lists of independent generators for valid states are the same for $G_n$ as $P_n$.

All the epistemic states for an elementary system are shown in Table 4.2. Recall that $\langle \mathcal{X} \rangle$ means the subgroup generated by $\mathcal{X}$, namely $\{\mathcal{I}, \mathcal{X}\}$, and we use the convention that $\langle \rangle = \{\mathcal{I}\}$. A few of the epistemic states for composite systems are shown in Table 4.3.

The pseudo-stabilizer notation allows the direct translation of some proofs about qubit stabilizers into proofs about the toy theory. One simply verifies the properties of the Pauli group $P_n$ used in the proof also apply to $G_n$. For example, the proof in §1.1.2 that a stabilizer subgroup of size $2^n$ corresponds to a pure state also applies in the pseudo-stabilizer case. Furthermore, the proof in [1, Proposition 1] that there are

$$2^n \prod_{k=0}^{n-1} (2^{n-k} + 1) \tag{4.7}$$

pure states on $n$ qubits also applies to the pure states of $n$ elementary systems in the toy theory.

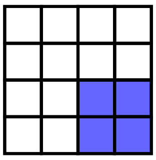
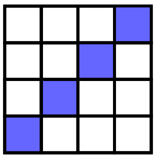
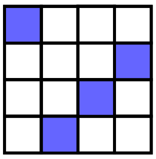
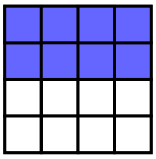
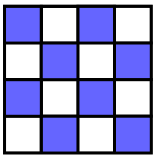
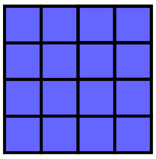
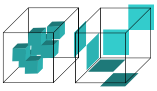
| Picture from [36] | Old | New | Pseudo-stabilizer |
|---|---|---|---|
|  | $(1\cdot3)\vee(1\cdot4)\vee(2\cdot3)\vee(2\cdot4)$ | $P_1=0,\ P_2=1$ | $\langle\mathcal{Z}_1,-\mathcal{Z}_2\rangle$ |
|  | $(1\cdot1)\vee(2\cdot2)\vee(3\cdot3)\vee(4\cdot4)$ | $Q_1+Q_2=0$,<br>$P_1+P_2=0$ | $\langle\mathcal{Z}_1\mathcal{Z}_2,\mathcal{X}_1\mathcal{X}_2\rangle$ |
|  | $(1\cdot2)\vee(2\cdot3)\vee(3\cdot4)\vee(4\cdot1)$ | $P_1+Q_2+P_2=1$,<br>$Q_1+Q_2=1$ | $\langle-\mathcal{Z}_1\mathcal{Y}_2,-\mathcal{X}_1\mathcal{X}_2\rangle$ |
|  | $(3\vee4)\cdot(1\vee2\vee3\vee4)$ | $P_1=1$ | $\langle-\mathcal{Z}_1\rangle$ |
|  | $[(1\vee3)\cdot(2\vee4)]\vee[(2\vee4)\cdot(1\vee3)]$ | $Q_1+Q_2=1$ | $\langle-\mathcal{X}_1\mathcal{X}_2\rangle$ |
|  | $(1\vee2\vee3\vee4)\cdot(1\vee2\vee3\vee4)$ | No knowledge | $\langle\rangle$ |
|  | $(1\cdot1\cdot1)\vee(1\cdot2\cdot2)\vee(2\cdot1\cdot2)\vee(2\cdot2\cdot1)\vee(3\cdot3\cdot3)\vee(3\cdot4\cdot4)\vee(4\cdot3\cdot4)\vee(4\cdot4\cdot3)$ | $Q_1+Q_2+Q_3=0$,<br>$P_1+P_2=0$,<br>$P_2+P_3=0$ | $\langle\mathcal{X}_1\mathcal{X}_2\mathcal{X}_3,\mathcal{Z}_1\mathcal{Z}_1,\mathcal{Z}_2\mathcal{Z}_3\rangle$ |

Table 4.3: Some epistemic states for composite systems. They are each composed of two elementary systems, except for the last which is composed of three. Refer to equations 52 and 124 in [36] to interpret the pictures.

## 4.2 Transformations

Only reversible transformations are considered here. In the old definition transformations are permutations of the ontic states which take any allowed epistemic state to an allowed epistemic state. In the case of transformations on an elementary system this is simply all $4! = 24$ possible permutations of the four ontic states, but for composite systems some permutations are not allowed as they result in states that violate the knowledge balance principle.

The new definition formalises this by defining transformations as "the group of symplectic affine transformations". The transformations of an $n$-system ontic state $m$ in $\Omega_n$ are those that can be written $m \to Sm + a$ for some $a$ in $\Omega_n$ and (necessarily invertible) $2n \times 2n$ matrix $S$ satisfying

$$S^T J_n S = J_n, \tag{4.8}$$

where as usual addition modulo 2 is implied. Such $S$ are known as a symplectic matrices. Note that in the qubit stabilizer case, Clifford group gates can also be associated with symplectic matrices [12].

Suppose that before the transformation we have knowledge that $F(m) = b$ for some canonical variable $F$ in $\Omega_n^*$, and bit $b$. Considering $F$ as a vector we can write this as $F^T m = b$. After the transformation $m \to Sm + a$ this becomes the knowledge $F^T(S^{-1}(m - a)) = b$, which can be re-arranged to $\tilde{F}^T m = b + \tilde{F}^T a$ where $\tilde{F} = S^{-1^T} F$. Suppose before the transformation the we have knowledge of $F$ and $G$ which therefore must satisfy $\{F, G\} = 0$. After the transformation we have knowledge of $\tilde{F}$ and $\tilde{G}$, and

$$\{\tilde{F}, \tilde{G}\} = \{S^{-1^T} F, S^{-1^T} G\} = F^T S^{-1} J_n S^{-1^T} G = F^T J_n G = \{F, G\} \tag{4.9}$$

where we have inverted both sides of $S^T J_n S = J_n$ and used $J_n^{-1} = J_n$ to find that $S^{-1} J_n S^{-1^T} = J_n$. In short, the condition (4.8) ensures that any valid knowledge before the transformation is still valid knowledge afterwards.

We can update a pseudo-stabilizer $g$ by changing the letters (the $g_k$ in (4.6)) such that the check vector $r(g)$ becomes $S^{-1^T} r(g)$ and flipping the sign if $r(g)^T S^{-1} a = 1$. It is easy to check that this is a group homomorphism that maps $-\mathcal{I}^n$ to $-\mathcal{I}^n$, and so much like the qubit stabilizer case we can specify it by its action on the remaining generators $\mathcal{X}_k$ and $\mathcal{Z}_k$.

For an elementary system the valid matrices $S$ are

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, S_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, S_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.10}$$

$$S_4 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, S_5 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, S_6 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{4.11}$$

These 6 options for $S$ combine with the four options for $a \in \Omega_1$ to give 24 transformations. These are exactly the 4! permutations of the ontic states, as shown in Table 4.4. Two of the transformations for a pair of elementary systems are shown in Table 4.5.

## 4.3 Measurements

In the old definition maximally informative measurements are specified by a partitioning of the ontic state space into pure epistemic states. The measurement outcome is determined solely by which part of the partition the ontic state was in. For example, a measurement can be carried out to determine whether the ontic state was (a) 1 or 2 or whether it was (b) 3 or 4. After the measurement the new epistemic state is the relevant part of the partition, for example $1 \vee 2$, a random disturbance of the ontic state ensuring that no further knowledge is available. There are also less informative measurements that partition the ontic state space into mixed epistemic states. The new epistemic state in this case is not uniquely defined by the knowledge balance principle, but a natural choice is the one with the highest possible classical fidelity with the previous ontic state (see [36]). I only consider that choice here.

