

Hardy is (almost) everywhere: nonlocality without inequalities for almost all entangled multipartite states

Samson Abramsky^a, Carmen M. Constantin^a, Shenggang Ying^b

^a*Department of Computer Science, University of Oxford
Wolfson Building, Parks Road, Oxford OX1 3QD, U.K.*

^b*Department of Computer Science, Tsinghua University, Beijing, China*

Abstract

We show that all n -qubit entangled states, with the exception of tensor products of single-qubit and bipartite maximally-entangled states, admit Hardy-type proofs of non-locality without inequalities or probabilities. More precisely, we show that for all such states, there are local, one-qubit observables such that the resulting probability tables are logically contextual in the sense of Abramsky and Brandenburger, this being the general form of the Hardy-type property. Moreover, our proof is constructive: given a state, we show how to produce the witnessing local observables. In fact, we give an algorithm to do this. Although the algorithm is reasonably straightforward, its proof of correctness is non-trivial. A further striking feature is that we show that $n + 2$ local observables suffice to witness the logical contextuality of any n -qubit state: two each for two for the parties, and one each for the remaining $n - 2$ parties.

1. Introduction

Non-locality theory is a central pillar of both quantum information and quantum foundations. The founding result, Bell's theorem, is Bell's fundamental contribution [1], which was subsequently given its standard form using the CHSH inequalities by Clauser, Horne, Shimony and Holt [2]. We shall refer to this as the Bell/CHSH argument. This original, and still standard, form of the argument made essential use of the probabilistic predictions

Email addresses: samson.abramsky@cs.ox.ac.uk (Samson Abramsky),
i.m.carmen@gmail.com (Carmen M. Constantin), yingshenggang@gmail.com
(Shenggang Ying)

of quantum mechanics to separate them from those of any local realistic theory. Subsequently, Greenberger, Horne, Zeilinger and Shimony [3, 4] gave an important strengthening of Bell’s theorem, to an **inequality-free** form, using only quantum predictions made with certainty. The “price” for this strengthened form of the theorem is that, while Bell’s argument applies to bipartite, 2-qubit systems, the GHZ argument needs at least 3 qubits. A further important step was taken by Hardy [5, 6], who showed that an inequality-free proof of Bell’s theorem could be given for a 2-qubit system. A curious feature of Hardy’s construction is that it works for any bipartite entangled state, **except** for the maximally entangled states. Subsequently, there have been many extensions and generalisations of these results, e.g. [7, 8, 9].

In [10], a general mathematical theory of non-locality and contextuality was developed in a sheaf-theoretic framework. In particular, the content of the Bell/CHSH, GHZ and Hardy arguments was abstracted into general properties of “empirical models”, *i.e.* of the probability tables summarising the empirical predictions for various combinations of measurements. An important point that was revealed by this analysis is that the Bell/CHSH, Hardy and GHZ arguments correspond to three different **strengths** or **degrees** of contextual behaviour. We have

$$\text{Bell} < \text{Hardy} < \text{GHZ}.$$

In the terminology of [10], the GHZ argument hinges on **strong contextuality**, which can be exhibited by certain quantum systems, and in particular in the usual n -qubit multipartite Bell scenarios, as soon as we have 3 or more parties. The Hardy argument hinges on a weaker property, which we termed **logical** or **possibilistic** contextuality in [10]. This property is in turn stronger than the probabilistic contextuality used in the Bell/CHSH argument. These notions form a proper hierarchy: strong contextuality is strictly stronger than logical contextuality, which is strictly stronger than probabilistic contextuality.

What distinguishes both logical and strong contextuality from probabilistic contextuality is that they rely only on the **support** of the probability distributions. That is, these forms of contextuality rely only on whether the probabilities of events are 0 (impossible) or > 0 (possible). In this sense, they are indeed inequality-free [4, 6], and even probability-free [11, 7]. The constructions in the Bell/CHSH, Hardy and GHZ arguments show that certain quantum states and local observables (*i.e.* one-qubit measurements which can be made by each of the parties on their part of the state) give

rise to probability tables which are, respectively, probabilistically, logically or strongly contextual.

This leads to the following natural question:

Given a quantum state, find the maximum level of contextuality it can exhibit for some choice of local observables.

In particular, we can ask:

For which quantum states can we find local observables which give rise to a logically contextual (inequality-free, probability-free) form of contextuality?

In this paper, we shall completely answer this question for n -qubit pure states, for all n . Our main result is as follows.

Theorem 1. *Let $|\psi\rangle$ be an n -qubit pure state. Then exactly one of the following two cases must hold:*

1. $|\psi\rangle$ can be written (up to permutation of tensor factors) as a product

$$|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle \quad (1)$$

where each $|\psi_i\rangle$ is either a 1-qubit state, or a 2-qubit maximally entangled state.

2. There are local observables such that the probability table arising from $|\psi\rangle$ and these local observables is logically contextual, that is, it admits a “Hardy paradox”, i.e. an inequality-free, probability-free proof of non-locality.

If (2) holds, then only $n + 2$ local observables are needed; two each for two of the parties, and one each for the other $n - 2$ parties. Moreover, there is an algorithm to decide which of the above cases holds, and which in case (2) explicitly computes the witnessing local observables. The complexity of this algorithm is $O(d \log^3 d)$ in the dimension $d = 2^n$.

The remainder of the paper is devoted to the detailed statement and proof of this theorem. In particular, while the algorithm is quite simple, the arguments justifying its correctness are non-trivial.

In Section 2, we define logical contextuality in the n -qubit setting. In Section 3, we state the main lemmas, and give the top-level proof of the

theorem from these lemmas. In Section 4, we prove the lemmas. In Section 5, we present the algorithm. Finally, in Section 6 we discuss some further directions.

Acknowledgements. We thank Chris Heunen for some helpful discussions and suggestions.

2. Logical Contextuality

Our setting consists of a general form of the probability models commonly studied in quantum information and quantum foundations. In these models, a number of agents each has the choice of one of several measurement settings. Moreover each measurement has a number of distinct outcomes. We consider here dichotomic measurements with possible outcomes labelled by the elements of $\mathbf{2} := \{+, -\}$.

Thus, recalling the terminology developed in [10, 12], a **measurement scenario** is given by a pair (X, \mathcal{U}) . Here X is a set of measurement labels, while \mathcal{U} is a family of subsets of X , giving the maximal sets of compatible measurements, or **contexts**.

An empirical model, or generalized probability table, on (X, \mathcal{U}) is a family $e = \{e_U\}_{U \in \mathcal{U}}$ of probability distributions e_U on the joint outcomes $\mathbf{2}^U$, one for each context U . The probability of obtaining any joint outcome $s \in \mathbf{2}^U$ is given by $e_U(s)$. The support of the model at U is the set $S(U) \subseteq \mathbf{2}^U$ of those joint outcomes $s \in \mathbf{2}^U$ for which $e_U(s) > 0$.

We can make a connection to logic by interpreting the $+$ outcome as *true* and the $-$ outcome as *false*. This allows us to think of the set of measurements X as a set of boolean variables. If U is a finite context, any subset of $\mathbf{2}^U$ can be defined by a propositional formula. For example, each joint outcome $s : U \rightarrow \mathbf{2}$ determines a propositional formula

$$\varphi_s = \bigwedge_{x \in U, s(x)=+} x \wedge \bigwedge_{x \in U, s(x)=-} \neg x$$

The only satisfying assignment of φ_s in $\mathbf{2}^U$ is s .

