Nonlocality and Impossible Machines

Paul Secular
*Imperial College London**
(Dated: May 2016)

Bell's theorem tells us that quantum mechanics is inherently nonlocal in the sense that there exist correlations between spatially separated, entangled systems which cannot be explained classically (unless we accept causal influences that propagate faster than light). However, it is possible to postulate "superquantum" correlations beyond those observed in Nature. This idea is encapsulated in a framework of impossible machines or "boxes". Here, I show how classical and quantum systems fit into this framework by deriving concrete examples. In the case of the CHSH Bell inequality, the most nonlocal impossible machine is called a PR-box. Finding reasons to explain why Nature should not permit PR-boxes may help in the search for a more intuitive, physical axiomatisation of quantum mechanics. I explore two promising, information-theoretic principles that rule out PR-boxes: nontrivial communication complexity [1] and information causality [2]. First I illustrate these proofs in terms of information-theoretic resources and then derive bounds on the number of such resources needed. Finally, I show how depolarization and nonlocality distillation protocols can convert one type of nonlocal box into another and discuss how this strengthens the previous results.

* DECLARATION: This project was carried out by Paul Secular under the supervision of Dr David Jennings, who provided regular guidance during meetings and via email. Dr Jennings suggested the research topics, but all notes, calculations, graphs, diagrams and computer code in this project are the sole work of Paul Secular.

CONTENTS

SUMMARY

"it is not unreasonable to imagine that information sits at the core of physics, just as it sits at the core of a computer"

—John Archibald Wheeler [3]

This thesis is based on an active research effort aimed at explaining quantum mechanics in terms of information-theoretic principles. The starting point is nonlocality, which can be viewed as a resource for communication tasks. I use the framework of generalised probabilistic theories known as "box world" where nonlocality can be more powerful than in quantum mechanics. I explore two information-theoretic principles that may help single out quantum nonlocality: nontrivial communication complexity and information causality.

In the introduction, key background concepts and results are summarised in the device-independent language of computer science and game theory. However, no prior knowledge of these areas is assumed. This part should be fully accessible to a general physicist with no background in quantum information. I start by discussing the motivation behind the project and then introduce the framework of nonlocal boxes. I then explain the no-signalling postulate and how it relates to special relativity. Following this, I describe the CHSH Bell inequality and Tsirelson's bound in game theoretic language.

Next I explore the key results of the project and construct explicit examples of both possible (classically or quantum mechanically realisable) and impossible machines. This part is by its nature more mathematical, although some steps are left to the appendices. I prove results as clearly as possible, expanding on those in the literature.

In section II I explain how nontrivial communication complexity rules out the most powerful nonlocal boxes and derive limits on the number or resources required. This appears to be a new result.

In section III I construct some quantum mechanical examples, using both projective and generalised measurements, so some exposure to advanced quantum mechanics is required. For example, familiarity with the density operator, partial trace, POVMs, and the Bloch sphere representation of a qubit are assumed.

In section IV I discuss the second of the key principles: information causality. Here some exposure to Shannon information theory would be useful, although not essential. I explain the proof of the Tsirelson bound and take a different approach to the literature by illustrating it graphically. This gives a tantalising glimpse of how it may one day be possible to derive quantum mechanics from more physically motivated axioms.

Finally, ongoing and future work are discussed in the concluding remarks, along with some important open questions.

Part I
Introduction

Unlike relativity, quantum theory is based on abstract axioms which seem to have no obvious physical motivation [4]. Finding more intuitive postulates to replace them is not just a pedagogical exercise, but may give us a far deeper insight into the workings of Nature. One starting point may be nonlocality, by which we mean correlations between distributed parties that cannot be explained classically. Bell's groundbreaking work [5] showed that such nonlocality is an intrinsic feature of quantum mechanics. It arises from a property called entanglement. However quantum mechanics cannot be derived from nonlocality alone. What other supplementary principles might be needed?

The primary aim of this work is to study some of the approaches that have been suggested to answer this question, and to understand them in terms of information-theoretic resources. We do this by considering "impossible machines". These are theoretical "black boxes" representing some no-go theorem. A famous example from quantum mechanics would be a box that can clone an arbitrary quantum state. Such a machine is impossible by the no-cloning theorem of Wootters and Zurek [6], and Dieks [7].

The power of this abstract approach is beautifully demonstrated in [8] where a hierarchy of impossible machines is considered in order to prove that the impossibility of superluminal communication implies the impossibility of classical teleportation of unknown quantum states. Since classical teleportation of unknown classical states is possible [9], the conclusion one can take from this proof is that quantum systems must comprise a radically different type of information that cannot be described classically. It's incredible that such a simple proof leads to a fundamental result underpinning the whole field of quantum information.

Although this research is motivated by foundational issues, progress may also have very practical applications for the burgeoning field of quantum computing. In particular, there could be implications for communication, cryptography, distributed computation, and interactive proof systems [10]. Nonlocality may also give us hints about theories of quantum gravity. Although there is currently no reason to think quantum mechanics is not the most fundamental description of Nature, there is a class of "almost-quantum" correlations that seems to arise naturally from both information-theoretic considerations [11] and from path integral approaches quantum gravity [12].

Box world

In the study of quantum nonlocality, the simplest, but also one of the most fundamental, types of experiment is one whose outcome depends on two experimenters who each make an independent measurement on some part of an entangled system. Often this is a pair of entangled qubits[1] (e.g. a singlet state). To facilitate our

[1] Qubit (short for quantum bit) is the abstract name used in quantum information theory for any two-dimensional quantum system, such as the spin of an electron. Analogously to classical information theory, where we usually choose the bit as our unit of information, qubits are a convenient choice of unit for quantum information. Here we adopt the "computational basis" convention in which our basis vectors are $|0\rangle$ and $|1\rangle$ and a general qubit is thus written as $\alpha|0\rangle + \beta|0\rangle$ where $\alpha, \beta \in C$.

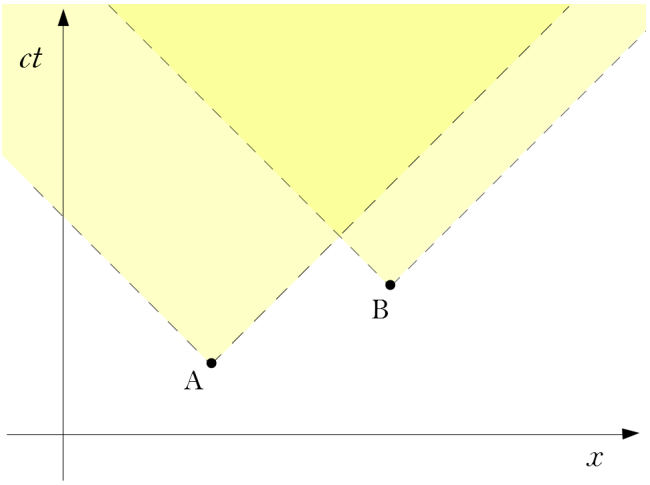

Figure 1. Spacelike separation between Alice and Bob's input/output events. Their forward light cones are shown as shaded regions bounded by dashed lines. Notice that both events are outside the other's cone.

discussion, we follow a long-established convention and name these two spatially-separated experimenters Alice and Bob. Such bipartite (i.e. two party) experiments can be realised in many ways, but the specific implementation details are not what interest us here. Instead, we abstract away the details and describe such experiments within a framework (sometimes referred to as box world [13]) of distributed "boxes". These are hypothetical machines with spatially separated input-output ports. Going back to our previous example then, we describe the singlet state as a bipartite box. When one party makes a local measurement on their qubit, we call their choice of measurement setting their "input". The result of the measurement is their "output". We normally assume that the outcome of a quantum mechanical measurement is obtained immediately after the measurement is made. We retain this assumption and define output events to occur simultaneously with their corresponding input events. For simplicity, we also assume that such input/output events are localised to a single point in space.

The strength of this framework is that it allows us to study nonlocality operationally in a theory-independent way. It can describe classical, quantum, and more general, "impossible" correlations. The hope is therefore to use it to learn something about what makes quantum mechanics special.

The no-signalling condition

Although a box allowing faster-than-light communication is an impossible machine, such boxes are usually excluded from discussions of nonlocality based on the fact that they are ruled out by special relativity. In this work, we therefore restrict ourselves to nonlocal boxes that do not permit superluminal signalling: a property known as "no-signalling". Imagine Alice and Bob both have one input-output port of a bipartite box. Importantly, their input/output events can be outside of each other's forward light cones (see figure 1), meaning that communication between them is impossible (hence "no signalling"). In information terms, this means that Alice and Bob's local outputs should give them no information about each other's inputs,

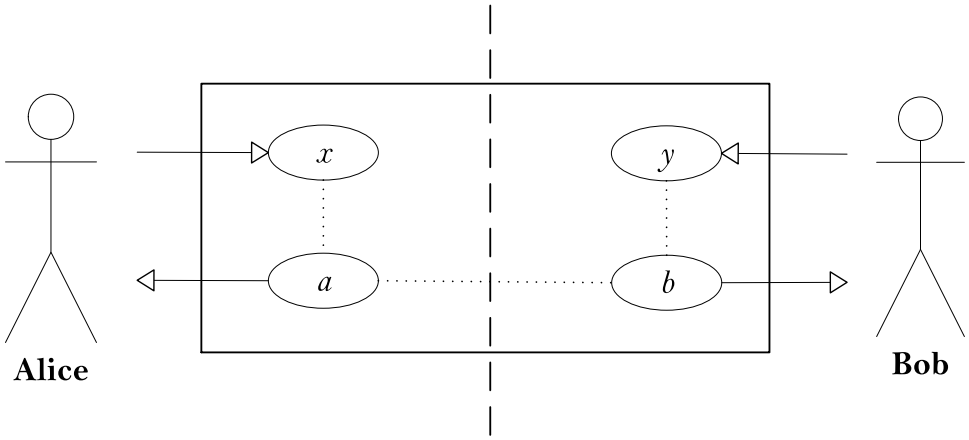

Figure 2. Schematic of a bipartite no-signalling box. The dotted lines represent possible statistical correlation between variables. Alice enters some input x and receives output a. Bob independently enters some input y and receives output b. Since the box is no-signalling, a can have no statistical dependence on y and b can have no statistical dependence on x.