In the new definition measurements consist of the determination of a jointly-knowable set of canonical variables. The highest classical fidelity update rule becomes the assumption that the values of any further jointly-knowable canonical variables are undisturbed by the measurement. In other words we assume that the disturbance on measurement is minimal subject to the knowledge balance principle.

The update rule considered here has the useful property that a maximally informative measurement, which seeks the value of $n$ jointly-knowable

| Old | New | Pseudo-stabilizer |
|---|---|---|
| (1)(2)(3)(4) | $S = S_1$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{X}$, $\mathcal{Z} \to \mathcal{Z}$ |
| (1)(2)(43) | $S = S_4$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{Y}$, $\mathcal{Z} \to \mathcal{Z}$ |
| (1)(32)(4) | $S = S_2$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{Z}$, $\mathcal{Z} \to \mathcal{X}$ |
| (1)(342) | $S = S_5$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{Y}$, $\mathcal{Z} \to \mathcal{X}$ |
| (1)(432) | $S = S_3$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{Z}$, $\mathcal{Z} \to \mathcal{Y}$ |
| (1)(42)(3) | $S = S_6$, $a = (0,0)$ | $\mathcal{X} \to \mathcal{X}$, $\mathcal{Z} \to \mathcal{Y}$ |
| (21)(3)(4) | $S = S_4$, $a = (1,0)$ | $\mathcal{X} \to -\mathcal{Y}$, $\mathcal{Z} \to \mathcal{Z}$ |
| (21)(43) | $S = S_1$, $a = (1,0)$ | $\mathcal{X} \to -\mathcal{X}$, $\mathcal{Z} \to \mathcal{Z}$ |
| (231)(4) | $S = S_3$, $a = (1,0)$ | $\mathcal{X} \to \mathcal{Z}$, $\mathcal{Z} \to -\mathcal{Y}$ |
| (2341) | $S = S_6$, $a = (1,0)$ | $\mathcal{X} \to -\mathcal{X}$, $\mathcal{Z} \to -\mathcal{Y}$ |
| (2431) | $S = S_2$, $a = (1,0)$ | $\mathcal{X} \to \mathcal{Z}$, $\mathcal{Z} \to -\mathcal{X}$ |
| (241)(3) | $S = S_5$, $a = (1,0)$ | $\mathcal{X} \to -\mathcal{Y}$, $\mathcal{Z} \to -\mathcal{X}$ |
| (321)(4) | $S = S_5$, $a = (0,1)$ | $\mathcal{X} \to -\mathcal{Y}$, $\mathcal{Z} \to \mathcal{X}$ |
| (3421) | $S = S_2$, $a = (0,1)$ | $\mathcal{X} \to -\mathcal{Z}$, $\mathcal{Z} \to \mathcal{X}$ |
| (31)(2)(4) | $S = S_6$, $a = (0,1)$ | $\mathcal{X} \to \mathcal{X}$, $\mathcal{Z} \to -\mathcal{Y}$ |
| (341)(2) | $S = S_3$, $a = (0,1)$ | $\mathcal{X} \to -\mathcal{Z}$, $\mathcal{Z} \to -\mathcal{Y}$ |
| (31)(42) | $S = S_1$, $a = (0,1)$ | $\mathcal{X} \to \mathcal{X}$, $\mathcal{Z} \to -\mathcal{Z}$ |
| (3241) | $S = S_4$, $a = (0,1)$ | $\mathcal{X} \to -\mathcal{Y}$, $\mathcal{Z} \to -\mathcal{Z}$ |
| (4321) | $S = S_6$, $a = (1,1)$ | $\mathcal{X} \to -\mathcal{X}$, $\mathcal{Z} \to \mathcal{Y}$ |
| (421)(3) | $S = S_3$, $a = (1,1)$ | $\mathcal{X} \to -\mathcal{Z}$, $\mathcal{Z} \to \mathcal{Y}$ |
| (431)(2) | $S = S_5$, $a = (1,1)$ | $\mathcal{X} \to \mathcal{Y}$, $\mathcal{Z} \to -\mathcal{X}$ |
| (41)(2)(3) | $S = S_2$, $a = (1,1)$ | $\mathcal{X} \to -\mathcal{Z}$, $\mathcal{Z} \to -\mathcal{X}$ |
| (4231) | $S = S_4$, $a = (1,1)$ | $\mathcal{X} \to \mathcal{Y}$, $\mathcal{Z} \to -\mathcal{Z}$ |
| (41)(32) | $S = S_1$, $a = (1,1)$ | $\mathcal{X} \to -\mathcal{X}$, $\mathcal{Z} \to -\mathcal{Z}$ |

Table 4.4: The reversible transformations for an elementary system in three notations. The first shows the permutations to the ontic states in cycle notation (defined in [36]). The second shows the matrix $S$ and state $a$. The third shows the action on the non-trivial generators of $G_1$.

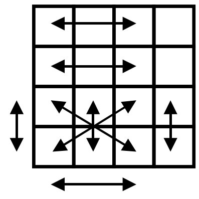
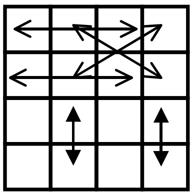
| Old | New | Pseudo-stabilizer |
|---|---|---|
|  | $S = \begin{pmatrix}1&0&1&0\\0&1&0&0\\0&0&1&0\\0&1&0&1\end{pmatrix},$<br>$a = (1,0,0,1)$ | $\mathcal{X}_1 \to -\mathcal{Y}_1$, $\mathcal{X}_2 \to \mathcal{X}_2$,<br>$\mathcal{Z}_1 \to \mathcal{Z}_1$, $\mathcal{Z}_2 \to -\mathcal{Y}_2$ |
|  | $S = \begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&1&1\end{pmatrix},$<br>$a = (0,0,0,0)$ | $\mathcal{X}_1 \to \mathcal{X}_1\mathcal{X}_2$, $\mathcal{X}_2 \to \mathcal{X}_2$,<br>$\mathcal{Z}_1 \to \mathcal{Z}_1$, $\mathcal{Z}_2 \to \mathcal{Z}_1\mathcal{Z}_2$ |

Table 4.5: Two reversible transformations for pairs of elementary systems in three notations. The first shows the permutations to the ontic states (see [36], from which the images are taken). The second shows the matrix $S$ and state $a$. The third shows the action on the non-trivial generators of $G_2$. The first transformation is separable: in the basis where the states are written $(q_1, p_1, q_2, p_2)$, $S$ is block-diagonal. The second transformation is analogous to a controlled-NOT (compare with the qubit case in [12]).

variables and therefore has $2^n$ outcomes, can be considered as a series of measurements with two outcomes. For example a measurement of $Q_1 + Q_2$ and $P_1 + P_2$ simultaneously (analogous to a Bell basis measurement) is the same a measurement of $Q_1 + Q_2$ and then a measurement of $P_1 + P_2$ since we have assumed that the former will not disturb the value of the latter.

Measurements in the pseudo-stabilizer notation can therefore by considered as a one or more "Pauli measurements" (i.e. measurements of a single canonical variable) which follow exactly the same procedure as measurements in the qubit stabilizer formalism.

