Extent of multiparticle quantum nonlocality

Nick S. Jones and Noah Linden
Department of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, United Kingdom

Serge Massar
Laboratoire d'Information Quantique and Centre for Quantum Information and Communication, C.P. 165/59, Av. F.D. Roosevelt 50, B-1050 Bruxelles, Belgium
(Received 29 August 2004; published 19 April 2005)

It is well known that entangled quantum states are nonlocal: the corrrelations between local measurements carried out on these states cannot be reproduced by local hidden variable models. Svetlichny, followed by others, showed that multipartite quantum states are more nonlocal than bipartite ones in the sense that even some nonlocal classical models with (super-luminal) communication between some of the parties cannot reproduce the quantum correlations. Here we study in detail the kinds of nonlocality present in quantum states. More precisely, we enquire what kinds of classical communication patterns cannot reproduce quantum correlations. By studying the extremal points of the space of all multiparty probability distributions, in which all parties can make one of a pair of measurements each with two possible outcomes, we find a necessary condition for classical nonlocal models to reproduce the statistics of all quantum states. This condition extends and generalizes work of Svetlichny and others in which it was showed that a particular class of classical nonlocal models, the "separable" models, cannot reproduce the statistics of all multiparticle quantum states. Our condition shows that the nonlocality present in some entangled multiparticle quantum states is much stronger than previously thought. We also study the sufficiency of our condition.

DOI: 10.1103/PhysRevA.71.042329 PACS number(s): 03.67.Mn, 03.65.Ud, 03.65.Ta, 02.10.Ox

I. INTRODUCTION

A natural way of characterizing the correlations present in entangled quantum states is to attempt to replicate their measurement statistics using classical models. Bell [1] showed that classical models respecting relativistic causality—often called local hidden variable (lhv) models—cannot always reproduce quantum statistics. If the classical model allows unlimited (superluminal) communication between the parties, then the quantum correlations can be reproduced trivially. More refined studies attempt to understand exactly what superluminal classical communication will reproduce the quantum correlations. In the bipartite case, efforts have concentrated on the number of bits of communication required to reproduce the quantum correlations [2–5]. In the three-party setting Svetlichny [6] showed that even allowing superluminal communication between arbitrary pairs of parties cannot reproduce the results of measurements performed on quantum states. Two papers independently generalized Svetlichny's result to an arbitrary number of parties [7,8].

In the present work we refine this analysis by showing that the class of classical correlations that cannot reproduce the quantum correlations is much larger than the separable class considered in [6–8]. We show that, not only is the nonlocality of multiparticle quantum correlations stronger than previously thought, but also there is a way of categorizing nonlocal correlations using graphs of communication.

Let us first recall Bell's idea. Consider m parties which each receive as input a measurement setting x_i and produce an output a_i. The probability of a certain outcome for a given set of settings is

$$P(\vec{a}|\vec{x}) = P(a_1,\ldots,a_m|x_1,\ldots,x_m). \tag{1}$$

One way to generate such correlations is for the parties to share an entangled quantum state. Depending on his input, each party then carries out a measurement on the quantum state. The result of the measurement is his output. We denote the quantum mechanical correlations obtained in this way by P_{qm}.

The most general way of generating correlations using only classical resources, without any signaling taking place between the parties, is for the parties to share a prior random variable λ (often called the hidden variable). Each party then chooses its outcome depending on its input and on λ. The set of correlations produced using such local hidden-variable models has the form

$$P_{\text{lhv}}(\vec{a}|\vec{x}) = \int d\lambda P(a_1|x_1,\lambda)P(a_2|x_2,\lambda)\cdots P(a_m|x_m,\lambda)\rho(\lambda), \tag{2}$$

where $\rho(\lambda)$ is a probability distribution over the variables λ

Note that neither the quantum nor lhv model allows signaling, since in neither of the models does any communication take place between the parties after they receive their input. Bell's central result was to show that some quantum correlations cannot be reproduced by lhv models. This is proved by introducing an inequality—called a "Bell inequality"—which must be satisfied by the lhv models but which is violated by the quantum correlations. See for instance [9,10] for all ''correlation inequalities'' in the multipartite case.

Svetlichny [6] generalized and refined Bell's result in the three-party setting by showing that some three-party quantum correlations cannot be reproduced classically, even if communication, about settings and results, is allowed between a pair of parties. The two parties in communication need not be fixed in advance, but can be chosen with probabilities p_i. The correlations considered by Svetlichny are thus of the form

$$P_{\rm Svet}(\vec{a}|\vec{x}) = \int d\lambda[p_1\rho_1(\lambda)P_1(a_1|x_1,\lambda)P_1(a_2,a_3|x_2,x_3,\lambda) + p_2\rho_2(\lambda)P_2(a_2|x_2,\lambda)P_2(a_1,a_3|x_1,x_3,\lambda) + p_3\rho_3(\lambda)P_3(a_3|x_3,\lambda)P_3(a_1,a_2|x_1,x_2,\lambda)]. \quad (3)$$

