Strict Hierarchy Among Bell Theorems ⋆

Gilles Brassard

Département d'informatique et de recherche opérationnelle
Université de Montréal
CP 6128, succursale centre-ville
Montréal (Québec), H3C 3J7 Canada

André Allan Méthot

Group of Applied Physics
Université de Genève
Rue de l'École-de-Médecine 20
1211 Genève 4, Switzerland

Abstract

As demonstrated by John Bell, quantum mechanics exhibits correlations in space-like separated bipartite systems that are impossible to reproduce by classical means. There are three levels of "Bell Theorems", depending on which aspects of the quantum correlations can or cannot be reproduced classically. The original "Bell Inequalities" (BI) require a perfect classical simulation of all quantum probabilities. With "Bell Theorems Without Inequalities" (BTWI), we ask the classical simulation to be able to produce precisely the outcomes that could occur according to quantum mechanics, but we do not worry about their exact probabilities. With "Pseudo-telepathy" (PT), we are satisfied if the classical simulation produces only outcomes allowed by quantum mechanics, but not necessarily all of them.

Bell's original proof of BI involved a maximally entangled 2×2 bipartite state such as the singlet state. Hardy proved that BTWI are possible in dimension 2×2, but his construction used a non-maximally entangled state. Here, we prove that no 2×2 maximally entangled state can serve to produce BTWI. Combining this with our earlier result that 2×2 entangled states cannot be used at all for the purpose of PT, it follows a strict hierarchy on the quantum resources that are required to exhibit the various levels of Bell Theorems.

Key words: Bell theorem, pseudotelepathy, nonlocality, POVM.
PACS: 03.65.Ud, 03.67.Mn

1 Introduction: Bell's theorem in theory and practice

In an attempt to prove that quantum mechanics is an incomplete theory, Albert Einstein, Boris Podolsky and Nathan Rosen introduced in 1935 the concept of *entanglement* (without naming it) [13]. They used entanglement to "prove" that some quantities whose measurement produce random results according to quantum mechanics must in fact be determined by what they called "elements of physical reality". Quantum mechanics is incomplete, they argued, because those elements do not "have a counterpart in the physical theory".

This launched a quest for what became known as *local hidden variables* (LHVs), whose existence would allow the predictions of quantum mechanics to be explained in purely local terms. For the better part of three decades, however, this topic was considered by most physicists to be as pointless as counting the number of angels that can dance on a pinhead. Indeed, such a classical theory, if found, would *by definition* be impossible to distinguish *by any experiment whatsoever* from the quantum-mechanical perspective. It was felt that to decide between two such theories, which predict exactly the same observable behaviour, would be better suited for a discussion among philosophers rather than physicists!

The face of the world changed forever when John Bell proved in 1964 that it is impossible for *any* LHV theory to agree with the predictions of quantum mechanics [4]. This opened the door for the possibility of experiments whose outcome would either disagree with quantum mechanics or with any possible LHV explanation of Nature of the sort that Einstein would have so dearly wished to see. The first realistic such experimental setup was proposed in 1969 by John F. Clauser, Michael A. Horne, Abner Shimony and Richard A. Holt (CHSH) [11]. A few years later, Clauser was able to vindicate quantum mechanics in a landmark experiment performed with Stuart J. Freedman [14]. The subsequent sequence of experiments by Alain Aspect, Jean Dalibard, Philippe Grangier and Gérard Roger were even more convincing in obtaining results that agree with the predictions of quantum mechanics in ways that contradict LHV theories [2,3,1]. Several increasingly fascinating experiments have taken place since then.

⋆ A preliminary version of this work has appeared in the Proceedings of the *Second International Conference on Quantum, Nano, and Micro Technologies* [7]. The proof of the main result given there was geometric in nature, whereas it is algebraic here.

Email addresses: brassard@iro.umontreal.ca (Gilles Brassard), andre.methot@gmail.com (André Allan Méthot).

URL: http://www.iro.umontreal.ca/∼brassard/en (Gilles Brassard).

In the CHSH scenario [11], a 2×2 maximally entangled state, such as the singlet state, is shared between two participants traditionally called Alice and Bob. The participants make independent random choices on how to measure their system and they record their result. If Alice and Bob are sufficiently distant and their timing sufficiently precise to make sure that the outcome of Alice's measurement cannot be influenced by the choice of measurement made by Bob (and vice versa) in any *local* realistic theory allowed by classical physics, the predictions of quantum mechanics can be compared with those allowed by classical LHV theories. More precisely, following Bell's footsteps, CHSH studied a specific operator (a formula that involves the mathematical expectation of the outcomes of various experiments performed independently by Alice and Bob). They were able to upper-bound the value of that operator in any LHV model of classical physics. Furthermore, they described a specific set of quantum local measurements for Alice and Bob whose outcome would violate this bound, which came to be known as "the Bell inequality" even though it was discovered by CHSH rather than Bell.

