

## All Reversible Dynamics in Maximally Nonlocal Theories are Trivial

David Gross,<sup>1,\*</sup> Markus Müller,<sup>2,3,†</sup> Roger Colbeck,<sup>4,5,‡</sup> and Oscar C. O. Dahlsten<sup>4,§</sup>

<sup>1</sup>*Institute for Theoretical Physics, Leibniz University Hannover, 30167 Hannover, Germany*

<sup>2</sup>*Institute of Mathematics, Technical University of Berlin, 10623 Berlin, Germany*

<sup>3</sup>*Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany*

<sup>4</sup>*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*

<sup>5</sup>*Institute of Theoretical Computer Science, ETH Zurich, 8092 Zurich, Switzerland*

(Received 6 November 2009; published 23 February 2010)

A remarkable feature of quantum theory is nonlocality (Bell inequality violations). However, quantum correlations are not maximally nonlocal, and it is natural to ask whether there are compelling reasons for rejecting theories in which stronger violations are possible. To shed light on this question, we consider post-quantum theories in which maximally nonlocal states (nonlocal boxes) occur. We show that reversible transformations in such theories are trivial: they consist solely of local operations and permutations of systems. In particular, no correlations can be created; nonlocal boxes cannot be prepared from product states and classical computers can efficiently simulate all such processes.

DOI: 10.1103/PhysRevLett.104.080402

PACS numbers: 03.65.Ta, 03.65.Ud

*Introduction.*—Quantum mechanics exhibits the remarkable feature of nonlocal correlations, as highlighted in Bell’s seminal paper [1]. Such correlations have (up to a few remaining loopholes) been extensively verified in experiments [2]. Aside from their theoretical importance, nonlocal correlations can be exploited for technological use: they are vital in entanglement-based quantum key distribution schemes [3], for example, where their presence can be used to guarantee security (see also [4] for a recent review).

While quantum mechanics violates Bell inequalities, it does not do so in the maximal possible way. There are conceivable devices, so-called nonlocal or Popescu-Rohrlich boxes, that permit even stronger correlations than quantum mechanics does, while respecting the no-signalling principle [5–7]. Such correlations are not observed in nature and the question arises as to whether other fundamental principles might be violated if they were to exist.

There has already been some progress towards answering this question. For example, the existence of nonlocal boxes would lead to some communication complexity problems becoming trivial [8,9], the possibility of oblivious transfer [10] and the lack of so-called information causality [11]. It has also been realized that in a theory in which maximally Bell violating correlations emerge, the set of possible dynamical transformations would be severely restricted compared to those allowed in quantum theory [12]. While a complete classification of the dynamics has remained elusive, it has been shown, for example, that entanglement swapping is impossible [13,14]. Furthermore, the question of the computational power of such a theory has been raised [12,14].

We work in the framework of generalized probabilistic theories [12,15–17], adopting the pragmatic operational view that the physical content of a theory is in the predicted

statistics of measurement outcomes given preparations and transformations. The framework makes minimal assumptions and allows for mathematical rigor. We consider a system composed of  $N$  subsystems. To each subsystem one of  $M \geq 1$  measurements may be applied, yielding one of  $K \geq 2$  outcomes (in the following, unless otherwise stated, we assume each subsystem has the same  $M$  and  $K$ ). The state space contains all nonsignaling correlations, corresponding to so-called generalized nonsignaling theory [12] or, more colloquially, boxworld.

Our main result (Theorem 1) is that (except in the case  $M = 1$  which corresponds to classical theory) the set of reversible transformations in boxworld is trivial: all such operations are a combination of local operations on a single system (which correspond to relabelings of measurements and their outcomes) and permutations of local systems (which correspond to relabelings of subsystems). This solves the aforementioned open problem concerning the computational power of boxworld in the case of reversible dynamics [12,14].

Another interesting consequence is that, in boxworld, measurements and dynamics are necessarily distinct physical processes, in the sense that a measurement cannot be seen as a reversible dynamics on the system comprising the state and measurement device (cf. quantum theory, where the measurement process can be seen as a unitary evolution from the point of view of an external observer).

We note that, in the case of a classical-boxworld hybrid system, Theorem 1 does not hold—we give an example of a CNOT operation on this system at the end of this Letter. However, for all types of system, including those where the number of measurements and outcomes differ between the subsystems, reversible dynamics map pure product states to pure product states; i.e., nonlocal states cannot be reversibly prepared from product states. This is our second main result (Theorem 2).

