Nonlocal correlations as an information-theoretic resource

Jonathan Barrett,[1,2,*] Noah Linden,[3,†] Serge Massar,[1,2,‡] Stefano Pironio,[1,2,§] Sandu Popescu,[4,5,‖] and David Roberts[4,¶]
[1]*Physique Théorique, C.P. 225, Université Libre de Bruxelles, Boulevard du Triomphe, 1050 Bruxelles, Belgium*
[2]*Centre for Quantum Information and Communication, C.P. 165/59, Université Libre de Bruxelles, Avenue F. D. Roosevelt 50, 1050 Bruxelles, Belgium*
[3]*Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*
[4]*H.H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*
[5]*Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom*
(Received 16 April 2004; published 2 February 2005)

It is well known that measurements performed on spatially separated entangled quantum systems can give rise to correlations that are nonlocal, in the sense that a Bell inequality is violated. They cannot, however, be used for superluminal signaling. It is also known that it is possible to write down sets of "superquantum" correlations that are more nonlocal than is allowed by quantum mechanics, yet are still nonsignaling. Viewed as an information-theoretic resource, superquantum correlations are very powerful at reducing the amount of communication needed for distributed computational tasks. An intriguing question is why quantum mechanics does not allow these more powerful correlations. We aim to shed light on the range of quantum possibilities by placing them within a wider context. With this in mind, we investigate the set of correlations that are constrained only by the no-signaling principle. These correlations form a polytope, which contains the quantum correlations as a (proper) subset. We determine the vertices of the no-signaling polytope in the case that two observers each choose from two possible measurements with d outcomes. We then consider how interconversions between different sorts of correlations may be achieved. Finally, we consider some multipartite examples.

DOI: 10.1103/PhysRevA.71.022101 PACS number(s): 03.65.Ud, 03.67.−a

I. INTRODUCTION

In a typical Bell-type experiment, two entangled particles are produced at a source and move apart to separated observers. Each observer chooses one from a set of possible measurements and obtains some outcome. The joint outcome probabilities are determined by the measurements and quantum state. One of the more striking features of quantum mechanics is that joint outcome probabilities can violate a Bell-type inequality [1], indicating that quantum mechanics is not, in Bell's terminology, locally causal. This prediction has been confirmed in numerous laboratory experiments [2].

Abstractly this scenario may be described by saying that the two observers have access to a black box. Each observer selects an input from a range of possibilities and obtains an output. The box determines a joint probability for each output pair given each input pair. It is clear that a quantum state provides a particular example of such a box, with input corresponding to measurement choice and output to measurement outcome. More generally, boxes can be divided into different types. Some will allow the observers to signal to one another via their choice of input and correspond to two-way classical channels, as introduced by Shannon [3]. Others will not allow signaling—it is well known, for example, that any box corresponding to an entangled quantum state will not. This is necessary for compatibility between quantum mechanics and special relativity. Of the nonsignaling boxes, some will violate a Bell-type inequality. The significance of this can be spelt out in information-theoretic terms: separated observers without the box, who have access to preshared classical random data but no other resources and, in particular, who cannot communicate, will not be able to simulate the box. We refer to any such box (and to the corresponding correlations) as nonlocal.

In general, these boxes can be viewed as an information-theoretic resource. This is obvious in the case of signaling boxes or classical channels. However, it is also known that nonlocal correlations arising from an entangled quantum state, even though they cannot be used directly for signaling, can be useful in reducing the amount of signaling that is needed in communication complexity scenarios below what could be achieved with only shared random data [4]. A local black box is, of course, simply equivalent to some shared random data, which in turn (depending on the precise nature of the problem) may be better than nothing [5].

A good question to ask now is, can any set of nonsignaling correlations be produced by measurements on some quantum state? The answer, in fact, is no. This was shown by Popescu and Rohrlich [6], who wrote down a set of correlations that return a value of 4 for the Clauser-Horne-Shimony-Holt (CHSH) expression [7], the maximum value algebraically possible, yet are nonsignaling. The maximum quantum value is given by Tsirelson's theorem as $2\sqrt{2}$ [8]. These should be compared with the maximum value obtainable by noncommunicating classical observers, which is 2. Popescu and Rohrlich concluded that quantum mechanics is only one of a class of nonlocal theories consistent with causality. In

*Electronic address: jbarrett@ulb.ac.be
†Electronic address: n.linden@bristol.ac.uk
‡Electronic address: smassar@ulb.ac.be
§Electronic address: spironio@ulb.ac.be
‖Electronic address: s.popescu@bristol.ac.uk
¶Electronic address: david.roberts@bris.ac.uk

terms of our boxes, there are some boxes that are nonsignaling but are more nonlocal than is allowed by quantum mechanics. It is interesting to note that from an information-theoretic point of view, some of these latter are very powerful. For example, van Dam has shown [9] that two observers who have access to a supply of Popescu-Rohrlich-type boxes would be able to solve essentially any two-party communication complexity problem with only a constant number of bits of communication. This should be contrasted with the quantum case, for which it is known that certain communication complexity problems require at least n bits of communication even if unlimited shared entanglement is available [10].

In this work, we investigate the set of nonsignaling boxes, considering them as an information-theoretic resource. Clearly this set includes those corresponding to measurements on quantum states as a subset. The motivation for studying the wider set is partly that it is interesting for its own sake. This is true even though no correlations other than quantum correlations have so far been observed in nature. Our findings are preliminary, but it is already clear that the set of nonsignaling boxes has interesting structure, and one finds analogies with other information-theoretic resources, in particular with the set of entangled quantum states. This work is not, however, purely academic. Another motivation is that a better understanding of the nature of quantum correlations can be gained by placing them in a wider setting. Only in this way, for example, can one hope to answer Popescu and Rohrlich's original question, of why quantum correlations are not more nonlocal than they are. More generally, a proper understanding of the information-theoretic capabilities of quantum mechanics includes an understanding of what cannot be achieved as well as what can.

This article is organized as follows. In Sec. II A, we introduce the convex polytope that describes the set of nonsignaling correlations. In Sec. II B, we examine more closely the particular case of correlations involving two possible inputs, obtaining all the vertices of the corresponding polytope. We then consider, in Sec. II C, how interconversions between these extreme points may be achieved using local operations. Section III is devoted to three-party correlations, and in Sec. III D, we examine how extremal correlations correlate to the environment. We conclude with some open questions in Sec. IV.

II. TWO-PARTY CORRELATIONS

A. Definitions

The no-signaling polytope. A bipartite correlation box (hereafter, just "box") is defined by a set of possible inputs for each of Alice and Bob, a set of possible outputs for each, and a joint probability for each output pair given each input pair. We denote Alice's and Bob's inputs X and Y, respectively, and their outputs a and b. The joint probability of getting a pair of outputs given a pair of inputs is $p_{ab|XY}$. Since $p_{ab|XY}$ are probabilities, they satisfy positivity,

$$p_{ab|XY} \geq 0 \quad \forall\, a,b,X,Y, \tag{1}$$

and normalization,

$$\sum_{a,b} p_{ab|XY} = 1 \quad \forall\, X,Y. \tag{2}$$

In this work we only consider nonsignaling boxes; i.e., we require that Alice cannot signal to Bob by her choice of X and vice versa. This means that the marginal probabilities $p_{a|X}$ and $p_{b|Y}$ are independent of Y and X, respectively:

$$\sum_{b} p_{ab|XY} = \sum_{b} p_{ab|XY'} \equiv p_{a|X} \quad \forall\, a,X,Y,Y', \tag{3}$$

$$\sum_{a} p_{ab|XY} = \sum_{a} p_{ab|X'Y} \equiv p_{b|Y} \quad \forall\, b,Y,X,X'. \tag{4}$$

A concrete example of a correlation box is an experiment with two spin-$\frac{1}{2}$ particles, with the inputs X and Y labeling Alice's and Bob's analyzer settings and the outputs a and b labeling the experimental outcomes. In a quantum experiment like this one, it is generally the case that the outcome of the measurement is obtained as soon as the measurement is performed. In addition, the entanglement is destroyed after the measurements, so that if the experiment is to be repeated a new entangled state is needed. We define boxes to have the same properties. Alice can select her input at any time and obtains her output immediately, and similarly Bob. There may of course be a time delay between Alice selecting her input and Bob selecting his input, but this makes no difference to the correlations. Further, after a box is used once, it is destroyed and to repeat the experiment a new box is needed.