The propositional formula whose set of satisfying assignments is the support of U is $\varphi_U := \bigvee_{s \in S(U)} \varphi_s$.

An empirical model is **logically contextual** if there exists some $U \in \mathcal{U}$ and some $s \in S(U)$ such that the formula

$$\Phi = \varphi_s \wedge \bigwedge_{V \in \mathcal{U} \setminus U} \varphi_V$$

is not satisfiable. This says that there is a possible joint outcome s which cannot be accounted for by any valuation on all the variables in X which is consistent with the support of the model. This immediately implies that there is no joint distribution on all the observables which marginalizes to yield the empirically observable probabilities [10]. As originally shown in the bipartite case by Fine [13], and in a very general form in [10], the non-existence of a joint distribution is equivalent to the usual definition of non-locality as given by Bell [1].

An n -qubit quantum state $|\psi\rangle$, together with a set of 1-qubit local observables for each i , $1 \leq i \leq n$, gives rise to an empirical model as above. An observable is represented by a 2×2 self-adjoint matrix M . The two possible outcomes correspond to its eigenvalues, which we label $+$ and $-$. We use labels of the form (M, i) , corresponding to a measurement of observable M performed by the i^{th} party, to denote the elements of the set X of all available measurements.

A set U of compatible measurements is given by a choice of observable at each of the n measurement sites. A given quantum state $|\psi\rangle$ determines a probability distribution on $\mathbf{2}^U$, as follows. If

$$U = \{(M_1, 1), (M_2, 2) \dots (M_n, n)\}$$

is a compatible set of measurements, the probability of obtaining the joint outcome $s : U \rightarrow \mathbf{2}$ is given by the Born rule:

$$e_U(s) = |\langle \mu_1 \rangle \otimes \langle \mu_2 \rangle \otimes \dots \otimes \langle \mu_n \rangle | \psi \rangle|^2,$$

where $|\mu_i\rangle$ is the eigenvector corresponding to the eigenvalue $s(M, i)$ of the observable (M, i) representing the measurement performed by the i^{th} party. The joint outcome s is in the support of the model if and only if this inner product is non-zero.

3. The structure of the argument

In this section, we shall state a number of main lemmas, and show how Theorem 1 is proved from these lemmas. The proofs of the lemmas will be given in the following section.

Notation. We shall write \mathcal{P}_n for the set of n -qubit pure states of the form (1). For succinctness, we shall write the tensor product of a ket $|\psi\rangle$ with a 1-qubit ket $|i\rangle$, $i = 0, 1$, as $|\psi\rangle|i\rangle$ rather than $|\psi\rangle \otimes |i\rangle$.

We shall prove Theorem 1 by induction on n . The case $n = 1$ is trivial. The $n = 2$ case is given by the following lemma.

Lemma 2 (The Base Case Lemma). *Every 2-qubit state $|\psi\rangle$ is either in \mathcal{P}_2 (i.e. it is either a product state, or a maximally entangled bipartite state), or there are two local observables for each party, which can be computed directly from the Schmidt decomposition of $|\psi\rangle$, and which witness the logical contextuality of $|\psi\rangle$.*

The following lemmas and corollary will be used in the induction step.

Lemma 3 (Going Up Lemma I). *If an n -qubit state $|\psi\rangle$ is logically contextual with some choice of local observables, then for any n -qubit state $|\theta\rangle$, and $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$ and $|\alpha|^2 + |\beta|^2 = 1$, the states*

$$\alpha|\psi\rangle|0\rangle + \beta|\theta\rangle|1\rangle, \quad \beta|\theta\rangle|0\rangle + \alpha|\psi\rangle|1\rangle$$

are also logically contextual with the same choice of observables, augmented with a single additional observable for the $n+1$ -th party.

Lemma 4 (Going Up Lemma II). *If $|\theta\rangle = \alpha|\psi\rangle + \beta|\phi\rangle$ is a logically contextual n -qubit state under some choice of local observables, then for any non-zero $x, y \in \mathbb{C}$ such that $|x\alpha|^2 + |y\beta|^2 = 1$, the state $|\omega\rangle = x\alpha|\psi\rangle|0\rangle + y\beta|\phi\rangle|1\rangle$ is also logically contextual under the same choice of observables, augmented with a single additional observable for the $n+1$ -th party.*

As an easy consequence of these lemmas, we have:

Corollary 5. *If $|\theta\rangle$ is logically contextual under some choice of local observables, so are $|\theta\rangle \otimes |\eta\rangle$ and $|\eta\rangle \otimes |\theta\rangle$ for any state $|\eta\rangle$, with the same choice of local observables for each party in $|\theta\rangle$, and a single observable for each party in $|\eta\rangle$.*

We now consider the induction step where we have an $n+1$ -qubit state $|\omega\rangle$, $n > 1$. We can write

$$|\omega\rangle = \alpha|\psi\rangle|0\rangle + \beta|\phi\rangle|1\rangle. \quad (2)$$

By the Going Up Lemma I, if either $|\psi\rangle$ or $|\phi\rangle$ are logically contextual, so is $|\omega\rangle$, and we are done.

Suppose now that $|\psi\rangle$ and $|\phi\rangle$ are both in \mathcal{P}_n . We consider the parameterised family of states

$$\tau(a) = a|\psi\rangle + \sqrt{1-a^2}|\phi\rangle, \quad a \in [0, 1]. \quad (3)$$

If for some a , $\tau(a)$ is logically contextual, so is $|\omega\rangle$, by the Going Up Lemma II. For the remaining case, we have the following rather remarkable result.

Lemma 6 (The Small Difference Lemma). *Let $|\psi\rangle$ and $|\phi\rangle$ be states in \mathcal{P}_n , and suppose that for all $a \in [0, 1]$, $\tau(a)$ is in \mathcal{P}_n , where $\tau(a)$ is defined by (3). Then $|\psi\rangle$ and $|\phi\rangle$ differ in at most one qubit.*

Applying this result to our decomposition (2) of $|\omega\rangle$, we have the following possibilities:

- $|\psi\rangle = |\phi\rangle$, in which case, from (2) and the bilinearity of the tensor product:

$$|\omega\rangle = \alpha|\psi\rangle|0\rangle + \beta|\psi\rangle|1\rangle = |\psi\rangle \otimes (\alpha|0\rangle + \beta|1\rangle).$$

Since by assumption $|\psi\rangle$ is in \mathcal{P}_n , $|\omega\rangle$ is in \mathcal{P}_{n+1} .

- $|\psi\rangle \neq |\phi\rangle$, in which case (up to permutation) we can write

$$\begin{aligned} |\psi\rangle &= |\Psi\rangle \otimes |\eta\rangle \\ |\phi\rangle &= |\Psi\rangle \otimes |\theta\rangle \end{aligned}$$

where $|\Psi\rangle$ is in \mathcal{P}_{n-1} and $|\eta\rangle$ and $|\theta\rangle$ are 1-qubit states. From this and (2), using the bilinearity of tensor product again we have

$$|\omega\rangle = |\Psi\rangle \otimes |\xi\rangle$$

where $|\xi\rangle$ is a 2-qubit state. We can apply the Base Case Lemma to $|\xi\rangle$ to conclude that $|\xi\rangle$ is either in \mathcal{P}_2 , in which case $|\omega\rangle$ is in \mathcal{P}_{n+1} , or $|\xi\rangle$ is logically contextual, in which case $|\omega\rangle$ is logically contextual by the corollary to the Going Up Lemma.