The fact that experimental realizations of the CHSH proposal agree sufficiently well with quantum mechanics to violate Bell's inequality in a way that is statistically significant when we run the experiment many times over gives evidence that no classical LHV theory can account for what can be observed about Nature. To be fair, however, it should be stressed that current experiments suffer from various loopholes that would allow LHV explanations, such as the *locality loophole*: if Alice and Bob are not sufficiently distant or if their timing is not sufficiently precise, a signal travelling at the speed of light emitted by the first system to be measured could reach the other system in time to influence the outcome of its measurement. Equally infamous is the *detection loophole*, which takes hold if the measurement apparatuses available to Alice and Bob are "allowed" to report failure at providing an outcome. The locality loophole was closed for the first time in Ref. [1], but that experiment left the detection loophole unchecked. This being said, those (and other) important experimental considerations are subtle issues that shall henceforth be ignored because they are not directly relevant to this theoretical paper.

The next section gives a formal definition of Bell Inequalities *à la* CHSH. It is then explained that two additional types of scenarios have been proposed, in which less and less is asked from purported LHV theories in their attempts to mimic quantum mechanics: Bell Theorems Without Inequalities and Pseudotelepathy. The question addressed in this paper is spelled out in Section 3: it concerns the *quantum resources* required to give rise to nonclassical correlations that are strong enough to refute local realism according to each type of scenario. We survey various quantum tools used in the rest of the paper in Section 4, such as positive-operator-valued measures, and we introduce our notion of positive and negative hemispheres. In Section 5, we prove our main result: even though the singlet state $|\Psi^-\rangle$ can be used to exhibit

Bell Inequalities, it is not sufficient as a quantum resource to exhibit a Bell Theorem Without Inequalities. Taking advantage of our earlier result [8], it follows a *strict hierarchy among Bell theorems*: the quantum resources that make Pseudotelepathy possible are a strict subset of those that make Bell Theorems Without Inequalities possible, which are in turn a strict subset of those that make Bell Inequalities possible.

2 Three levels of Bell theorems

Let Alice and Bob hold parts of some physical system, which they are going to measure. Let us denote Alice's and Bob's measurements by $x \in X$ and $y \in Y$, respectively, and their outcomes by $a \in A$ and $b \in B$, respectively. Quantum mechanics defines a joint conditional probability distribution $\Pr_{\mathsf{QM}}[a, b \mid x, y]$ on the outcomes given the measurements, whereas any LHV theory defines its own probability distribution $\Pr_{\mathsf{LHV}}[a, b \mid x, y]$. What makes a LHV theory *local* is the existence of some λ (the *hidden* variable) such that the probability for Alice to obtain a depends only on x and λ, but not on Bob's y, and similarly for Bob's probability to obtain b, which depends only on y and λ, but not on Alice's x. Moreover, $\Pr_A[a \mid x, \lambda]$ must be independent from $\Pr_B[b \mid y, \lambda]$ once λ is fixed:

$$\Pr_{\mathsf{LHV}}[a, b \mid x, y, \lambda] = \Pr_A[a \mid x, \lambda] \; \Pr_B[b \mid y, \lambda] \, .$$

In this case, $\Pr_{\mathsf{LHV}}[a, b \mid x, y]$ is given by a weighted summation (or an integral in the case that λ is a continuous variable) of $\Pr_{\mathsf{LHV}}[a, b \mid x, y, \lambda]$ for all possible values of hidden variable λ.

The point of the original Bell Inequalities [4] (BI), including the version due to CHSH [11], is that no LHV theory can define a probability distribution that agrees with quantum mechanics when the bipartite system shared by Alice and Bob and the set of measurements performed by them is chosen appropriately:

$$(\nexists \mathsf{LHV})\, (\forall\, a, b, x, y) \;\; \Pr_{\mathsf{LHV}}[a, b \mid x, y] = \Pr_{\mathsf{QM}}[a, b \mid x, y] \, . \qquad (1)$$

It must be emphasized that a physical bipartite apparatus that exhibits statistical correlations in accordance with $\Pr_{\mathsf{QM}}$ (or sufficiently close to it) can be used to convince an observer that LHV theories are unsustainable *even if that observer is ignorant of quantum mechanics* (provided the experimental setup is designed with sufficient care to make it impossible for Alice and Bob to communicate—thus preventing the locality loophole—and that the quantum apparatus is sufficiently responsive to prevent also the locality loophole, not an easy task indeed!). Yet, despite David Mermin's excellent essay on "Quantum mysteries for anybody" [22], it remains tricky to explain to a non-scientist why no LHV theory could possibly be so that $\Pr_{\mathsf{LHV}} = \Pr_{\mathsf{QM}}$.