Here $P(a_2,a_3|x_2,x_3,\lambda)$ and the two terms like it can be *any* probability distribution [it need not separate into $P_1(a_2|x_2,\lambda)P_1(a_3|x_3,\lambda)$ [11]]. The main result of Svetlichny [6] is to show that there are quantum states (e.g., the Greenberger-Horne-Zeilinger state $1/\sqrt{2}(|000\rangle+|111\rangle)$; see [12] for a proof that this is the optimal state for this purpose) with $P_{qm}(\vec{a}|\vec{x})$ such that no distribution $P_{\rm Svet}(\vec{a}|\vec{x})$ can be found such that $P_{qm}=P_{\rm Svet}$ for all $\vec{x}$. This is proved by introducing an inequality — called a "Sveltichny inequality" — which must be satisfied by correlations of the form of Eq. (3) but can be violated by quantum correlations. Thus even allowing some nonlocal—i.e., superluminal—classical communication between pairs of parties, one cannot reproduce all three-party quantum correlations.

References [7,8] extended the hybrid local-nonlocal model of Svetlichny to the m-party setting. They allowed arbitrary communication within disjoint subsets of the parties, but each subset was independent of the settings and results of other subsets. In Collins *et al.* [8] these correlations were termed "separable." It was shown that the GHZ state has measurement statistics which cannot be reproduced by these models. Sets of correlations which are not separable will be called "inseparable."

In the present paper we will show that there are many inseparable correlations which cannot reproduce quantum correlations. More precisely we will define a class of "partially paired" (PP) correlations $P_{\rm PP}(\vec{a}|\vec{x})$, which include the separable correlations and some inseparable correlations, and categorize them in terms of networks of communication. Using a generalized Svetlichny inequality, taken from [8], we show that this class cannot reproduce all quantum correlations. We will also show that the complement of this class, the "totally paired" (TP) correlations, $P_{\rm TP}(\vec{a}|\vec{x})$ maximally violate these inequalities. These generalized Svetlichny inequalities therefore cannot discriminate between models in TP and quantum correlations. The class TP, unlike PP, is thus a good candidate for a classical description of all quantum correlations.

Our results are based on two advances: first we have formulated an intuitive graphical means of classifying multiparticle correlations, allowing us to define PP and TP; second, we developed a deeper understanding of the multiparticle Svetlichny inequality and the set of all similar inequalities. Both of these advances promise results beyond the scope of this paper.

In the following we first introduce some of the basic conceptual tools we shall use in our work (Secs. II and III), we then sketch our results in the four-party case (Secs. IV and V) before providing full proofs for m-parties (Secs. VI and VII).

II. GEOMETRY OF THE SPACE OF CORRELATIONS

Consider m parties, each of which receives an input x_i and produces an output a_i. The outputs can be correlated to the inputs in an arbitrary way. Hence the most general way of describing such a situation, independently of any underlying physical model, is by a set of probability distributions $P(\vec{a}|\vec{x})$. The starting point of our investigation is to describe, in detail, the geometry of the set of such probability distributions.

The set of probability distributions is characterized by the normalization conditions

$$\sum_{\vec{a}} P(\vec{a}|\vec{x}) = 1 \quad (4)$$

and the positivity conditions

$$P(\vec{a}|\vec{x}) \geqslant 0. \quad (5)$$

Therefore the set of possible probability distributions is a convex polytope. This polytope belongs to the subspace defined by Eq. (4). Its facets are given by the equality in (5).

It is useful to find the extreme points of this polytope. These are the probability distributions which saturate a maximum of the positivity conditions, Eq. (5), while satisfying the normalization conditions, Eq. (4). It is easy to see that the extreme points are the probability distributions such that, for each $\vec{x}$, there is a unique $\vec{a}=\vec{a}(\vec{x})$ with $P(\vec{a}|\vec{x})=1$ [and therefore if $\vec{a}\neq\vec{a}(\vec{x})$, then $P(\vec{a}|\vec{x})=0$]. Thus there is a one-to-one correspondence between the set of extreme points and the functions $\vec{a}(\vec{x})$ from the inputs to the outputs. Any particular $\vec{a}(\vec{x})$ defines an extreme point. We call the extreme points *deterministic models* since there is no randomness in this case: the output is completely fixed by the input. We will soon associate a graph with families of extreme points.

Subspaces of the space of all distributions satisfying normalization and positivity can be defined by taking all convex combinations of a subset of the extreme points. This is a natural construction because, if a physical model can produce certain extreme points, then it can produce any convex combination of these extreme points simply by randomly choosing which extreme point to realize.

A first interesting example is the subspace defined by lhv models. The corresponding extreme points are of the form $a_i(\vec{x})=a_i(x_i)$: the measurement results of party i depend only on the settings of that party.