i.e. information cannot be instantaneously transferred from one point in space to another. This can be described mathematically by conditional probabilities on Alice and Bob's outputs, A and B, due to two corresponding inputs, X and Y, plus any shared initial conditions λ (e.g. correlated bits):

$$\begin{cases} P(A = a \mid X = x, Y = y, \lambda) = P(A = a \mid X = x, \lambda) \\ P(B = b \mid X = x, Y = y, \lambda) = P(B = b \mid Y = y, \lambda) \end{cases} \tag{1}$$

Taking the dependence on λ as implicit from now on, we can rewrite (1) in self-explanatory shorthand notation as

$$\begin{cases} P\left(a \mid x, y\right) = P\left(a \mid x\right) \\ P(b \mid x, y) = P\left(b \mid y\right) \end{cases} \tag{2}$$

(for a derivation, see appendix A). In other words, a no-signalling box is one where an independent choice of input at one port cannot affect the output statistics at any other port (see figure 2).

Classical physics and quantum mechanics[2] both obey the no-signalling condition, but it is a result of fundamental importance that no-signalling alone does not constrain nonlocal correlations to just those permitted by quantum mechanics [15]. In particular, no-signalling places no restriction on the relationship between A and B so, as we will see, they can be correlated in ways impossible in quantum mechanics. We refer to any no-signalling box with correlated outputs that cannot be simulated classically as a nonlocal box.

Nonlocal games

If we consider a multipartite experiment whose final outcome is restricted to take just one of two values, then it can be described in the language of game theory

[2] It's very interesting to ask why non-relativistic quantum mechanics should obey no-signalling. Kennedy [14] argues that it is due to the tensor product axiom.

as a simple, n-player, co-operative game, where the outcomes correspond to winning/losing. In this work we consider the following type of game:

Alice and Bob share one or more bipartite boxes and are each simultaneously asked a secret question. Once these questions have been asked, Alice and Bob are no longer allowed to communicate with each other. The final result of the game depends on the combination of their (possibly correlated) outcomes/responses.

The probability of Alice and Bob winning is known as the value of the game. In the quantum information literature, this scenario (which can be generalised to n parties) is sometimes referred to as a nonlocal game [10]. In computer science terms, winning the game can be thought of as distributively computing the answer to a decision problem (a question with a yes/no answer), and in logic/mathematics a decision problem can be written as a Boolean function of two or more variables where the outcome represents true/false or one/zero. Hence we see an important unity between what might naively seem to be disparate areas of research.

Alice's bits and Bob's

A binary bipartite box is one whose inputs and outputs can only take on one of two values, which we can therefore represent as bits. We can write out its probability distribution as a 4×4 matrix, where each matrix element corresponds to $P(a, b \mid x, y)$ for a particular combination of $a, b, x, y \in \{0, 1\}$ (we omit the commas to save space):

$$\begin{pmatrix} P(00 \mid 00) & P(01 \mid 00) & P(10 \mid 00) & P(11 \mid 00) \\ P(00 \mid 01) & P(01 \mid 01) & P(10 \mid 01) & P(11 \mid 01) \\ P(00 \mid 10) & P(01 \mid 10) & P(10 \mid 10) & P(11 \mid 10) \\ P(00 \mid 11) & P(01 \mid 11) & P(10 \mid 11) & P(11 \mid 11) \end{pmatrix}. \tag{3}$$

Bell inequalities and Tsirelson bounds

In terms of the game theory description we have adopted, what Bell's theorem tells us is that there exist games which Alice and Bob can win with a higher probability if they share quantum entanglement than if they use a classical strategy. Bell inequalities express the maximum classical value of these games. One of the most widely cited Bell inequalities is the CHSH inequality, so-called for Clauser, Horne, Shimony, and Holt who expanded on Bell's original result [16]. In fact there are eight Bell inequalities for the case of binary bipartite games, but they are all equivalent in the sense that any one of them can be transformed into any other by local reversible operations [17]. We therefore need only consider one of these inequalities and thus choose the following canonical form as the winning condition for the CHSH game:

$$a \oplus b = x \cdot y. \tag{4}$$

Here, and in the rest of this work, we use the $\oplus$ symbol to represent addition modulo 2 (which is equivalent to an XOR operation[3]). The corresponding CHSH inequality can thus be written in game-theoretic terms as:

$$P(a \oplus b = x \cdot y \mid x, y) \leq \frac{3}{4} \tag{5}$$

[3] This means that $0 \oplus 0 = 0$, $0 \oplus 1 = 1 \oplus 0 = 1$, and $1 \oplus 1 = 0$.

Tsirelson (sometimes transliterated as Cirel'son) [18] proved that quantum mechanics instead satisfies:

$$P(a \oplus b = x \cdot y \mid x, y) \leq \frac{1}{2} + \frac{1}{2\sqrt{2}} \tag{6}$$

In general, the quantum value of a nonlocal game is known as a Tsirelson bound. These bounds can be seen as the quantum equivalent of Bell inequalities. It turns out, however, that this isn't the full story. Nonlocal correlations beyond those that occur in quantum mechanics can also be defined. These are sometimes referred to as "superquantum" correlations. In the case of the CHSH game, any correlations that give a value of $\frac{1}{2}+\frac{1}{2\sqrt{2}} < P(a \oplus b = x \cdot y \mid x, y) \leq 1$ must be superquantum (although there also exist correlations not realisable in quantum mechanics that give a CHSH game value arbitrarily close to $\frac{3}{4}$). Despite the difference in nonlocal correlation strength, it has been shown that no-signalling theories with such correlations share many important properties with quantum mechanics [19]. This begs the question: why is quantum mechanics not *more* nonlocal? [15]

PR-Boxes

In the context of the CHSH game, the most powerful superquantum machine is called a PR-box[4] (named after Popescu and Rohrlich for their influential 1994 paper [15]). As Popescu and Rohrlich discovered, PR-boxes do not allow superluminal signalling (for proof, see appendix B), and yet are impossible in both classical physics and quantum mechanics. They are defined by the following probability distribution:

$$\begin{pmatrix} P(00 \mid 00) & P(01 \mid 00) & P(10 \mid 00) & P(11 \mid 00) \\ P(00 \mid 01) & P(01 \mid 01) & P(10 \mid 01) & P(11 \mid 01) \\ P(00 \mid 10) & P(01 \mid 10) & P(10 \mid 10) & P(11 \mid 10) \\ P(00 \mid 11) & P(01 \mid 11) & P(10 \mid 11) & P(11 \mid 11) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \tag{7}$$

which we can express more concisely as

$$\begin{aligned} P(a, b \mid x, y) &= \tfrac{1}{2} \text{ if } a \oplus b = x \cdot y \\ &= 0 \text{ otherwise.} \end{aligned} \tag{8}$$

This latter form makes it clear that PR-boxes have a CHSH game value of 1 (i.e. they give a maximal violation of the CHSH inequality) since there is zero probability that $a \oplus b \neq x \cdot y$, no matter that the values of x and y.

The van Dam theorem

PR-boxes turn out to be a very useful tool in the study of nonlocality [20–23]. In particular, they have a special relationship with communication complexity [24]. This is the measure of how many classical bits need to be communicated from Alice to Bob if they are to win a game with 100% probability based solely on Bob's final response. In such scenarios we assume unlimited local computational resources but require minimal communication from Alice to Bob. In information processing terms,

[4] Sometimes these are referred to as nonlocal boxes in the literature, but here we reserve that term for a more general class of box.

we say that "communication is expensive". An important result due to van Dam [1] is that PR-boxes trivialise the communication complexity of decision problems. That is to say, no matter how many bits they are given as input, Alice and Bob can solve any distributed decision problem with the communication of just one bit so long as PR-boxes are available. While this may be a reason to rule out PR-boxes in Nature, it doesn't explain the CHSH Tsirelson bound.

Information causality

Another approach to finding an information-theoretic principle that constrains Nature to quantum mechanics is "information causality". This was introduced by Pawłowski *et al.* in [2] and can be thought of as a natural generalisation of the no-signalling condition. Roughly speaking, the principle states that if Alice sends Bob m bits of information, then Bob should at most be able to retrieve no more than m bits of Alice's dataset. When $m = 0$, it is easy to see that this reduces to the usual no-signalling condition. That is to say, Bob can have no information at all about any of Alice's bits in the absence of signalling. Although the proof is fairly involved (making use of a new, model-independent information measure which reduces to the Shannon and von Neumann measures in the classical and quantum cases respectively), the remarkable result is that the CHSH Tsirelson bound is recovered[5]. Although this does not uniquely identify quantum mechanics from the set of all possible no-signalling theories, it is a very promising step towards such a goal.

[5] Interestingly, a recent preprint by Carmi and Moskovich [25] claims that another generalisation of the no-signalling principle, which they call "statistical no-signalling", can also limit CHSH correlations to the Tsirelson bound.