## 4.4 Computation in the toy theory

As observed in [36] the toy theory can be efficiently simulated on a classical computer simply by sampling from the set of ontic states compatible with a prepared epistemic states and then applying the transformations and measurements to that ontic state. The epistemic state can also be computed efficiently by tracking the pseudo-stabilizers. We can still ask if the theory even permits universal classical computation. As reviewed in Chapter 3, in the case of qubit stabilizers the computational power depends on whether a circuit or measurement-based model is used, and so we consider each case separately here.

### 4.4.1 Circuit computation

Consider preparing a set of elementary systems in $\mathcal{Z}$ and/or $-\mathcal{Z}$ states, applying some sequence of allowed transformations and then doing a $\mathcal{Z}$ measurement on each system. This can be considered a computational circuit in the toy theory. From §4.2 it is clear that for a fixed circuit we can efficiently simulate the computation by application of matrices and addition of vectors modulo 2. On the other hand a circuit of NOT and controlled-NOT gates can be simulated with a toy theory circuit using a gate that sends $\mathcal{Z} \rightarrow -\mathcal{Z}$ for NOTs and a gate that sends $\mathcal{Z}_1 \rightarrow \mathcal{Z}_1$, $\mathcal{Z}_2 \rightarrow \mathcal{Z}_1\mathcal{Z}_2$ for controlled-NOTs. Therefore circuit computation in the toy theory is equivalent to a "parity computer", or more formally is complete for the complexity class $\oplus\mathbf{L}$ (for the definition see [1] where the same statement is proven for qubit stabilizer circuits).

### 4.4.2 Measurement-based computation

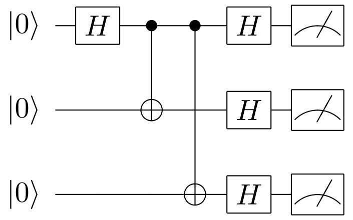
In §3.2 we say that a parity computer can do full classical computation if it is able to perform Pauli measurements on certain stabilizer states. Can measurements of some state in the toy theory similarly promote parity computations to universal classical computation?

The toy theory, by construction, has an LHV model (the ontic states). Therefore the result of [30] would suggest that the answer to the above question is no. However, that result assumes there are only two choices of measurement, whereas in the toy model we have a choice of three measurement (four including the trivial $\mathcal{I}$ measurement).

Suppose the measurement choice is made at each site by using two bits $(i_1, i_2)$, say with $(0, 0)$ corresponding to $\mathcal{I}$, $(1, 0)$ to $\mathcal{X}$, $(0, 1)$ to $\mathcal{Z}$ and $(1, 1)$ to $\mathcal{Y}$. Then for each of the four ontic states $(p, q)$ the measurement outcome bit is a linear function $pi_1 + qi_2$ of the input bits. Since linear functions of linear functions are linear, the overall function computed by such measurements controlled by a parity computer is also linear (this is the same argument used in [30]). It is easy to check that any one-to-one correspondence between two bits is automatically a linear function. Therefore this argument will still apply if the correspondence between input bits and measurements is made in a different way.

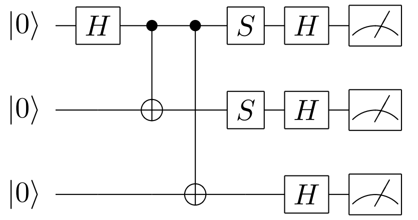
# 5 LHV models for quantum circuits

In the case of qubit stabilizers, we saw that the measurement-based model was more powerful than the circuit model, and that the power of the measurement-based model required quantum non-locality. Do qubit stabilizers in the circuit model also involve quantum non-locality? The direct answer to this question is yes, since using only Clifford group gates we can prepare the GHZ state (2.1) then implement effective $X$ measurements (of the first type discussed in §1.4) of each qubit using the circuit



(5.1)

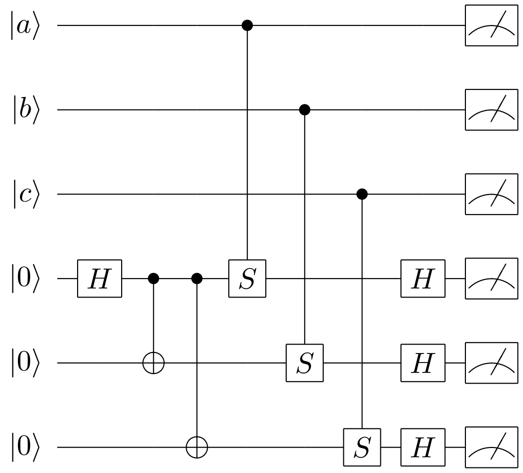
and similarly we can implement effective $-Y_1$, $-Y_2$ and $X_3$ measurements with



(5.2)

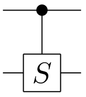
An effective $-Y_1$ measurement can be turned into an effective $Y_1$ measurement by multiplying the result by $-1$. Similar circuits can be devised for $Y_1, X_2, Y_3$ and $X_1, Y_2, Y_3$. It was shown in §2.2 that quantum non-locality is involved in these measurements.

Perhaps the more interesting question is this: do Clifford group circuits *need* to involve quantum non-locality? The demonstration of non-locality given above requires circuits four circuits to be considered together. Let us instead only consider one circuit at a time. We could combine the above four circuits into the following single circuit



(5.3)

where $a$, $b$ and $c$ are 0 or 1 representing the choice between $X$ and $Y$ measurements respectively, and



(5.4)

is a controlled-$S$ transformation that does nothing if the first qubit is in the state $|0\rangle$ and applies an $S$ transformation if the first qubit is in the state $|1\rangle$. This circuit alone involves non-locality by the arguments above. However, unlike the four circuits we started with, it is not a circuit of Clifford group gates, since controlled-$S$ transformations are not in the Clifford group (this is shown in §A.1.2).

Hence the possibility remains that any particular circuit of Clifford group gates admits an LHV model, and indeed this will be proven below. I begin by defining exactly what I mean by an LHV model for a quantum circuit.

## 5.1 Definition

Consider a quantum circuit $C$ consisting of unitary gates on $n$ qubits. We only consider the case where the input to the circuit is in the computational basis, and the output is measured in the computational basis.

A local hidden variable model for that circuit $C$ is a specification of the following.

1. For each qubit $i$, a set of local “hidden variable” states $S_i$.

2. For each qubit $i$ and $b \in \{0, 1\}$, a probability distribution over $S_i$ which representations preparation of $|b\rangle$ on that qubit. Here locality means that the distribution does not depend on what states are prepared on the other qubits.

3. For each gate $U$ in the circuit, taken to act on qubits $i_1, \dots, i_k$, and each local state input $(\lambda_{i_1}, \dots, \lambda_{i_k})$, a probability distribution over output local states which represents the change to the local state due to the gate. Here locality means that the distribution does not depend on the qubits not acted on by the gate, and furthermore that the local state of such qubits is not affected by the gate.

4. For each qubit $i$, a joint probability distribution $p(\lambda_i, b)$ on $S_i$ and $\{0, 1\}$ which gives the probability of obtaining outcome $b$ when the qubit is measured in the computational basis at the end of the circuit and its local state is $\lambda_i$. Here locality means that the measurement outcome depends on nothing but the local state of the qubit being measured.

We require that the model reproduces the quantum mechanical predictions for the joint probabilities of the measurement outcomes for every input state in $\{|0\rangle, |1\rangle\}^{\otimes n}$.