A second example subspace is provided by the separable correlations considered by Svetlichny and in [7,8]. The corresponding extreme points can be characterized as follows. For each extreme point there is a partition of the set of all parties into two subsets—say, $\{1,\ldots,k\}$ and $\{k+1,\ldots,m\}$—

such that $a_i(\vec{x})=a_i(x_1,\ldots,x_k)$ if $i\in\{1,\ldots,k\}$ and $a_i(\vec{x})=a_i(x_{k+1},\ldots,x_m)$ if $i\in\{k+1,\ldots,m\}$. In the former situation, party i has results a_i independent of the settings of the set $\{k+1,\ldots,m\}$ and dependent on the settings $(x_1,\ldots,x_k)$.

Note that the formulation given by Svetlichny [see Eq. (3)] and by [7,8] may seem more general than this since in their formulation the outputs of any party in set $\{1,\ldots,k\}$ can depend on the inputs *and* outputs of all parties $1,\ldots,k$ whereas above we have only allowed the outputs of any party in set $\{1,\ldots,k\}$ to depend on the inputs of all parties $1,\ldots,k$.

Let us now show that this is not the case and that the two formulations are equivalent. For definiteness we will focus on the three-party case, Eq. (3), but the argument immediately extends to an arbitrary number of parties. Consider the set of probabilities $P_1(a_1|x_1\lambda), P_1(a_2,a_3|x_2,x_3,\lambda)$, etc., appearing in Eq. (3). The key point to note is that we can take these probabilities to be deterministic strategies. Indeed we have just argued that any probability can be written as a convex combination of extremal probabilities: $P_1(a_2,a_3|x_2,x_3,\lambda)=\Sigma_\mu p_\lambda(\mu)P_\mu^{\text{ext}}(a_2,a_3|x_2,x_3)$, where $p_\lambda(\mu)$ is a probability distribution over μ and where $P_\mu^{\text{ext}}(a_2,a_3|x_2,x_3)$ equals zero except if $a_2=a_2(x_2,x_3)$ and $a_3=a_3(x_2,x_3)$. One can now suppose that the variables μ, which specify the weighting of each extremal probability, are included in the lhv variable λ; i.e., the lhv tells the parties in each set what deterministic strategy they must use. We have now proved our claim since, in the case of extremal correlations, each output depends only on the inputs of the parties, not on their outputs. Thus in the case of separable correlations, letting the outputs of the parties in each set depend only on the inputs of the parties in their set is completely general: there is no need to also let them depend on the outputs of the parties in the set.

We now introduce new subspaces of the probability distributions which form the basis for the present analysis.

III. CLASSIFYING PROBABILITY DISTRIBUTIONS USING COMMUNICATION PATTERNS

We will classify the extremal points by the settings that each variable a_i depends on. This dependence can be represented using directed graphs. An arrow from party i to j means that a_j depends on the setting x_i. If there is no arrow from i to j, a_j is independent of x_i (changing the value of x_i only, leaves a_j unaltered). The graph is directed as one might have $a_i(x_j)$ but a_j independent of x_i (an arrow $j\rightarrow i$ but no arrow $j\leftarrow i$). We call such a graph a *communication pattern*.

Any such communication pattern can be associated naturally with a model in which (i) each party receives its input, (ii) if there is an arrow $i\rightarrow j$, i sends its input to j, and (iii) the parties which receive arrows produce their measurement results conditional on the list of inputs sent to them and their own input.

Steps (ii) and (iii) should be thought of as a single-shot mail strategy: all of the settings are posted at the same time, along the appropriate arrow, and they are received at the same time. On receipt, the parties immediately generate their results: there is no communication about the results obtained.