A different form of Bell Theorem *Without Inequalities* (BTWI) was discovered in the early 1970s (but unpublished) by Simon Kochen [21], then again in 1983 by Peter Heywood and Michael L. G. Redhead [20], subsequently simplified by Harvey R. Brown and George Svetlichny [9] and independently rediscovered and made much more popular in 1989 by Daniel M. Greenberger, Michael A. Horne and Anton Zeilinger [17] (see also Ref. [16]). In the traditional Bell theorems [4,11], the argumentation was centred on the impossibility to reproduce *exactly* with LHV theories the correlations made possible by quantum mechanics. According to the BTWI paradigm, it suffices to consider which outcomes can or cannot occur, without taking account of the actual probabilities that are involved. In other words, a bipartite system shared by Alice and Bob and a set of measurements performed by them exhibits a BTWI if

$$(\nexists \mathsf{LHV})\,(\forall\, a,b,x,y)\ \ \Pr_{\mathsf{LHV}}[a,b \mid x,y]=0 \ \Leftrightarrow\ \Pr_{\mathsf{QM}}[a,b \mid x,y]=0\,. \qquad (2)$$

When a physical quantum-mechanical bipartite apparatus exhibits statistical correlations in accordance with a BTWI, it is usually much easier to convince a non scientist that the observed phenomenon would be inexplicable in classical terms, compared to an apparatus based on traditional BI.

The most beautiful example of BTWI has been presented by Lucien Hardy [18] in 1992 and popularized by Mermin [23]. The main interest of Hardy's approach is that it can be implemented quantum-mechanically with a single pair of entangled qubits, albeit not in a maximally entangled state. Even though what follows was not explicit in Hardy's paper [18]—but see Ref. [24] for a detailed argument in which "0", "1", "I" and "H" should be substituted for "+", "−", "σ_z" and "σ_x", respectively—a BTWI is obtained if Alice and Bob share the state

$$|\Gamma\rangle = \tfrac{1}{\sqrt{3}}|01\rangle + \tfrac{1}{\sqrt{3}}|10\rangle + \tfrac{1}{\sqrt{3}}|11\rangle$$

and if each one of them chooses at random either to measure his or her qubit in the computational basis ($|0\rangle$ vs. $|1\rangle$) or in the Hadamard basis ($\mathsf{H}|0\rangle = \tfrac{1}{\sqrt{2}}|0\rangle + \tfrac{1}{\sqrt{2}}|1\rangle$ vs. $\mathsf{H}|1\rangle = \tfrac{1}{\sqrt{2}}|0\rangle - \tfrac{1}{\sqrt{2}}|1\rangle$).

Let us denote a measurement in the computational basis by I (for "identity") and a measurement in the Hadamard basis by H, and let us denote the outcome of the measurement by "0" for $|0\rangle$ or $\mathsf{H}|0\rangle$ and by "1" for $|1\rangle$ or $\mathsf{H}|1\rangle$, respectively. Again, we write $\Pr_{\mathsf{QM}}[a,b \mid x,y]$ to denote the probability, according to quantum mechanics, that Alice and Bob get outcomes a and b, respectively, when they perform measurements x and y, respectively. It turns out that

$$\begin{aligned} &\Pr_{\mathsf{QM}}[0,0 \mid I,I\,] &&= 0 \\ &\Pr_{\mathsf{QM}}[1,1 \mid I,H\,] &&= 0 \\ &\Pr_{\mathsf{QM}}[1,1 \mid H,I\,] &&= 0 \\ \text{but } &\Pr_{\mathsf{QM}}[1,1 \mid H,H\,] &&\neq 0\,. \end{aligned}$$

On the other hand, *any* LHV theory for which

$$\begin{aligned} &\Pr_{\mathsf{LHV}}[1,1 \mid I,H\,] = 0 \\ \text{and } &\Pr_{\mathsf{LHV}}[1,1 \mid H,I\,] = 0 \end{aligned}$$

precludes having simultaneously

$$\begin{aligned} &\Pr_{\mathsf{LHV}}[0,0 \mid I,I\,] &&= 0 \\ \text{and } &\Pr_{\mathsf{LHV}}[1,1 \mid H,H\,] &&\neq 0\,. \end{aligned}$$

It follows that

$$(\forall\, a,b,x,y)\;\; \Pr_{\mathsf{LHV}}[a,b \mid x,y] = 0 \;\Leftrightarrow\; \Pr_{\mathsf{QM}}[a,b \mid x,y] = 0$$

is impossible for any LHV theory, and therefore the condition in Equation (2) is fulfilled: this is indeed a case of BTWI, which serves to illustrate how easy it can be to prove that the behaviour observed from deploying an ideal quantum-mechanical pair of gadgets would be impossible to explain in local realistic classical terms.