Part II
Results

I. A SIGNALLING BOX

Although we will be considering no-signalling boxes in this work, it can be illustrative to see a simple counter-example. I here construct an interesting impossible box that breaks the no-signalling condition. Consider the following probability distribution:

$$\begin{pmatrix} P(00 \mid 00) & P(01 \mid 00) & P(10 \mid 00) & P(11 \mid 00) \\ P(00 \mid 01) & P(01 \mid 01) & P(10 \mid 01) & P(11 \mid 01) \\ P(00 \mid 10) & P(01 \mid 10) & P(10 \mid 10) & P(11 \mid 10) \\ P(00 \mid 11) & P(01 \mid 11) & P(10 \mid 11) & P(11 \mid 11) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{9}$$

It is straightforward to verify that:

- $(A = a)$ and $(Y = y)$ are independent events,
- $(B = b)$ and $(X = x)$ are independent events,
- $(B = b, Y = y)$ and $(X = x)$ are independent events.

However, $(A = a, X = x)$ and $(Y = y)$ are *not* independent events. In fact, $x \oplus a \oplus 1 = y$ for all x, y, a. So this box would allow Bob to signal perfectly to Alice. Notice also that:

- $(B = b)$ and $(Y = y)$ are independent events, since

$$P(b \mid x, y) = P(b \mid y) = P(b) = \frac{1}{2},$$

  so that Alice cannot signal to Bob;

- $(A = a)$ and $(B = b)$ are independent events, meaning that Alice and Bob's outputs are uncorrelated.

II. PR-BOXES

Proof of the van Dam theorem

I now explain van Dam's proof [1] that PR-boxes trivialise communication complexity. In other words, I show how Alice and Bob can solve any decision problem (i.e. calculate any Boolean function) with just a single bit of communication if they use PR-boxes. Unlike the proof given in [1], I also show how the protocol can be carried out to minimise the number of PR-boxes required.

Proof. Suppose some Boolean function $f(x_1 x_2 \ldots x_n, y_1 y_2 \ldots y_m)$ depends on n bits x_i known only to Alice, and m bits y_j known only to Bob. Without loss of generality, we can simplify the mathematics by letting $n = m$.

Boolean functions can be expressed in a number of different ways, but one useful and intuitive form is the disjunctive normal form [26]. This expresses the function as a sum (modulo 2) of products of the input bits and their complements (i.e. x_i, y_j, $\bar{x}_i$ and $\bar{y}_j$, where $\bar{x}_i = x_i \oplus 1$). These terms can be thought of as elementary indicator functions. Factorising gives us:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n} Q_k(\mathbf{y}) P_k(\mathbf{x}) \tag{10}$$

or:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n} Q_k'(\mathbf{x}) P_k'(\mathbf{y}), \tag{11}$$

where the Q_k and Q_k' functions are monomials and the P_k and P_k' are polynomials. Since $x^n = x$ for any bit x, there are 2^n distinct monomials that can be formed from n bits[6]. van Dam's proof works in exactly the same way whichever of the above forms is used, but the number of PR-boxes required can be minimised by choosing the form comprising the least number of monomials (since some of these terms may be identically zero). In van Dam's algorithm, each of the remaining terms in the sum corresponds to the use of one PR-box. However, as we will show, it is possible to reduce the number of PR-boxes by one if there is a term of the form $Q_k P_k$ where $Q_k \equiv 1$.

Let us assume that f is in the form given by (10). Alice calculates the values of the P_k, and Bob the Q_k. Alice then takes the kth PR-box and inputs the value of P_k, while Bob inputs the corresponding value of Q_k. The outputs they receive will be a_k and b_k, respectively. These will be random but, as we know from the definition of a PR-box, they are related by:

$$a_k \oplus b_k = Q_k(\mathbf{y}) P_k(\mathbf{x}). \tag{12}$$

Hence, (5) becomes:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n} a_k \oplus b_k = \left(\sum_{k=1}^{2^n} a_k \right) \oplus \left(\sum_{k=1}^{2^n} b_k \right). \tag{13}$$

Notice how the PR-boxes have converted products into sums. Thus, Alice can simply calculate $\sum_{k=1}^{2^n} a_k$ and send the result (a single bit) to Bob. Bob calculates $\sum_{k=1}^{2^n} b_k$, to which he adds the bit from Alice, giving $f(\mathbf{x}, \mathbf{y})$. Hence by using PR-boxes, Alice and Bob can indeed compute the value of any Boolean function with just one bit of communication. □

PR-boxes as resources

Being no-signalling, PR-boxes do not allow Alice and Bob to communicate. However, they can be seen as an information processing resource, since they facilitate the calculation of distributed functions of the form $f(x_1 x_2 \ldots x_n, y_1 y_2 \ldots y_n)$. An interesting question to ask, therefore, is how do PR-boxes compare to classical bits of communication as a resource for calculating such a function? If Alice simply sends

[6] One way to see this is from the fact that $\sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n$

all n of her bits to Bob, then it is trivial to see that no PR-boxes at all are required. However, I now derive bounds on the number of PR-boxes Alice and Bob need if they are restricted in how many bits they can communicate.

If Alice is only allowed to send one bit to Bob, then we have an upper bound of 2^n from van Dam's proof. However, I here show that we can do slightly better by deriving an upper bound of $(2^n - 1)$.

Proof. Consider first a function written in the form 10. We can rewrite this as:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n-1} Q_k(\mathbf{y}) P_k(\mathbf{x}) \oplus 1 \cdot P_{2^n}(\mathbf{x}). \tag{14}$$

Using PR-boxes as before for the first $(2^n - 1)$ terms, we get:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n-1} a_k \oplus b_k \oplus P_{2^n}(\mathbf{x}) = \left(\sum_{k=1}^{2^n-1} a_k \oplus P_{2^n}(\mathbf{x}) \right) \oplus \left(\sum_{k=1}^{2^n-1} b_k \right). \tag{15}$$

So Alice still only needs to send one bit to Bob in order for him to compute f. Similarly, if our function is in the form

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n} Q'_k(\mathbf{x}) P'_k(\mathbf{y}), \tag{16}$$

then we write

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n-1} Q'_k(\mathbf{x}) P'_k(\mathbf{y}) \oplus 1 \cdot P'_{2^n}(\mathbf{y}) \tag{17}$$

and again use $(2^n - 1)$ PR-boxes to get

$$f(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{2^n-1} a'_k \oplus b'_k \oplus P'_{2^n}(\mathbf{y}) = \left(\sum_{k=1}^{2^n-1} a'_k \right) \oplus \left(\sum_{k=1}^{2^n-1} b'_k \oplus P'_{2^n}(\mathbf{y}) \right). \tag{18}$$

So our upper bound for the number of PR-boxes needed is actually $(2^n - 1)$ as claimed. □

Obviously Alice and Bob can hypothetically use as many PR-boxes as they like, but any boxes above this upper bound are redundant. Of course, the minimum number required actually depends on the particular form of f. In the most trivial case, we may have

$$f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}), \tag{19}$$

in which case no PR-boxes are required at all and no bits need to be sent from Alice to Bob. However, if f does have some explicit $\mathbf{x}$ dependence then Alice will always need to send at least one bit since PR-boxes are no signalling: the output Bob gets from his half of a PR-box is random, so it is useless to him unless he knows the output from Alice's half. The simplest, non-trivial case would require Alice and Bob to use one PR-box and one bit of communication. I illustrate this with an example.

Example 1. Consider:

$$f(x_1x_2, y_1y_2) = x_1y_1y_2 \oplus x_2. \tag{20}$$

This can be computed if Alice sends her 2 bits to Bob. Alternatively, let Alice input x_1 into a PR-box and Bob input y_1y_2. Label their corresponding outputs as a and b. Then

$$f(x_1x_2, y_1y_2) = a \oplus b \oplus x_2. \tag{21}$$

So Alice need only send $a \oplus x_2$ (one bit) to Bob in order for him to compute f. Hence we have shown that this function can be computed using 1 bit of classical communication and a single PR-box (rather than 3 boxes as suggested by our upper bound).

I now show that, in general, a linear increase in communication complexity gives an exponential decrease in the maximum number of PR-boxes necessary (up to the no-signalling limit of one bit).

Proof. Assume Alice is allowed to communicate $k+1$ bits to Bob, where $k+1 < n$. In other words, she is allowed k additional bits of communication with respect to the previous case. If she sends Bob the last k bits of her bit string, then effectively they now only need to use PR-boxes plus one additional bit of communication to calculate a distributed function of the form $f(x_1x_2 \ldots x_{n-k}, y_1y_2 \ldots y_{n-k})$. Hence, from our previous result, the upper bound on the number of PR-boxes becomes $(2^{n-k} - 1)$. □

III. NOISY PR-BOXES

A useful, single parameter generalisation of the PR-box is the isotropic "noisy" PR-box [27]. This is defined as a box which acts like a perfect PR-box with probability E and like a completely random box with probability $(1 - E)$. In other words $(1 - E)$ characterises how "noisy" the box is. The CHSH game value of such a box is then given by

$$P(a \oplus b = x \cdot y) = E \cdot 1 + (1 - E) \cdot \frac{1}{2} = \frac{1}{2}(1 + E), \tag{22}$$

and its probability distribution is:

$$\frac{1}{4}\begin{pmatrix} (1+E) & (1-E) & (1-E) & (1+E) \\ (1+E) & (1-E) & (1-E) & (1+E) \\ (1+E) & (1-E) & (1-E) & (1+E) \\ (1-E) & (1+E) & (1+E) & (1-E) \end{pmatrix}. \tag{23}$$

By varying E between 0 and 1, we can make an isotropic box behave as a classical, quantum or superquantum device. Clearly, any such box can be simulated using a PR-box and classical randomness. But what makes this such a useful canonical form is that any bipartite box can be used to simulate an isotropic "noisy" PR-box with the same CHSH game value (henceforth, CHSH value) [19]. This means that any no-go theorem holding for a noisy PR-box, must also apply to all non-isotropic boxes with the same CHSH value.