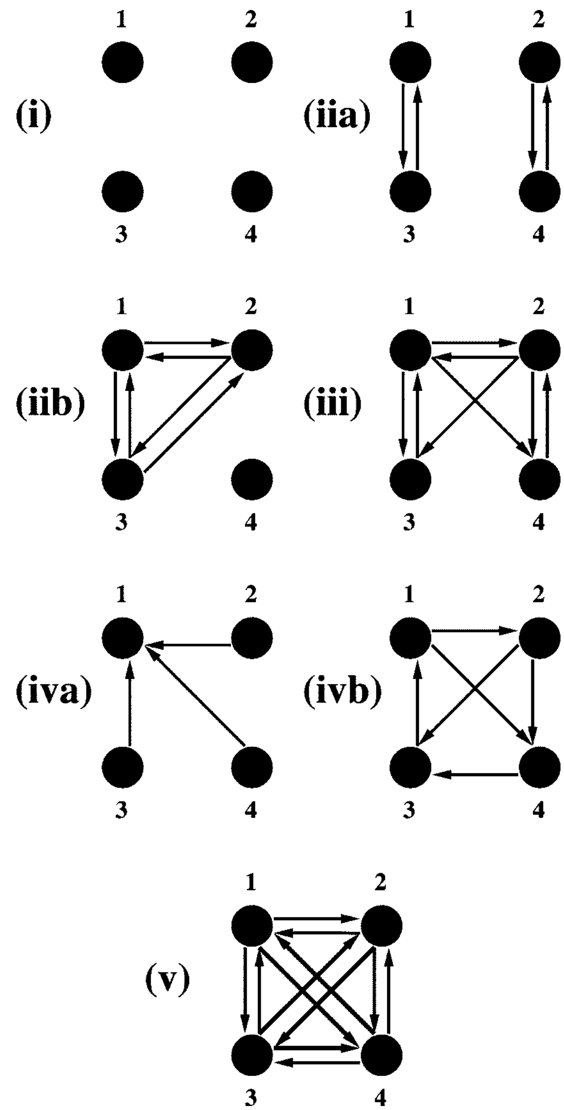

FIG. 1. (i) This graph represents the set of extreme points identifiable with lhv models—i.e., $a_i(x_i)$. Bell considered convex combinations of these points. (ii) (a) and (b) These graphs represent functions $\vec{a}(\vec{x})$ with the structure $a_1(x_1,\wedge,x_3,\wedge)$, $a_3(x_1,\wedge,x_3,\wedge)$, $a_2(\wedge,x_2,\wedge,x_4)$, $a_4(\wedge,x_2,\wedge,x_4)$ and $a_1(x_1,x_2,x_3,\wedge)$, $a_3(x_1,x_2,x_3,\wedge)$, $a_2(x_1,x_2,x_3,\wedge)$, $a_4(\wedge,\wedge,\wedge,x_4)$, respectively. Reference [8] considers correlations which are convex combinations of extremal points represented by separable graphs of this form and show that such models have $\langle S_4\rangle\leqslant 2$ (see Sec. IV). (iii) This graph represents functions $\vec{a}(\vec{x})$ which have $a_1(x_1,x_2,x_3,\wedge)$, $a_3(x_1,x_2,x_3,\wedge)$, $a_2(x_1,x_2,\wedge,x_4)$, $a_4(x_1,x_2,\wedge,x_4)$. Parties 3 and 4 are separated in the sense that no party receives arrows from both. We will show that convex combinations of these extreme points yield a value of $\langle S_4\rangle\leqslant 2$. We call graphs of type (i)–(iii) "partially paired." (iv) Examples of graphs in which there is always a party which knows the settings of any pair — we call these "totally paired." Certain convex combinations of the extremal points represented by these graphs can achieve the algebraic maximum $\langle S_4\rangle=4$, as described in the text. (a) $a_1(x_1,x_2,x_3,x_4)$, $a_3(\wedge,\wedge,x_3,\wedge)$, $a_2(\wedge,x_2,\wedge,\wedge)$, $a_4(\wedge,\wedge,\wedge,x_4)$ and (b) $a_1(x_1,\wedge,x_3,\wedge)$, $a_3(\wedge,x_2,x_3,x_4)$, $a_2(x_1,x_2,\wedge,\wedge)$, $a_4(x_1,x_2,\wedge,x_4)$. (v) This graph represents the set of all extreme points and all possible four-party graphs are its subgraphs.

It is now clear why the arrows are directed: i sending a letter to j does not imply that j mails i. Note that this single-shot mailing has to be superluminal if the measurements take place simultaneously and at spatially separated locations.

Note that any particular graph represents many extreme points. Indeed each extreme point corresponds to a unique function $\vec{a}=\vec{a}(\vec{x})$ whereas each graph only determines the variables the functions $\vec{a}$ depend on.

A formalization of the above will prove useful. A given graph G represents a set E_G of extreme points. Each point is identifiable with a different function $\vec{a}(\vec{x})$. These extreme points can be combined in convex combinations to make different distributions $P(\vec{a}|\vec{x})$ for each $\vec{x}$. The set of all such correlations produced by convex combinations of the extremal points E_G will be called C_G: the set of correlations of type G. It is important to note that we define the set E_G as including the extremal points represented by all subgraphs of G. For example, the graph F, in which all points send arrows to all others [Fig. 1(e)], represents the set of all extreme points. Hence C_F is the space of all correlations.

Different models, such as lhv models [1] or separable models [6–8], can be associated with different classes of graphs. This is illustrated in the case of four parties in Fig. 1. The notation $a_1(x_1,\wedge,\wedge,x_4)$ is to be read as "party 1's outcome is independent of the settings of parties 2,3 but dependent on the settings of 1,4"; i.e., a_1 is unaltered by changes in x_2,x_3.

IV. FOUR-PARTY CASE: THE SVETLICHNY INEQUALITY

Our study of multiparticle nonlocality is based on combining the classification of correlations in terms of communication patterns with the Svetlichny inequality [6] and its generalization described in [7,8]. We now illustrate this connection in the case of four parties.

From now on we restrict ourselves to the case where each party's input is a single bit $x_i\in\{0,1\}$ and each party's output is a single bit $a_i\in\{0,1\}$. In order to make contact with earlier work we introduce an alternative notation.