At the turn of the century, an even stronger type of Bell theorems emerged out of the computer science perspective of Gilles Brassard, Richard Cleve and Alain Tapp [6], which became known as *pseudotelepathy games* (PT). In these games, each player is isolated from the other player (or the other players in the case of a multi-player game) in such a way that it is impossible for them to communicate while the game is in progress. (This can be enforced by requesting very quick answers from the players and putting sufficient distance between them to preclude any attempt at communicating to reach the intended receiver quickly enough—at the speed of light—to carry information that might allow cheating.) The players are then asked questions (say x for Alice and y for Bob) and requested to respond (say a from Alice and b from Bob) in such a way that some nontrivial relation is satisfied between the pair of questions and the pair of responses. Mathematically, the game is characterized by a relation $R \subseteq X \times Y \times A \times B$ and the players win if $(x,y,a,b) \in R$. We speak of pseudotelepathy when the game can be won with certainly between players who share adequate entanglement, whereas it could not be won all the time in a classical world. A comprehensive survey of pseudotelepathy games can be found in Ref. [5].

The connexion between pseudotelepathy games and Bell theorems is that the former is equivalent to having sets X and Y of measurements on each side of some entangled state $|\Psi\rangle$ and sets of outcomes A and B so that

$$(\nexists \mathsf{LHV})\, (\forall\, a,b,x,y)\;\; \Pr_{\mathsf{QM}}[a,b \mid x,y] = 0 \;\Rightarrow\; \Pr_{\mathsf{LHV}}[a,b \mid x,y] = 0\,, \tag{3}$$

where $\Pr_{\mathsf{QM}}[a,b \mid x,y]$ denotes the probability of obtaining outcome (a,b) when quantum state $|\Psi\rangle$ is subject to joint measurement (x,y). Indeed, if we define relation R as

$$R = \{(x,y,a,b) \mid \Pr_{\mathsf{QM}}[a,b \mid x,y] \neq 0\}\,,$$

then any pair of outcomes (a,b) produced with nonzero probability by quantum mechanics given the pair of measurements (x,y) will always be so that $(x,y,a,b) \in R$, as it should, but for any LHV there will be at least one pair of responses (a,b) produced with nonzero probability on some pair of questions (x,y) so that $(x,y,a,b) \notin R$.

3 Statement of the problem

It suffices to inspect the three key equations, (1), (2) and (3), to realize that any manifestation of a PT game is a special case of a BTWI and any manifestation of a BTWI is a special case of a BI. How about the other way round?

The question addressed in this paper concerns the *quantum resources* required to give rise to nonclassical correlations that are strong enough to refute local realism according to each flavour of Bell's theorem. We prove a strict hierarchy among them in the sense that any pure entangled state on two qubits suffices to obtain a BI, whereas such a state must be *non maximally entangled* in order to obtain a BTWI; this is our main theorem. In addition, we had previously proven with Tapp that no entangled state on two qubits can serve to win a PT game [8].

More formally, we define the following three sets of quantum states.

$$\begin{aligned} BI &= \{|\Psi\rangle \mid \text{we can construct a } \mathsf{BI} \text{ with } |\Psi\rangle\}, \\ BTWI &= \{|\Psi\rangle \mid \text{we can construct a } \mathsf{BTWI} \text{ with } |\Psi\rangle\}, \\ PT &= \{|\Psi\rangle \mid \text{we can construct a } \mathsf{PT} \text{ game with } |\Psi\rangle\}. \end{aligned} \tag{4}$$

It is obvious by definition that

$$PT \subseteq BTWI \subseteq BI\,.$$

As our main theorem, we prove $BTWI \neq BI$. Given that we had already proven $PT \neq BTWI$ [8], the strict hierarchy follows:

$$PT \subsetneq BTWI \subsetneq BI\,.$$

Before we proceed, it is important to point out that Hardy had "shown that it is possible to demonstrate nonlocality for two particles without using inequalities for all entangled states except maximally entangled states such as the singlet state" [19]. It may be tempting to think that this is a claim that maximally entangled states cannot be used to obtain a BTWI (which is our main theorem!). However, Hardy does not say this at all. All he says in the penultimate paragraph of Ref. [19] is that "the above nonlocality proof [his own, of course] will not go through" when using maximally entangled states. But the fact that his technique fails is not a proof that there cannot be some other approach that might have been able to obtain a BTWI from a singlet state. And indeed the fact that this is not possible is the main contribution of our paper.

4 Quantum tools

In this section, we review positive-operator-valued measures (POVMs), state some known results and then define and prove some notions that will be useful for our purpose. POVMs are the most general type of measurement allowed by quantum mechanics. A POVM is a collection of *POVM elements*, each of which is a positive matrix M_i, such that $\sum_i M_i = \mathbb{1}$, the identity matrix. When applied on a state ρ, a POVM returns the value i with probability $\Pr[i] = \mathrm{Tr}(M_i \rho)$. In general, there could be a quantum state leftover after the measurement, but this is irrelevant for the purpose of this paper. We now rewrite POVMs in a form proposed in Ref. [10], present the definitions of positive and negative hemispheres and conclude the section with a technical lemma.