At this point, we have established (1) and (2) of Theorem 1, but it seems that we require an infinite search to determine if there exists some $a \in [0, 1]$ for which $\tau(a)$ is logically contextual.

However, the following lemma shows that we only need to test a fixed, finite number of values for a to determine this.

Lemma 7 (The 21 Lemma). *With the same notation as in the Small Difference Lemma, suppose that $\tau(a)$ is in \mathcal{P}_n for 21 distinct values of a in $[0, 1]$. Then $\tau(a)$ is in \mathcal{P}_n for all $a \in [0, 1]$.*

Thus this lemma allows us to determine which case applies on the basis of a finite number of tests.

In the next section, we shall give proofs of these lemmas. We shall then give an explicit algorithm in Section 5 to complete the proof of Theorem 1.

4. Proofs of the lemmas

Firstly, we collect a few useful basic properties.

4.1. Background lemmas

We consider relations $\sim = \{\sim_n\}_{n \in \mathbb{N}}$, where \sim_n is an equivalence relation on n -qubit states. We say that \sim is **LC invariant** if for all n -qubit states $|\psi\rangle, |\phi\rangle$:

- If $|\psi\rangle \sim_n |\phi\rangle$, then $|\psi\rangle \in \mathcal{P}_n$ iff $|\phi\rangle \in \mathcal{P}_n$.
- If $|\psi\rangle \sim_n |\phi\rangle$, then $|\psi\rangle$ is logically contextual iff $|\phi\rangle$ is logically contextual.

Lemma 8. *The relation induced by permutation of tensor factors is LC invariant.*

Lemma 9. *The relation of LU equivalence is LC invariant. Here LU equivalence refers to the relation induced by the action of local (1-qubit) unitaries.*

The following result will be used in the proof of the Small Difference lemma. It refers to the “partial inner product” operation described e.g. in [14, p. 129].¹

Lemma 10. *Let $|\phi\rangle$ be a state in $\mathcal{H} \otimes \mathcal{K}$. For any states $|\eta\rangle$ in \mathcal{H} and $|\theta\rangle$ in \mathcal{K} , if for all $|\eta^\perp\rangle$ orthogonal to $|\eta\rangle$, $\langle \eta^\perp | \phi \rangle = \mathbf{0}$, and for all $|\theta^\perp\rangle$ orthogonal to $|\theta\rangle$, $\langle \theta^\perp | \phi \rangle = \mathbf{0}$, then (up to global phase) $|\phi\rangle = |\eta\rangle \otimes |\theta\rangle$.*

Proof. We extend $|\eta\rangle$ into an orthonormal basis $|\eta_1\rangle, \dots, |\eta_n\rangle$ with $|\eta\rangle = |\eta_1\rangle$, and similarly extend $|\theta\rangle$ into $|\theta_1\rangle, \dots, |\theta_m\rangle$ with $|\theta\rangle = |\theta_1\rangle$. Then $B = \{|\eta_i\rangle \otimes |\theta_j\rangle\}_{i,j}$ forms an orthonormal basis of $\mathcal{H} \otimes \mathcal{K}$. Note that

$$\mathcal{H} \otimes \mathcal{K} = (|\eta\rangle \otimes |\theta\rangle)^{\perp\perp} \oplus S^{\perp\perp}$$

where $S = \{|\eta_i\rangle \otimes |\theta_j\rangle \mid (i, j) \neq (1, 1)\}$. Hence $S^{\perp\perp} = (|\eta\rangle \otimes |\theta\rangle)^\perp$. By our assumption and the defining property of the partial inner product [14, Equation (6.47)], $S \subseteq |\phi\rangle^\perp$. Hence

$$|\phi\rangle^{\perp\perp} \subseteq S^\perp = S^{\perp\perp\perp} = (|\eta\rangle \otimes |\theta\rangle)^{\perp\perp}.$$

Since these are one-dimensional subspaces, this implies that, up to global phase, $|\phi\rangle = |\eta\rangle \otimes |\theta\rangle$. \square

We now turn to detailed proofs of the main lemmas. For convenience, we shall repeat the statements of the lemmas.

¹This is actually the application of a linear map to a vector under Map-State duality [15].

4.2. The Base Case lemma

Lemma 2 (The Base Case Lemma). *Every 2-qubit state $|\psi\rangle$ is either in \mathcal{P}_2 (i.e. it is either a product state, or a maximally entangled bipartite state), or there are two local observables for each party, which can be computed directly from the Schmidt decomposition of $|\psi\rangle$, and which witness the logical contextuality of $|\psi\rangle$.*

Proof. This is essentially Hardy's construction in [6]. Using the Schmidt decomposition, every two-particle entangled state can be written in the form

$$|\psi\rangle = \alpha|+\rangle_1|+\rangle_2 + \beta|-\rangle_1|-\rangle_2$$

for an appropriate choice of basis states $|\pm\rangle_i$ for each particle i , and normalized non-zero real constants α and β .

The logical contextuality of $|\psi\rangle$ is witnessed by a set of four dichotomic observables, two for each of the two parties, namely U_i and D_i , $i = 1, 2$. These observables can be defined as $U_i = |u_i\rangle\langle u_i|$ and $D_i = |d_i\rangle\langle d_i|$ where

$$|u_i\rangle = \frac{1}{\sqrt{|\alpha| + |\beta|}}(\beta^{\frac{1}{2}}|+\rangle_i + \alpha^{\frac{1}{2}}|-\rangle_i) \quad (4)$$

$$|d_i\rangle = \frac{1}{\sqrt{|\alpha|^3 + |\beta|^3}}(\beta^{\frac{3}{2}}|+\rangle_i - \alpha^{\frac{3}{2}}|-\rangle_i) \quad (5)$$

Hardy's paper also explains why it is not possible to run this particular non-locality argument when either α or β are equal to zero (product states), or when $|\alpha| = |\beta|$ (maximally entangled states), that is, when the state $|\psi\rangle$ belongs to \mathcal{P}_2 .

There is one subtle remaining point. To show that the dichotomy asserted in the lemma is strictly disjoint, we must show that in the maximally entangled case, there is **no** choice of local observables which can give rise to logical contextuality. This is shown for the case where each party has the same two local observables as Theorem 3.5 in [16], and more generally for any finite sets of local observables as Theorem 2.6.5 in [17]. \square

4.3. The Going Up lemmas

The proofs of the two Going Up lemmas are quite similar. We shall prove the second, which is somewhat harder.

Lemma 4 (Going Up Lemma II). *If $|\theta\rangle = \alpha|\psi\rangle + \beta|\phi\rangle$ is a logically contextual n -qubit state under some choice of local observables, then for any non-zero $x, y \in \mathbb{C}$ such that $|x\alpha|^2 + |y\beta|^2 = 1$, the state $|\omega\rangle = x\alpha|\psi\rangle|0\rangle + y\beta|\phi\rangle|1\rangle$ is also logically contextual under the same choice of observables, augmented with a single additional observable for the $n+1$ -th party.*

Proof. Since $|\theta\rangle$ is logically contextual there must be some context U' and some $s' \in S(U')$ such that the formula $\Psi = \varphi_{s'} \wedge \bigwedge_{U \in \mathcal{U} \setminus U'} \varphi_U$ is not satisfiable. We will show that it is possible to construct a similar unsatisfiable formula in order to prove the logical non-locality of $|\omega\rangle$.