First of all, it is useful to suppose that the outputs have values $+1,-1$ instead of 0,+1. In this case we denote the output of party i by A_i with the correspondence $A_i=(-1)^{a_i}$.

Second, we denote by a superscript x_i on A_i (i.e., $A_i^{x_i}$) the value of the input of party i. This traditional notation is good in the case of lhv models, since each output A_i is uniquely determined by its input. But it is somewhat unnatural for other models where A_i can depend on the inputs of all parties and not only on x_i. However, the product of four outputs such as $A_1^1A_2^1A_3^0A_4^0$ is well defined in this notation since all inputs are specified. The expression $A_1^1A_2^1A_3^0A_4^0$ denotes the product of the outputs of parties 1, 2, 3, and 4 given that the inputs are $(x_1,x_2,x_3,x_4)=(1,1,0,0)$.

We will be interested in the Svetlichny polynomials which are specific combinations of all products $A_1^{x_1}A_2^{x_2}A_3^{x_3}A_4^{x_4}$. In the case of four parties the Svetlichny polynomial is

$$S_4=\frac{1}{4}\Big[A_1^0A_2^0(-A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0+A_3^1A_4^1)+A_1^1A_2^0(A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0-A_3^1A_4^1)+A_1^0A_2^1(A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0-A_3^1A_4^1)-A_1^1A_2^1(-A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0+A_3^1A_4^1)\Big]=\sum_{\vec{x}}F_4(\vec{x})A_1^{x_1}A_2^{x_2}A_3^{x_3}A_4^{x_4}. \quad (6)$$

The expectation value of the Svetlichny polynomial is obtained by taking the average over all possible inputs and outputs weighted by the corresponding probabilities. Explicitly this can be written as

$$\langle S_4\rangle=\sum_{\vec{x}}F_4(\vec{x})(-1)^{a_1}(-1)^{a_2}(-1)^{a_3}(-1)^{a_4}P(\vec{a}|\vec{x}). \quad (7)$$

The basis of our analysis will be to compare the maximum value of the Svetlichny polynomial attained by different models. Note that since $\langle S_4\rangle$ is a linear function of the probabilities $P(\vec{a}|\vec{x})$, its maximum value when the $P(\vec{a}|\vec{x})$ belong to a convex space will be attained on the extreme points of this space. This is the main justification for classifying correlations according to their extreme points: Bell-type expressions obtain their maximum value on the extreme points.

It is easy to show that in the case of lhv models $\langle S_4\rangle\le 2$. Collins *et al.* [8] show that for separable models, $\langle S_4\rangle\le 2$, and for certain quantum states the value of $\langle S_4\rangle=2^{3/2}$ [12]. We will show that much more general nonlocal models than the separable models considered in [7,8] also have $\langle S_4\rangle\le 2$. Thus quantum mechanical states can exhibit even more general types of nonlocality than previously anticipated.

V. FOUR-PARTY CASE: COMMUNICATION PATTERNS AND THE SVETLICHNY INEQUALITY

Collins *et al.* [8] showed that the set of correlations of the type in Figs. 1(iia) and (iib) describe statistics of "separable" physical models that cannot simulate all quantum states. We will now show that, despite being much more correlated than (ii), no extreme point represented by graph (iii) has a larger maximum value for S_4. Below is a sketch of a proof, details being saved for the m-party setting.

Consider a deterministic setting $\vec{a}=\vec{a}(\vec{x})$, characterized by the graph (iii) where $a_1=a_1(x_1,x_2,x_3,\wedge)$, $a_3=a_3(x_1,x_2,x_3,\wedge)$, $a_2=a_2(x_1,x_2,\wedge,x_4)$, $a_4=a_4(x_1,x_2,\wedge,x_4)$. Noting the form of the a_i, the following term from (6),

$$\frac{1}{4}A_1^0A_2^0(-A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0+A_3^1A_4^1), \quad (8)$$

can be rewritten as

$$\frac{1}{4}\Big[-(-1)^{a_1(0,0,0,\wedge)+a_2(0,0,\wedge,0)+a_3(0,0,0,\wedge)+a_4(0,0,\wedge,0)}+(-1)^{a_1(0,0,0,\wedge)+a_2(0,0,\wedge,1)+a_3(0,0,0,\wedge)+a_4(0,0,\wedge,1)}+(-1)^{a_1(0,0,1,\wedge)+a_2(0,0,\wedge,0)+a_3(0,0,1,\wedge)+a_4(0,0,\wedge,0)}+(-1)^{a_1(0,0,1,\wedge)+a_2(0,0,\wedge,1)+a_3(0,0,1,\wedge)+a_4(0,0,\wedge,1)}\Big]. \quad (9)$$