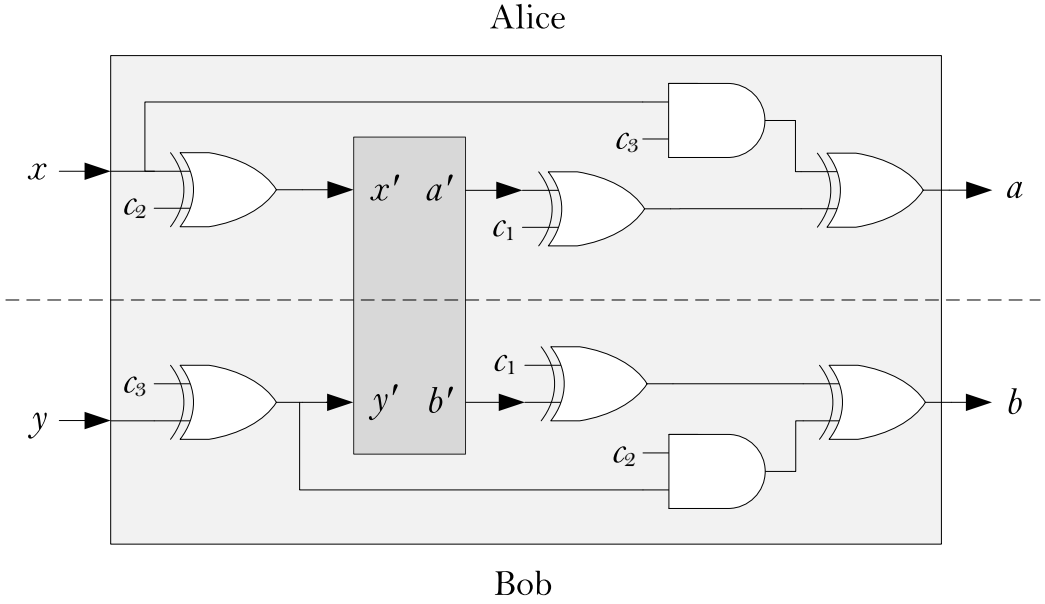

Figure 3. Depolarization procedure illustrated as a series of logic gates. Using three bits of shared classical randomness (c_1, c_2, c_3) and a bipartite box (shown in dark grey), Alice and Bob can simulate an isotropic box (shown in light grey) with the same overall CHSH value.

Using the idea of noisy PR-boxes, Brassard *et al.* were able to derive a bound of $P(a \oplus b = x \cdot y) = \frac{3+\sqrt{6}}{6}$ for trivial communication complexity [22]. Not quite the Tsirelson bound, but curiously close. Results like these have led to PR-boxes being considered as a natural measure of nonlocality [28]. However it has been shown that they cannot be the most general such measure, even in the bipartite case [20, 29]. Another important result is that, although entanglement and nonlocality can both be viewed as resources [17], they are not equivalent [30, 31].

Depolarization

The depolarization procedure was introduced by Masanes, Acin, and Gisin in [19] (note the American spelling convention which I retain). It converts the probability distribution of any bipartite nonlocal box into the isotropic form of a noisy PR-box using local operations and 3 bits of shared (classical) randomness, c_1, c_2, c_3. Here I expand on the description given by Masanes, Acin, and Gisin by illustrating explicitly how Alice and Bob carry out this procedure. I also prove that it gives the desired result.

The simplest way to understand depolarization is as follows:

1. Alice and Bob transform their inputs x and y to x' and y' as follows:

$$\begin{aligned} x' &= x \oplus c_2 \\ y' &= y \oplus c_3 \end{aligned} . \tag{24}$$

2. They then input x' and y' into their (non-isotropic) box and receive outputs a' and b'.

3. Finally, they transform a' and b' to a and b which are their (isotropic) outputs:

$$\begin{aligned} a &= a' \oplus c_1 \oplus x \cdot c_3 \\ b &= b' \oplus c_1 \oplus (y \oplus c_3) \cdot c_2 \end{aligned} \tag{25}$$

Although (24) and (25) were derived independently by the present author from the description given in [19], the depolarization procedure is given in a similar algebraic form in [32] and [33] (where it is referred to as "twirling"). I illustrate this transformation in figure 3 as a series of logic gates. The first thing to notice is that these transformations are involutory (their own inverse), which means:

$$\begin{aligned} x &= x' \oplus c_2 \\ y &= y' \oplus c_3 \\ a' &= a \oplus c_1 \oplus x \cdot c_3 \\ b' &= b \oplus c_1 \oplus (y \oplus c_3) \cdot c_2 \end{aligned} \tag{26}$$

Hence:

$$\begin{aligned} a' \oplus b' = x' \cdot y' &\Leftrightarrow (a \oplus c_1 \oplus x \cdot c_3) \oplus (b \oplus c_1 \oplus (y \oplus c_3) \cdot c_2) = (x \oplus c_2) \cdot (y \oplus c_3) \\ &\Leftrightarrow a \oplus x \cdot c_3 \oplus b \oplus y \cdot c_2 \oplus c_3 \cdot c_2 = x \cdot y \oplus y \cdot c_2 \oplus x \cdot c_3 \oplus c_3 \cdot c_2 \\ &\Leftrightarrow a \oplus b = x \cdot y \end{aligned} \tag{27}$$

In other words, depolarization leaves the CHSH equation invariant. This is because each of the eight possible combinations of (c_1, c_2, c_3) corresponds to one of the equation's eight symmetries.

We therefore have that $P(a \oplus b = x \cdot y \mid x, y) = P(a' \oplus b' = x' \cdot y' \mid x, y) = P(a' \oplus b' = x' \cdot y')$ (see appendix D for proof). This means that, no matter what the value of x and y, the probability of winning the CHSH game is now independent of the inputs. This would still hold true without the first shared bit, c_1, but using this gives uniform marginals for a and b. Hence the depolarization procedure converts any box into isotropic form as claimed.

Although I do not prove it here, it should also be clear from the above argument that an isotropic box cannot be converted into a nonisotropic one whilst leaving the CHSH value invariant. However, I now construct a simple example to show that Alice and Bob can use an isotropic box to simulate a nonisotropic one with a lower CHSH value, if one of them uses local randomness to probabilistically transform their output.

Let Alice and Bob use a PR-box. If Bob finds his input and output are $y = 0$ and $b = 1$ respectively he flips his output b with probability q. Clearly the transformed probability distribution becomes:

$$\frac{1}{2}\begin{pmatrix} 1 & 0 & q & (1-q) \\ 1 & 0 & 0 & 1 \\ 1 & 0 & q & (1-q) \\ 0 & 1 & 1 & 0 \end{pmatrix}, \tag{28}$$

which is not of the isotropic form of equation (23).

Examples of bipartite boxes

In this section, I describe five concrete examples of noisy PR-boxes (summarised in table I).

Description	$P(a,b \mid x,y)$	$P(a \oplus b = x \cdot y)$	E
Completely uncorrelated box	$\frac{1}{4}$	$\frac{1}{2}$	0
Almost uncorrelated box	$\frac{13}{48}$ if $a \oplus b = x \cdot y$ $\frac{11}{48}$ otherwise	$\frac{13}{24}$	$\frac{1}{12}$
Partially correlated classical box	$\frac{5}{16}$ if $a \oplus b = x \cdot y$ $\frac{3}{16}$ otherwise	$\frac{5}{8}$	$\frac{1}{4}$
Maximally correlated classical box	$\frac{3}{8}$ if $a \oplus b = x \cdot y$ $\frac{1}{8}$ otherwise	$\frac{3}{4}$	$\frac{1}{2}$
CHSH quantum box	$\frac{2+\sqrt{2}}{8}$ if $a \oplus b = x \cdot y$ $\frac{2-\sqrt{2}}{8}$ otherwise	$\frac{2+\sqrt{2}}{4}$	$\frac{\sqrt{2}}{2}$
PR-box	$\frac{1}{2}$ if $a \oplus b = x \cdot y$ 0 otherwise	1	1

Table I. Examples of isotropic noisy PR-boxes. Two different ways of classifying them are shown: $P(a \oplus b = x \cdot y)$ is the CHSH value, and E is the noise parameter defined in equation (22).

1. *Completely uncorrelated box*

The most trivial case is a completely noisy PR-box with $E = 0$. This is equivalent to Alice and Bob simply tossing fair coins to decide their outputs, which clearly leads to completely uncorrelated results. Moreover, their own inputs and outputs are also completely uncorrelated. So there is a uniform probability distribution for a and b, $P(a,b \mid x,y) = \frac{1}{4}$, which leads to a CHSH value of:

$$P(a \oplus b = x \cdot y) = \frac{1}{2} \tag{29}$$

By "coins", what we really mean is that Alice and Bob each have a local random bit. Classical bits can be encoded quantum mechanically as qubits. The uncorrelated box can thus be described by an ensemble of qubit product states such as:

$$\rho = \frac{1}{4}\left(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|\right). \tag{30}$$

Notice that no matter what basis Alice and Bob measure in, they will always get a uniformly random uncorrelated result.