A geometric intuition behind this result is as follows. The state space of the theory is a convex polytope, and reversible transformations must map it to itself. They therefore correspond to symmetries of the polytope. The polytope is in some way *stellated*, with the vertices corresponding to maximally nonlocal states having a different character from local ones. They are, hence, not connected by symmetries of the polytope. A two-dimensional caricature is shown in Fig. 1.

*Boxworld.*—Recall that we have a system comprising  $N$  subsystems and, on each subsystem, one of  $M$  possible measurements can be applied (corresponding to different measurement devices), yielding one of  $K$  possible outcomes. The local measurements are denoted  $\{X_0, \dots, X_{M-1}\}$ . A measurement on the entire system made up of local measurements can then be described by a string  $A_1 \dots A_N$ , where  $A_i \in \{X_0, \dots, X_{M-1}\}$  specifies the measurement applied to the  $i$ th subsystem. Similarly, the corresponding outcomes are denoted  $a_1 \dots a_N$ , with  $a_i \in \{0, \dots, K-1\}$ . Measurement-outcome pairs are called effects, e.g., a measurement of  $X_1$  giving outcome 3. A state is then a function  $P: (a_1 \dots a_N | A_1 \dots A_N) \mapsto [0, 1]$ , which gives the probability of the effect that  $A_1 \dots A_N$  is measured and gives outcomes  $a_1 \dots a_N$ . More general measurements are possible: a measurement is a collection of effects for which the sum of the outcome probabilities over the collection is 1 when acting on any state. Such measurements include procedures whereby the measurement performed on a particular subsystem depends on the outcomes of previous measurements, convex combinations of such procedures and more [14]. However, the statistics of the local measurements  $A_1 \dots A_N$  are sufficient to uniquely determine the outcome probabilities of all measurements, and hence can be used to specify the state. This nontrivial assumption is known as the local observability principle [18].

Furthermore, the subsystems can be spatially separated, and hence we require that  $P$  satisfies the nonsignaling conditions, i.e., that  $\sum_{a_i=0}^{K-1} P(a_1, \dots, a_i, \dots, a_N | A_1, \dots, A_i, \dots, A_N)$  is independent of  $A_i$ . This implies that the marginal distribution on some set of subsystems is independent of the choice of measurement(s) on other subsystems.

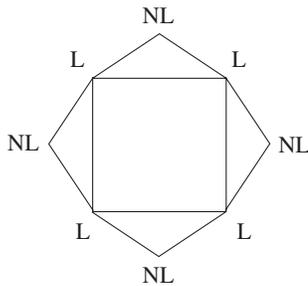


FIG. 1. Two-dimensional caricature of the (normalized) boxworld state space formed by stellating a square. Local vertices are denoted by L and nonlocal ones by NL. No symmetries of this object take L states to NL states or vice versa.

Boxworld is a physical theory whose state space consists of any  $P$  subject to: (i)  $P$  takes values in  $[0, 1]$ ; (ii)  $P$  is normalized in the obvious sense; and (iii)  $P$  satisfies the nonsignaling conditions. The constraints (i)–(iii) are such that the state space is a convex polytope.

We first deal with the special case  $M = K = 2$  (the case of so-called *gbits* [12]). The corresponding state spaces (defined below) contain interesting nonlocal states, for example, nonlocal boxes with maximal Bell violating correlations. We label the two measurements  $X_0 = X$  and  $X_1 = Z$ .

*Mathematical framework.*—We work in the generalized probabilistic framework (see, e.g., [12,15–17]). Here, states are represented as vectors embedded in a real vector space. Effects will also be represented as vectors, such that the probabilities of outcomes will be given by inner products between the relevant vectors. We begin with the case of a single system ( $N = 1$ ). We choose three linearly independent vectors  $X, Z, \underline{1} \in \mathbb{R}^3$ . The vector  $X$  is identified with  $(1|X)$ , which is the effect that the  $X$  measurement gives outcome 1. We define a vector  $\neg X := \underline{1} - X$  and associate it with  $(0|X)$ . The prefix  $\neg$  may be interpreted as a negation. Lastly, the  $\neg Z$  effect is defined analogously as  $\neg Z := \underline{1} - Z$ . Because  $X, Z, \underline{1}$  are linearly independent, for every state  $P$ , there is a unique vector  $s \in \mathbb{R}^3$  representing  $P$  in the sense that  $\langle X, s \rangle = P(1|X)$ ,  $\langle Z, s \rangle = P(1|Z)$  and  $\langle \underline{1}, s \rangle = 1$ . It follows that  $\langle \neg X, s \rangle = P(0|X)$  and likewise for  $Z$ . We will refer to the set  $\mathcal{P}^{(1)} = \{X, \neg X, Z, \neg Z\}$  as the single-site extremal effects. (Note that the quantum analogue of our effect vectors are projectors, and the inner product is analogous to the Hilbert-Schmidt scalar product, mapping states,  $\rho$ , and projectors,  $\Pi$ , to probabilities,  $\text{Tr}(\rho\Pi)$ .)