We will always consider that the number of possible inputs and outputs is finite. Since the above constraints are all linear, the set of boxes with a given number of inputs and outputs is a polytope, which we denote by $\mathcal{P}$. It is easy to see that the set is convex—if two boxes each satisfy the constraints, then a probabilistic mixture of them (defined in the obvious manner) will also do so.

The local polytope. In general, the set of nonsignaling boxes can be divided into two types: local and nonlocal. A box is local if and only if it can be simulated by noncommunicating observers with only shared randomness as a resource. This means that we can write

$$p_{ab|XY} = \sum_{\lambda} p_{\lambda} p_{a|X}(\lambda) p_{b|Y}(\lambda), \tag{5}$$

where λ is the value of the shared random data and p_{λ} is the probability that a particular value of λ occurs. We have that $p_{a|X}(\lambda)$ is the probability that Alice outputs a given that the shared random data was λ and the input was chosen to be X and similarly for $p_{b|Y}(\lambda)$.

We recall what is known about the set of local boxes (see, for instance, [11,12]). This set is itself a convex polytope, with vertices corresponding to local deterministic boxes (all $p_{a|X}$, $p_{b|Y}$ are 0 or 1). The positivity conditions of Eq. (1) are trivial facets of this polytope, while nontrivial facets correspond to Bell-type inequalities. Violation of the latter implies that a point lies outside the local polytope and that the corresponding box is therefore nonlocal. We denote the local polytope by $\mathcal{L}$.

Quantum mechanical correlations. Finally, there is a third set of interest: the correlations obtainable by measurements

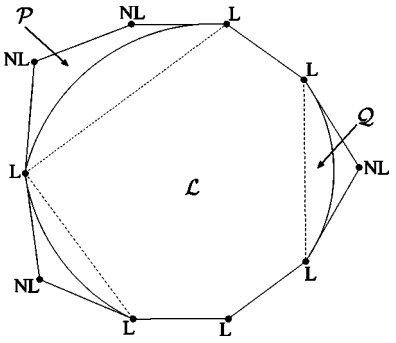

FIG. 1. A schematic representation of the space of nonsignaling correlation boxes. The vertices are labeled L and NL for local and nonlocal. Bell inequalities are the facets represented in dashed lines. The set bounded by these is $\mathcal{L}$. The region accessible to quantum mechanics is $\mathcal{Q}$. A general nonsignaling box $\in \mathcal{P}$.

on bipartite quantum states. We denote this set $\mathcal{Q}$ (where $\mathcal{Q}$ is defined for a fixed number of measurement settings and outcomes). The set $\mathcal{Q}$ is investigated in Refs. [8,12–15]. It is convex but is not a polytope as the number of extremal points is not finite. Since the correlations allowed by quantum mechanics can violate Bell inequalities, $\mathcal{Q}$ is nonlocal. However, as they violate the CHSH inequality only up to Tsirelson's bound of $2\sqrt{2}$ [6,8], they form a proper subset of the no-signaling polytope. Overall, we have that $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{P}$. This situation is illustrated in Fig. 1.

B. Two-input no-signaling polytope

1. Two outputs

Having defined the objects that we are interested in, we begin by considering in detail the simple case in which Alice and Bob are each choosing from two inputs, each of which has two possible outputs. We write $X,Y,a,b \in \{0,1\}$. The probabilities $p_{ab|XY}$ thus form a table with 2^4 entries, although these are not all independent due to the constraints of Sec. II A. The dimension of the polytope is found by subtracting the number of independent constraints from 2^4 and turns out to be 8. To understand the polytope $\mathcal{P}$, we wish to find its vertices. These will be boxes that satisfy all of the constraints and saturate a sufficient number of the positivity constraints to be uniquely determined. In the next subsection, we present an argument that allows us to find all the vertices of the two-input *d*-output polytope. Here we simply state the results for the simple two-input two-output case.

We find that there are 24 vertices, which may be divided into two classes: those corresponding to local boxes and those corresponding to nonlocal boxes. Local vertices are simply the local deterministic boxes, which assign a definite value to each of Alice's and Bob's inputs. There are thus 16 local vertices, which can be expressed as

$$p_{ab|XY} = \begin{cases} 1, & a = \alpha X \oplus \beta, \\ & b = \gamma Y \oplus \delta, \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where $\alpha,\beta,\gamma,\delta \in \{0,1\}$. Here and throughout, $\oplus$ denotes addition modulo 2.

The eight nonlocal vertices may be expressed compactly as

$$p_{ab|XY} = \begin{cases} 1/2, & a \oplus b = XY \oplus \alpha X \oplus \beta Y \oplus \gamma, \\ 0, & \text{otherwise}, \end{cases} \tag{7}$$

where $\alpha,\beta,\gamma \in \{0,1\}$. We will refer to these boxes as Popescu-Rohrlich (PR) boxes.

By using reversible local operations Alice and Bob can convert any vertex in one class into any other vertex within the same class. There are two types of reversible local operations. Alice may relabel her inputs, $X \to X \oplus 1$, and she may relabel her outputs (conditionally on the input), $a \to a \oplus \alpha X \oplus \beta$. Bob can perform similar operations. Thus up to local reversible transformations, each local vertex is equivalent to the vertex setting $\alpha=0$, $\beta=0$, $\gamma=0$, and $\delta=0$—i.e,

$$p_{ab|XY} = \begin{cases} 1, & a = 0 \text{ and } b = 0, \\ 0, & \text{otherwise}. \end{cases} \tag{8}$$

Each nonlocal vertex is equivalent to

$$p_{ab|XY} = \begin{cases} 1/2, & a \oplus b = XY, \\ 0, & \text{otherwise}. \end{cases} \tag{9}$$

We note that if we allow irreversible transformations on the outputs, we may convert any nonlocal vertex into a local vertex.

For the case of two inputs and two outputs, it is well known that the only nontrivial facets of the local polytope $\mathcal{L}$ correspond to the CHSH inequalities [16]. There is an important connection between the CHSH inequalities and the nonlocal vertices of $\mathcal{P}$. In order to explain this, we first recall explicitly the CHSH inequalities. Let $\langle ij \rangle$ be defined by

$$\langle ij \rangle = \sum_{a,b=0}^{1} (-1)^{a+b} p_{ab|X=i,Y=j}. \tag{10}$$

Then the nontrivial facets of $\mathcal{L}$ are equivalent to the following inequalities.

$$B_{\alpha\beta\gamma} \equiv (-1)^{\gamma}\langle 00 \rangle + (-1)^{\beta+\gamma}\langle 01 \rangle + (-1)^{\alpha+\gamma}\langle 10 \rangle + (-1)^{\alpha+\beta+\gamma+1}\langle 11 \rangle \leqslant 2, \tag{11}$$

where $\alpha,\beta,\gamma \in \{0,1\}$. For each of the eight Bell expressions $B_{\alpha\beta\gamma}$, the algebraic maximum is $B_{\alpha\beta\gamma}=4$. We find that for each choice of α, β, and γ the correlations defined by Eq. (7) return a value for the corresponding Bell expression of $B_{\alpha\beta\gamma}=4$. Thus there is a one-to-one correspondence between the nonlocal vertices of $\mathcal{P}$ and the nontrivial facets of $\mathcal{L}$, with each vertex violating the corresponding CHSH inequality up to the algebraic maximum. These extremal correlations describe in a compact way the logical contradiction in the CHSH inequalities.