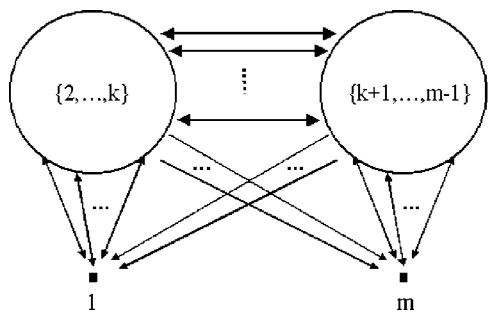

FIG. 2. Graphical representation of the most general Partially Paired (PP) m-party graph. One circle contains parties $\{2,\ldots,k\}$ and the other circle contains parties $\{k+1,\ldots,m-1\}$. All parties within a circle can send and receive arrows from all other parties within the same circle. Each party in a circle can send and receive arrows to all parties in the other circle. Party 1 can send and receive arrows from parties $\{2,\ldots,k\}$. Party 1 can receive arrows from (but does not send arrows to) parties $\{k+1,\ldots,m-1\}$ Party m can receive and send arrows to parties $\{k+1,\ldots,m-1\}$. Party m can receive arrows from (but does not send arrows to) parties $\{2,\ldots,k\}$. There are no arrows between parties 1 and m. With these restrictions there is no party that receives arrows both from party 1 and party m. The subgraph composed only of the arrows emanating from parties 1 and m is therefore separable. The graph thus belongs to the PP class.

One can now find the maximum value of this expression. Defining $a_1(0,0,x_3,\wedge)+a_3(0,0,x_3,\wedge)\bmod 2=\beta_1(x_3)$ and $a_2(0,0,\wedge,x_4)+a_4(0,0,\wedge,x_4)\bmod 2=\beta_2(x_4)$ (this approach will be reused later) this expression simplifies to

$$\frac{1}{4}\left[-(-1)^{\beta_1(0)+\beta_2(0)}+(-1)^{\beta_1(0)+\beta_2(1)}+(-1)^{\beta_1(1)+\beta_2(0)}+(-1)^{\beta_1(1)+\beta_2(1)}\right]. \tag{10}$$

One now notes that, whatever the value of functions β_i, this expression is $\leqslant\frac{1}{2}$. Similarly each of the three other terms in S_4,

$$\frac{1}{4}A_1^1A_2^0(A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0-A_3^1A_4^1),$$

$$\frac{1}{4}A_1^0A_2^1(A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0-A_3^1A_4^1),$$

$$-\frac{1}{4}A_1^1A_2^1(-A_3^0A_4^0+A_3^0A_4^1+A_3^1A_4^0+A_3^1A_4^1), \tag{11}$$

can have a maximum value of $\frac{1}{2}$ and can do so simultaneously: the maximum value of S_4 is thus 2. By convexity, even a probabilistic mix of all strategies of the form (i), (iia), (iib), and (iii) still cannot exceed $S_4=2$.

Thus very strongly correlated graphs, representing inseparable probability distributions, can still fail to exceed $S_4=2$. If two parties in a graph are separated from each other or "unpaired," such that no party knows the settings of both, then no matter how correlated the remainder of the state, S_4 cannot exceed 2.

This observation motivates the definition of "partially paired" graphs given in Fig. 2 and in definition 1 below. Indeed it will be shown that, despite being highly connected, convex combinations of extremal points defined by "partially paired" graphs cannot replicate all quantum statistics.

By contrast, graphs in which, for any given pair of settings, there is always a party with results dependent on that pair [such as Figs. 1(iva) and 1(ivb)] can represent extreme points which reach the maximum value of S_4—namely, 4. It is straightforward to show that the graph of Fig. 1(iva) with $a_1(x_1,x_2,x_3,x_4)$ can represent a function $\vec{a}(\vec{x})$ which will obtain the algebraic maximum. The graph in Fig. 1(ivb), despite having apparently less communication than (iii), also has extremal points which reach this maximum. Recalling, for graph (ivb), that $a_1(x_1,\wedge,x_3,\wedge)$, $a_3(\wedge,x_2,x_3,x_4)$, $a_2(x_1,x_2,\wedge,\wedge)$, $a_4(x_1,x_2,\wedge,x_4)$, Eq. (8) becomes

$$\frac{1}{4}\Big[-(-1)^{a_1(0,\wedge,0,\wedge)+a_2(0,0,\wedge\wedge)+a_3(\wedge,0,0,0)+a_4(0,0,\wedge,0)}+(-1)^{a_1(0,\wedge,0,\wedge)+a_2(0,0,\wedge,\wedge)+a_3(\wedge,0,0,1)+a_4(0,0,\wedge,1)}+(-1)^{a_1(0,\wedge,1,\wedge)+a_2(0,0,\wedge,\wedge)+a_3(\wedge,0,1,0)+a_4(0,0,\wedge,0)}+(-1)^{a_1(0,\wedge,1,\wedge)+a_2(0,0,\wedge,\wedge)+a_3(\wedge,0,1,1)+a_4(0,0,\wedge,1)}\Big]. \tag{12}$$