The  $N$ -subsystem extremal effects  $\mathcal{P}^{(N)}$  are defined to be the tensor products  $A_1 \otimes \dots \otimes A_N$ , where  $A_i \in \mathcal{P}^{(1)}$  (the reason for this definition is that it recovers the full set of nonsignaling distributions for the state space, as will be shown in Lemma 1). We further define the identity on  $N$  sites,  $\underline{1}^{(N)} := \underline{1} \otimes \dots \otimes \underline{1}$ . A central object is the convex cone [19]  $\mathcal{K}^{(N)}$  generated by  $\mathcal{P}^{(N)}$ . This cone is the collection of all vectors which can be written as a linear combination of elements of  $\mathcal{P}^{(N)}$  with non-negative coefficients. To any convex cone  $\mathcal{K}$ , one can associate a dual cone  $\mathcal{K}^* = \{s | \langle A, s \rangle \geq 0 \forall A \in \mathcal{K}\}$ . We will identify this with the set of unnormalized states. Our interest in cones and duality stems from the following lemma, which characterizes the state space of boxworld in terms of the cone  $\mathcal{K}^{(N)}$ .

*Lemma 1.*—Let  $\mathcal{S}^{(N)}$  be the set of vectors  $s$  in the dual cone  $(\mathcal{K}^{(N)})^*$  which satisfy  $\langle \underline{1}^{(N)}, s \rangle = 1$ . The space of (normalized) states in boxworld can be represented by  $\mathcal{S}^{(N)}$ .

*Proof.*—We use the notation  $\neg^0 A := A$  and  $\neg^1 A := \neg A$  for  $A \in \{X, Z\}$ . The vectors  $s \in \mathcal{S}^{(N)}$  will henceforth be called states; they satisfy  $\langle \underline{1}^{(N)}, s \rangle = 1$  and  $\langle B, s \rangle \geq 0$  for all  $B \in \mathcal{K}^{(N)}$ . To every state  $s$ , we associate a probability dis-

tribution  $P \equiv P(a_1, \dots, a_N | A_1, \dots, A_N)$  by  $\langle -^{a_1}A_1 \otimes \dots \otimes -^{a_N}A_N, s \rangle$  for  $A_i \in \{X, Z\}$  and  $a_i \in \{0, 1\}$ . First we show that every such  $P$  is a valid nonsignaling probability distribution. By definition,  $P$  is non-negative. To verify normalization, note that  $\mathbb{1}^{(N)} = \sum_{x \in \{0,1\}^N} -^{x_1}A_1 \otimes \dots \otimes -^{x_N}A_N$  for any choices of  $A_i \in \{X, Z\}$ , so that  $\sum_{a_1, \dots, a_N} P(a_1, \dots, a_N | A_1, \dots, A_N) = \langle \mathbb{1}^{(N)}, s \rangle = 1$ .  $P$  is also nonsignaling since  $\sum_{a_i} P(a_1, \dots, a_N | A_1, \dots, A_N)$  is the sum of  $\langle -^{a_1}A_1 \otimes \dots \otimes A_i \otimes \dots \otimes -^{a_N}A_N, s \rangle$  and  $\langle -^{a_1}A_1 \otimes \dots \otimes -A_i \otimes \dots \otimes -^{a_N}A_N, s \rangle$  which equals  $\langle -^{a_1}A_1 \otimes \dots \otimes \mathbb{1} \otimes \dots \otimes -^{a_N}A_N, s \rangle$ , and is hence independent of  $A_i$ .