The $n + 1$ -th party is assigned a single observable $B = B(x, y)$, whose eigenvectors are $|b_+\rangle = \bar{y}|0\rangle + \bar{x}|1\rangle$ and $|b_-\rangle = x|0\rangle - y|1\rangle$. The observable B is given by the self-adjoint matrix

$$B(x, y) = \begin{pmatrix} -|x|^2 + |y|^2 & 2x\bar{y} \\ 2\bar{x}y & |x|^2 - |y|^2 \end{pmatrix} \quad (6)$$

For any n -qubit state $|\mu\rangle$ we have

$$\begin{aligned} \langle \mu | \langle b_+ | \cdot | \omega \rangle &= x\alpha \langle \mu | \psi \rangle \langle b_+ | 0 \rangle + y\beta \langle \mu | \phi \rangle \langle b_+ | 1 \rangle \\ &= xy\alpha \langle \mu | \psi \rangle + xy\beta \langle \mu | \phi \rangle \\ &= xy(\alpha \langle \mu | \psi \rangle + \beta \langle \mu | \phi \rangle) \\ &= xy \langle \mu | \theta \rangle \end{aligned}$$

which implies

$$|\langle \mu | \theta \rangle|^2 = 0 \Leftrightarrow |\langle \mu | \langle b_+ | \cdot | \omega \rangle|^2 = 0 \quad (7)$$

The augmented set of allowed measurements $\tilde{X} = X \cup (B, n + 1)$ is then covered by the family of compatible subsets

$$\tilde{\mathcal{U}} := \{\tilde{U} = U \cup (B, n + 1) \mid U \in \mathcal{U}\}.$$

Let $T(\tilde{U})$ denote the support of $\tilde{U} \in \tilde{\mathcal{U}}$. Equation (7) implies that $S(U) = \{s \mid s+ \in T(\tilde{U})\}$ where $s+ : \tilde{U} \rightarrow \mathbf{2}$ extends s by mapping $(B, n + 1)$ to $+$. As a side remark, note that the presence of an analogously defined section $\sigma-$ in $T(\tilde{U})$ does not necessarily imply the presence of σ in $S(U)$.

Now let $t' := s'+$. We have $t' \in T(\tilde{U}')$ and we can define the analogue of the proposition $\varphi_{s'}$ as

$$\begin{aligned} \varphi_{t'} &= \bigwedge_{x \in \tilde{U}', t'(x)=+} x \wedge \bigwedge_{x \in \tilde{U}', t'(x)=-} \neg x \\ &= \bigwedge_{x \in U', s'(x)=+} x \wedge \bigwedge_{x \in U', s'(x)=-} \neg x \wedge z_{n+1} \\ &= \varphi_{s'} \wedge z_{n+1} \end{aligned}$$

where z_{n+1} denotes the boolean variable corresponding to the outcome on the $n + 1$ -qubit. Recall that $+$ stands for true and $-$ stands for false.

It also holds that

$$\begin{aligned}
\varphi_{\tilde{U}} &= \bigvee_{t \in T(\tilde{U})} \varphi_t = \bigvee_{t \in T(\tilde{U}), t=s+} \varphi_t \vee \bigvee_{t \in T(\tilde{U}), t=s-} \varphi_t \\
&= \left(\bigvee_{s \in S(U)} (\varphi_s \wedge z_{n+1}) \right) \vee \left(\bigvee_{t \in T(\tilde{U}), t=\sigma-} \varphi_\sigma \wedge \neg z_{n+1} \right) \\
&= \left(\left(\bigvee_{s \in S(U)} \varphi_s \right) \wedge z_{n+1} \right) \vee \left(\underbrace{\left(\bigvee_{t=\sigma- \in T(\tilde{U})} \varphi_\sigma \right)}_{\gamma_U} \wedge \neg z_{n+1} \right) \\
&= (\varphi_U \wedge z_{n+1}) \vee (\gamma_U \wedge \neg z_{n+1})
\end{aligned}$$

We can now define the formula Ω which specifies the joint outcome $t' = s' + \in \tilde{U}'$ as well as the joint outcomes within the supports of all compatible sets of measurements $\tilde{U} \neq \tilde{U}'$. This formula is just the conjunction of Ψ and z_{n+1} . In order to show that $|\omega\rangle$ is logically contextual, it suffices to show that Ω has no satisfiable assignment. This is indeed the case, as the fact that Ψ is not satisfiable implies that Ω is also not satisfiable, thus completing our proof. Indeed, we have

$$\begin{aligned}
\Omega &= \varphi_{t'} \wedge \bigwedge_{\tilde{U} \in \tilde{\mathcal{U}} \setminus \tilde{U}'} \varphi_{\tilde{U}} \\
&= (\varphi_{s'} \wedge z_{n+1}) \wedge \bigwedge_{\tilde{U} \in \tilde{\mathcal{U}} \setminus \tilde{U}'} [(\varphi_{\tilde{U}} \wedge z_{n+1}) \vee (\gamma_U \wedge \neg z_{n+1})] \\
&= \bigwedge_{\tilde{U} \in \tilde{\mathcal{U}} \setminus \tilde{U}'} [(\varphi_{s'} \wedge z_{n+1}) \wedge ((\varphi_{\tilde{U}} \wedge z_{n+1}) \vee (\gamma_U \wedge z_{n+1}))] \\
&= \bigwedge_{\tilde{U} \in \tilde{\mathcal{U}} \setminus \tilde{U}'} [(\varphi_{s'} \wedge z_{n+1}) \wedge (\varphi_{\tilde{U}} \wedge z_{n+1})] \\
&= \varphi_{s'} \wedge \left(\bigwedge_{\tilde{U} \in \tilde{\mathcal{U}} \setminus \tilde{U}'} \varphi_{\tilde{U}} \right) \wedge z_{n+1} = \Psi \wedge z_{n+1}
\end{aligned}$$

□

The proof of the Going Up Lemma I follows by a similar argument, where the observables are augmented by the Z measurement for the $n+1$ -th party.

The Going Up lemmas have the following useful corollary.

Corollary 5. *If $|\theta\rangle$ is logically contextual under some choice of local observables, so are $|\theta\rangle \otimes |\eta\rangle$ and $|\eta\rangle \otimes |\theta\rangle$ for any state $|\eta\rangle$, with the same choice of local observables for each party in $|\theta\rangle$, and a single observable for each party in $|\eta\rangle$.*

Proof. We argue by induction on the number of qubits in $|\eta\rangle$. If $|\eta\rangle = \alpha|0\rangle + \beta|1\rangle$, then by bilinearity of the tensor product,

$$|\theta\rangle \otimes |\eta\rangle = \alpha|\theta\rangle|0\rangle + \beta|\theta\rangle|1\rangle$$

and we can apply the Going Up Lemma I.

For the inductive case, we can write

$$|\eta\rangle = \alpha|\eta_0\rangle|0\rangle + \beta|\eta_1\rangle|1\rangle$$

and by bilinearity

$$|\theta\rangle \otimes |\eta\rangle = \alpha|\theta\rangle|\eta_0\rangle|0\rangle + \beta|\theta\rangle|\eta_1\rangle|1\rangle.$$

By induction hypothesis, $|\theta\rangle|\eta_0\rangle$ and $|\theta\rangle|\eta_1\rangle$ are logically contextual, and we can apply the Going Up Lemma I again to conclude. \square

4.4. The Small Difference lemma

Lemma 6 (The Small Difference Lemma). *Let $|\psi\rangle$ and $|\phi\rangle$ be states in \mathcal{P}_n , and suppose that for all $a \in [0, 1]$, $\tau(a)$ is in \mathcal{P}_n , where $\tau(a)$ is defined by (3). Then $|\psi\rangle$ and $|\phi\rangle$ differ in at most one qubit.*

Proof. Firstly, we note that each state $|\psi\rangle$ in \mathcal{P}_n has an **entanglement type**, which can be described by a graph on n vertices with an edge from i to j when the corresponding qubits of $|\psi\rangle$ are maximally entangled. There are finitely many such graphs, and we can partition \mathcal{P}_n into P_1, \dots, P_M according to the entanglement type. All the states in each P_l are LU equivalent.