If we set $a_3(\wedge,0,1,0)=a_4(0,0,\wedge,0)=1$ and all of the other terms above equal to zero, the expression takes the value 1. It is relatively easy to find by inspection a set of a_i such that S_4 reaches its algebraic maximum of 4. Indeed defining $a_1(x_1,x_3)=x_1x_3+x_1$, $a_2(x_1,x_2)=x_1x_2+x_2$, $a_3(x_2,x_3,x_4)=x_2x_3+x_3x_4+x_3$, and $a_4(x_1,x_2,x_4)=x_1x_4+x_2x_4+x_4+1$, obtains the maximum value for S_4. In this case it may be seen that

$$\sum_{i=1}^{4}a_i(\vec{x})=\sum_{i<j}^{4}x_ix_j+\sum_{i=1}^{4}x_i+1. \tag{13}$$

This form will prove significant later.

A deterministic strategy with functional form represented by Fig. 1(iii) is nonlocal: it can be pictured as requiring faster than light correlations. Nonetheless, it cannot always replicate quantum statistics as we have described above. Conversely, as we described, extremal points represented by graphs like (iva) and (ivb) can reach the algebraic maximum of S_4—namely, 4.

Note that individual extremal points represented by graphs like (iva) and (ivb) are signaling: by looking at the outputs of one party one can learn about the inputs of other parties. We will show below, however, that there exist convex combinations of extremal points represented by graphs like (iva) and (ivb) which are no-signaling and which reach the algebraic maximum of S_4. This is an important remark since special relativity only permits no-signaling correlations.

Before moving to the m-party generalization, we review: we have identified two kinds of four-party correlations. The first are correlations represented by graphs (i)—(iii) and the second represented *only* by graphs like (iv). The former reach a value of $\langle S_4\rangle\leqslant 2$ the latter can reach $S_4=4$. For all quantum states $\langle S_4\rangle\leqslant 2\sqrt{2}$.

VI. *m*-PARTY SVETLICHNY POLYNOMIAL

Having sketched proofs for the four-party Svetlichny polynomial S_4, we can provide general proofs for the m-party setting with polynomials S_m. We first recall the definition of S_m, as formulated in [8].

We again consider the situation where each party's input $x_i \in \{0,1\}$ is a single bit and each party's output $a_i \in \{0,1\}$ is a single bit. We will use the notation introduced at the beginning of Sec. IV. As in [8] we construct S_m from the Mermin polynomials. (The Mermin inequality is just one of the correlation inequalities described in [9,10] where all correlation inequalities for n parties, each having two inputs and two outputs, were characterized). Let the two-party Mermin polynomial be

$$M_2 = \frac{1}{2}(A_1^0A_2^0 + A_1^1A_2^0 + A_1^0A_2^1 - A_1^1A_2^1) \tag{14}$$

and define the new notation

$$M_m \equiv \sum_{\vec{x}} F_m(\vec{x})A_1^{x_1}A_2^{x_2}\cdots A_m^{x_m}, \tag{15}$$

where $\vec{x}=(x_1,\ldots,x_m)$ and also

$$M_m' \equiv \sum_{\vec{x}} F_m(\vec{x})A_1^{\overline{x_1}}A_2^{\overline{x_2}}\cdots A_m^{\overline{x_m}} = \sum_{\vec{x}} F_m(\overline{\vec{x}})A_1^{x_1}A_2^{x_2}\cdots A_m^{x_m}, \tag{16}$$

where $\overline{x_i}=x_i+1 \bmod 2$. M_m is generated from M_{m-1} by the recursion relation

$$M_m = \frac{1}{2}\left[M_{m-1}(A_m^0 + A_m^1) + M_{m-1}'(A_m^0 - A_m^1)\right]. \tag{17}$$

Using this twice yields

$$M_{m+2} = \frac{1}{2}\left[M_m(M_2 + M_2') + M_m'(M_2 - M_2')\right]. \tag{18}$$

The recursion relation (18) can be written in terms of F_m as

$$F_{m+2}(\vec{x}) = \frac{1}{2}\{F_m(\vec{x})[F_2(x_{m+1},x_{m+2}) + F_2(\overline{x_{m+1}},\overline{x_{m+2}})] + F_m(\overline{\vec{x}}) \times[F_2(x_{m+1},x_{m+2}) - F_2(\overline{x_{m+1}},\overline{x_{m+2}})]\}. \tag{19}$$

Following [8] we define the Svetlichny polynomials as

$$S_m = \begin{cases} M_m, & \text{if } m \text{ is even,} \\ \frac{1}{2}(M_m + M_m'), & \text{if } m \text{ is odd.} \end{cases} \tag{20}$$

We also define

$$S_m = \sum_{\vec{x}} \mu_m(\vec{x})A_1^{x_1}\cdots A_m^{x_m}, \tag{21}$$

which implies

$$\mu_m(\vec{x}) = \begin{cases} F_m(\vec{x}), & \text{if } m \text{ is even,} \\ \frac{1}{2}[F_m(\vec{x}) + F_m(\overline{\vec{x}})], & \text{if } m \text{ is odd.} \end{cases} \tag{22}$$

In order to have a useful characterization of S_m we will obtain an explicit expression for μ_m and F_m.