2. d outputs

We now generalize the results of the preceding section. Again we have two parties, Alice and Bob, who choose from two inputs X and $Y \in \{0,1\}$ and receive outputs a and b with a joint probability $p_{ab|XY}$. We denote the number of distinct outputs associated with inputs X and Y by d_X^A and d_Y^B. If

Alice's input is X, for example, then $a \in \{0, \ldots, d_X^A - 1\}$.

Theorem 1. The nonlocal vertices of $\mathcal{P}$ for two input settings and d_X^A and d_Y^B outputs are equivalent under reversible local relabeling to

$$p_{ab|XY} = \begin{cases} 1/k, & (b-a) \bmod k = XY, \\ & a,b \in \{0, \ldots, k-1\}, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

for each $k \in \{2, \ldots, \min_{X,Y}(d_X^A, d_Y^B)\}$.

We note that the case $d_X^A = d_Y^B = 2$ gives the PR correlations we found previously. If $d_X^A = d_Y^B = k = d$, then the vertex violates the d-dimensional generalization of the CHSH inequality [17] up to its algebraic maximum. We call such a box a d-box (a more complete name would specify that the number of parties and the number of inputs per party are each two, but this simple name will do for our purposes).

Proof of Theorem 1. A probability table $p_{ab|XY}$ is a vertex of $\mathcal{P}$ if and only if it is the unique solution of Eqs. (1), (2), (3), and (4) with $\dim(\mathcal{P})$ of the positivity inequalities (1) replaced with equalities.

It will be useful to distinguish two kinds of extremal points: partial-output vertices and full-output vertices. Partial-output vertices are vertices for which at least one of the $p_{a|X} = 0$ or $p_{b|Y} = 0$. They can be identified with vertices of polytopes $\mathcal{P}'$ with fewer possible outputs: $d_X'^A < d_X^A$ or $d_Y'^B < d_Y^B$. Conversely, the vertices of a polytope $\mathcal{P}'$, with $d_X'^A < d_X^A$ or $d_Y'^B < d_Y^B$ can be extended to vertices of $\mathcal{P}$ by mapping the outcomes of X' and Y' to a subset of the outcomes of X and Y, and by assigning a zero probability $p_{a|X} = 0$ and $p_{b|Y} = 0$ to extra outcomes. Full-output vertices are vertices for which all $p_{a|X} \neq 0$ and $p_{b|Y} \neq 0$—i.e., for which all outputs contribute nontrivially to $p_{ab|XY}$. Thus the extremal points of a given two-setting polytope consist of the full-output vertices of that polytope and, by iteration, of all the full-output vertices of two-settings polytopes with fewer outcomes. Hence in the following, we need construct only the full-output vertices for a polytope characterized by d_X^A and d_Y^B.

The joint probabilities $p_{ab|XY}$ form a table of $\Sigma_{X,Y} d_X^A d_Y^B$ entries. These are not all independent because of the normalization and no-signaling conditions. There are four normalization equalities expressed by Eq. (2) and $\Sigma_X d_X^A + \Sigma_Y d_Y^B$ no-signaling equalities expressed by Eqs. (3) and (4). But for each value of X, the no-signaling condition for one of Alice's outputs can be deduced from the conditions of normalization and no-signaling for the $d_X^A - 1$ other outputs. A similar argument applies for each value of Y and Bob's outputs. Hence Eqs. (2), (3), and (4) form a set of only $4 + \Sigma_X (d_X^A - 1) + \Sigma_Y (d_Y^B - 1) = \Sigma_X (d_X^A) + \Sigma_Y (d_Y^B)$ linearly independent equations. The dimension of the no-signaling polytope is thus

$$\dim(\mathcal{P}) = \sum_{X,Y=0}^{1} d_X^A d_Y^B - \sum_{X=0}^{1} d_X^A - \sum_{Y=0}^{1} d_Y^B. \tag{13}$$

This is the number of entries in the table $p_{ab|XY}$ that must be set to zero to obtain a vertex. Moreover, to obtain a full-output vertex, these must be chosen so that neither $p_{a|X} = 0$ nor $p_{b|Y} = 0$. If we fix a particular pair of inputs (X,Y), then no more than $d_X^A d_Y^B - \max(d_X^A, d_Y^B)$ probabilities may be set to zero; otherwise, there will be fewer than $\max(d_X^A, d_Y^B)$ probabilities $p_{ab|XY} > 0$, and thus one of Alice's or one of Bob's outcomes will not be output for these values of X and Y. Because of the no-signaling conditions, it will not be output for the other possible pairs of inputs, so the vertex will be a partial-output one. Overall, the maximal number of allowed zero entries for a full-output vertex is

$$Z = \sum_{X,Y} [d_X^A d_Y^B - \max(d_X^A, d_Y^B)]. \tag{14}$$

Such a vertex is thus possible if $\dim(\mathcal{P}) \leqslant Z$ or

$$\sum_{X=0}^{1} d_X^A + \sum_{Y=0}^{1} d_Y^B \geqslant \sum_{X,Y=0}^{1} \max(d_X^A, d_Y^B). \tag{15}$$

This condition is fulfilled (with equality) only for $d_X^A = d_Y^B = d$, $\forall X, Y \in \{0,1\}$.

We can thus restrict our analysis to d-outcome polytopes. The extremal points of more general ones, where $d_X^A \neq d_Y^B$, will be the full-output extremal points of d-outcome polytopes for $d = 2, \ldots, \min_{X,Y}(d_X^A, d_Y^B)$.

Using $d_X^A = d_Y^B = d$, $\forall X, Y \in \{0,1\}$, in the discussion before Eq. (14), it follows that the dimension of a d-outcome polytope is $4d(d-1)$ and that for a given pair of inputs exactly $d(d-1)$ probabilities must be assigned the value zero or, equivalently, that d probabilities must be >0. We can therefore write the probabilities as

$$p_{ab|XY} \begin{cases} > 0 & \text{if } b = f_{XY}(a), \\ = 0 & \text{otherwise}, \end{cases} \tag{16}$$

where $f_{XY}(a)$ is a permutation of the d outcomes. Indeed, if $f_{XY}(a)$ is not a permutation, then at least one of Bob's outcomes will not be output.

We can relabel Alice's outcomes for $X = 0$ so that $f_{01}(a) = a$, those of Bob for $Y = 0$ so that $f_{00}(a) = a$, and finally those of Alice for $X = 1$ so that $f_{10}(a) = a$. In other words,

$$p_{ab|XY} \begin{cases} > 0 & \text{if } (b-a) \bmod d = 0, \\ = 0 & \text{otherwise}, \end{cases} \tag{17}$$

for $(X,Y) \in \{(0,0),(0,1),(1,0)\}$. It remains to determine f_{11}. It must be chosen so that the probability table $p_{ab|XY}$ is uniquely determined—i.e., so that specific values are assigned to the probabilities different from zero. In fact, it is easy to show that this can only be the case if the permutation f_{11} is of order d—i.e., $f_{11}^k(a) = a$ only for $k = 0 \bmod d$.