As before, we can write $|\psi\rangle$, up to permutation of tensor factors, as

$$|\psi\rangle = |\psi_1\rangle \otimes \dots \otimes |\psi_k\rangle \tag{8}$$

where each $|\psi_i\rangle$ is either a 1-qubit state, or a 2-qubit maximally entangled state.

We recall the definition of the parameterised family of states $\tau(a)$, $a \in [0, 1]$:

$$\tau(a) = a|\psi\rangle + g(a)|\phi\rangle$$

where $g(a) = \sqrt{1 - a^2}$.

Now for each $\delta \in [0, 1]$, we define a set

$$R_\delta = \{\tau(a) \mid 1 - \delta < a \leq 1\}$$

Under our assumption on $\tau(a)$, each set R_δ , which is infinite, is partitioned among the sets P_1, \dots, P_M . Also, $\delta < \delta'$ implies $R_\delta \supset R_{\delta'}$. Hence, by an application of König's infinity lemma [18], we can conclude that there is l with $1 \leq l \leq M$, and an infinite increasing sequence $\{a_i\}$ with supremum 1, such that $\tau(a_i)$ is in P_l for all i .

Since all the states $\tau(a_i)$ are LU-equivalent, we can express them in terms of a representative state $|\Theta\rangle \in P_l$ as

$$\tau(a_i) = U_{a_i}^1 \otimes U_{a_i}^2 \otimes \dots \otimes U_{a_i}^n |\Theta\rangle \quad (9)$$

Since $\mathbf{U}(2)^n$ is compact, there is a convergent subsequence $\{U_{b_i}^1 \otimes \dots \otimes U_{b_i}^n\}_{b_i}$, whose limit as $i \rightarrow \infty$ is $W^1 \otimes \dots \otimes W^n$. The limit of the corresponding subsequence $\{a_{b_i}\}$ is still 1. Hence $|\psi\rangle$ is also a member of P_l , as

$$|\psi\rangle = \lim_{i \rightarrow \infty} \tau(a_i) = \lim_{i \rightarrow \infty} \tau(a_{b_i}) = W^1 \otimes \dots \otimes W^n |\Theta\rangle.$$

This, together with Equation (9), implies that we can express each $\tau(a_i)$ as

$$\tau(a_i) = [(U_{a_i}^1 W^{1\dagger}) \otimes \dots \otimes (U_{a_i}^n W^{n\dagger})] |\psi\rangle.$$

Equivalently, using Equation 8 and the definition of $\tau(a_i)$, we can obtain a family of equations, one for each a_i :

$$\begin{aligned} a_i |\psi\rangle + g(a_i) |\phi\rangle &= [(U_{a_i}^1 W^{1\dagger}) \otimes \dots \otimes (U_{a_i}^n W^{n\dagger})] |\psi\rangle \\ &= |\psi_1^i\rangle \otimes \dots \otimes |\psi_k^i\rangle \end{aligned} \quad (10)$$

where $|\psi_j^i\rangle$ is the state obtained after the LU transformation of $|\psi_j\rangle$, $1 \leq j \leq k$.

We shall now make use of the “partial inner product” operation described e.g. in [14, p. 129]. We will use this operation to probe the components of (10).

By Lemma 10, if for all j , and for any state $|\psi_j^\perp\rangle$ orthogonal to $|\psi_j\rangle$, the application of $|\psi_j^\perp\rangle$ to $|\phi\rangle$ results in a null vector, we must have $|\phi\rangle = |\psi\rangle$, and the lemma is proved.

Otherwise, assume there is some j , and some $|\psi_j^\perp\rangle$, such that $\langle \psi_j^\perp | \phi \rangle$ is a non-zero vector. For ease of notation, we take $j = k$.

Applying $\langle \psi_k^\perp |$ on both sides of Equation (10), we obtain

$$g(a_i) \langle \psi_k^\perp | \phi \rangle = |\psi_1^i\rangle \otimes \dots \otimes |\psi_{k-1}^i\rangle \underbrace{\langle \psi_k^\perp | \psi_k^i \rangle}_{\neq 0} \quad (11)$$

$$\underbrace{\langle \psi_k^\perp | \phi \rangle}_{\text{constant vector}} = \epsilon_i |\psi_1^i\rangle \otimes \dots \otimes |\psi_{k-1}^i\rangle \quad (12)$$

The fact that the LHS of (12) is constant implies that for any i and j we must have $\epsilon_i = \epsilon_j$. We write ϵ for this common value. We must also have $|\psi_t^i\rangle = |\psi_t^j\rangle$ for all $t < k$, and we write $|\psi_t'\rangle$ for the common value. In fact we have

$$|\psi_t\rangle = \lim_{i \rightarrow \infty} |\psi_t^i\rangle = |\psi_t'\rangle, \quad 1 \leq t < k \quad (13)$$

This means that we can rewrite Equation (12) as

$$\langle \psi_k^\perp | \phi \rangle = \epsilon |\psi_1\rangle \otimes \dots \otimes |\psi_{k-1}\rangle.$$

A similar analysis will apply to any state $|\eta\rangle$ orthogonal to $|\psi_k\rangle$ for which $\langle \eta | \phi \rangle$ is non-zero. Using Equation (6.48) from [14], and bilinearity of the tensor product, we obtain

$$|\phi\rangle = x |\phi_2\rangle \otimes |\psi_k\rangle + \epsilon |\psi_1\rangle \otimes \dots \otimes |\psi_{k-1}\rangle \otimes |\xi\rangle \quad (14)$$

for some $|\phi_2\rangle$ and $|\xi\rangle$. We can use Equation (13) to rewrite Equation (10) as

$$a_i |\psi\rangle + g(a_i) |\phi\rangle = |\psi_1\rangle \otimes \dots \otimes |\psi_{k-1}\rangle \otimes |\psi_k^i\rangle.$$

For any $j < k$, if we apply any state $|\psi_j^\perp\rangle$, orthogonal to $|\psi_j\rangle$, to the above equation we obtain $\langle \psi_j^\perp | \phi \rangle = 0$. Together with Equation (14), this implies that $|\phi_2\rangle = |\psi_1\rangle \otimes \dots \otimes |\psi_{k-1}\rangle$. So $|\phi\rangle$ and $|\psi\rangle$ can differ by at most one component.

These components cannot be two-qubit maximally entangled states, since by assumption all linear combinations $\tau(a)$ of $|\psi\rangle$ and $|\phi\rangle$, with $a \in [0, 1]$, belong to \mathcal{P}_n , and hence, using bilinearity again, the corresponding linear combinations of these components would also have to be maximally entangled, yielding a contradiction.

Thus we conclude that $|\psi\rangle$ and $|\phi\rangle$ can differ at most in a one-qubit component. \square

4.5. The 21 lemma

We shall need some elementary facts about partial traces (see e.g. [19]):

- A pure state in $\mathcal{H} \otimes \mathcal{K}$ is a product state if and only if tracing out over \mathcal{H} results in a pure state.
- Tracing out over one party of a maximally entangled bipartite state yields a maximally mixed state.
- A mixed state ρ is pure if and only if $\text{Tr} \rho^2 = 1$.