2. *Almost uncorrelated box*

Let Alice and Bob share the following mixed quantum state:

$$\rho = \frac{1}{2}\left(|00\rangle\langle 00| + |1+\rangle\langle 1+|\right), \tag{31}$$

meaning that their local reduced states are:

$$\rho_A = Tr_B(\rho) = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) \tag{32}$$

$$\rho_B = Tr_A(\rho) = \frac{1}{2}\left(|0\rangle\langle 0| + |+\rangle\langle +|\right). \tag{33}$$

Using the POVM formalism, we will describe Alice and Bob's measurements by the measurement operators, $\mathfrak{m}_{a|x} = \left\{M_{a|x}\right\}_a$ and $\mathfrak{m}_{b|y} = \left\{M_{b|y}\right\}_b$, where:

$$\mathfrak{m}_{a|0} = \left\{M_{a|0}\right\}_a = \left\{|0\rangle\langle 0|, |1\rangle\langle 1|\right\} \tag{34}$$

$$\mathfrak{m}_{a|1} = \left\{M_{a|1}\right\}_a = \left\{|+i\rangle\langle +i|, |-i\rangle\langle -i|\right\} \tag{35}$$

$$\mathfrak{m}_{b|0} = \left\{M_{b|0}\right\}_b = \left\{\sqrt{\frac{2}{3}}|0\rangle\langle 0|, \left(\sqrt{\frac{1}{3}}|0\rangle\langle 0| + |1\rangle\langle 1|\right)\right\} \tag{36}$$

$$\mathfrak{m}_{b|1} = \left\{M_{b|1}\right\}_b = \left\{|+i\rangle\langle +i|, |-i\rangle\langle -i|\right\}. \tag{37}$$

It is easy to show that:

$$\begin{aligned} P(a \oplus b = x \cdot y \mid x, y) &= \tfrac{2}{3} \text{ if } x = y = 0 \\ &= \tfrac{1}{2} \text{ otherwise.} \end{aligned} \tag{38}$$

Thus $P(a \oplus b = x \cdot y) = \frac{13}{24}$ and, if we put the box into isotropic form, we find a value of $E = \frac{1}{12}$.

3. Partially correlated classical box

Consider now a mixed quantum state given by the following density matrix:

$$\rho = \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 11|\right). \tag{39}$$

Presume Alice and Bob can both make projective measurements in either the Z or X basis. Clearly if they both measure in the Z basis they will get perfectly correlated results (otherwise, the results will be completely random). Let this case correspond to $x = y = 0$ and let the state after measurement be $|ab\rangle\langle ab|$. Then:

$$\begin{aligned} P(a \oplus b = x \cdot y \mid x, y) &= 1 \text{ if } x = y = 0 \\ &= \tfrac{1}{2} \text{ otherwise.} \end{aligned} \tag{40}$$

Alice and Bob choosing between the X and Z basis with a 50-50 probability is equivalent to $p(x, y) = \frac{1}{4}$, giving $P(a \oplus b = x \cdot y) = \frac{5}{8}$. After depolarization, this state is equivalent to a noisy PR-box with $E = \frac{1}{4}$. Such a box can also be realised classically as follows:

Let Charlie secretly prepare two containers. Each container has two compartments with doors labelled 0 and 1. Inside the two "0" compartments Charlie places a piece of paper with the same number written on it: either a 0 or a 1. He picks this number at random (say by tossing a fair coin). He then takes the first box. Inside its "1" compartment he places a piece of paper with a random number written on it

(again, a 0 or a 1). Finally, he takes the second box and places a piece of paper inside its "1" compartment with a third random number (0 or 1) written on it. Charlie then closes all the doors, gives one box to Alice and one to Bob. These classical boxes give exactly the same operational behaviour as the above mixed state.

4. Maximally correlated classical box

But Alice and Bob can do better, even without using quantum mechanics, if Charlie prepares the boxes slightly differently. As well as putting correlated random numbers, a_0, in the "0" compartments, he could also put anti-correlated random numbers, a_1 and $\bar{a_1}$, in the "1" compartments. Now Alice and Bob will find the following probabilities:

$$\begin{aligned} P(a \oplus b = x \cdot y \mid x, y) &= 1 \text{ if } x = y \\ &= \tfrac{1}{2} \text{ if } x \neq y. \end{aligned} \tag{41}$$

Since $p(x, y) = \frac{1}{4}$, we have $P(a \oplus b = x \cdot y) = \frac{3}{4}$. Interestingly, if we allow our players to inspect both compartments (something which is perfectly possible, classically) then by employing a suitable local randomisation procedure, this can be made equivalent to a box with $P(a \oplus b = x \cdot y \mid x, y) = \frac{3}{4}$ for all x and y (see appendix C); in other words, a noisy PR-box with $E = \frac{1}{2}$, which is the classical limit [2]. Such a box can also be realised quantum mechanically using two qubits for Alice and two qubits for Bob:

$$\begin{aligned} \rho &= \tfrac{1}{4} \sum_{a_0, a_1} \left(\left| a_0 a_1 \right\rangle \otimes \left| a_0 \bar{a_1} \right\rangle \right) \left(\left\langle a_0 a_1 \right| \otimes \left\langle a_0 \bar{a_1} \right| \right) \\ &= \tfrac{1}{4} \sum_{a_0, a_1} \left| a_0 a_1 a_0 \bar{a_1} \right\rangle \left\langle a_0 a_1 a_0 \bar{a_1} \right| \\ &= \tfrac{1}{4} \left(\left| 0001 \right\rangle \left\langle 0001 \right| + \left| 0100 \right\rangle \left\langle 0100 \right| + \left| 1011 \right\rangle \left\langle 1011 \right| + \left| 1110 \right\rangle \left\langle 1110 \right| \right). \end{aligned} \tag{42}$$

The same correlations are found if Alice and Bob share two entangled qubits in a Bell state and make measurements in the Pauli X and Z bases. For example,

$$\left| \Phi^- \right\rangle = \frac{1}{\sqrt{2}} \left(\left| 00 \right\rangle - \left| 11 \right\rangle \right) = \frac{1}{\sqrt{2}} \left(\left| +- \right\rangle + \left| -+ \right\rangle \right) \tag{43}$$

clearly gives the desired results if we let the Z basis correspond to an input of 0 and the X basis correspond to an input of 1. However, by allowing Alice and Bob to employ different measurement bases, such an entangled state allows them to saturate the classical bound.

5. CHSH quantum box

Let Alice and Bob share the following Bell state[7]:

$$\left| \Phi^+ \right\rangle = \frac{1}{\sqrt{2}} \left(\left| 00 \right\rangle + \left| 11 \right\rangle \right) = \frac{1}{\sqrt{2}} \left(\left| ++ \right\rangle + \left| -- \right\rangle \right). \tag{44}$$

[7] Note that any Bell state can be transformed into any other simply by Alice or Bob applying a Pauli gate to their qubit.

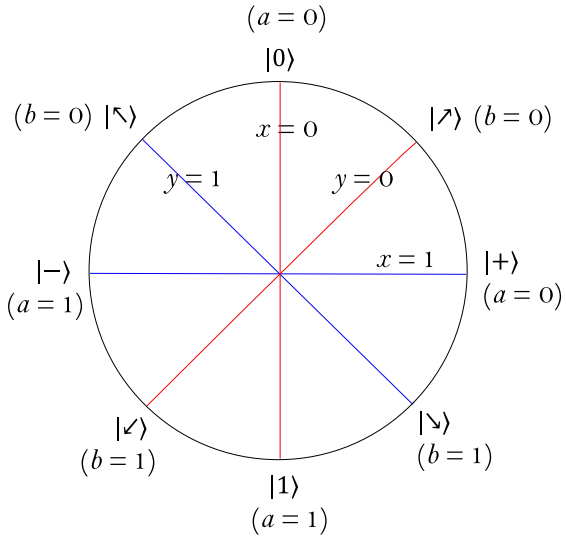

Figure 4. X-Z plane of the Bloch sphere. The measurement bases which allow Alice and Bob to maximise the value of the CHSH game are shown along with the corresponding values of x, y, a, and b. The bases are chosen such that there is always an angle of $\frac{\pi}{4}$ between the Bloch vectors corresponding to Alice and Bob's winning outputs, meaning that $P(a \oplus b = x \cdot y \mid x, y) = \cos^2(\frac{\pi}{8})$ for all x and y.

Alice will again choose to make her measurements in either the Z or X basis ($x = 0$ or $x = 1$, respectively). But this time we let Bob make his measurements in two different bases:

$$\left\{ |\nearrow\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle,\ |\swarrow\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle - \cos\left(\frac{\pi}{8}\right)|1\rangle \right\} \tag{45}$$

$$\left\{ |\nwarrow\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle - \sin\left(\frac{\pi}{8}\right)|1\rangle,\ |\searrow\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle + \cos\left(\frac{\pi}{8}\right)|1\rangle \right\} \tag{46}$$

As can best be seen geometrically from the Bloch sphere representation (figure 4), this gives a CHSH value of:

$$P(a \oplus b = x \cdot y \mid x, y) = \cos^2\left(\frac{\pi}{8}\right) = \frac{2+\sqrt{2}}{4} \quad \forall (x, y), \tag{47}$$

which is the Tsirelson bound. Hence we have a value of $E = \frac{\sqrt{2}}{2}$, which is larger than the maximal classical value by a factor of $\sqrt{2}$.