The only remaining freedom in the relabeling of the outcomes so that property (17) is conserved is to relabel simultaneously the outputs for all four possible inputs. We can relabel them globally so that $f_{11}(a) = (a+1) \bmod d$. This implies that $p_{ab|11} = 1/d$ if $(b-a) \bmod d = 1$. This completes the proof. ■

C. Resource conversions

In the preceding section we found all the vertices of the no-signaling polytope for bipartite, two-input boxes. As described in the Introduction, the ethos adopted in this work is

TABLE I. Comparison of information-theoretic resources.

Resource	Instantiation	Quantitative measure
Shared random data	Random variables	Mutual information
Shared secret data	Random variables	Secrecy rate
Entanglement	Quantum states	Entanglement cost
Nonlocality	Boxes	Classical simulation cost

that boxes (in particular, nonlocal boxes) can be regarded as an information-theoretic resource and investigated as such. Useful comparisons can be drawn with other information-theoretic resources, including shared random data [18], shared secret data [19,20], and entanglement [21]. In each case, there is a convex set of possible states and a notion of interconversion between different states. There is also a notion of interconversion between different resources. Each resource is useful for some task(s) and can be quantified via some measure(s). Some of this is illustrated in Table I. Note that the quantitative measures given are not the only possibilities. Note also that even if the given measure vanishes, a useful resource may still be present. Thus uncorrelated random variables can still be useful (as local randomness), as can separable quantum states (for various things) and as can local boxes (as local or shared randomness).

In light of this, it is natural to ask, what interconversions between boxes are possible and what would be a good measure of the nonlocality of a box? To the second question, several answers suggest themselves, such as the amount of classical communication needed to simulate the box (given that the only other resource is shared random data) and the degree of violation of Bell inequalities [22]. In this work, however, we concentrate on the first question—partly because it is independently interesting and partly because an understanding of possible interconversions is a prerequisite for a good understanding of quantitative measures.

The problem that we consider, then, is whether one can simulate one type of box using one or more copies of another type as a resource. Local operations such as relabeling are of course allowed. As nonlocality is the resource that we have in mind, it is also natural to allow the parties free access to local boxes (i.e., to local and shared randomness). We note, however, that neither local nor shared randomness can help if the box to be simulated is a vertex[1]; thus, none of the protocols we describe below make use of this. We make the assumption that communication between the parties is not allowed.

In general, outputs for one box can be used as inputs for another box. This allows nontrivial protocols to be constructed. As an interesting logical possibility, we note that the temporal order in which each party uses the boxes need not be the same and that this allows loops to be constructed that would be ill defined if it were not for the no-signaling condition. (Thus if signaling boxes were to be considered, our stipulation that outputs be obtained immediately after inputs would have to be altered.) Such a loop is illustrated in Fig. 2. In all of the protocols presented below, however, the parties use the boxes in the same temporal order.

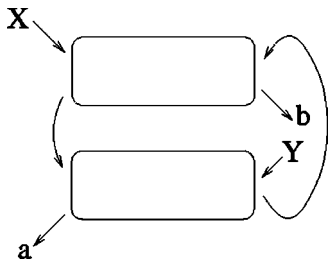

FIG. 2. An example of how two parties that are given two boxes may process locally their inputs and outputs. They result in simulating another type of box with inputs X,Y and outcomes a,b. Note that due to the no-signaling condition, the parties can use their two boxes with a different time ordering.

In the following, we will describe three simple examples. We show that given a d-box and a d'-box, we can simulate a dd'-box. We will also show that given a dd'-box, we can simulate one d-box. Finally, an unlimited supply of d-boxes can simulate a d'-box to arbitrarily high precision. In addition, we will describe a negative result: it is not in general possible to go *reversibly* from n d-boxes to m d'-boxes, where $d \neq d'$. Although we only prove this for exact transformations, we believe a similar result should hold even if transformations need only be exact in an asymptotic limit. It follows from this that d- and d'-boxes are ultimately inequivalent resources and that in our context, it is inappropriate to suppose that they can be characterized by a single numerical measure of nonlocality.[2]

Suppose first, then, that Alice and Bob have one d-box and one d'-box and they wish to simulate one dd'-box. Simulate means that for each value of $X \in \{0,1\}$, a procedure should be defined for Alice, using the d- and d'-boxes, which eventually enables her to determine the value of an output $a \in \{0,\ldots,dd'-1\}$. Similarly for Bob, for each value of Y

[1]This is easy to see. For each value of the local or shared randomness, one can write down the box that is simulated, conditioned on that value occurring. The box simulated by the overall protocol is then the average of these conditional boxes, with the average taken over the possible values of the randomness. But if this box is a vertex, then each of the conditional boxes must be the same vertex, and the protocol could have been carried out without the randomness.

[2]Similar considerations apply to the other resources we have mentioned. In the case of entanglement, for example, reversible interconversions are not in general possible for mixed states; thus, there is no unique measure of entanglement for mixed states. In the case of shared random data, interconversions by local operations are rather limited and provide no very meaningful measure of shared randomness. However, if one expands the set of operations that Alice and Bob are allowed, then the picture changes. Thus, in the case of shared random data, allowing that Alice and Bob can communicate classically, while demanding that the communication must be subtracted at the end, gives an operational meaning to the mutual information [18]. Inspired by this, it may be interesting to consider conversions between boxes, with classical communication allowed but subtracted at the end or, indeed, conversions between entangled quantum states with quantum communication allowed but subtracted at the end. We do not pursue these questions here.

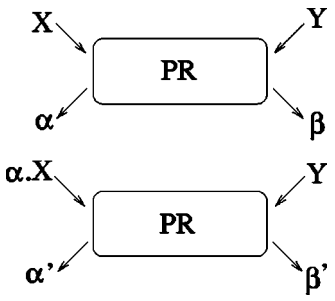

FIG. 3. Making a 4-box from two PR boxes. Alice inputs X into the first box and αX into the second, while Bob inputs Y into both boxes. Alice's output is given by $a=2\alpha'+\alpha$ and Bob's by $b=2\beta'+\beta$.

there is an eventual output b. The joint probabilities for a and b should satisfy Eq. (12) (with dd' inserted instead of d where necessary).

Protocol 1: 1 d-box and 1 d′-box → 1 dd′-box.

Alice. Alice inputs X into the d-box, obtaining outcome α. She then inputs X into the d'-box if $\alpha=d-1$ and inputs 0 into the d'-box otherwise, obtaining an output α'. Alice's output for the protocol is $a=\alpha' d+\alpha$.

Bob. Bob inputs Y into the d-box, obtaining output β, and inputs Y into the d'-box, obtaining output β'. His output for the protocol is then $b=\beta' d+\beta$.

Protocol 1 is illustrated in Fig. 3 for the case $d=d'=2$. We indicate briefly why this protocol works. Recall that a dd'-box satisfies $(b-a)\text{mod } dd'=XY$. Write $a=\alpha' d+\alpha$ and $b=\beta' d+\beta$, where α can take values $\alpha=0,\ldots,d-1$, α' can take values $\alpha'=0,\ldots,d'-1$, and so on. We see that the condition satisfied by a dd'-box is equivalent to

$$\beta-\alpha \bmod d = XY,$$

$$\beta'-\alpha' \bmod d' = \begin{cases} XY, & \alpha=d-1, \\ 0, & \text{otherwise}. \end{cases} \tag{18}$$

Protocol 1 is designed precisely to satisfy this condition. It is then not difficult to check that the correct probabilities are reproduced.