Since every aspect of the model is allowed to depend on the circuit $C$, it may be a very poor candidate for the description of physical reality. It does however show that no Bell inequality (see Chapter 2) can be violated by the circuit. Conversely, any single circuit that demonstrates the violation of a Bell inequality, for example (5.3), does not admit an LHV model.

## 5.2 “Almost classical” circuits

As a warm-up to circuits of Clifford gates consider the following simple example of circuits that admit LHV models.

Let $C$ consist of a round of single qubit gates $U^{(1)}, U^{(2)}, \dots U^{(n)}$ applied to qubits $1, \dots, n$, followed by a series of gates that are each either diagonal matrices or permutation matrices in the computational basis (this includes, for example the $X$, $Z$, $S$, controlled-NOT and Toffili gates, but not the Hadamard gate $H$). In [13] such circuits were related to probabilistic classical computation, and in particular it was shown that such circuits can be efficiently simulated classically. The algorithm for doing so can easily be converted into an LHV model for the circuits.

The model is as follows. The local state for each qubit is a classical bit, $S_i = \{0,1\}$. Preparation of $|0\rangle$ (resp. $|1\rangle$) sets the bit to 0 (resp. 1). If the input bit to one of the single qubit gates $U$ is $x$ then the output bit is $y$ with probability $|U_{yx}|^2$. The diagonal gates have no effect in the model. The action of the permutation gates is found by treating the input to each gate as a binary number and applying the permutation to give the output bits as a binary number. Measurement consists of reading out the bit.

This model is easily seen to give the same probabilities as quantum mechanics. Let $f : \{0,1\}^n \to \{0,1\}^n$ be the one-to-one function computed by the permutation gates. Suppose the input to the circuit was $|b_1\rangle \otimes \cdots \otimes |b_n\rangle$, denoted $|b_1, \ldots, b_n\rangle$. Then the measurement results will be given by $f(r_1, \ldots, r_n)$ where $(r_1, \ldots, r_n)$ is selected randomly with probability $\left|U^{(1)}_{r_1 b_1}\right|^2 \cdots \left|U^{(n)}_{r_n b_n}\right|^2$.

Meanwhile the quantum state after the single qubit round would be

$$\sum_{(r_1,\ldots,r_n)\in\{0,1\}^n} U^{(1)}_{r_1 b_1} \cdots U^{(n)}_{r_n b_n} |r_1, \ldots, r_n\rangle \tag{5.5}$$

and after the remaining gates it would be

$$\sum_{(r_1,\ldots,r_n)\in\{0,1\}^n} e^{i\theta_{r_1,\ldots,r_n}} U^{(1)}_{r_1 b_1} \cdots U^{(n)}_{r_n b_n} |f(r_1, \ldots, r_n)\rangle , \tag{5.6}$$

where the $e^{i\theta_{r_1,\ldots,r_m}}$ are phase factors due to the diagonal gates. When measured in the computational basis this gives the same probabilities as the LHV model.

This construction is easily generalized to circuits where the first rounds consists of larger gates, for example $U^{(12)}$ applied to qubits 1 and 2, $U^{(34)}$ applied to qubits 3 and 4 etc. One could go as far as an $n$-qubit unitary applied across the entire input, but note that the locality conditions are somewhat trivial in that case.

## 5.3 Clifford circuits

Consider a circuit $C$ consisting only of gates in the Clifford group defined in §1.2. Then we will show that $C$ admits an LHV model. We will further show that the LHV model can be viewed as a classical circuit with only NOT and controlled-NOT gates along with additional input bits selected uniformly at random. This perhaps gives some insight into the result of [1] (that simulation of Clifford circuits is $\oplus\mathbf{L}$-complete) but does not imply it since I make no claims about the computational complexity of finding the LHV model for a given circuit.

We can form a new circuit $C'$ where all the gates in $C$ have been individually decomposed into $H$, $S$ and (not necessarily nearest-neighbour) controlled-NOT gates. It is easy to see that an LHV model for $C'$ gives rise to an LHV model for $C$, in other words decomposing the gates can only make the locality requirements stronger. We therefore assume that $C$ has already been decomposed into this form.

I begin by discussing a model that doesn't quite work, and will then show how it can be fixed for every circuit. This LHV model for $C$ is just Spekken's toy theory. The local state of each qubit is an elementary system of the theory, so $S_i = \Omega_1 = (\mathbb{Z}_2)^2$. Preparation consists of preparing the ontic state $\mathcal{Z}$ for $|0\rangle$ and $-\mathcal{Z}$ for $|1\rangle$. $H$ gates apply the transformation $\mathcal{X} \to \mathcal{Z}$, $\mathcal{Z} \to \mathcal{X}$. $S$ gates apply $\mathcal{X} \to \mathcal{Y}$, $\mathcal{Z} \to \mathcal{Z}$. A controlled-NOT gate with control 1 and target 2 applies $\mathcal{X}_1 \to \mathcal{X}_1\mathcal{X}_2$, $\mathcal{X}_2 \to \mathcal{X}_2$, $\mathcal{Z}_1 \to \mathcal{Z}_1$, $\mathcal{Z}_2 \to \mathcal{Z}_1\mathcal{Z}_2$. The first two transformations can be found in Table 4.4, the final one in Table 4.5. Measurement is simply a $\mathcal{Z}$ measurement.

We can compare ($m$ applied to) the pseudo-stabilizer generators for this model to the stabilizer generators for the qubit circuit. After the preparation stage we can take them be identical: $\pm\mathcal{Z}_1$ and $\pm Z_1$ (where the sign depends on the input to the first qubit), $\pm\mathcal{Z}_2$ and $\pm Z_2$ etc. Gates have the same effect, up to a possible minus sign. For example, recall from (1.5) that $H$ sends the qubit stabilizer $Y \to -Y$. However, in the LHV model it sends $\mathcal{Y} = \mathcal{X}\mathcal{Z} \to \mathcal{Z}\mathcal{X} = \mathcal{Y}$. After all the gates have been applied the pseudo-stabilizer generators will therefore be identical up to signs. Let $\mathcal{S}$ denote the pseudo-stabilizer subgroup, and $S$ the qubit stabilizer subgroup, at that point.

Recall that $m : G_n \to P_n$ was defined in §4.1 as the natural function with $m(\mathcal{X}) = X$ etc. Define the "computational subgroup" of $G_n$ as $\langle -\mathcal{I}^{\otimes n}, \mathcal{Z}_1, \dots, \mathcal{Z}_k \rangle$, for example the two-qubit computational subgroup is $\{\pm\mathcal{I}^{\otimes 2}, \pm\mathcal{Z}_1, \pm\mathcal{Z}_2, \pm\mathcal{Z}_1\mathcal{Z}_2\}$. Similarly, define the "computational subgroup" of $P_n$ as $\langle -I^{\otimes n}, Z_1, \dots, Z_k \rangle$. We see that these subgroups play very nicely with $m$:

- $g$ is in the computational subgroup of $G_n$ if and only if $m(g)$ is in the computational subgroup of $P_n$, indeed $m$ gives a one-to-one correspondence between the computational subgroups; and
- If $g$ and $h$ are both in the computational subgroup of $G_n$ then $m(gh) = m(g)m(h)$ (this isn't always true for more general $g$ and $h$).