Nonlocality distillation

It is a well known result in quantum information theory that entanglement can be distilled [34]. Similarly, it turns out that in box world nonlocality can sometimes be distilled. What we mean by this is that there exist protocols by which a number of nonlocal boxes of a particular type can be used to simulate one or more different boxes which are *more* nonlocal. The first such protocol was discovered by Forster,

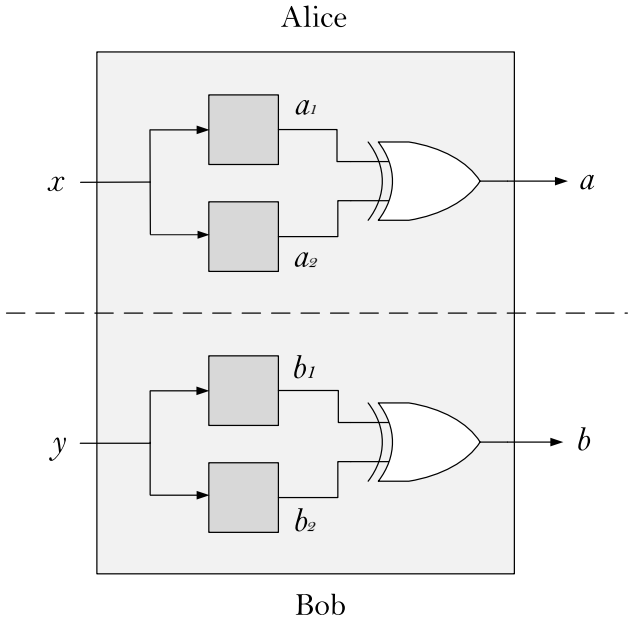

Figure 5. For certain types of nonlocal box, the distillation procedure given in [35] allows Alice and Bob to simulate a stronger box using just two copies. They enter their inputs into both boxes and then take the sum modulo 2 of their local outputs. These become the transformed outputs.

Winkler, and Wolf in [35]. I illustrate this protocol in figure 5 for the case of two nonlocal boxes being used to distil one stronger box.

If some class of nonlocal boxes can be used to distil even a single box that trivialises communication complexity (i.e. a box with a $P(a \oplus b = x \cdot y) \geq \frac{3+\sqrt{6}}{6}$), then that would give a reason to rule out such a class as being physically realisable. It turns out that the protocol in [35] does not allow boxes to be distilled with a CHSH value equal to or greater than $\frac{7}{8}$. However, other protocols have been discovered that do allow certain classes of nonlocal box to be ruled out by nontrivial communication complexity. For example, a protocol due to Brunner and Skrzypczyk [32] even works for some nonisotropic superquantum boxes that have a CHSH value arbitrarily close to the classical limit of $\frac{3}{4}$. Interestingly, no protocol found to date works with isotropic boxes. In fact there are partial results which suggest distillation using these boxes may be impossible [33, 36]. Finding a general proof of this is still an open question [37] but, if true, it would mean that isotropic boxes (and the nonisotropic boxes they can simulate, e.g. (28)) are inequivalent to other classes of isotropic box. In turn this would suggest that not all superquantum boxes can be ruled out by nontrivial communication complexity alone.

IV. INFORMATION CAUSALITY

We now turn our focus to the principle of information causality, which, if taken as an axiom alongside no-signalling, restricts us to nonlocal boxes having CHSH values no greater than those allowed by quantum mechanics. The principle requires that: "the information gain that Bob can reach about a previously unknown to him data set of Alice, by using all his local resources and m classical bits communicated

by Alice, is at most m bits" [2].

Generalised mutual information

The next step is to ask how Bob's information about Alice's N bits can be quantified. This is actually quite subtle, since it involves defining a mutual information measure in the context of generalised probabilistic theories which is compatible with the classical expression. It turns out that such a measure can be defined and is upper bounded by the classical expression, which in turn is lower bounded by

$$I_m \equiv N - \sum_i h(p_i), \tag{48}$$

where p_i is the probability of Bob correctly calculating Alice's ith bit, and $h(p)$ is the classical binary entropy [2] (that is, the Shannon entropy of a Bernoulli process):

$$h(p) = -p \log_2(p) - (1-p) \log_2(1-p). \tag{49}$$

Hence the information causality condition can be written mathematically as:

$$I_m \leq m. \tag{50}$$

In order to understand how this allows us to recover the Tsirelson bound, we frame the scenario as a communication complexity game.

Accessing Alice's bits

Bob is asked the value of one of Alice's N bits. Bob is not allowed to signal to Alice and communication is "expensive" so they want to restrict the number of bits Alice must send to Bob. This situation is equivalent to Bob being given some non-negative integer y and being asked to calculate the function $f = x_y$ where x_y is the yth bit of Alice's string. Classically, this would require Alice to send Bob all N of her bits since, without communication, there is no way she can know which bit Bob requires. With less than N bits of communication from Alice, the best Bob can do is to guess the answer, meaning he has a 50% chance of answering correctly. On the other hand, we have seen that if Alice and Bob share enough PR-boxes then Bob can compute this function correctly 100% of the time with just a single bit of communication from Alice. I now explain explicitly how they do this and show what happens if Alice and Bob use noisy PR-boxes instead. That is to say, I will show how Bob's probability of correctly calculating f depends on the parameter E.

Obviously when Alice has only one bit, no PR-boxes are necessary. The simplest non-trivial case is where Alice has two bits, x_0 and x_1, and Bob has one bit, y. The function Bob wants to calculate can be written in terms of elementary indicator functions as

$$f(\mathbf{x}, y) = x_y = (y \oplus 1) \cdot x_0 \oplus y \cdot x_1 \tag{51}$$

and rearranged to give

$$f = y(x_0 \oplus x_1) \oplus x_0. \tag{52}$$

So, if Alice and Bob respectively input $(x_0 \oplus x_1)$ and y into a single PR-box, they will get corresponding outputs a and b where $a \oplus b = y \cdot (x_0 \oplus x_1)$. In terms of these outputs, f can be written as

$$f = a \oplus x_0 \oplus b, \tag{53}$$

which means Alice need only send Bob the value of $a \oplus x_0$ (a single bit) for him to calculate f (by adding b to the bit Alice sends him). If the PR-box in this example is replaced by a noisy PR-box, then it is clear Bob will give the correct answer with probability $\frac{1}{2}(1 + E)$, since this is the probability that the box will give the same output as a noiseless PR-box.

Now let Alice have four bits: x_{00}, x_{01}, x_{10}, x_{11}. To specify which of these bits Bob should recover, he requires two bits: y_0, y_1. Bob needs to calculate $f(\mathbf{x}, \mathbf{y}) = x_{y_0 y_1}$. We start by writing this as

$$f = (y_0 \oplus 1)(y_1 \oplus 1)x_{00} \oplus (y_0 \oplus 1)y_1 x_{01} \oplus y_0(y_1 \oplus 1)x_{10} \oplus y_0 y_1 x_{11}. \tag{54}$$

Clearly y_0 specifies whether the bit Bob requires is in the left half or the right half of Alice's bit string, and y_1 specifies whether the required bit is the left or right bit of this sub-string. Recognising this, Alice and Bob can compute the function by recursively applying the previous protocol. That this is possible is seen most clearly by arranging (54) into the same form as (52):

$$f = y_0\Big([y_1(x_{00} \oplus x_{01}) \oplus x_{00}] \oplus [y_1(x_{10} \oplus x_{11}) \oplus x_{10}]\Big) \oplus [y_1(x_{00} \oplus x_{01}) \oplus x_{00}]\,. \tag{55}$$

To carry out this computation, Alice will need to use three PR-boxes since she has no information about Bob's bits. However, because only one of the two terms inside the square brackets needs to be evaluated, Bob will only need two of these boxes. Here I illustrate the algorithm step by step:

- Alice and Bob arrange their PR-boxes into two levels. On the first level there are two PR-boxes and on the second there is one.
- Alice starts with the first level. She inputs $(x_{00} \oplus x_{01})$ into the left box and $(x_{10} \oplus x_{11})$ into the right box, receiving corresponding outputs a_0 and a_1.
- Bob chooses either the left box or the right box on the first level, depending whether the bit he is interested in is in the left half or right half of Alice's bit string, respectively.
- Into his chosen box, Bob inputs y_1. This equals the position of the bit he wants in the corresponding sub-string. He receives output b_1.
- Rearranging f gives:

$$\begin{aligned} f &= y_0\Big([b_1 \oplus a_0 \oplus x_{00}] \oplus [b_1 \oplus a_1 \oplus x_{10}]\Big) \oplus [[b_1 \oplus a_0] \oplus x_{00}] \\ &= y_0\Big(a_0 \oplus x_{00} \oplus a_1 \oplus x_{10}\Big) \oplus y_0\Big(b_1 \oplus b_1\Big) \oplus [b_1 \oplus a_0 \oplus x_{00}] \\ &= y_0\Big(a_0 \oplus x_{00} \oplus a_1 \oplus x_{10}\Big) \oplus b_1 \oplus a_0 \oplus x_{00} \end{aligned} \tag{56}$$

- Alice now enters $(a_0 \oplus x_{00} \oplus a_1 \oplus x_{10})$ into the PR-box on the second level. She receives output a.