We note next that it is easy to convert one dd'-box into one d-box.

Protocol 2: 1 dd′-box→1 d-box.

Alice. Alice inputs X into the dd'-box, obtaining an output α. Her output for the protocol is then $a=\alpha \bmod d$.

Bob. Bob inputs Y into the dd'-box, obtaining an output β. His output for the protocol is $b=\beta \bmod d$.

Again, it is not difficult to check that $(b-a) \bmod d=XY$ and that the correct probabilities are reproduced.

Now we show how n d-boxes can be used to simulate a d'-box to arbitrarily high precision. This is done using a combination of Protocols 1 and 2.

Protocol 3: n d-boxes⇝1 d′-box.

Alice and Bob begin by using the n d-boxes to simulate a d^n-box, as per Protocol 1. Call the outputs for the d^n-box α and β. They satisfy $(\beta-\alpha)\text{mod } d^n=XY$. Alice and Bob now use Protocol 2 to obtain something close to a d'-box: the final outputs are $a=\alpha \bmod d'$ and $b=\beta \bmod d'$.

If $d^n=kd'$ for some positive integer k, this protocol works exactly. Otherwise, one can calculate the errors resulting in Protocol 3. Denote by k the largest integer such that $kd' \leqslant d^n$. Now we have that if $X=0$ or $Y=0$, then $(b-a)$ mod $d'=0$ as required. However, the probabilities are skewed by an amount $\propto 1/k \approx d'/d^n$. If $X=Y=1$, then the probabilities are skewed in a similar manner. But in addition we have that if $b=d^n-1$, then $(b-a)\text{mod } d'=1$ is not satisfied with probability $1/d^n$. The important thing here is that all errors tend to zero exponentially fast as n becomes large.

We have seen several examples of how interconversions between nonlocal extremal boxes are possible using only local operations. It is also interesting to consider how boxes may be simulated using only classical communication (CC) and shared random (SR) data—i.e., without other boxes. For example, we can see that one d-box may be simulated with one bit of one-way communication and $\log_2 d$ bits of shared randomness.

Protocol 4: 1 bit CC and $\log_2 d$ *bits* SR→*1 d-box.*

Alice and Bob share a random variable $\alpha \in \{0,\ldots,d-1\}$, where α takes all its possible values with equal probability $1/d$.

Alice. Alice sends her input X to Bob and outputs $a=\alpha$.

Bob. Bob, knowing X and α, outputs $b=(\alpha+XY) \bmod d$.

This protocol is optimal regarding the amount of one-way communication exchanged. This is a consequence of the following lemma, which places a lower bound on the amount of communication needed to simulate boxes. The lemma is used in the proof of Theorem 2, our final main result for this section.

Lemma 1. The simulation of n d-boxes using one-way communication requires at least n bits of communication if shared randomness is available and $n+n\log_2 d$ bits without shared randomness.

Proof. Note that this bound can be achieved using Protocol 4 for each of the n boxes, replacing if necessary $n\log_2 d$ bits of shared randomness by $n\log_2 d$ bits of communication from Alice to Bob.

Let us show that this amount of communication is necessary. Suppose first that both parties have access to shared random data and that communication is allowed from Alice to Bob. Bob's output is thus $b=b(\vec{\mathbf{Y}},C,r)$ where $\vec{\mathbf{Y}}=Y_1,\ldots,Y_n$ are the joint inputs for Bob, C is the communication, and r the shared data. Note simply that for Alice, there are 2^n possible joint inputs into n d-boxes. If Alice is sending fewer than n bits, there will be at least one pair of joint inputs for which her communication is the same. Call them $\vec{\mathbf{X}}_1$ and $\vec{\mathbf{X}}_2$. A careful examination of the definition of a d-box reveals that there will be at least one joint input of Bob's into the n boxes such that his output must be different according to whether Alice's input was $\vec{\mathbf{X}}_1$ or $\vec{\mathbf{X}}_2$. Thus $<n$ bits of communication are not sufficient.

If Alice and Bob do not have access to shared randomness, then Bob's output is of the form $b=b(\vec{\mathbf{Y}},C)$. The proof then follows by an argument similar to the one used above, noting that for Alice there are $2^{n+n\log_2 d}$ possible joint input-output pairs $(\vec{\mathbf{X}},\vec{\mathbf{A}})$. ■

These types of considerations will help us to establish the final result of this section.

Theorem 2. It is in general impossible, using local reversible operations, exactly to transform n d-boxes into m d'-boxes.

The theorem follows from the following two lemmas.

Lemma 2. Using n d-boxes, Alice and Bob can exactly simulate at most n d'-boxes, for $d \geqslant d'$.

Lemma 3. Using n d'-boxes, Alice and Bob can exactly simulate at most $n(1+\log_2 d')/(1+\log_2 d) < n$ d-boxes for $d' \leqslant d$.

Proof. We prove Lemma 2 as follows. We know that we can simulate n d-boxes with n bits of communication and $n \log_2 d$ bits of shared randomness. Suppose that there were a protocol using only local operations that could convert n d-boxes into N d' boxes, for some $d' \leqslant d$, where $N > n$. Then, by combining the simulation of the d-boxes with the protocol for their conversion, we would have constructed a protocol for simulating N d'-boxes using only n bits of communication, in contradiction with Lemma 1. The proof of Lemma 3 is very similar. Note that we can simulate n d'-boxes with $n+n \log_2 d'$ bits of classical communication and no shared randomness. Suppose that there were a protocol that converts n d'-boxes into N d-boxes, for some $d \geqslant d'$, where $N > n(1+\log_2 d')/(1+\log_2 d)$. As argued above, it follows from the fact that d-boxes are vertices that this protocol would not need any additional shared randomness. Then we would have constructed a protocol for simulating N d-boxes using only $n+n \log_2 d'$ bits of communication and no shared randomness, again in contradiction with Lemma 1. ■

III. THREE-PARTY CORRELATIONS

A. Definitions

In this section, we generalize the considerations of the previous sections to consider tripartite correlations. As before, we consider that correlations are produced by a black box with specified inputs and outputs, but now the box is assumed to be shared between three separated parties A, B, and C.

The no-signaling polytope. A box is defined by joint probability distributions $p_{abc|XYZ}$, which satisfy positivity,

$$p_{abc|XYZ} \geqslant 0 \quad \forall\, a,b,c,X,Y,Z, \tag{19}$$

normalization,

$$\sum_{a,b,c} p_{abc|XYZ} = 1 \quad \forall\, X,Y,Z, \tag{20}$$

and no-signaling. With three parties it is possible to imagine various types of communication, and correspondingly there are different types of no-signaling conditions. Obviously we require that A cannot signal to B or C (and cyclic permutations). We should also, however, require the stronger condition that if the systems B and C are combined, then A cannot signal to the resulting composite system BC. This is expressed by

$$\sum_{a} p_{abc|XYZ} = \sum_{a} p_{abc|X'YZ} \quad \forall\, b,c,Y,Z,X,X', \tag{21}$$

where, again, we include cyclic permutations. Finally, note that if systems A and B are combined, the resulting composite system AB should not be able to signal to C. This type of condition does not require a separate statement, however, as it already follows from Eq. (21). Indeed, using the fact that A cannot signal to BC and that B cannot signal to AC, we deduce

$$\begin{aligned}\sum_{a,b} p_{abc|XYZ} &= \sum_{a,b} p_{abc|X'YZ} \quad \forall\, b,c,X,X',Y,Z \\ &= \sum_{a,b} p_{abc|X'Y'Z} \quad \forall\, c,X,X',Y,Y',Z,\end{aligned} \tag{22}$$

which is the condition that AB cannot signal to C. Hence the only conditions we need to impose on a tripartite box are those of Eqs. (19), (20), and (21). The set of boxes satisfying these conditions is the polytope $\mathcal{P}$.