Lemma 7 (The 21 Lemma). *With the same notation as in the Small Difference Lemma, suppose that $\tau(a)$ is in \mathcal{P}_n for 21 distinct values of a in $[0, 1]$. Then $\tau(a)$ is in \mathcal{P}_n for all $a \in [0, 1]$.*

Proof. If a state belongs to \mathcal{P}_n then all partial traces over $n-1$ parties result either in a pure state or in the maximally mixed state $\frac{1}{2}I_2$.

We can express $|\psi\rangle$ and $|\phi\rangle$ as

$$\begin{aligned} |\psi\rangle &= \sum_{\sigma_i} a_{\sigma_i}^0 |\sigma_i\rangle |0\rangle + \sum_{\sigma_i} a_{\sigma_i}^1 |\sigma_i\rangle |1\rangle \\ |\phi\rangle &= \sum_{\sigma_i} b_{\sigma_i}^0 |\sigma_i\rangle |0\rangle + \sum_{\sigma_i} b_{\sigma_i}^1 |\sigma_i\rangle |1\rangle \end{aligned}$$

where the σ_i index the elements of the computational basis on $n-1$ qubits.

This means we can write the density matrix corresponding to $|\tau(a)\rangle = a|\phi\rangle + b|\psi\rangle$, with $b = g(a)$, as

$$\begin{aligned} |\tau(a)\rangle\langle\tau(a)| &= \sum_{\sigma_i, \sigma_j} (aa_{\sigma_i}^0 + bb_{\sigma_i}^0)(\overline{aa_{\sigma_j}^0} + \overline{bb_{\sigma_j}^0}) |\sigma_i\rangle\langle\sigma_j| \otimes |0\rangle\langle 0| \\ &+ \sum_{\sigma_i, \sigma_j} (aa_{\sigma_i}^0 + bb_{\sigma_i}^0)(\overline{aa_{\sigma_j}^1} + \overline{bb_{\sigma_j}^1}) |\sigma_i\rangle\langle\sigma_j| \otimes |0\rangle\langle 1| \\ &+ \sum_{\sigma_i, \sigma_j} (aa_{\sigma_i}^1 + bb_{\sigma_i}^1)(\overline{aa_{\sigma_j}^0} + \overline{bb_{\sigma_j}^0}) |\sigma_i\rangle\langle\sigma_j| \otimes |1\rangle\langle 0| \\ &+ \sum_{\sigma_i, \sigma_j} (aa_{\sigma_i}^1 + bb_{\sigma_i}^1)(\overline{aa_{\sigma_j}^1} + \overline{bb_{\sigma_j}^1}) |\sigma_i\rangle\langle\sigma_j| \otimes |1\rangle\langle 1| \end{aligned}$$

The partial trace over the first $n - 1$ qubits of $|\tau(a)\rangle$ is given by

$$\begin{aligned}
\rho_n &= \text{Tr}_{n-1} |\tau(a)\rangle \langle \tau(a)| = \sum_{\sigma_i} \langle \sigma_i | \tau(a) \rangle \langle \tau(a) | \sigma_i \rangle \\
&= \sum_{\sigma_i} (aa_{\sigma_i}^0 + bb_{\sigma_i}^0)(a\overline{a}_{\sigma_i}^0 + b\overline{b}_{\sigma_i}^0) |0\rangle \langle 0| + \sum_{\sigma_i} (aa_{\sigma_i}^0 + bb_{\sigma_i}^0)(a\overline{a}_{\sigma_i}^1 + b\overline{b}_{\sigma_i}^1) |0\rangle \langle 1| \\
&\quad + \sum_{\sigma_i} (aa_{\sigma_i}^1 + bb_{\sigma_i}^1)(a\overline{a}_{\sigma_i}^0 + b\overline{b}_{\sigma_i}^0) |1\rangle \langle 0| + \sum_{\sigma_i} (aa_{\sigma_i}^1 + bb_{\sigma_i}^1)(a\overline{a}_{\sigma_i}^1 + b\overline{b}_{\sigma_i}^1) |1\rangle \langle 1| \\
&= |0\rangle \langle 0| \sum_{\sigma_i} (a^2 a_{\sigma_i}^0 \overline{a}_{\sigma_i}^0 + (1 - a^2) b_{\sigma_i}^0 \overline{b}_{\sigma_i}^0 + a\sqrt{1 - a^2} (a_{\sigma_i}^0 \overline{b}_{\sigma_i}^0 + b_{\sigma_i}^0 \overline{a}_{\sigma_i}^0)) \\
&\quad + |0\rangle \langle 1| \sum_{\sigma_i} (a^2 a_{\sigma_i}^0 \overline{a}_{\sigma_i}^1 + (1 - a^2) b_{\sigma_i}^0 \overline{b}_{\sigma_i}^1 + a\sqrt{1 - a^2} (a_{\sigma_i}^0 \overline{b}_{\sigma_i}^1 + b_{\sigma_i}^0 \overline{a}_{\sigma_i}^1)) \\
&\quad + |1\rangle \langle 0| \sum_{\sigma_i} (a^2 a_{\sigma_i}^1 \overline{a}_{\sigma_i}^0 + (1 - a^2) b_{\sigma_i}^1 \overline{b}_{\sigma_i}^0 + a\sqrt{1 - a^2} (a_{\sigma_i}^1 \overline{b}_{\sigma_i}^0 + b_{\sigma_i}^1 \overline{a}_{\sigma_i}^0)) \\
&\quad + |1\rangle \langle 1| \sum_{\sigma_i} (a^2 a_{\sigma_i}^1 \overline{a}_{\sigma_i}^1 + (1 - a^2) b_{\sigma_i}^1 \overline{b}_{\sigma_i}^1 + a\sqrt{1 - a^2} (a_{\sigma_i}^1 \overline{b}_{\sigma_i}^1 + b_{\sigma_i}^1 \overline{a}_{\sigma_i}^1))
\end{aligned}$$

The partial trace ρ_n is equal to the maximally mixed state if and only if

$$\begin{aligned}
1/2 &= \sum_{\sigma_i} (a^2 a_{\sigma_i}^0 \overline{a}_{\sigma_i}^0 + (1 - a^2) b_{\sigma_i}^0 \overline{b}_{\sigma_i}^0 + a\sqrt{1 - a^2} (a_{\sigma_i}^0 \overline{b}_{\sigma_i}^0 + b_{\sigma_i}^0 \overline{a}_{\sigma_i}^0)) \\
0 &= \sum_{\sigma_i} (a^2 a_{\sigma_i}^0 \overline{a}_{\sigma_i}^1 + (1 - a^2) b_{\sigma_i}^0 \overline{b}_{\sigma_i}^1 + a\sqrt{1 - a^2} (a_{\sigma_i}^0 \overline{b}_{\sigma_i}^1 + b_{\sigma_i}^0 \overline{a}_{\sigma_i}^1)) \\
0 &= \sum_{\sigma_i} (a^2 a_{\sigma_i}^1 \overline{a}_{\sigma_i}^0 + (1 - a^2) b_{\sigma_i}^1 \overline{b}_{\sigma_i}^0 + a\sqrt{1 - a^2} (a_{\sigma_i}^1 \overline{b}_{\sigma_i}^0 + b_{\sigma_i}^1 \overline{a}_{\sigma_i}^0)) \\
1/2 &= \sum_{\sigma_i} (a^2 a_{\sigma_i}^1 \overline{a}_{\sigma_i}^1 + (1 - a^2) b_{\sigma_i}^1 \overline{b}_{\sigma_i}^1 + a\sqrt{1 - a^2} (a_{\sigma_i}^1 \overline{b}_{\sigma_i}^1 + b_{\sigma_i}^1 \overline{a}_{\sigma_i}^1))
\end{aligned}$$

Each of these equations yields a polynomial of degree 4 in a . Indeed, each equation has the form

$$cab + q(a) = d$$

where c and d are constants, $b = \sqrt{1 - a^2}$, and $q(a)$ is a quadratic polynomial in a . We can write this as

$$cab = -q(a) + d$$

and square both sides to obtain

$$c^2 a^2 (1 - a^2) = (-q(a) + d)^2$$

which is a quartic polynomial in a . Hence, there can be at most 4 values of a in $[0, 1]$ for which the partial trace ρ_n is equal to a maximally mixed state.