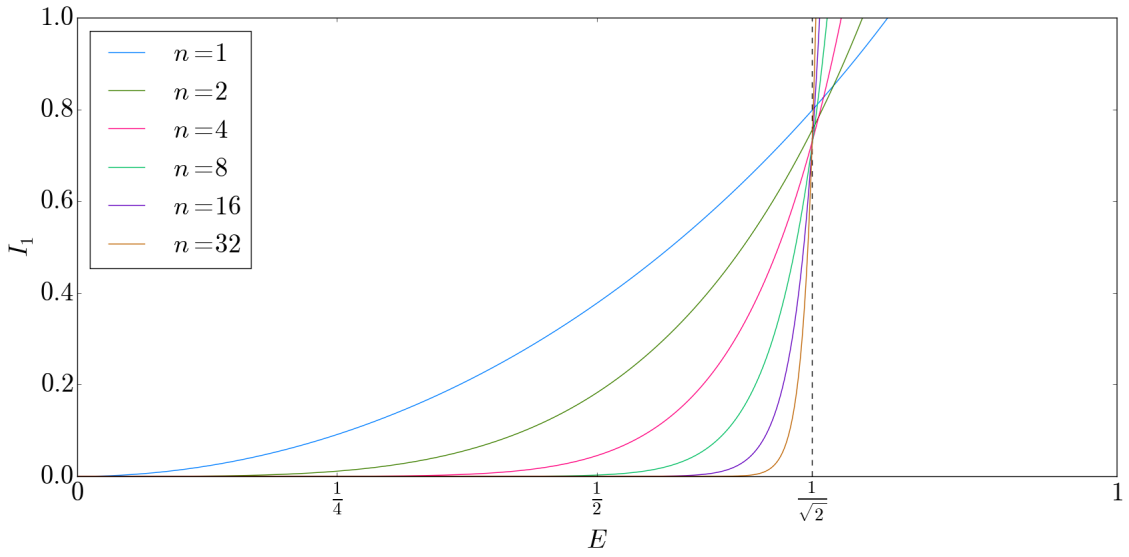

Figure 6. Visual illustration of the proof due to Pawłowski *et al.* [2] that, if Nature obeys the principle of information causality, we are restricted to noisy PR-boxes with a maximum value of $E = \frac{1}{\sqrt{2}}$. In other words, we recover the Tsirelson bound for the CHSH game. The value of I_1 is plotted against the noisy PR-box parameter E for various lengths of Alice's string ($N = 2^n$) . Information causality says that $I_1 \leq 1$. But this must be true no matter how many bits Alice has. So, letting n approach infinity, we see that information causality is violated by all isotropic boxes with $E > \frac{1}{\sqrt{2}}$, and hence by all nonlocal boxes with a CHSH value greater than $\frac{2+\sqrt{2}}{4}$ (since they can be depolarized into isotropic form).

- Bob enters y_0 into the PR-box on the second level. The value of this bit identifies whether Bob is interested in the left half ($y_0 = 0$) or right half ($y_0 = 1$) of Alice's string. He receives output b_0.
- Alice sends Bob a single bit equal to $(a \oplus a_0 \oplus x_{00})$.
- Bob now adds this bit from Alice to his two PR-box outputs. This final sum gives the value of f:
$$f = a \oplus b_0 \oplus b_1 \oplus a_0 \oplus x_{00}. \tag{57}$$
It's important to note that if $y_0 = 0$ then f is actually independent of a_1, whereas if $y_0 = 1$ then f is independent of a_0 (this follows from the simple fact that $a_0 \oplus a_0 = 0$). Therefore the final result only depends on the outputs of the two PR-boxes used by Bob. The third PR-box is only needed by Alice because she does not know which bit Bob is interested in.

This algorithm can readily be generalised, analogously to a binary search. If Bob has n bits and Alice has $N = 2^n$ bits, the protocol will involve n levels. A total of $(2^n - 1)$ PR-boxes are required, though the final result only depends on the n boxes chosen by Bob (one per level). The fact that only n (rather than $N - 1$) PR-boxes contribute to the value of f is important when we consider Alice and Bob using this same algorithm with noisy PR-boxes.

When a noisy PR-box is used, there is a probability of $\frac{1}{2}(1 - E)$ that it will not return the expected result, meaning $a \oplus b = x.y \oplus 1$. However, since $1 \oplus 1 = 0$, two such errors occurring in a sum of outputs will mean that the errors "cancel" each

other out. Therefore, the above algorithm will give the correct answer whenever an even number of errors occur. In the case of $n = 2$, this means that Bob's answer will be correct if either zero or two errors have occurred. I now show how we can use this to calculate the probability of Bob answering correctly as a function of E.

The probability of $a_0 \oplus b_1$ giving the expected result is $\frac{1}{2}(1 + E)$. But the probability of $a \oplus b_0$ giving the expected result is also $\frac{1}{2}(1 + E)$. Since f depends on $a_0 \oplus b_1 \oplus a \oplus b_0$, and since the outputs of the boxes are independent, the probabilities multiply. So the probability of zero errors is $\left[\frac{1}{2}(1 + E)\right]^2$. Similarly, the probability of two errors is $\left[\frac{1}{2}(1 - E)\right]^2$. Hence the total probability of Bob calculating f correctly is:

$$\left[\frac{1}{2}(1 + E)\right]^2 + \left[\frac{1}{2}(1 - E)\right]^2 = \frac{1}{4}(1 + 2E + E^2 + 1 - 2E + E^2) = \frac{1}{2}(1 + E^2) \quad (58)$$

This argument can be extended and the expression for the general case is $\frac{1}{2}(1 + E^n)$, as shown in [2].

The Tsirelson bound

Using the above result in equation (50) with $m = 1$, we thus have a condition on E. From this, the Tsirelson bound (corresponding to $E = \frac{1}{\sqrt{2}}$) can be recovered. I illustrate this graphically in figure 6, where I have plotted I_1 against E for various lengths of Alice's string $N = 2^n$.

Part III
Concluding remarks

In this work, we have looked at nonlocality through an abstract lens. The power of this approach has been shown through the proofs of two surprising results:

1. The van Dam theorem: systems that maximally saturate the CHSH inequality allow for two-party, distributed computation of functions of *any number* of bits with the requirement of just a single bit of communication between the parties. Neither quantum nor classical physics can make communication complexity trivial, and it is reasonable to believe that Nature herself would not allow such a phenomenon.

2. Information causality: systems with a CHSH game value larger than that allowed by quantum mechanics allow one party access to more than one bit of another party's data for each classical bit communicated to them. Again, this seems to fly in the face of our intuitions about information.

This line of work gives us clues about what a physically-motivated set of quantum mechanical axioms may look like. In fact, many other promising candidates have been suggested, including: macroscopic locality [38], nonlocality swapping [39], and local orthogonality [40].

We have also looked at these nonlocality results in terms of resources and have characterised some different types of nonlocal box. In particular we have looked at the difference between isotropic and nonisotropic boxes, and shown that all boxes can be put into isotropic form (whilst the converse is not possible).

From a practical perspective, we consider no-go theorems and quantify resources, as it helps us learn what can and what cannot be physically achieved with quantum computers. For example, this may lead us towards more practical cryptography protocols or even toward a quantum internet.

Finally, there remains the possibility that work on nonlocality may give us clues about a quantum theory of gravity. This, in turn, might tell us even more about the foundations of quantum mechanics.

Open questions and future work

This is an exciting area of research that is still relatively new, and there remain many open questions. In this work I have only been able to scratch the surface. Most noticeable is the fact that I have only looked at nonlocality in the simplest case: binary bipartite games. Extending results to more parties is difficult, since it introduces complexities not present in the bipartite case. Some results are known, but there is no general algorithm known for finding Bell inequalities in more complicated scenarios, let alone characterising nonlocality and the quantum boundary.

One part of this project is still in progress. That is, I have been trying to quantify nonlocality with respect to the information causality game. To do this I have been considering a less than optimum number of perfect PR-boxes in the protocol and comparing this to a full number of imperfect boxes. I have been considering replacing the "missing" PR-boxes with three different types of "free" resource: randomness, classical correlations and quantum correlations.

Possible future extensions to the present work include:

1. Looking at the depolarization of quantum systems with generalised measurements. Whilst it is easy to see how a quantum state can be depolarized when projective measurements are considered (the local states become maximally mixed), this is not the case with generalised measurements, since the measurements themselves may have to be transformed.

2. Identifying the class of all nonisotropic boxes which can be simulated by isotropic ones, and exploring why distillation fails for these machines.

3. Looking at the proof of information causality in terms of generalised entropy, instead of generalised mutual information, as per Al-Safi and Short [41].

4. Trying to clarify whether information causality can single out the full quantum set of correlations in the bipartite case, or whether it only allows us to recover the "almost quantum" set.

ACKNOWLEDGMENTS

First and foremost, the author would like to thank Dr David Jennings for his helpful supervision and feedback throughout the project. The author would also like to thank Dr Sania Jevtic for useful feedback following the literature review and project viva. Finally, the author acknowledges the assistance of Tom Laird in developing the signalling box example (9).

---

[1] W. van Dam, Natural Computing **12**, 9 (2012).
[2] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Nature **461**, 1101 (2009).
[3] J. A. Wheeler, *Geons, black holes, and quantum foam : a life in physics* (Norton, New York, 1998).
[4] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955).
[5] J. S. Bell, Physics **1**, 195 (1964).
[6] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
[7] D. Dieks, Physics Letters A **92**, 271 (1982).
[8] R. F. Werner, in *Quantum Information*, Springer Tracts in Modern Physics No. 173 (Springer Berlin Heidelberg, 2001) pp. 14–57, dOI: 10.1007/3-540-44678-8_2.
[9] O. Cohen, Fluctuation and Noise Letters **06**, C1 (2006).
[10] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, in *19th IEEE Annual Conference on Computational Complexity, 2004. Proceedings* (2004) pp. 236–249.
[11] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, Nature Communications **6**, 6288 (2015).
[12] F. Dowker, J. Henson, and P. Wallden, New Journal of Physics **16**, 033033 (2014).
[13] A. J. Short and J. Barrett, New Journal of Physics **12**, 033034 (2010).
[14] J. B. Kennedy, Philosophy of Science **62**, 543 (1995).
[15] S. Popescu and D. Rohrlich, Foundations of Physics **24**, 379 (1994).
[16] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Physical Review Letters **23**, 880 (1969).

[17] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Physical Review A **71**, 022101 (2005).

[18] B. S. Cirel'son, Letters in Mathematical Physics **4**, 93 (1980).

[19] L. Masanes, A. Acin, and N. Gisin, Physical Review A **73**, 012112 (2006).

[20] J. Barrett and S. Pironio, Physical Review Letters **95**, 140401 (2005).