Locality conditions. In the tripartite case, as well as different types of no-signaling condition, there are different types of locality condition. First, a box is fully local if the probabilities can be written in the form

$$p_{abc|XYZ} = \sum_{\lambda} p_\lambda p_{a|X}(\lambda) p_{b|Y}(\lambda) p_{c|Z}(\lambda). \tag{23}$$

The set of such boxes is a convex polytope denoted $\mathcal{L}$. Second, we say that a box is two-way local if either there exists a bipartition of the parties—say, AB versus C—such that the composite system AB is local versus C or if the box can be written as a convex combination of such boxes—i.e.,

$$\begin{aligned}p_{abc|XYZ} = {} & p_{12} \sum_{\lambda_{12}} p_{\lambda_{12}} p_{ab|XY}(\lambda_{12}) p_{c|Z}(\lambda_{12}) \\ & + p_{13} \sum_{\lambda_{13}} p_{\lambda_{13}} p_{ac|XZ}(\lambda_{13}) p_{b|Y}(\lambda_{13}) \\ & + p_{23} \sum_{\lambda_{23}} p_{\lambda_{23}} p_{bc|YZ}(\lambda_{23}) p_{a|X}(\lambda_{23}),\end{aligned} \tag{24}$$

where $p_{12}+p_{23}+p_{13}=1$. The set of such boxes is again a convex polytope, denoted $L2$. Finally, any box that cannot be written in this form demonstrates genuine three-way nonlocality. We have that $\mathcal{L} \subset \mathcal{L}2 \subset \mathcal{P}$ and also that $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{P}$.

In the following, we restrict our attention to the case $a,b,c,X,Y,Z \in \{0,1\}$. We find the vertices of the polytope $\mathcal{P}$ and point out some connections with three-party Bell-type inequalities. Finally we consider some examples of interconversions, in particular of how to construct tripartite boxes using PR boxes as a resource.

B. Two inputs and two outputs

For the tripartite boxes with two inputs and two outputs per observer, Eq. (20) expresses eight normalization constraints, and Eq. (21) expresses $3 \times 16 = 48$ no-signaling constraints. However, as in the bipartite case, there is also some further redundancy; there turn out to be 38 independent constraints. Therefore the dimension of this polytope is $\dim \mathcal{P} = 2^6 - 38 = 26$.

Finding the vertices of a polytope given its facets is the so-called "vertex enumeration problem" for which several algorithms are available, although they are efficient only for low dimensional problems. We determined the extreme

points of our three-party polytope, with both *Porta* [23] and *cdd* [24]. It turns out that there are 46 classes of vertices, where vertices within one class are equivalent under local relabeling operations and permutations of the parties. These 46 classes of extreme points can be divided into three categories: local, two-way local and three-way nonlocal.

Local vertices. This category contains boxes for which *A*'s, *B*'s and *C*'s outputs are all deterministic. They all belong to the same class under reversible local operations, a representative of which is

$$p_{abc|XYZ} = \begin{cases} 1, & a=0, b=0, c=0, \\ 0, & \text{otherwise}. \end{cases} \tag{25}$$

Two-way local vertices. In view of the preceding discussion for bipartite correlations, there is only one class of extremal two-way local correlations that are not fully local. This is because if a box is a vertex, there can be only one term in the decomposition on the right-hand side of Eq. (24). Then it follows from Theorem 1 that this term must describe a PR box shared between two parties, along with a deterministic outcome for the third party. Thus any box of this type is equivalent under local relabelings and permutations of parties to

$$p_{abc|XYZ} = \begin{cases} 1/2, & a \oplus b = XY \text{ and } c=0, \\ 0, & \text{otherwise}. \end{cases} \tag{26}$$

Three-way nonlocal vertices. This category contains genuine three-party nonlocal extremal correlations. It is much more complex than the two above, since it comprises 44 different classes of vertices. Out of these, we mention 3 classes of particular interest. The first class can be expressed as

$$p_{abc|XYZ} = \begin{cases} 1/4, & a \oplus b \oplus c = XY \oplus XZ, \\ 0, & \text{otherwise}. \end{cases} \tag{27}$$

If we imagine that *B* and *C* form a composite system with input $Y \oplus Z$ and output $b \oplus c$, then this is a PR box shared between *A* and *BC*. We refer to them as "X(Y+Z)" boxes.

Correlations in the second class are equivalent to

$$p_{abc|XYZ} = \begin{cases} 1/4, & a \oplus b \oplus c = XY \oplus YZ \oplus XZ, \\ 0, & \text{otherwise}. \end{cases} \tag{28}$$

We call them "Svetlichny" correlations (for reasons explained below).

Finally, the third class contains what we call "XYZ" correlations.

$$p_{abc|XYZ} = \begin{cases} 1/4, & a \oplus b \oplus c = XYZ, \\ 0, & \text{otherwise}. \end{cases} \tag{29}$$

The XYZ correlations are special because, as van Dam pointed out to us [25], they can be used to solve any three-party communication complexity problem with only 1 bit broadcast by each party. He also pointed out that they have a natural generalization to *n* parties: $a_1 \oplus a_2 \oplus \cdots \oplus a_n = X_1 X_2 \ldots X_n$, where $X_i \in \{0,1\}$ is the input of party *i* and $a_i \in \{0,1\}$ the output of party *i*. These *n*-party correlations can be used to solve any *n*-party communication complexity problem with 1 bit broadcast by each party. They can be constructed from a supply of PR boxes.

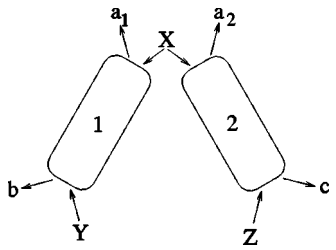

FIG. 4. Making an X(Y+Z) box from two PR boxes. Alice outputs $a = a_1 \oplus a_2$, Bob outputs *b* and Charles outputs *c*.

We conclude this section with some remarks on these correlation vertices and known multipartite Bell-type inequalities. First, each of the X(Y+Z), XYZ, and Svetlichny boxes violates the Mermin-Klyshko inequality [26,27] up to the algebraic maximum. Second, we recall that inequalities can be written down that detect genuine three-way nonlocality. One such is the Svetlichny inequality [28]. If we define $\langle ijk \rangle$ by

$$\langle ijk \rangle = \sum_{a,b,c} (-1)^{a+b+c} p_{abc|X=i,Y=j,Z=k}, \tag{30}$$

then the Svetlichny inequality is

$$M = \langle 000 \rangle + \langle 001 \rangle + \langle 010 \rangle + \langle 100 \rangle - \langle 011 \rangle - \langle 101 \rangle - \langle 110 \rangle - \langle 111 \rangle \leqslant 4. \tag{31}$$

Any local or two-way local box must satisfy this inequality. Quantum mechanically we can obtain $M = 4\sqrt{2}$ using a Greenberger-Horne-Zeilinger (GHZ) state [29] (although note that different measurements are needed from those that produce the well known GHZ paradox [30]). X(Y+Z) boxes do not violate the Svetlichny inequality as written (although they must violate some Svetlichny-type inequality as they are three-way nonlocal). Svetlichny boxes give $M=8$, the algebraic maximum of the expression (hence their name); XYZ correlations again do not violate the Svetlichny inequality as written, but return $M=6$ after the local relabeling $a \to a \oplus X$, $b \to b \oplus Y$, $c \to c \oplus Z \oplus 1$.