On the other hand, ρ_n is equal to a pure state if and only if $\text{Tr}\rho_n^2 = 1$. By a similar analysis, this condition turns out to be equivalent to a polynomial equation of degree 16 in a . Unless the polynomial is degenerate, *i.e.* the coefficients cancel so that the equation reduces to $1 = 1$, there can be at most 16 values of a for which the partial trace ρ_n is equal to a pure state.

Therefore, if there are more than $4 + 16 = 20$ values of a in $[0, 1]$ for which the linear combination $\tau(a)$ belongs to \mathcal{P}_n , we can conclude that one of the polynomial equations above was degenerate, hence $\tau(a) \in \mathcal{P}_n$ for all values of a . \square

Remark. If $|\psi\rangle$ and $|\phi\rangle$ differ by one qubit, the partial trace of $\tau(a)$ for any value of a will yield a pure state, and hence we will always have $\text{Tr}\rho_n^2 = 1$. Thus the polynomial will indeed be degenerate in this case. This shows the necessity for the Small Difference Lemma.

5. The algorithm

We now give an explicit, albeit informal description of the algorithm which follows straightforwardly from our results.

We begin with a subroutine which we will use to test if a state is in \mathcal{P}_n .

SUBROUTINE **Test** \mathcal{P}_n

Input n -qubit quantum state $|\theta\rangle$

Output Either

Yes, and entanglement type of $|\theta\rangle$, or

No

1. Compute the $n - 1$ partial traces ρ_i over $n - 1$ qubits of $|\theta\rangle$. If any ρ_i is not a maximally mixed state, compute $\text{Tr}\rho_i^2$. If $\text{Tr}\rho_i^2 \neq 1$, return **No**. We now have the list $\{i_1, \dots, i_k\}$ of indices for which the maximally mixed state was returned.
2. For each i_p in the list, find its “partner” i_q by computing the partial traces ρ_{i_p, i_q} over $n - 2$ qubits, and then testing if $\text{Tr}\rho_{i_p, i_q}^2 = 1$. If we cannot find the partner for some i_p , return **No**.
3. Otherwise, we return **Yes**. We also have the complete entanglement type of $|\theta\rangle$, and we have computed all the single-qubit components. \square

ALGORITHM

Input An n -qubit state $|\omega\rangle$
Output Either
 Yes if $|\omega\rangle$ is logically contextual,
 together with a list of $n + 2$ local observables, or
 No if $|\omega\rangle$ is in \mathcal{P}_n .

Base Cases

1. If $n = 1$, output **No**.
2. If $n = 2$, apply the Hardy procedure of the Base Case Lemma to the Schmidt decomposition of $|\omega\rangle$.

Recursive Case: $n + 1, n > 1$

1. We apply $\text{Test}\mathcal{P}_{n+1}$ to $|\omega\rangle$. If $|\omega\rangle$ is in \mathcal{P}_{n+1} , return **No**.
2. Otherwise, we write

$$|\omega\rangle = \alpha|\psi\rangle|0\rangle + \beta|\phi\rangle|1\rangle.$$

Explicitly, if $|\omega\rangle$ is represented by a 2^{n+1} -dimensional complex vector

$$\sum_{\sigma \in \{0,1\}^{n+1}} a_{\sigma} |\sigma\rangle$$

in the computational basis, we can define

$$\alpha = \sqrt{\sum_{\sigma \in \{0,1\}^n} |a_{\sigma 0}|^2}, \quad \beta = \sqrt{\sum_{\sigma \in \{0,1\}^n} |a_{\sigma 1}|^2}$$

$$|\psi\rangle = \frac{1}{\alpha} \sum_{\sigma \in \{0,1\}^n} a_{\sigma 0} |\sigma\rangle, \quad |\phi\rangle = \frac{1}{\beta} \sum_{\sigma \in \{0,1\}^n} a_{\sigma 1} |\sigma\rangle.$$

3. We apply $\text{Test}\mathcal{P}_n$ to $|\psi\rangle$. If $|\psi\rangle$ is not in \mathcal{P}_n , we proceed recursively with $|\psi\rangle$, and then extend the observables using the construction of the Going Up Lemma I.
4. Otherwise, we proceed similarly with $|\phi\rangle$.

5. Otherwise, both $|\psi\rangle$ and $|\phi\rangle$ are in \mathcal{P}_n .

For a in $(0, 1)$, we define

$$\tau(a) := a|\psi\rangle + \sqrt{1-a^2}|\phi\rangle.$$

For 19 distinct values in $(0, 1)$, we assign these values to a , and apply $\text{Test}\mathcal{P}_n$ to $\tau(a)$.

If we find a value of a for which $\tau(a)$ is not in \mathcal{P}_n , we use that value to compute the local observable $B(\frac{\alpha}{a}, \frac{\beta}{\sqrt{1-a^2}})$ for the $n+1$ -th party, as specified in the Going Up Lemma II, and continue the recursion with the n -qubit state $\tau(a)$.

6. Otherwise, by the 21 Lemma and the Small Difference Lemma, the only remaining case is where $|\psi\rangle$ and $|\phi\rangle$ differ in one qubit. We have these qubits $|\psi_1\rangle$, $|\phi_1\rangle$ from our previous applications of $\text{Test}\mathcal{P}_n$. In this final case, we can write $|\omega\rangle$ as

$$|\omega\rangle = |\Psi\rangle \otimes |\xi\rangle$$

where $|\Psi\rangle$ is in \mathcal{P}_{n-1} , and $|\xi\rangle$ is a 2-qubit state. Moreover, we have

$$|\xi\rangle = \alpha|\psi_1\rangle|0\rangle + \beta|\phi_1\rangle|1\rangle.$$

7. We apply the Base Case procedure to $|\xi\rangle$, which we know cannot be maximally entangled, by Step 1. We output **Yes**, together with the two local observables for each party produced by the Hardy construction, and the $n-2$ local observables for $|\Psi\rangle$ produced by the Corollary to the Going Up lemmas. \square

The above algorithm of course involves computation over the real and complex numbers. More precisely, with the usual coding of complex numbers as pairs of reals, we require the field operations and comparison tests on real numbers. For simplicity, we discuss the complexity of the algorithm in the Blum-Shub-Smale model of computation [20], where we assume that arbitrary real numbers can be stored, and the above operations performed, with unit cost. Thus the input size of an n -qubit state is the dimension $d = 2^n$.