Lemma 1 *Any POVM can be rewritten in such a way that all its elements are proportional to one-dimensional projectors.*

Proof. Even though this lemma is proven in Ref. [10], we give a proof below for completeness. Consider a POVM $\{M_i\}$. From the spectral decomposition theorem, each of the M_i's can be written as $M_i = \sum_j b_{ij} P_{ij}$, where the b_{ij}'s are real, $0 < b_{ij} \leq 1$ and each P_{ij} is a one-dimensional projector. We can then construct a new POVM by putting together all the $b_{ij} P_{ij}$ as elements. These new elements are positive matrices and $\sum_{ij} b_{ij} P_{ij} = \sum_i M_i = \mathbb{1}$. To simulate the effect of the original POVM, we interpret the new POVM outcomes as follows: if the outcome ij is obtained when the new POVM is applied, we pretend that the outcome was simply i. □

For notational convenience, we shall henceforth avoid the double indexes needed in the proof of the above lemma and consider that any POVM is already in the form $\{M_i\}$ where each $M_i = \gamma_i P_i$ with $0 < \gamma_i \leq 1$ a real number and P_i a projector.

Furthermore, any projector on a qubit can be represented as a unit vector $\vec{v}$ in three-dimensional real space $\mathbb{R}^3$

$$\vec{v} = (x, y, z) = \Big(\sin(2\theta)\cos(\phi), \sin(2\theta)\sin(\phi), \cos(2\theta)\Big), \tag{5}$$

with $0 \leq \theta \leq \pi/2$ and $0 \leq \phi < 2\pi$, which can be seen as a point on the Bloch sphere. When we describe a POVM as a collection of projectors P_i and positive real numbers γ_i, let $\vec{v}_i$ (for each i) be the point on the Bloch sphere that corresponds to P_i according to Equation (5). It is shown in Ref. [10] that the POVM condition $\sum_i \gamma_i P_i = \mathbb{1}$ is equivalent to $\sum_i \gamma_i \vec{v}_i = \vec{0}$ and $\sum_i \gamma_i = 2$.

If $\vec{v} = (x_1, y_1, z_1)$ and $\vec{w} = (x_2, y_2, z_2)$ are two points in $\mathbb{R}^3$, we denote their dot product (aka inner product, scalar product) as

$$\vec{v} \cdot \vec{w} = x_1 x_2 + y_1 y_2 + z_1 z_2 \, .$$

Recall that $\vec{v}$ and $\vec{w}$ are *orthogonal* (aka perpendicular) if and only if $\vec{v} \cdot \vec{w} = 0$.

Definition 1 (Positive and negative hemispheres) *Let $\vec{p}$, $\vec{q}$ and $\vec{r}$ be any three linearly independent points on the Bloch sphere. We divide the Bloch ball (i.e. the Bloch sphere plus its interior) between the* positive *and the* negative hemispheres *(formally, we should speak of "hemiballs") as follows. The positive hemisphere consist of all the points $\vec{v}$ in the Bloch ball such that one of the following three conditions apply:*

(1) $\vec{p} \cdot \vec{v} > 0$, or
(2) $\vec{p} \cdot \vec{v} = 0$ and $\vec{q} \cdot \vec{v} > 0$, or
(3) $\vec{p} \cdot \vec{v} = 0$ and $\vec{q} \cdot \vec{v} = 0$ and $\vec{r} \cdot \vec{v} > 0$.

The negative hemisphere is defined similarly, replacing "$>$" with "$<$".

A geometric definition of the positive and negative hemispheres, as well as a geometric demonstration of Lemma 2 below, is given in the preliminary version of this paper [7]. Here, we favour an algebraic treatment. Note that each point in the Bloch ball belongs to either the positive or the negative hemisphere (never both), except for the origin $\vec{0}$, which belongs to neither. This is because the only way for a point $\vec{v}$ in the Bloch ball to belong to neither hemisphere is if $\vec{p} \cdot \vec{v} = \vec{q} \cdot \vec{v} = \vec{r} \cdot \vec{v} = 0$, meaning that $\vec{v}$ is simultaneously orthogonal to $\vec{p}$, $\vec{q}$ and $\vec{r}$, which is only possible in $\mathbb{R}^3$ for $\vec{v} = \vec{0}$ because $\vec{p}$, $\vec{q}$ and $\vec{r}$ are linearly independent.