To show that every nonsignaling distribution has an associated state, note that there is a unique vector  $s \in (\mathbb{R}^3)^{\otimes N}$  such that  $\langle -^{a_1}A_1 \otimes \dots \otimes -^{a_N}A_N, s \rangle = P(a_1, \dots, a_N | A_1, \dots, A_N)$  for  $-^{a_i}A_i \in \{X, Z, -X\}$  (since these effects span the space). It is then easy to see that the no-signalling property enforces consistency also in the case that  $-^{a_i}A_i = -Z$  for some  $i$ . Non-negativity and normalization follow directly from the corresponding statements for the probability distribution  $P$ .  $\square$

*Transformations.*—First note that all allowed dynamical transformations in general probabilistic theories (reversible or not) are linear—this follows from the fact that they have to respect convex combinations, which correspond to probabilistic mixtures. For a general proof of this fact see [12].

The allowed transformations,  $T$ , are defined to be linear maps with the property that for all  $s \in \mathcal{S}^{(N)}$ ,  $Ts \in \mathcal{S}^{(N)}$ . A transformation is reversible if both  $T$  and  $T^{-1}$  are allowed transformations. It follows that a reversible transformation maps the state space  $\mathcal{S}^{(N)}$  bijectively onto itself. Since  $T$  is a linear map, it is also the case that  $T$  maps extremal states to extremal states.

Note that the states  $s \in \mathcal{S}^{(N)}$  themselves do not have a physical meaning—only their scalar products with effects do, i.e.,  $\langle A, s \rangle$  (which are probabilities). Since  $\langle A, Ts \rangle = \langle T^\dagger A, s \rangle$ , the dynamics may equivalently be specified by means of the adjoint map  $T^\dagger$ . (In quantum theory, the analogue is passing from the Schrödinger to the Heisenberg picture.) Then:

*Lemma 2.*—Adjoint reversible transformations  $T^\dagger$  map the cone of effects  $\mathcal{K}^{(N)}$  bijectively onto itself. Moreover, they map the set of extremal effects,  $\mathcal{P}^{(N)}$ , onto itself.

*Proof.*—The first claim is a straightforward consequence of the fact that  $(\mathcal{K}^{(N)})^{**} = \mathcal{K}^{(N)}$  [19].

We turn to the second statement. Since it is a convex cone,  $\mathcal{K}^{(N)}$  is completely characterized by its extremal rays. By linearity,  $T^\dagger$  maps the extremal rays of  $\mathcal{K}^{(N)}$  onto themselves. From the definition of  $\mathcal{K}^{(N)}$ , we know that the cone is the convex hull of the  $4^N$  rays formed by all  $A \in \mathcal{P}^{(N)}$ . One can verify that these are indeed the extremal rays. Therefore, for every  $A \in \mathcal{P}^{(N)}$ , there exists an  $A' \in \mathcal{P}^{(N)}$  and a non-negative number  $\lambda$  such that  $T^\dagger(A) = \lambda A'$ . To see that  $\lambda$  must equal 1, observe that for every  $B \in$

$\mathcal{P}^{(N)}$ , there exist (product) states  $s_0, s_1 \in \mathcal{S}^{(N)}$  such that  $\langle B, s_0 \rangle = 0$  and  $\langle B, s_1 \rangle = 1$ . Since this holds, in particular, for both  $A$  and  $A'$ , it follows that  $\lambda = 1$ .  $\square$

*Orthogonal representation of transformations.*—There are  $4^N$  extremal effects, and thus  $4^N!$  permutations acting on  $\mathcal{P}^{(N)}$ . We go on to show that only a tiny fraction of those is actually realizable in boxworld. It will be convenient to use a specific representation of  $X, Z$  and  $\mathbb{1}$ : We set  $X = (1/2, 1/\sqrt{2}, 0)$ ,  $Z = (1/2, 0, 1/\sqrt{2})$  and  $\mathbb{1} = (1, 0, 0)$ .

*Lemma 3.*—With respect to the representation above, it holds that any reversible transformation  $T$  is orthogonal, i.e., on  $N$  subsystems,  $T^\dagger T = \mathbb{1}_{3^N}$ , where  $\mathbb{1}_d$  is the  $d$ -dimensional identity matrix.