The $\text{Test}\mathcal{P}_n$ subroutine performs $n-1$ partial traces over $n-1$ qubits. Each such partial trace involves computing the 4 entries of a matrix, where each entry is a sum over 2^{n-1} products. It also computes $O(n^2)$ partial traces over $n-2$ qubits, each of which involves computing 16 entries, each a sum over 2^{n-2} products. For an n -qubit input, at each level of the recursion,

we call the subroutine a number of times bounded by a constant, and the recursion terminates in $O(n)$ steps. Thus we obtain a complexity bound of $O(d \log^3 d)$ operations, in the input size d . Of course, in practice the limiting factor is the exponential size of the classical representation of a quantum state.

The algorithm has been implemented in MathematicaTM, and has been tested on input states of up to 10 qubits.

6. Final Remarks

Our results provide a complete answer to the question we posed in the Introduction, as to which quantum states give rise to logical contextuality, for the case of pure n -qubit states. There are natural generalisations to qudits and mixed states, which we leave to future work.

Given the paucity of structural results on multipartite entanglement, we feel that the perspective offered by the question of the highest degree of contextuality which states can achieve may be fruitful. This is particularly so in the light of recent results showing that contextuality is a key ingredient enabling quantum computation [21, 22]. The fact that our results apply to **all** multipartite entangled states, with certain bipartite exceptions, is striking. It shows that the probability- and inequality-free, logical formulation of contextuality and non-locality is not a rare occurrence, but in fact arises for almost all states. We also note that these qualitative arguments can easily be turned into quantitative ones. It is shown in [12] how any instance of logical contextuality gives rise to a Bell inequality based on logical consistency conditions, which allows for quantitative, robust experimental tests. Our results also add force to the argument, which we have made elsewhere [16], that the bipartite case, while by far the most studied in non-locality theory, may actually be anomalous. This certainly proves to be the case for logical contextuality, as our results show.

Moreover, if we consider the notion of **strong contextuality**, also introduced in [10], we obtain additional evidence in this direction. We recall that an empirical model is strongly contextual if the set of formulas corresponding to the supports for the various contexts are not simultaneously satisfiable; that is, there is no assignment whatsoever to all the measurement variables which is consistent with the constraints imposed by the model. This is clearly stronger than logical contextuality, which only requires that there is **some** local joint outcome which cannot be extended to the whole set of variables. Strong contextuality is a natural and salient notion with a number of equivalent characterisations. Again, the bipartite case proves anomalous. In

fact, the only strongly contextual bipartite models are the PR-boxes [23, 10], which are of course not quantum realizable [24]. By contrast, for all $n > 2$, the n -partite GHZ states are strongly contextual [10].

This leads us on to the **main question** which is the natural next challenge following on from the one we have addressed in this paper:

For which quantum states can we find local observables which give rise to a strongly contextual empirical model?

This question remains open, and appears difficult. It has an interesting relation to the question of “All-versus-Nothing” arguments [25], which have recently been studied in the sheaf-theoretic approach [26], and shown to be related to the cohomological witnesses for contextuality previously introduced in [27]. All currently known examples of strong contextuality arising in quantum mechanics come from All-versus-Nothing arguments. Determining whether this is true in general is another challenging problem, which may hold the key to the main question.

References

References

- [1] J. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics* 1 (3) (1964) 195–200.
- [2] J. Clauser, M. Horne, A. Shimony, R. Holt, Proposed experiment to test local hidden-variable theories, *Physical Review Letters* 23 (15) (1969) 880–884.
- [3] D. M. Greenberger, M. A. Horne, A. Zeilinger, Going beyond Bell’s theorem, in: *Bells theorem, quantum theory and conceptions of the universe*, Springer, 1989, pp. 69–72.
- [4] D. Greenberger, M. Horne, A. Shimony, A. Zeilinger, Bell’s theorem without inequalities, *American Journal of Physics* 58 (1990) 1131.
- [5] L. Hardy, Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories, *Physical Review Letters* 68 (20) (1992) 2981.
- [6] L. Hardy, Nonlocality for two particles without inequalities for almost all entangled states, *Physical Review Letters* 71 (11) (1993) 1665–1668.

- [7] A. Cabello, Bell's theorem without inequalities and without probabilities for two observers, *Physical review letters* 86 (10) (2001) 1911–1914.
- [8] Z. Wang, D. Markham, Nonlocality of symmetric states, *Physical Review Letters* 108 (21) (2012) 210407.
- [9] S. Mansfield, T. Fritz, Hardy's non-locality paradox and possibilistic conditions for non-locality, *Foundations of Physics* 42 (5) (2012) 709–719.
- [10] S. Abramsky, A. Brandenburger, The sheaf-theoretic structure of non-locality and contextuality, *New Journal of Physics* 13 (2011) 113036.
- [11] J. Zimba, R. Penrose, On Bell non-locality without probabilities: more curious geometry, *Studies in History and Philosophy of Science Part A* 24 (5) (1993) 697–720.
- [12] S. Abramsky, L. Hardy, Logical Bell inequalities, *Physical Review A* 85 (6) (2012) 062114.
- [13] A. Fine, Hidden variables, joint probability, and the Bell inequalities, *Physical Review Letters* 48 (5) (1982) 291–295.
- [14] B. Schumacher, M. Westmoreland, *Quantum Processes, Systems and Information*, Cambridge University Press, 2010.
- [15] S. Abramsky, B. Coecke, Categorical quantum mechanics, in: K. Engesser, D. M. Gabbay, D. Lehmann (Eds.), *Handbook of Quantum Logic and Quantum Structures: Quantum Logic*, Elsevier, 2008, pp. 261–324.
- [16] S. Abramsky, C. M. Constantin, A classification of multipartite states by degree of non-locality, in: B. Coecke, M. Hoban (Eds.), *Proceedings 10th International Workshop on Quantum Physics and Logic 2013*, Vol. 171 of *Electronic Proceedings in Theoretical Computer Science*, 2014, pp. 10–25.
- [17] S. Mansfield, The mathematical structure of non-locality and contextuality, Ph.D. thesis, University of Oxford (2013).
- [18] A. Levy, *Basic Set Theory*, Dover Books, 2012.
- [19] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

- [20] L. Blum, M. Shub, S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, *Bulletin (New Series) of the American Mathematical Society* 21 (1) (1989) 1–46.
- [21] R. Raussendorf, Contextuality in measurement-based quantum computation, *Physical Review A* 88 (2) (2013) 022322.
- [22] M. Howard, J. Wallman, V. Veitch, J. Emerson, Contextuality supplies the ‘magic’ for quantum computation, *Nature* 510 (7505) (2014) 351–355.
- [23] R. Lal, A sheaf-theoretic approach to cluster states, private communication (2011).
- [24] S. Popescu, D. Rohrlich, Quantum nonlocality as an axiom, *Foundations of Physics* 24 (3) (1994) 379–385.
- [25] N. Mermin, Simple unified form for the major no-hidden-variables theorems, *Physical Review Letters* 65 (27) (1990) 3373–3376.
- [26] S. Abramsky, R. S. Barbosa, K. Kishida, R. Lal, S. Mansfield, Contextuality, Cohomology and Paradox, arXiv preprint arXiv:1502.03097.
- [27] S. Abramsky, S. Mansfield, R. Soares Barbosa, The cohomology of non-locality and contextuality, in: B. Jacobs, P. Selinger, B. Spitters (Eds.), *Proceedings 8th International Workshop on Quantum Physics and Logic 2011*, Vol. 95 of *Electronic Proceedings in Theoretical Computer Science*, 2012, pp. 1–14. doi:10.4204/EPTCS.95.1.