Lemma 2 *Any POVM whose elements are proportional to projectors contains at least one element in the positive hemisphere and one element in the negative hemisphere.*

Proof. Let $\{\gamma_i \vec{v}_i\}$ be the elements of the POVM and recall that $\sum_i \gamma_i \vec{v}_i = \vec{0}$ and $0 < \gamma_i \leq 1$ for each i. By bilinearity of the dot product,

$$\sum_i \gamma_i (\vec{p} \cdot \vec{v}_i) = \vec{p} \cdot \left(\sum_i \gamma_i \vec{v}_i \right) = \vec{p} \cdot \vec{0} = 0. \tag{6}$$

Because the γ_i's are strictly positive, this is only possible if (1) there exist j and k such that $\vec{p} \cdot \vec{v}_j > 0$ and $\vec{p} \cdot \vec{v}_k < 0$, or (2) $\vec{p} \cdot \vec{v}_i = 0$ for each i. In case (1), $\gamma_j \vec{v}_j$ belongs to the positive hemisphere and $\gamma_k \vec{v}_k$ belongs to the negative hemisphere, which proves the lemma. Case (2) can only occur with probability 0 if $\vec{p}$ is chosen at random (as it will be in the proof of Theorem 1). Let us nevertheless consider this case for completeness. For a reason similar to Equation (6), $\sum_i \gamma_i (\vec{q} \cdot \vec{v}_i) = 0$. Again, this implies that either (2.1) there exist j and k such that $\vec{q} \cdot \vec{v}_j > 0$ and $\vec{q} \cdot \vec{v}_k < 0$, or (2.2) $\vec{q} \cdot \vec{v}_i = 0$ for each i. In case (2.1), $\gamma_j \vec{v}_j$ belongs to the positive hemisphere (because we also have $\vec{p} \cdot \vec{v}_j = 0$) and $\gamma_k \vec{v}_k$ belongs to the negative hemisphere, which proves the lemma. In case (2.2), $\sum_i \gamma_i (\vec{r} \cdot \vec{v}_i) = 0$. This time, however, none of the $\vec{r} \cdot \vec{v}_i$ can be 0 because $\vec{p}$, $\vec{q}$ and $\vec{r}$ are linearly independent, $\vec{p} \cdot \vec{v}_i = \vec{q} \cdot \vec{v}_i = 0$ and the $\vec{v}_i$'s are not $\vec{0}$. The only possibility that remains is that there exist j and k such that $\vec{r} \cdot \vec{v}_j > 0$ and $\vec{r} \cdot \vec{v}_k < 0$. In that case, $\gamma_j \vec{v}_j$ belongs to the positive hemisphere (because we also have $\vec{p} \cdot \vec{v}_j = \vec{q} \cdot \vec{v}_j = 0$) and $\gamma_k \vec{v}_k$ belongs to the negative hemisphere, which proves the lemma. To summarize, we have established in every possible case the existence of some $\gamma_j \vec{v}_j$ in the positive hemisphere and some other $\gamma_k \vec{v}_k$ in the negative hemisphere, as required. □

5 A strict hierarchy of Bell theorems

Theorem 1 $|\Psi^-\rangle = \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle \notin BTWI$.

Proof. To prove that $|\Psi^-\rangle \notin BTWI$, we must design an explicit LHV model for Alice and Bob with the property that, whenever they receive the description $\{\vec{a}_i\}$ (for Alice) and $\{\vec{b}_i\}$ (for Bob) of POVMs to pretend they apply on their (pretended) halves of a $|\Psi^-\rangle$ state, they can simultaneously

(1) avoid producing a forbidden outcome, and
(2) generate with a non-zero probability each non-forbidden outcome.

By "forbidden", we mean an outcome that would never be produced if Alice and Bob had a shared $|\Psi^-\rangle$ at their disposal and if they chose their outcome by actually performing the required measurements on their respective shares.

Let us first establish the correlations of $|\Psi^-\rangle$, in particular which outcomes are forbidden according to quantum mechanics. If Alice measures her half of the state according to $\{\vec{a}_i\}$ and Bob to $\{\vec{b}_j\}$, the joint probability distribution is

$$\Pr[i,j] = \frac{|\vec{a}_i|\,|\vec{b}_j| - \vec{a}_i \cdot \vec{b}_j}{2}. \tag{7}$$

Consequently, the only zeros in the probability distribution come when Alice's POVM element points in the same direction as Bob's.

We are now ready to present our LHV model. Alice and Bob agree on the following strategy: while they are still together, they choose at random linearly independent vectors $\vec{p}$, $\vec{q}$ and $\vec{r}$ on the Bloch sphere, thus defining a positive and a negative hemispheres (according to Definition 1). Then, they are separated and will no longer be able to communicate. Nevertheless, their random division of the Bloch ball makes them ready to simulate the quantum correlations, where "simulate" has to be taken with its BTWI meaning, according to points (1) and (2) as spelled out at the beginning of this proof.