In Appendix A.3 I show that it is sufficient for an LHV model to have the correct expectation values for all observables in the computational subgroup. Both quantum mechanics and the toy theory predict an expectation values of 1 for observables in the state subgroup, $-1$ for observables whose minus is in the subgroup, and 0 for all other observables. Therefore it suffices that applying $m$ to the computational subgroup elements of $\mathcal{S}$ gives the computational subgroup elements of $S$. By the argument above this will certainly be true up to signs.

If the signs are incorrect then the model can always be fixed by flipping some of the preparation procedures, i.e. so that preparing $|0\rangle$ prepares $-\mathcal{Z}$ and $|1\rangle$ prepares $\mathcal{Z}$ for some of the qubits. This flips the signs of the generators and therefore of the computational subgroup. There is always a combination of flips that will make the signs in the computational subgroup of the pseudo-stabilizers match the signs in the computational subgroup of the qubit stabilizers. By the bullet points above it is enough to check a list of generators for each.

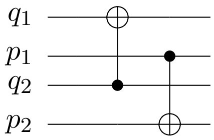
The necessary flips can be found inductively. Write the independent generators of $\mathcal{S}$ as $g_1, g_2, \dots$. We will define a list of independent generators for the computational subgroup elements of $\mathcal{S}$ as $c_1, c_2, \dots$. Start with an empty list of generators.

Assume we have considered all the generators up to $g_k$ and written an independent list of computational subgroup generators up to $c_l$ with the correct signs. If $g_{k+1}$ is computational we set $c_{l+1} = g_{k+1}$ and flip the sign of $g_{k+1}$ if necessary to ensure that $m(c_{l+1})$ is in $S$. Note that $c_{l+1}$ is independent of $c_1, \ldots, c_l$ since $g_{k+1}$ is independent of $g_1, \ldots, g_k$ and $c_1, \ldots, c_l$ are in $\langle g_1, \ldots, g_k \rangle$. If $g_{k+1}$ is not in the computational subgroup then we check if there is a $h$ in $\langle g_1, \ldots, g_k \rangle$ with $g_{k+1}h$ computational. In that case we set $c_{l+1} = g_{k+1}h$, which is independent of $g_1, \ldots, g_k$ and therefore of $c_1, \ldots, c_l$. Again we flip the sign of $g_{k+1}$ if necessary to ensure that $m(c_{l+1})$ is in $S$. (If there is another element $h'$ in $\langle g_1, \ldots, g_k \rangle$ that also has $g_{k+1}h'$ computational, this doesn't matter since $g_{k+1}h'$ is in $\langle c_1, \ldots, c_{l+1} \rangle$ because $g_{k+1}hg_{k+1}h' = hh'$ which is in $\langle c_1, \ldots, c_l \rangle$.) If there is no such $h$ then we ignore $g_{k+1}$.

Once the correct sign flips have been found for the $|0\rangle^{\otimes n}$ input, it will work for any input. This is because changing the sign of one of the input generators (e.g. $Z_k \to -Z_k$) has the same effect on the output generators (e.g. $f_C(Z_k) \to f_C(-Z_k) = -f_C(Z_k)$ where $f_C$ is the transformation implemented by the entire circuit) for both qubit stabilizers and the LHV model.

The LHV model can be thought of as a classical circuit as follows. Each qubit corresponds to two classical bits (representing $q$ and $p$). Preparation sets the value of the second bit to 0 for $|0\rangle$ and 1 for $|1\rangle$. This is immediately followed by a NOT gate if it was necessary to flip the sign of the corresponding generator when constructing the model. The first bit is set uniformly at random.

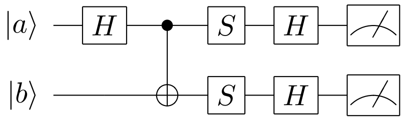
$H$ gates swap the two bits, which can be done using three consecutive controlled-NOT gates applied in alternating directions. $S$ gates apply a controlled-NOT controlled by the second bit onto the first. Quantum controlled-NOT gates controlled by the first qubit become two classical controlled-NOT gates:



(5.7)

### 5.3.1 Example

Here is an example of the entire process for constructing the LHV model. Consider the circuit



(5.8)

where $a$ and $b$ are each 0 or 1.

First we examine what the quantum mechanical predictions for this circuit. The stabilizer generators at the input stage are $(-1)^a Z_1$ and $(-1)^b Z_2$. Applying the transformations in the circuit we find

$$Z_1 \to X_1 \to X_1X_2 \to Y_1Y_2 \to Y_1Y_2, \tag{5.9}$$

$$Z_2 \to Z_2 \to Z_1Z_2 \to Z_1Z_2 \to X_1X_2, \tag{5.10}$$

and so just before the measurement the stabilizer subgroup $S$ is generated by $(-1)^a Y_1Y_2$ and $(-1)^b X_1X_2$. Since $Y_1Y_2X_1X_2 = -Z_1Z_2$ the entire subgroup is

$$S = \left\{ I^{\otimes 2}, (-1)^a Y_1Y_2, (-1)^b X_1X_2, (-1)^{1+a+b} Z_1Z_2 \right\}. \tag{5.11}$$

The only element in the computational subgroup is $(-1)^{1+a+b}Z_1Z_2$. This shows that each of the computational basis measurements at the end can give either outcome with equal probability (since neither $\pm Z_1$ nor $\pm Z_2$ are in $S$), but the two outcomes will be the same if $a \neq b$ or opposite if $a = b$ (since multiplying the two outcomes is an effective measurement of $Z_1Z_2$).

Applying the corresponding transformations to the pseudo-stabilizers of the LHV model we find

$$\mathcal{Z}_1 \to \mathcal{X}_1 \to \mathcal{X}_1\mathcal{X}_2 \to \mathcal{Y}_1\mathcal{Y}_2 \to \mathcal{Y}_1\mathcal{Y}_2, \tag{5.12}$$

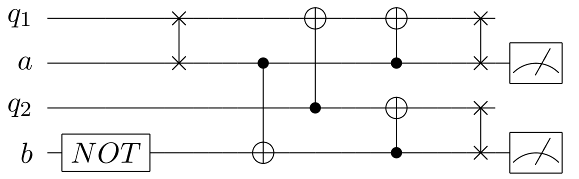
$$\mathcal{Z}_2 \to \mathcal{Z}_2 \to \mathcal{Z}_1\mathcal{Z}_2 \to \mathcal{Z}_1\mathcal{Z}_2 \to \mathcal{X}_1\mathcal{X}_2, \tag{5.13}$$

and so just before the measurement the pseudo-stabilizer subgroup $\mathcal{S}$ is generated by $(-1)^a\mathcal{Y}_1\mathcal{Y}_2$ and $(-1)^b\mathcal{X}_1\mathcal{X}_2$. Since $\mathcal{Y}_1\mathcal{Y}_2\mathcal{X}_1\mathcal{X}_2 = \mathcal{Z}_1\mathcal{Z}_2$ the entire subgroup is

$$\mathcal{S} = \left\{ \mathcal{I}^{\otimes 2}, (-1)^a\mathcal{Y}_1\mathcal{Y}_2, (-1)^b\mathcal{X}_1\mathcal{X}_2, (-1)^{a+b}\mathcal{Z}_1\mathcal{Z}_2 \right\}. \tag{5.14}$$

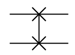
As promised the subgroups are identical up to sign differences. We now follow the inductive procedure to fix the signs. We have $g_1 = (-1)^a \mathcal{Y}_1\mathcal{Y}_2$ and $g_2 = (-1)^b \mathcal{X}_1\mathcal{X}_2$. $g_1$ is not computational and there is no $h$ in $\langle\rangle = \{\mathcal{I}^{\otimes 2}\}$ with $g_1 h$ computational so we ignore $g_1$. $g_2$ is not computational but there is $h = (-1)^a \mathcal{Y}_1\mathcal{Y}_2$ in $\langle g_1 \rangle$ with $g_2 h = (-1)^{a+b} \mathcal{Z}_1\mathcal{Z}_2$ computational. The sign is wrong compared to the qubit stabilizer and so we introduce a flip to the second qubit: preparing $|0\rangle$ will prepare $-\mathcal{Z}_2$ whilst $|1\rangle$ prepares $\mathcal{Z}_2$. We now have $c_1 = (-1)^{a+(1+b)} \mathcal{Z}_1\mathcal{Z}_2$ and so the computational subgroups are identical as required for the model to work.