From the fact that some quantum states violate the Svetlichny inequality, we can conclude that in the two-input two-output case, $\mathcal{Q} \not\subset L2$. From the fact that bipartite correlations can be more nonlocal than quantum mechanics allows, we can also conclude that $\mathcal{L}2 \not\subset \mathcal{Q}$.

C. Simulating tripartite boxes

We consider how we may simulate some of these tripartite boxes, using a supply of PR boxes as a resource. We will give three examples, showing how to simulate an X(Y+Z) box with two PR boxes, a Svetlichny box with three PR boxes, and an XYZ box with three PR boxes.

First, suppose that two PR boxes are shared, with box 1 between Alice and Bob and box 2 between Alice and Charles. The following protocol shows how the three observers may simulate one X(Y+Z) box (see Fig. 4).

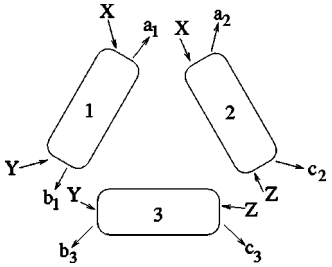

FIG. 5. Making a Svetlichny box from three PR boxes. Alice outputs $a=a_1\oplus a_2$, Bob outputs $b=b_1\oplus b_3$, and Charles outputs $c=c_2\oplus c_3$.

Protocol 5: 2 PR boxes→1 X(Y+Z) box.

Alice. Alice inputs X into box 1 and box 2, obtaining outputs a_1 and a_2. She then outputs $a=a_1\oplus a_2$.

Bob. Bob inputs Y into box 1, obtaining output b.

Charles. Charles inputs Z into box 2 obtaining output c.

The protocol works because

$$a\oplus b\oplus c=a_1\oplus a_2\oplus b\oplus c=XY\oplus XZ. \tag{32}$$

Suppose now that three PR boxes are shared, with box 1 between Alice and Bob, box 2 between Alice and Charles, and box 3 between Bob and Charles. Protocol 6 (summarized in Fig. 5) allows them to simulate one Svetlichny box.

Protocol 6: 3 PR boxes→1 Svetlichny box.

Alice. Alice inputs X into both box 1 and box 2, obtaining a_1 and a_2. Her final output is $a=a_1\oplus a_2$.

Bob. Bob inputs Y into both box 1 and box 3, obtaining b_1 and b_3. His final output is $b=b_1\oplus b_3$.

Charles. Charles inputs Z into both box 2 and box 3, obtaining c_2 and c_3. His final output is $c=c_2\oplus c_3$.

This works because

$$\begin{aligned} a\oplus b\oplus c&=a_1\oplus b_1\oplus b_3\oplus c_3\oplus a_2\oplus c_2\\ &=XY\oplus YZ\oplus XZ. \end{aligned} \tag{33}$$

Protocol 7 (summarized in Fig. 6) shows how to simulate one XYZ box using three PR boxes.

Protocol 7: 3 PR boxes→1 XYZ box.

Alice. Alice inputs X into box 1, obtaining an output a_1. She then inputs a_1 into box 2, obtaining output a_2. Alice's output for the protocol is $a=a_2$.

Bob. Bob inputs Y into box 1, obtaining an output b_1. He then inputs b_1 into box 3, obtaining output b_3. Bob's output for the protocol is $b=b_3$.

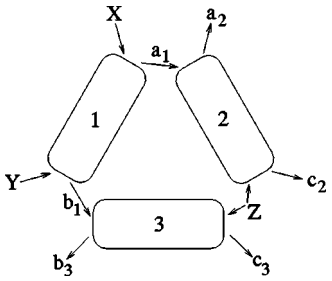

FIG. 6. Making an XYZ box from three PR boxes. Alice outputs $a=a_2$, Bob outputs $b=b_3$, and Charles outputs $c=c_2\oplus c_3$.

Charles. Charles inputs Z into both boxes 2 and 3, obtaining outputs c_2 and c_3. Charles' output for the protocol is $c=c_2\oplus c_3$.

The protocol works because

$$a\oplus b\oplus c=a_2\oplus b_3\oplus c_2\oplus c_3=Za_1\oplus Zb_1=XYZ. \tag{34}$$

Finally, we note that it is of course possible to perform conversions among tripartite boxes. For example, it is easy to see how to make one Svetlichny box using two XYZ boxes. The protocol is obvious once it is realized that a Svetlichny box is locally equivalent to a box defined by Eq. (28) with $XY\oplus YZ\oplus XZ$ on the right-hand side replaced by $XYZ\oplus(1\oplus X)(1\oplus Y)(1\oplus Z)$. We omit the details.

D. Nonlocality and the environment

Suppose that we have some three-party no-signaling distribution $p_{abe|XYE}$ with parties A, B, and E. We will show that if the reduced probability distribution $p_{ab|XY}=\Sigma_e p_{abe|XYE}$ is a vertex of the bipartite no-signaling polytope, then the composite system AB is local versus E. This is analogous to the result that pure quantum states cannot be entangled with a third party or the environment. It means that extremal nonlocal correlations cannot be correlated to any other system. (Note that this raises interesting new possibilities for cryptography. These are investigated in Ref. [31].)

By Bayes' theorem,

$$\begin{aligned} p_{abe|XYE}&=p_{ab|XYEe}p_{e|XYE}\\ &=p_{ab|XYEe}p_{e|E}, \end{aligned} \tag{35}$$

where we have used the fact that AB cannot signal to E to deduce the second equality. The condition that E cannot signal to AB implies

$$\begin{aligned} p_{ab|XY}&=\sum_e p_{abe|XYE}\quad \forall E\\ &=\sum_e p_{ab|XYEe}p_{e|E}\quad \forall E. \end{aligned} \tag{36}$$

For each value E, the last equality provides a convex decomposition of $p_{ab|XY}$ in terms of non-signaling correlations, with e playing the role of the shared randomness. Since we supposed that $p_{ab|XY}$ is extremal, this decomposition is unique and $p_{ab|XYEe}=p_{ab|XY}\ \forall e,E$. We then deduce

$$p_{abe|XYE}=p_{ab|XY}p_{e|E}, \tag{37}$$

i.e., that AB is uncorrelated with E.

A natural question that we leave as an open problem is whether the converse is true: if $p_{ab|XY}$ is in the interior of the no-signaling polytope, is it always possible to extend it to a tripartite distribution $p_{abe|XYE}$ such that AB is nonlocal versus E? (It is always possible, if $p_{ab|XY}$ is not a vertex, to write it as $p_{ab|XY}=\Sigma_e p_{abe|XYE}$, where E takes the single value $E=0$. One can also require that E take several values, in such a way that $p_{abe|XYE}$ is nonsignaling. What is nontrivial is the requirement that $p_{abe|XYE}$ is nonlocal in the partition AB versus E. We do not know if this is possible in general.)

IV. DISCUSSION AND OPEN QUESTIONS

In conclusion, we have defined nonsignaling correlation boxes and investigated their potential as an information-theoretic resource. Once the structure of the set of such boxes is understood as a convex polytope, it is clear that there are analogies with other information-theoretic resources, in particular the resource of shared quantum states (with nonlocality taking the place of entanglement). With this in mind, we have shown how various interconversions between boxes are possible. The set of multipartite boxes in particular appears very rich. Finally, we furthered the analogy with quantum states by demonstrating how nonlocality is monogamous, in much the same way that entanglement is monogamous. We finish with some open questions.