Upon receiving a description of the POVM $\{\vec{a}_i\}$ that she is to simulate (where each $\vec{a}_i$ is proportional to a projector), Alice selects at random some $\vec{a}_i$ in the positive hemisphere. Similarly, upon reception of $\{\vec{b}_j\}$, Bob selects at random some $\vec{b}_j$ in the *negative* hemisphere. This is always possible for them because we have proved in Lemma 2 that such elements always exist. Having made those choices independently, Alice and Bob produce outcomes i and j, respectively.

It is obvious from the construction of the model that it will never produce elements pointing in the same direction, which would be in the same hemisphere, since the outcomes of Alice and Bob correspond systematically to POVM elements that are in distinct hemispheres. Consequently, our classical LHV construction never produces an outcome that has a zero probability of occurring according to quantum mechanics. However, we still need to prove that *every* outcome possible according to quantum mechanics has nonzero probability of being produced.

This proof is also simple and intuitive. Consider any element $\vec{a}_i$ in Alice's POVM and any element $\vec{b}_j$ in Bob's, so that $\vec{a}_i$ and $\vec{b}_j$ do not point in the same direction. It is easy to see that if $\vec{p}$ is chosen at random on the Bloch sphere, the probability that $\vec{p} \cdot \vec{a}_i > 0$ and $\vec{p} \cdot \vec{b}_j < 0$ is nonzero. If such a $\vec{p}$ is chosen by Alice and Bob at the onset of the classical process, there is a nonzero probability that Alice and Bob will independently select those $\vec{a}_i$ and $\vec{b}_j$ among their POVM elements that lie in the positive and negative hemispheres, respectively. Hence, it is possible for the pair (i,j) to be produced as outcome, as it should. □

Corollary 1 *$PT \subsetneq BTWI \subsetneq BI$.*

Proof. We already know from Ref. [8] that $PT \subsetneq BTWI$ and it follows immediately from Equation (4) that $BTWI \subseteq BI$. Hence, all that remains to prove is that $BTWI \neq BI$. But it has been known since Bell's seminal paper [4] that $|\Psi^-\rangle \in BI$, which allows us to conclude thanks to Theorem 1. □

6 Conclusions

We have proved that a maximally entangled state on a pair of qubits, while non-local enough to yield a violation of a Bell inequality, cannot be used to derive a Bell Theorem Without Inequalities. This shows that BTWI provide a stronger refutation of local realism than the usual versions of Bell's theorem, including CHSH [11], since entangled states that can be used to exhibit the latter are not necessarily capable of exhibiting the former:

$$BTWI \subsetneq BI\,.$$

The picture is thus complete since we already knew [8] that

$$PT \subsetneq BTWI\,.$$

It should be pointed out that the above inclusion signs have a semantically different meaning than in the usual computational complexity theory. For example, we know that *P* is included in *NP* and that, should the inclusion ever be proved to be strict, *NP* will have been shown to be strictly more powerful than *P*. In our notation, a strict inclusion of *BTWI* in *BI* means that there exists a quantum state that yields correlations strong enough to construct a Bell theorem around them, while no construction of a Bell theorem without inequalities can exist with this same state. It follows that Bell theorems without inequalities are harder to satisfy, from a non-local correlation point of view, and are therefore a stronger refutation of local realism.

The technique used in the proof of Theorem 1 is surprising, since we exhibit a structure of correlations of non-maximally entangled states of two qubits that cannot be found in its maximally entangled counterpart. This result is akin to the discussion in Ref. [25] where the authors discuss many similar cases. This technique further emphasizes that non-maximally entangled states are harder to simulate by LHVs than their maximally entangled counterpart. The simple, albeit stunning, conclusion from this fact is that non-maximally entangled states are more non-local than maximally entangled states [25], at least in this sense!

It is interesting to note that the proofs given by Heywood and Redhead [20] and by Greenberger, Horne and Zeilinger [17], as most of the early proofs of Bell theorems without inequalities, while presented in Section 2 as historical examples of BTWI, are in fact pseudotelepathy games. However, the authors of those papers were not aware that they were actually constructing something stronger than what they were claiming to have achieved in their papers. The only construction known to the authors in which a Bell theorem without inequalities is genuinely not a pseudotelepathy game is the one given by Hardy [18], which was inevitable in the light of Ref. [8] since his construction used a single pair of entangled qubits.

As this research shows, a detailed study of the different no-go theorems can teach us fundamental truths about Nature. We know that entangled pairs of qubits are sufficient to create correlations for Bell theorems with or without inequalities, whereas we need at least an entangled pair of *qutrits* or *three* entangled qubits [8] for pseudotelepathy to take hold. We also know that bipartite binary measurements are sufficient for Bell theorems with or without inequalities [18], whereas we need at least one ternary measurement or a third participant for pseudotelepathy [12]. Furthermore, two measurement settings (i.e. two different questions) per participant suffice to give rise to Bell theorems with or without inequalities, whereas three questions per participant are required in the bipartite scenario (or two questions per participant in the tripartite scenario) for pseudotelepathy [15].