Following the prescription above we can write the LHV model as a classical circuit



(5.15)

where $q_1$ and $q_2$ are independently set uniformly at random,



(5.16)

represents the swapping of two bits and the meters simply output the classical bit that they receive. Note that the two qubit lines (each represented by two classical bits) interact only at at the controlled-NOT gate. Since classical circuits do not have any quantum non-locality this ensure that the model is local. It can be checked directly that the output probabilities are the same as the quantum mechanical predictions.

# 6 Conclusions and outlook

The pseudo-stabilizer notation emphasizes the close similarities between Spekkens' toy theory and the qubit stabilizer formalism. Recent work [11] has shown that the only GHZ-style correlations possible in theories of a certain type are the toy theory and qubit ones. It would be interesting to seek a set of axioms that are satisfied by both the toy theory and qubit stabilizers. Hopefully a choice from two different additional axioms would then give either the toy theory (for example from an axiom that the theory must satisfy all Bell inequalities) or qubit stabilizers. Such axioms would preferably be of a physical and/or operational flavour with the mathematical structure (for example the correspondence between states and subgroups) being derived from them.

The pseudo-stabilizer notation may also make it easier to consider extensions of the toy theory. For example, here I followed [36] in only considering reversible transformations, but there may be some irreversible transformations that can be added in a natural way. One could also consider more dramatic changes, for example adding explicitly non-local interactions during measurement in order to violate Bell inequalities.

We have seen that the computational power of qubit stabilizers and the toy theory in a circuit model is equal, whilst in a measurement-based model the former is more powerful. This difference has been shown to coincide with a difference in locality: qubit stabilizer circuits admit LHV models whereas freely chosen Pauli measurements on certain stabilizer states do not. In contrast the toy theory is always, by construction, local.

An interesting class of quantum circuits that can be efficiently classically simulated are circuits of "nearest neighbour matchgates" [21]. I have not yet been able to determine whether all such circuits admit an LHV model in the sense considered here. If they do, this would provide further evidence of a link between non-locality and computational speed-up.

If quantum non-locality makes quantum mechanics more computationally powerful than local theories, would theories with "more" non-locality than quantum mechanics be even more powerful? Such theories predict correlations that, whilst still not permitting signalling, violate the quantum Tsirelson's bound [10]. A striking example of such correlations is the PR or non-local box [29]. One such theory is the "Generalized Non-Signalling Theory" of [3]. The computational power of such theories may depend on the computational model, for example a circuit model versus a measurement-based model.

Quantum non-locality is arguably the most remarkable feature of quantum mechanics, and quantum computation is perhaps its most exciting application. This report provides a small contribution to the evidence that they are two sides of the same coin.

# A Appendices

## A.1 Linking the stabilizer formalism to the Hilbert space formalism

### A.1.1 States

The stabilizer state $\rho_S$ was defined as the state for which the expectation values of observables in $S$ is 1, whilst the expectation value of other observables is 0. In the Hilbert space formalism a state is given by a trace one positive operator known as the density operator [28, Section 2.4]. The density operator $\rho_S$ is given by

$$\rho_S = \frac{1}{2^n} \sum_{g \in S} g. \tag{A.1}$$

*Proof.* Let $A$ and $B$ be two observables in $P_n$. Then it is easy to check that

- If $A = B$ then $\mathrm{tr}(AB) = \mathrm{tr}(I^{\otimes n}) = 2^n$,
- If $A = -B$ then $\mathrm{tr}(AB) = \mathrm{tr}(-I^{\otimes n}) = -2^n$,
- Otherwise $\mathrm{tr}(AB) = 0$.

Then for any observable $A$ in $P_n$ we can use the linearity of the trace to see that the expectation value for $A$

$$\mathrm{tr}(A\rho_S) = \frac{1}{2^n} \sum_{g \in S} \mathrm{tr}(Ag) = \begin{cases} 1 & \text{if } A \in S \\ -1 & \text{if } -A \in S \\ 0 & \text{otherwise} \end{cases}. \tag{A.2}$$

We also need to check that $\rho_S$ is a valid density operator. We have already checked that it has trace 1, since $I^{\otimes n} \in S$ (the identity is in any subgroup) and so we have that the expectation value $\mathrm{tr}(I^{\otimes n}\rho_S) = \mathrm{tr}(\rho_S)$

| Density operator $\rho_S$ | Bloch vector | Stabilizer $S$ |
|---|---|---|
| $\lvert+\rangle\langle+\rvert$ where $\lvert+\rangle = \frac{1}{\sqrt{2}}(\lvert0\rangle + \lvert1\rangle)$ | $(1,0,0)$ | $\langle X\rangle$ |
| $\lvert-\rangle\langle-\rvert$ where $\lvert-\rangle = \frac{1}{\sqrt{2}}(\lvert0\rangle - \lvert1\rangle)$ | $(-1,0,0)$ | $\langle -X\rangle$ |
| $\lvert0\rangle\langle0\rvert$ | $(0,0,1)$ | $\langle Z\rangle$ |
| $\lvert1\rangle\langle1\rvert$ | $(0,0,-1)$ | $\langle -Z\rangle$ |
| $\lvert i\rangle\langle i\rvert$ where $\lvert i\rangle = \frac{1}{\sqrt{2}}(\lvert0\rangle + i\lvert1\rangle)$ | $(0,1,0)$ | $\langle Y\rangle$ |
| $\lvert -i\rangle\langle -i\rvert$ where $\lvert -i\rangle = \frac{1}{\sqrt{2}}(\lvert0\rangle - i\lvert1\rangle)$ | $(0,-1,0)$ | $\langle -Y\rangle$ |
| $\frac{1}{2}I$ (the maximally mixed state) | $(0,0,0)$ | $\langle\rangle$ |

Table A.1: The stabilizer states for a single qubit. The Bloch vector $(x, y, z)$ is such that $\rho_S = \frac{I+xX+yY+zZ}{2}$, see [28, Exercise 2.72].