Nonlocal vertices and Bell inequalities. We saw in Sec. II B 1 that for the two-input two-output polytope there is a one-to-one correspondence between extremal nonlocal correlations and facet Bell inequalities (nontrivial facets of the local polytope). One might wonder whether this one-to-one correspondence holds in general. It appears, however, that for more complicated situations, involving more possible inputs or outputs, it does not. It would be interesting to investigate what is the precise relation between nonlocal vertices and facet Bell inequalities. This might help understand further the geometrical structure of nonlocal correlations.

Other vertices. We have given a complete characterization of two-input extremal nonlocal boxes in the bipartite case and presented some examples in the tripartite case. In general, one might also consider extremal boxes involving more inputs, more outcomes, or more parties.

For instance, a natural way to generate more complex boxes is by taking products of simpler ones. Suppose Alice and Bob have access to two boxes $p^0_{a_0b_0|X_0Y_0}$ and $p^1_{a_1b_1|X_1Y_1}$, where for simplicity we consider that there are M possible inputs and d possible outputs for each box. If Alice inputs X_0 and X_1 in each of the two boxes and outputs $a=da_1+a_0$ and similarly for Bob, they have now produced a nonlocal box with M^2 inputs and d^2 outputs $p_{ab|XY}=p^0_{a_0b_0|X_0Y_0}p^1_{a_1b_1|X_1Y_1}$, where $X=MX_1+X_0$ and similarly for Y. If the two original boxes were extremal for the (M,d) polytope, will the product be extremal for the (M^2,d^2) polytope? In the case of quantum states, the analogous result of course holds—a product of two pure states is itself a pure state. We have been able to show that in the case of boxes, the result holds provided that we restrict ourselves to extremal boxes with the following property: the output of one party is uniquely determined when the two inputs and the other party's output are specified. This is true for all the vertices presented in this paper. Plausibly it is true for all vertices, but this is not proven.

Interconversions. We have so far been able to achieve only a limited set of interconversions between extremal boxes. This is especially true for the three party case, where there are 46 classes of vertices and we have investigated only 5 of these. Understanding what kinds of interconversions between extremal boxes are possible is necessary to appraise their relative power as an information-theoretic resource.

The motivation is also to answer the general question of whether there exist inequivalent types of nonlocal correlations. Note, for instance, that the three-way nonlocal correlations of Eqs. (27), (29), and (28) cannot be reduced to two-way nonlocal ones using only local operations. This follows from the fact that the outcomes for two out of the three parties are totally independent of one another (unless the outcome of the third party is communicated to them). In this sense genuinely tripartite extremal correlations and bipartite extremal correlations belong to inequivalent classes. Are there inequivalent classes of bipartite extremal correlations? In other words, are there two bipartite extremal boxes such that one cannot simulate the other even approximately, no matter how many copies are available?

Another problem is whether all bipartite and multipartite correlations can be constructed using PR boxes, as is the case for all the extremal boxes presented in this paper (and thus also for probabilistic mixtures of them). PR boxes could then be viewed as the unit of nonlocal correlation, in analogy with the bit, qubit, and ebit, which are the units of classical and quantum information-theoretic resources.

Interior points. We have only considered conversions between extremal probability distributions. It would be interesting to consider the interior points of the polytope, which include quantum correlations. In particular we would like to find out if distillation of such mixed correlations is possible—i.e., if given a number of copies of a mixed box we can by local operations obtain some number of extremal boxes. Note that Tsirelson's bound [13] shows that the quantum correlations $\mathcal{Q}$ are a proper subset of the set of all nonsignaling correlations $\mathcal{P}$. Thus it is impossible to distill correlations in $\mathcal{Q}$ to extremal correlations. But apart from this, we do not know of any constraint on possible distillation of nonlocal correlations.

Finally, one could consider distillation in a new context, where we allow some communication between the parties but account for it at the end of the protocol (as noted above, an analogous approach was considered in Ref. [18] in the context of classical distillation of shared randomness). Alternatively, following Ref. [20], one could introduce a new element: that of secrecy. Suppose that inputs and outputs are considered to be secret and that Alice and Bob have a supply of noisy (that is nonextremal) boxes. Can Alice and Bob distill a supply of extremal boxes, whose inputs and outputs are also secret, via public communication?

As we outlined in the Introduction, nonlocal extremal correlations can be a very powerful resource for communication complexity problems. This will also be the case for correlations that can be distilled to these with no or little communication. On the other hand, Tsirelson's bound and results in communication complexity [10] put limits on the power of quantum mechanics as a resource in distributed tasks. A better understanding of the possible interconversions between nonlocal correlations might bring an information-theoretic explanation of these limitations.

Note added. After the completion of this work we were made aware of a work by Tsirelson [32] that contains some of the results presented here, in particular those of Secs. II A and II B 1.

ACKNOWLEDGMENTS

We would like to thank Wim van Dam, Andreas Winter, and Harry Buhrman for useful discussions. We acknowledge financial support from the Communauté Française de Belgique under Grant No. ARC 00/05-251, from the IUAP programme of the Belgian Government under Grant No. V-18, and from the EU through project RESQ (Grant No. IST-2001-37559).

[1] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[2] A. Aspect, Nature (London) **398**, 189 (1999).
[3] C. Shannon, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman (University of California Press, Berkeley, CA, 1961), Vol. 1, pp. 611–644.
[4] R. Cleve and H. Buhrman, Phys. Rev. A **56**, 1201 (1997).
[5] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, Cambridge, England, 1997).
[6] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[7] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[8] B. S. Cirel'son, Lett. Math. Phys. **4**, 83 (1980).
[9] W. van Dam, Ph.D. thesis, University of Oxford, Department of Physics (2000), available at http://web.mit.edu/vandam/www/publications.html
[10] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp, Lect. Notes Comput. Sci. **1509**, 61 (1999).
[11] R. F. Werner and M. M. Wolf, Quantum Inf. Comput. **1**, 1 (2001).
[12] I. Pitowsky, *Quantum Probability, Quantum Logic*, Lecture Notes in Physics Vol. 321 (Springer, Heidelberg, 1989).
[13] B. S. Cirel'son, J. Sov. Math. **36**, 557 (1987).
[14] L. J. Landau, Found. Phys. **18**, 449 (1988).
[15] Ll. Masanes, e-print quant-ph/0309137.
[16] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[17] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[18] R. Ahlswede and I. Csiszár, IEEE Trans. Inf. Theory **44**, 225 (1998).
[19] R. Ahlswede and I. Csiszár, IEEE Trans. Inf. Theory **39**, 1121 (1993).
[20] D. Collins and S. Popescu, Phys. Rev. A **65**, 032321 (2002).
[21] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[22] S. Pironio, Phys. Rev. A **68**, 062102 (2003).
[23] Thomas Christof and Andreas Löbel, URL http://www.zib.de/Optimization/Software/Porta/index.html
[24] Komei Fukuda, URL http://www.cs.mcgill.ca/~fukuda/soft/cdd_home/cdd.html
[25] W. van Dam (private communication).
[26] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).
[27] A. V. Belinski and D. N. Klyshko, Phys. Usp. **36**, 653 (1993).
[28] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[29] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht, 1989), p. 69.
[30] P. Mitchell, S. Popescu, and D. Roberts, e-print quant-ph/0202009.
[31] J. Barrett, L. Hardy, and A. Kent, e-print quant-ph/0405101.
[32] B. S. Tsirelson, Hadronic J. Suppl. **8**, 329 (1993).