As the above paragraph indicates, most separation results previously known highlighted the difference between Bell theorems (with or without inequalities) and pseudotelepathy games. In sharp contrast, this paper separates the original Bell theorems from those without inequalities.

Acknowledgements

G. B. is supported in part by Canada's Natural Sciences and Engineering Research Council, the Canada Research Chair program, the Canadian Institute for Advanced Research and the Quantum*Works* Network. At the time of this research, A. A. M. was supported in part by the European Commission under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848 and by the Swiss NCCR Quantum Photonics. Part of this research was carried out while A. A. M. was visiting G. B. at the Université de Montréal.

References

[1] A. Aspect, J. Dalibard and G. Roger, "Experimental test of Bell's inequalities using time-varying analyzers", *Physical Review Letters* **49**:1804–1807, 1982.

[2] A. Aspect, P. Grangier and G. Roger, "Experimental tests of realistic local theories via Bell's theorem", *Physical Review Letters* **47**:460–463, 1981.

[3] A. Aspect, P. Grangier and G. Roger, "Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell's inequalities", *Physical Review Letters* **49**:91–94, 1982.

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

[5] G. Brassard, A. Broadbent and A. Tapp, "Quantum pseudo-telepathy", *Foundations of Physics* **35**:1877–1907, 2005.

[6] G. Brassard, R. Cleve and A. Tapp, "Cost of exactly simulating quantum entanglement with classical communication", *Physical Review Letters* **83**:1874–1877, 1999.

[7] G. Brassard, and A. A. Méthot, "Strict hierarchy of Bell theorems", *Proceedings of Second International Conference on Quantum, Nano, and Micro Technologies (ICQNM 2008)*, Sainte Luce, Martinique, pp. 98–103, February 2008.

[8] G. Brassard, A. A. Méthot and A. Tapp, "Minimal bipartite state dimension required for pseudo-telepathy", *Quantum Information and Computation* **5**:275–284, 2005.

[9] H. R. Brown and G. Svetlichny, "Nonlocality and Gleason's lemma, Part I: Deterministic theories", *Foundations of Physics* **20**:1379–1387, 1990.

[10] N. Cerf, N. Gisin and S. Massar, "Classical teleportation of a quantum bit", *Physical Review Letters* **84**:2521–2524, 2000.

[11] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, "Proposed experiment to test local hidden-variable theories", *Physical Review Letters* **23**:880–884, 1969.

[12] R. Cleve, P. Høyer, B. Toner and J. Watrous, "Consequences and limits of nonlocal strategies", *Proceedings of 19th IEEE Conference on Computational Complexity*, pp. 236–249, 2004.

[13] A. Einstein, B. Podolsky and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?", *Physical Review* **47**:777–780, 1935.

[14] S. J. Freedman and J. F. Clauser, "Experimental test of local hidden-variable theories", *Physical Review Letters* **28**:938–941, 1972.

[15] N. Gisin, A. A. Méthot and V. Scarani, "Pseudo-telepathy: Input cardinality and Bell-type inequalities", *International Journal of Quantum Information* **5**:525–534, 2007.

[16] D. M. Greenberger, M. A. Horne, A. Shimony and A. Zeilinger, "Bell's theorem without inequalities", *American Journal of Physics* **58**:1131–1143, 1990.

[17] D. M. Greenberger, M. A. Horne and A. Zeilinger, "Going beyond Bell's theorem", in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht), pp. 69–72, 1989.

[18] L. Hardy, "Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories", *Physical Review Letters* **68**:2981–2984, 1992.

[19] L. Hardy, "Nonlocality for two particles without inequalities for almost all entangled states", *Physical Review Letters* **71**:1665–1668, 1993.

[20] P. Heywood and M. L. G. Redhead, "Nonlocality and the Kochen-Specker paradox", *Foundations of Physics* **13**:481–499, 1983.

[21] S. Kochen, Personal communication to Abner Shimony [16, footnote 5, page 1142], c. 1970.

[22] N. D. Mermin, "Bringing home the atomic world: Quantum mysteries for anybody", *American Journal of Physics* **49**:940–943, 1981.

[23] N. D. Mermin, "Quantum mysteries refined", *American Journal of Physics* **62**:880–887, 1994.

[24] A. A. Méthot, "On local-hidden-variable no-go theorems", *Canadian Journal of Physics* **84**:633–638, 2006.

[25] A. A. Méthot and V. Scarani, "An anomaly of non-locality", *Quantum Information and Computation* **7**:157–170, 2007.