| Density operator $\rho_S$ | Stabilizer $S$ |
|---|---|
| $\lvert01\rangle\langle01\rvert$ | $\langle Z_1, -Z_2\rangle$ |
| $\lvert\Phi^+\rangle\langle\Phi^+\rvert$ where $\lvert\Phi^+\rangle = \frac{1}{\sqrt{2}}(\lvert00\rangle + \lvert11\rangle)$ | $\langle Z_1Z_2, X_1X_2\rangle$ |
| $\lvert\psi\rangle\langle\psi\rvert$ where $\lvert\psi\rangle = \frac{1}{2}(\lvert00\rangle - i\lvert01\rangle + i\lvert10\rangle - \lvert11\rangle)$ | $\langle -Z_1Y_2, -X_1X_2\rangle$ |
| $(\lvert1\rangle\langle1\rvert) \otimes \frac{1}{2}I = \frac{1}{2}(\lvert10\rangle\langle10\rvert + \lvert11\rangle\langle11\rvert)$ | $\langle -Z_1\rangle$ |
| $\frac{1}{2}(\lvert+-\rangle\langle+-\rvert + \lvert-+\rangle\langle-+\rvert)$ | $\langle -X_1X_2\rangle$ |
| $\frac{1}{4}I^{\otimes 2}$ (maximally mixed state) | $\langle\rangle$ |
| $\lvert\mathrm{GHZ}\rangle\langle\mathrm{GHZ}\rvert$ where $\lvert\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(\lvert000\rangle + \lvert111\rangle)$ | $\langle X_1X_2X_3, Z_1Z_2, Z_2Z_3\rangle$ |

Table A.2: Some stabilizer states for composite systems. They are each composed of two qubits, except for the last which is composed of three. $\lvert01\rangle$ is shorthand for $\lvert0\rangle \otimes \lvert1\rangle$, $\lvert+-\rangle$ for $\lvert+\rangle \otimes \lvert-\rangle$ etc.

of this (trivial) observable is 1. We also need that $\rho_S$ is positive. Since it is certainly Hermitian we need only check that its eigenvalues are non-negative. In [9, Section 2] it is shown that $\frac{2^n}{|S|}\rho_S$ (where $|S|$ is the number of elements of $S$) is a projector, and since the eigenvalues of a projector are all 0 or 1 we have that the eigenvalues of $\rho_S$ are all 0 or $\frac{|S|}{2^n}$. This observation also gives the rank of $\rho_S$ as $\frac{2^n}{|S|}$, in particular $\rho_S$ is rank one (i.e. a pure state) if and only if $|S| = 2^n$. (Hence the definition in §1.1.2 corresponds exactly to the standard Hilbert space definition.) □

### A.1.2 Transformations

In the Hilbert space formalism a reversible transformation is represented by a unitary matrix $U$ that sends the density operator $\rho$ to $U\rho U^\dagger$. The Clifford group on $n$ qubits is the set of unitary matrices $U$ with $UgU^\dagger$ in $P_n$ for all $g$ in $P_n$. The function $f_U$ on $P_n$ is then given by $f_U(g) = UgU^\dagger$. The homogeneity and group homomorphism properties of $f_U$ follow directly from the definition. By inspection of (A.1) we see that $U\rho_S U^\dagger = \rho_{f_U(S)}$.

The unitary matrices for the Hadamard, phase and controlled-NOT transformations are

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \text{ and } CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{A.3}$$

respectively. We can then check directly that $f_H(X) = HXH^\dagger = Z$ and so on. That these three transformations generate the entire Clifford group (ignoring irrelevant global phases) is proven in, for example, [9, Section 7].

Finally, we note that the controlled-$S$ gate

$$CS = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{pmatrix} \tag{A.4}$$

is not in the Clifford group, since for example $CSX_1CS^\dagger = \frac{1}{2}(X_1 + X_1Z_2 + Y_1 + Y_1Z_2)$ which is not in the Pauli group.

### A.1.3 Measurements

In the Hilbert space formalism an observable is represented by a Hermitian matrix $O$, and the state after measurement is found using projection operators. In our case the eigenvalues of $O$ are $\pm 1$ and so we can write $O = O_+ - O_-$ where $O_\pm$ are projection operators onto the $\pm 1$ eigenspaces of $O$. For Pauli observables $O_\pm = \frac{I^{\otimes n} \pm O}{2}$, and so the new state is given by

$$\frac{O_\pm \rho_S O_\pm}{\frac{1}{2}} = 2 \sum_{g \in S} O_\pm g O_\pm . \tag{A.5}$$

Note that for any $g \in P_n$

$$2O_\pm g O_\pm = \frac{g \pm gO \pm Og + OgO}{2} = \begin{cases} g \pm Og & \text{if } g, O \text{ commute} \\ 0 & \text{if } g, O \text{ anticommute} \end{cases} . \tag{A.6}$$

Hence the stabilizer subgroup $S$ is updated according the rules given in §1.3.

## A.2 Stabilizer generators versus pseudo-stabilizer generators

Let $\{g_1, g_2, \ldots, g_l\}$ be a subset of $G_n$. Then the following are equivalent:

1. $\{g_1, g_2, \ldots, g_l\}$ are indepedent generators of a jointly-knowable subgroup of $G_n$ that does not contain $-\mathcal{I}^n$.

2. $\{m(g_1), m(g_2), \ldots, m(g_l)\}$ are independent generators of a subgroup of $P_n$ that does not contain $-I$.

"Independent generators" means the list of generators is minimal, i.e. none of them can be written as a product of the rest.

*Proof.* The second statement is equivalent to "$m(g_1), m(g_2), \ldots, m(g_l)$ commute, have linearly independent check vectors, and each square to $I^{\otimes n}$" [9, Equation 54]. As already noted, the check vectors for $g$ and $m(g)$ are the same and commuting in $P_n$ is equivalent to being jointly-knowable in $G_n$. Furthermore $m(g)^2 = I^{\otimes n}$ since the only elements of $P_n$ that don't square to $I^{\otimes n}$ are those with phases $\alpha = i$ or $-i$, which aren't in the range of $m$.

Hence the second statement is equivalent to "$g_1, g_2, \ldots, g_l$ are jointly-knowable and have linearly independent check vectors". We have that if $a, b, c \in G_n$ are jointly knowable then so are $ab$ and $c$. Therefore $g_1, g_2, \ldots, g_l$ are jointly-knowable if and only if $\langle g_1, g_2, \ldots, g_l \rangle$ are jointly knowable.

Suppose $\langle g_1, g_2, \ldots, g_l \rangle$ contains $-\mathcal{I}^{\otimes n}$. Then the check vector of $-\mathcal{I}^{\otimes n}$ can be written is a linear combination of the check vectors of $g_1, g_2, \ldots, g_l$.

But the check vector of $-\mathcal{I}^{\otimes n}$ is 0 and so the check vectors of $g_1, g_2, \ldots, g_l$ are linearly dependent. Suppose that the first statement holds but the check vectors of $g_1, g_2, \ldots, g_l$ are linearly dependent. Then the check vector of one of them, say of $g_1$, can be written as a linear combination of the others. That means that either $g_1$ or $-g_1$ can be written as a product of the others. The first possibility contradicts the assumption of independent generators, and since $g_1(-g_1) = -\mathcal{I}^{\otimes n}$ the second contradicts the assumption that the subgroup does not contain $-\mathcal{I}^{\otimes n}$. □

## A.3 Joint probabilities as expectation values of products

Rather than examining the joint probability distribution of measurement outcomes directly, it is sometimes easier to check that an LHV model predicts the correct expectation values for all the "products of $Z$ operators". Here I show that this condition is sufficient (it is obviously necessary). For all $1 \leq k \leq n$ let $\hat{Z}_k$ be a measurement outcome 1 or $-1$. Then we clearly have

$$\left(\frac{1+\hat{Z}_k}{2}\right) = \begin{cases} 1 & Z_k = 1 \\ 0 & Z_k = -1 \end{cases}. \tag{A.7}$$