[21] N. J. Cerf, N. Gisin, S. Massar, and S. Popescu, Physical Review Letters **94**, 220403 (2005).

[22] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Physical Review Letters **96**, 250401 (2006).

[23] V. Scarani, arXiv:quant-ph/0603017v2 **844**, 309 (2006), arXiv: quant-ph/0603017.

[24] W. van Dam, *Nonlocality & communication complexity*, Ph.D. thesis, Faculty of Physical Sciences, University of Oxford (1999).

[25] A. Y. Carmi and D. Moskovich, arXiv:1507.07514 [quant-ph] (2015), arXiv: 1507.07514.

[26] Y. Crama and P. L. Hammer, *Boolean Functions: Theory, Algorithms, and Applications*, 1st ed. (Cambridge University Press, Cambridge ; New York, 2011).

[27] M. Fitzi, E. Hänggi, V. Scarani, and S. Wolf, Journal of Physics A: Mathematical and Theoretical **43**, 465305 (2010).

[28] N. S. Jones and L. Masanes, Physical Review A **72**, 052312 (2005).

[29] F. Dupuis, N. Gisin, A. Hasidim, A. A. Méthot, and H. Pilpel, Journal of Mathematical Physics **48**, 082107 (2007).

[30] A. Broadbent and A. A. Méthot, Theoretical Computer Science **358**, 3 (2006).

[31] N. Brunner, N. Gisin, and V. Scarani, New Journal of Physics **7**, 88 (2005).

[32] N. Brunner and P. Skrzypczyk, Physical Review Letters **102**, 160403 (2009).

[33] A. J. Short, Physical Review Letters **102**, 180502 (2009).

[34] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Physical Review A **53**, 2046 (1996).

[35] M. Forster, S. Winkler, and S. Wolf, Physical Review Letters **102**, 120401 (2009).

[36] M. Forster, Physical Review A **83**, 062114 (2011).

[37] B. Lang, T. Vértesi, and M. Navascués, Journal of Physics A: Mathematical and Theoretical **47**, 424029 (2014).

[38] M. Navascués and H. Wunderlich, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **466**, 881 (2010).

[39] P. Skrzypczyk, N. Brunner, and S. Popescu, Physical Review Letters **102**, 110402 (2009).

[40] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Nature Communications **4**, 2263 (2013).

[41] S. W. Al-Safi and A. J. Short, Physical Review A **84**, 042323 (2011).

[42] L. A. Khalfin and B. S. Tsirelson, in *Symposium on the foundations of modern physics*, Vol. 85 (Singapore: World Scientific, 1985) p. 441.

Appendix A: Statistical independence and no-signalling

Here I prove that the no-signalling condition (2) can be derived from the following three conditions:

1. independence of Alice and Bob's inputs, i.e. $(X = x)$ and $(Y = y)$ are independent events;

2. the combination of Alice's input and output is independent of Bob's input, i.e. $(A = a, X = x)$ and $(Y = y)$ are independent events;

3. the combination of Bob's input and output is independent of Alice's input, i.e. $(B = b, Y = y)$ and $(X = x)$ are independent events.

Proof. Assuming $P(X) \neq 0$ and $P(Y) \neq 0$, we have from the first two conditions:

$$P(A|X,Y) = \frac{P(A,X,Y)}{P(X,Y)} = \frac{P(A,X)\cdot P(Y)}{P(X)\cdot P(Y)} = \frac{P(A,X)}{P(X)} = P(A|X). \tag{A1}$$

Similarly, from the first and third conditions it follows that

$$P(B|X,Y) = P(B|Y). \tag{A2}$$

□

These assumptions also imply that:

- $(A = a)$ and $(Y = y)$ are independent events,

- $(B = b)$ and $(X = x)$ are independent events.

Proof. Assuming $P(Y) \neq 0$,

$$\begin{aligned} P(A|Y) &= \sum_x \frac{P(A,X,Y)}{P(Y)} = \sum_x \frac{P(A|X,Y)\cdot P(X,Y)}{P(Y)} \\ &= \sum_x P(A|X)\cdot P(X) \\ &= P(A), \end{aligned} \tag{A3}$$

where the second line follows from the independence of X and Y, and the no-signalling condition. But, by definition,

$$P(A|Y) = \frac{P(A,Y)}{P(Y)},$$

so $P(A,Y) = P(A)\cdot P(Y)$.

Applying the same argument to B and X gives $P(B,X) = P(B)\cdot P(X)$. □

Appendix B: PR-boxes are no-signalling

In [15], Popescu and Rohrlich show that PR-boxes are no-signalling. However, an equivalent result was first given by Khalfin and Tsirelson in [42]. Here I give an independent proof of this important result.

Proof. From the definition of a PR-box (8), we have that:

$$P(a|xy) = \sum_b P(a,b|x,y) = \frac{1}{2} + 0 = \frac{1}{2} \tag{B1}$$

Next we note that:

$$P(a,x) \;=\; \textstyle\sum_y P(a,x,y) \;=\; \sum_y P(a|x,y)\cdot P(x,y) \tag{B2}$$

Inserting (B1) into (B2), we find:

$$P(a,x) = \sum_y \frac{1}{2}\cdot P(x,y) = \frac{1}{2}\cdot P(x)\sum_y P(y|x) = \frac{1}{2}\cdot P(x) \tag{B3}$$

Hence:

$$P(a|x) = \frac{P(a,x)}{P(x)} = \frac{\frac{1}{2}\cdot P(x)}{P(x)} = \frac{1}{2} = P(a|x,y) \tag{B4}$$

By symmetry we must also have:

$$P(b|y) = P(b|x,y) \tag{B5}$$

□

Appendix C: Local randomisation

Here I illustrate with an explicit example that if a "classical" box (i.e. one that can be simulated using local hidden variables) were reusable, then it could be converted by local randomisation alone into isotropic form. In particular, I show that a box described by

$$\begin{aligned} P(a\oplus b = x\cdot y \mid x,y) \;&= 1 \text{ if } x = y \\ &= \tfrac{1}{2} \text{ if } x \neq y, \end{aligned} \tag{C1}$$

can be converted into one with

$$P(a\oplus b = x\cdot y \mid x,y) = \frac{3}{4} \quad \forall (x,y). \tag{C2}$$

Let Alice and Bob share the mixed state described by the following density matrix:

$$\rho = \frac{1}{4}\left(\left|0001\right\rangle\!\left\langle 0001\right| + \left|0100\right\rangle\!\left\langle 0100\right| + \left|1011\right\rangle\!\left\langle 1011\right| + \left|1110\right\rangle\!\left\langle 1110\right|\right), \tag{C3}$$

and assume they make their measurements in the Pauli Z basis. This box is thus described by (C1).

However, in order to decide his output, let Bob now measure both qubits (this is allowed by the assumption that the box can be used twice) and call the outcomes b_0 and b_1. Then, with a 50% probability (e.g. by tossing a fair coin), let Bob flip one of these bits. To decide which, he calculates $f = b_0 \oplus b_1$ and then flips b_f. He still returns b_y as his output, but now there is a chance it may have been flipped. It is easy to see that the final probability distribution is now equivalent to that of the following ensemble:

$$\begin{aligned}\rho = \tfrac{1}{8}\big(&\big|0001\big\rangle\big\langle 0001\big| + \big|0100\big\rangle\big\langle 0100\big| + \big|1011\big\rangle\big\langle 1011\big| + \big|1110\big\rangle\big\langle 1110\big| \\ +&\big|0000\big\rangle\big\langle 0000\big| + \big|0110\big\rangle\big\langle 0110\big| + \big|1001\big\rangle\big\langle 1001\big| + \big|1111\big\rangle\big\langle 1111\big|\big)\end{aligned} \tag{C4}$$

Straightforward calculation shows that for any x and y, this ensemble gives a CHSH game value of $\frac{3}{4}$ as desired.

Appendix D: Depolarized boxes are isotropic

Here I show that the probability of winning the CHSH game using a "depolarized" box is independent of Alice and Bob's inputs, x and y. I use the result, proven in section III, that $a' \oplus b' = x' \cdot y' \iff a \oplus b = x \cdot y$, where a, b, x and y are related to a', b', x' and y' by equations (24) and (25).

Proof.

$$\begin{aligned}P(a \oplus b = x \cdot y \mid x, y) &= P(a' \oplus b' = x' \cdot y' \mid x, y) \\ &= P(a' \oplus b' = 0 \mid X' = 0, Y' = 0) \cdot P(X' = 0, Y' = 0 | x, y) \\ &\quad + P(a' \oplus b' = 0 \mid X' = 0, Y' = 1) \cdot P(X' = 0, Y' = 1 | x, y) \\ &\quad + P(a' \oplus b' = 0 \mid X' = 1, Y' = 0) \cdot P(X' = 1, Y' = 0 | x, y) \\ &\quad + P(a' \oplus b' = 1 \mid X' = 1, Y' = 1) \cdot P(X' = 1, Y' = 1 | x, y) \\ &= P(C_2 = x, C_2 = y) \cdot P(a' \oplus b' = 0 \mid 0, 0) \\ &\quad + P(C_2 = x, C_2 = y \oplus 1) \cdot P(a' \oplus b' = 0 \mid 0, 1) \\ &\quad + P(C_2 = x \oplus 1, C_2 = y) \cdot P(a' \oplus b' = 0 \mid 1, 0) \\ &\quad + P(C_2 = x \oplus 1, C_2 = y \oplus 1) \cdot P(a' \oplus b' = 1 \mid 1, 1) \\ &= \frac{1}{4} \sum_{x', y'} P(a' \oplus b' = x' \cdot y' | x', y') \\ &= P(a' \oplus b' = x' \cdot y')\end{aligned}$$

□