*Proof.*—First observe that with this choice,  $\sum_{A \in \mathcal{P}^{(1)}} |A\rangle\langle A| = \mathbb{1}_3$  and hence (since  $\mathcal{P}^{(N)}$  factorizes)  $\sum_{A \in \mathcal{P}^{(N)}} |A\rangle\langle A| = \mathbb{1}_{3^N}$ . Then, since  $T^\dagger$  permutes the extremal effects,  $T^\dagger T = T^\dagger (\sum_{A \in \mathcal{P}^{(N)}} |A\rangle\langle A|) T = \sum_{A \in \mathcal{P}^{(N)}} |A\rangle\langle A| = \mathbb{1}_{3^N}$ .  $\square$

The fact that  $T$  (and thus  $T^\dagger$ ) is orthogonal, gives rise to a host of invariants. If one picks any two extremal effects  $Q, R \in \mathcal{P}^{(N)}$ , then clearly their inner product is a conserved quantity:  $\langle Q, R \rangle = \langle T^\dagger Q, T^\dagger R \rangle$ . However,  $|\langle Q, R \rangle| = 4^{-N} 3^{N-d_H(Q,R)}$ , where  $d_H(Q, R)$  is the Hamming distance between  $Q$  and  $R$ , i.e., the number of places at which  $Q$  and  $R$  differ. Thus the Hamming distance of extremal effects is a conserved quantity:  $d_H(Q, R) = d_H(T^\dagger Q, T^\dagger R)$ . It is well-known in the theory of error correction [20] that the set of maps on finite strings which preserve the Hamming distance is highly restricted: the group of those maps is generated by local transformations and permutations of sites only. Thus  $T^\dagger$  acts as such an operation on  $\mathcal{P}^{(N)}$ . Moreover, since the states in  $\mathcal{P}^{(N)}$  span the entire space, the action on this set is sufficient to completely specify  $T^\dagger$ .

Furthermore, it is straightforward to see that the set of allowed local operations comprises exchanging  $X$  and  $Z$  (relabeling measurements), exchanging  $X$  and  $-X$  (relabeling the outcome upon input  $X$ ), exchanging  $Z$  and  $-Z$  (relabeling the outcome upon input  $Z$ ) and combinations thereof.

*Main results.*—Combining all the previous results proves the following theorem in the special case of  $M = 2$  measurements with  $K = 2$  outcomes (the general proof is slightly more involved but analogous).

*Theorem 1.*—Every reversible transformation on a system comprising  $N$  subsystems in boxworld, with  $M \geq 2$  measurements at every subsystem each having  $K \geq 2$  outcomes, is a permutation of subsystems, followed by local relabelings of measurements and their outcomes.

We go on to prove:

*Theorem 2.*—In boxworld, every reversible transformation maps pure product states to pure product states. This is true even if the system is coupled to an arbitrary number of classical systems, and if the number of devices and outcomes varies from subsystem to subsystem.

We need to slightly extend the notion of outcome vectors to the general case. We denote the set of extremal effects

for the  $i$ th subsystem by  $\mathcal{P}^i = \{X_m^i(k)\}$ , where  $m$  labels the measurements (the number of different  $m$ s may depend on  $i$ ) and  $k$  the corresponding outcomes (the number of different  $k$ s may depend on  $m$  and on  $i$ ). The only relation these vectors satisfy is  $\sum_k X_m^i(k) = \underline{1}^i$ , where  $\underline{1}^i$  represents the identity.

The identity on the full system is then  $\underline{1}^{(N)} := \underline{1}^1 \otimes \dots \otimes \underline{1}^N$ , and the extremal effects are  $\mathcal{P}^{(N)} := \mathcal{P}^1 \otimes \dots \otimes \mathcal{P}^N$ . The convex cone  $\mathcal{K}^{(N)}$  and the state space  $\mathcal{S}^{(N)}$  are defined analogously to the binary case previously described. The statements and proofs of Lemmas 1 and 2 remain valid in this more general case.

*Proof.*—To complete the proof of Theorem 2, note that a state  $s \in \mathcal{S}^{(N)}$  is a pure product state (that is, of the form  $s = s_1 \otimes \dots \otimes s_N$ , where all  $s_i$  are pure) if and only if  $\langle A, s \rangle \in \{0, 1\}$  for all extremal effects  $A \in \mathcal{P}^{(N)}$ . Suppose that  $s$  is a pure product state and  $T$  a reversible transformation, then  $\langle A, Ts \rangle = \langle T^\dagger A, s \rangle \in \{0, 1\}$  for all  $A \in \mathcal{P}^{(N)}$ , which proves that  $Ts$  must also be a pure product state.  $\square$

Note that Theorem 1 does not, in general, apply to the case of site-dependent numbers of measurements. For example, suppose that we have two sites, where the first has two binary measurements,  $X$  and  $Z$ , and the second allows only a single binary measurement,  $Y$ . (In other words, a gbit is coupled to a classical bit.) It is then straightforward to construct a reversible CNOT operation, where the classical bit is the control bit. For example, there is an adjoint reversible transformation that acts as  $A \otimes Y \mapsto A \otimes Y, A \otimes \neg Y \mapsto \neg A \otimes \neg Y$  for all  $A \in \{X, Z, \neg X, \neg Z\}$ .