Hence for any probability distribution $p$ on the $Z_k$ we have

$$\begin{aligned} p(\hat{Z}_1 = 1, \ldots, \hat{Z}_n = 1) &= \left\langle \frac{1+\hat{Z}_1}{2} \cdots \frac{1+\hat{Z}_n}{2} \right\rangle \\ &= \frac{1 + \left\langle \hat{Z}_1 \right\rangle + \cdots + \left\langle \hat{Z}_n \right\rangle + \left\langle \hat{Z}_1 \hat{Z}_2 \right\rangle + \cdots + \left\langle \hat{Z}_1 \cdots \hat{Z}_n \right\rangle}{2^n}. \end{aligned} \tag{A.8}$$

Similarly

$$\left(\frac{1-\hat{Z}_k}{2}\right) = \begin{cases} 0 & Z_k = 1 \\ 1 & Z_k = -1 \end{cases}, \tag{A.9}$$

and so, for example with $n = 2$

$$p(\hat{Z}_1 = 1, \hat{Z}_2 = -1) = \left\langle \frac{1+\hat{Z}_1}{2} \frac{1-\hat{Z}_2}{2} \right\rangle = \frac{1 + \left\langle \hat{Z}_1 \right\rangle - \left\langle \hat{Z}_2 \right\rangle - \left\langle \hat{Z}_1 \hat{Z}_2 \right\rangle}{4}. \tag{A.10}$$

# Bibliography

[1] S. Aaronson and D. Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70(5):052328, 2004. Available from: `http://dx.doi.org/10.1103/PhysRevA.70.052328`.

[2] J. Anders and D. E. Browne. Computational power of correlations. *Phys. Rev. Lett.*, 102(5):050502, 2009. Available from: `http://dx.doi.org/10.1103/PhysRevLett.102.050502`.

[3] J. Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75(3):032304, 2007. Available from: `http://dx.doi.org/10.1103/PhysRevA.75.032304`.

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

[5] J. S. Bell. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy.* Cambridge University Press, Cambridge, 2nd edition, 2004.

[6] G. Blaylock. The EPR paradox, Bell's inequality, and the question of locality. *American Journal of Physics*, 78(1):111–120, 2010. Available from: `http://link.aip.org/link/?AJP/78/111/1`.

[7] D. Bohm. *Quantum Theory.* Prentice-Hall, New Jersey, 1951.

[8] H. J. Briegel and R. Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86(5):910–913, 2001. Available from: `http://dx.doi.org/10.1103/PhysRevLett.86.910`.

[9] C. M. Caves. Stabilizer formalism for qubits. Internal Report, 2006. Available from: `http://info.phys.unm.edu/~caves/reports/reports.html`.

[10] B. S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4:93–100, 1980. Available from: `http://dx.doi.org/10.1007/BF00417500`.

[11] B. Coecke, B. Edwards, and R. W. Spekkens. Phase groups and the origin of non-locality for qubits. 2010. Available from: `http://arxiv.org/abs/1003.5005`.

[12] J. Dehaene and B. De Moor. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Phys. Rev. A*, 68(4):042318, 2003. Available from: `http://dx.doi.org/10.1103/PhysRevA.68.042318`.

[13] M. V. den Nest. Classical simulation of quantum computation, the Gottesman-Knill theorem, and slightly beyond. *Quant. Inf. Comp.*, 10(3&4):0258–0271, 2010. Available from: `http://www.rintonpress.com/journals/qiconline.html#v10n34`.

[14] D. Deutsch. Quantum computational networks. *Proc. R. Soc. A*, 425(1868):73–90, 1989. Available from: `http://dx.doi.org/10.1098/rspa.1989.0099`.

[15] D. P. DiVincenzo. Quantum gates and circuits. *Proc. R. Soc. A*, 454(1969):261–276, 1998. Available from: `http://dx.doi.org/10.1098/rspa.1998.0159`.

[16] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47(10):777–780, 1935. Available from: `http://dx.doi.org/10.1103/PhysRev.47.777`.

[17] D. Gottesman. Class of quantum error-correcting codes saturating the quantum hamming bound. *Phys. Rev. A*, 54(3):1862–1868, 1996. Available from: `http://dx.doi.org/10.1103/PhysRevA.54.1862`.

[18] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, page 69. Kluwer Academic, Dordrecht, 1989.

[19] L. Hardy. Quantum ontological excess baggage. *Stud. Hist. Phil. Sci. B*, 35(2):267–276, 2004. Available from: `http://dx.doi.org/10.1016/j.shpsb.2003.12.001`.

[20] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69(6):062311, 2004. Available from: http://dx.doi.org/10.1103/PhysRevA.69.062311.

[21] R. Jozsa and A. Miyake. Matchgates and classical simulation of quantum circuits. *Proc. R. Soc. A*, 464:3089–3106, 2008. Available from: http://dx.doi.org/10.1098/rspa.2008.0189.

[22] T. Maudlin. *Quantum Non-Locality and Relativity.* Blackwell, Massachusetts, 2nd edition, 2002.

[23] T. Maudlin. What Bell proved: A reply to Blaylock. *American Journal of Physics*, 78(1):121–125, 2010. Available from: http://link.aip.org/link/?AJP/78/121/1.

[24] N. D. Mermin. *Boojums All The Way Through: Communicating Science In a Prosaic Age.* Cambridge University Press, Cambridge, 1990.

[25] N. D. Mermin. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.*, 65(27):3373–3376, 1990. Available from: http://dx.doi.org/10.1103/PhysRevLett.65.3373.

[26] N. D. Mermin. What's wrong with these elements of reality? *Physics Today*, 43(6):9–11, 1990. Available from: http://link.aip.org/link/?PTO/43/9/1.

[27] M. A. Nielsen. Cluster-state quantum computation. *Rep. Math. Phys.*, 57(1):147–161, 2006. Available from: http://dx.doi.org/10.1016/S0034-4877(06)80014-5.

[28] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

[29] S. Popescu and D. Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24:379–385, 1994. Available from: http://dx.doi.org/10.1007/BF02058098.

[30] R. Raussendorf. Quantum computation, discreteness, and contextuality. 2009. Available from: http://arxiv.org/abs/0907.5449.

[31] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86(22):5188–5191, 2001. Available from: `http://dx.doi.org/10.1103/PhysRevLett.86.5188`.

[32] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68(2):022312, 2003. Available from: `http://dx.doi.org/10.1103/PhysRevA.68.022312`.

[33] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, 1978. Available from: `http://doi.acm.org/10.1145/359340.359342`.

[34] Y. Shi. Both Toffoli and controlled-NOT need little help to universal quantum computing. *Quant. Inf. Comp.*, 3(1):084–092, 2003. Available from: `http://www.rintonpress.com/journals/qiconline.html#v3n12`.

[35] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comp.*, 26(5):1484–1509, 1997. Available from: `http://link.aip.org/link/?SMJ/26/1484/1`.

[36] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75(3):032110, 2007. Available from: `http://dx.doi.org/10.1103/PhysRevA.75.032110`.

[37] R. W. Spekkens. The power of epistemic restrictions in reconstructing quantum theory. Perimiter Institute Recorded Seminar Archive, 2009. Available from: `http://pirsa.org/09080009/`.

[38] S. J. van Enk. A toy model for quantum mechanics. *Foundations of Physics*, 37(10):1447–1460, 2007. Available from: `http://dx.doi.org/10.1007/s10701-007-9171-3`.