In the case of a system composed of several classical subsystems, Theorem 1 also does not hold—the dynamics in such a case is nontrivial. Nevertheless, Theorem 2 does apply to this case—it remains impossible to prepare entangled states from separable ones.

*Conclusions.*—We have shown that the set of reversible operations in boxworld is trivial: the only possible operations relabel subsystems, local measurements and their outcomes. In particular, there is no boxworld analogue of an entangling unitary in quantum theory, one cannot reversibly prepare nonlocal states from separable ones, nor perform useful computations reversibly.

In addition, the results have consequences for the interplay between dynamics and measurements in boxworld: suppose we have a system comprising a particle,  $A$ , and two observers,  $B$  and  $C$ , initially in an uncorrelated tripartite product state. In quantum theory, if  $B$  measures  $A$ , but  $C$  does not take part in the interaction, then  $C$  can model the corresponding dynamics by a unitary transformation on the  $AB$  system. That is,  $C$  can view the whole interaction as reversible while retaining the ability to correctly predict the outcome probabilities of any future measurements. (Theories with such a property might be called fundamentally reversible.) In boxworld, on the other hand, this is not true:  $B$ 's measurement on  $A$  would have to create correlations between  $A$  and  $B$ , but this could never be achieved by

a reversible transformation. Hence  $C$  would have to model the  $AB$  measurement using irreversible dynamics, even if  $C$  did not take part in the interaction itself.

It would be interesting to extend our result to explore which state spaces are compatible with fundamentally reversible theories in this sense, or with theories that are transitive, i.e., that every pure state can be reversibly mapped to any other. This property has been used by Hardy as an axiom for quantum theory [16]. Both conditions seem to strongly restrict the possible geometry of the state space, and an interesting open question is how non-local such theories can be.

We are grateful to H. Barnum, V. Scholz, and R. Werner for interesting discussions and to N. Harrigan for comments which improved the presentation. O. D. would like to thank M. Leifer for introducing him to generalized probabilistic theories, and MM would like to thank C. Witte for the same reason and for the quadratic cocktail model of a gbit. R. C. and O. D. acknowledge support from the Swiss National Science Foundation (Grant No. 200021-119868). DG's research is supported by the EU (CORNER).

\*david.gross@itp.uni-hannover.de

†mueller@math.tu-berlin.de

‡colbeck@phys.ethz.ch

§dahlsten@phys.ethz.ch

- [1] J. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987).
- [2] A. Aspect, *Nature (London)* **398**, 189 (1999).
- [3] A. K. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).
- [4] V. Scarani *et al.*, *Rev. Mod. Phys.* **81**, 1301 (2009).
- [5] L. A. Khalfin and B. S. Tsirelson, *Symposium on the Foundations of Modern Physics* (World Scientific, Singapore, 1985), p. 441.
- [6] B. S. Tsirelson, *Hadronic J. Suppl.* **8**, 329 (1993).
- [7] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
- [8] W. van Dam, arXiv:quant-ph/0501159.
- [9] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, *Phys. Rev. Lett.* **96**, 250401 (2006).
- [10] H. Buhrman *et al.*, *Proc. R. Soc. A* **462**, 1919 (2006).
- [11] M. Pawłowski *et al.*, *Nature (London)* **461**, 1101 (2009).
- [12] J. Barrett, *Phys. Rev. A* **75**, 032304 (2007).
- [13] A. J. Short, S. Popescu, and N. Gisin, *Phys. Rev. A* **73**, 012101 (2006).
- [14] A. J. Short and J. Barrett, arXiv:0909.2601.
- [15] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, arXiv:quant-ph/0611295.
- [16] L. Hardy, arXiv:quant-ph/0101012.
- [17] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
- [18] G. M. D'Ariano, arXiv:quant-ph/0701217.
- [19] C. D. Aliprantis and R. Tourky, *Cones and Duality* (American Mathematical Society, Providence, 2007).
- [20] P. Kaski and P. R. J. Östergård, *Classification Algorithms for Codes and Designs* (Springer, New York, 2006).