Lemma 1. Let $q=\lceil m/2\rceil$ (where $\lceil\ \rceil$ indicates rounding up to the next nearest integer). Then,

$$\mu_m(\vec{x}) = \frac{1}{2^q}(-1)^{[\sum_{i<j}^m x_ix_j + (q+1)\sum_{i=1}^m x_i + (q^2-q)/2]} \tag{23}$$

and

$$F_{2k+1}(\vec{x}) = \frac{1}{2}F_{2k}(x_1,\ldots,x_{2k})(1 + (-1)^{\sum_{i=1}^{2k}x_i + x_{2k+1} + k}). \tag{24}$$

As an example, note that in the four-party case one finds

$$\mu_4(\vec{x}) = \frac{1}{4}(-1)^{[\sum_{i<j}^4 x_ix_j + \sum_{i=1}^4 x_i + 1]}, \tag{25}$$

which [combined with Eq. (21)] reproduces Eq. (6). Note that the exponent in Eq. (25) is identical to Eq. (13).

We now turn to the proof of lemma 1.

Proof of Lemma 1.

Proof of Eq. (23) *for* $m=2k$, k *integer*. In this case $\mu_{2k} = F_{2k}$. One easily checks that lemma 1 is true for $k=1$ when

$$F_2(x_1,x_2) = \frac{1}{2}(-1)^{x_1x_2},$$

$$F_2(\overline{x_1},\overline{x_2}) = \frac{1}{2}(-1)^{x_1x_2+x_1+x_2+1}.$$

We now proceed by induction. We suppose lemma 1 is true for k and will show it is true for $k+1$. From Eq. (23)

$$F_{2k}(\overline{\vec{x}}) = \frac{1}{2^k}(-1)^{[\sum_{i<j}^{2k}(x_i+1)(x_j+1) + (k+1)\sum_{i=1}^{2k}(x_i+1) + (k^2-k)/2]}, \tag{26}$$

where $F_{2k}(\vec{x})=F_{2k}(x_1,\ldots,x_{2k})$ and $q=k$. Equation (26) can be written as

$$F_{2k}(\overline{\vec{x}}) = \frac{1}{2^k}(-1)^{[\sum_{i<j}^{2k}x_ix_j + k\sum_{i=1}^{2k}x_i + (k^2+k)/2]} \tag{27}$$

using the identities

$$\sum_{i<j}^{2k}\overline{x_ix_j} = \sum_{i<j}^{2k}x_ix_j + \sum_{i}^{2k}x_i + k \bmod 2 \tag{28}$$

and $(-1)^{2\alpha}=1$ and $(-1)^{-\alpha}=(-1)^{\alpha}$ for α integer. Inserting $F_2(x_{2k+1},x_{2k+2})$, $F_2(\overline{x_{2k+1}},\overline{x_{2k+2}})$, $F_{2k}(\overline{\vec{x}})$, and $F_{2k}(\vec{x})$ into Eq. (19) and noting that

$$F_{2k}(\vec{x}) = (-1)^{\sum_{i=1}^{2k}x_i + k}F_{2k}(\overline{\vec{x}}), \tag{29}$$

the right-hand side of Eq. (19) becomes

$$F_{2k+2}(\vec{x}) = \frac{1}{4}F_{2k}(\vec{\bar{x}})(-1)^{x_{2k+1}x_{2k+2}}[1+(-1)^{x_{2k+1}+x_{2k+2}} + (-1)^{\Sigma_{i=1}^{2k}x_i+k} + (-1)^{x_{2k+1}+x_{2k+2}+\Sigma_{i=1}^{2k}x_i+k+1}] = \frac{1}{2}F_{2k}(\vec{\bar{x}})(-1)^{(x_{2k+1}+x_{2k+2})(\Sigma_{i=1}^{2k}x_i+k)+x_{2k+1}x_{2k+2}}, \quad (30)$$

where we have used the identity

$$(-1)^{(ab)} = \frac{1}{2}[1+(-1)^b+(-1)^a+(-1)^{a+b+1}], \quad (31)$$

for a,b integer. Equation (30) can be rewritten as

$$F_{2k+2}(\vec{x}) = \frac{1}{2^{k+1}}(-1)^{\Sigma_{i<j}^{2k}x_ix_j+(x_{2k+1}+x_{2k+2})(\Sigma_{i=1}^{2k}x_i)} \times(-1)^{x_{2k+1}x_{2k+2}+(k+2)\Sigma_{i=1}^{2k}x_i+(k+2)(x_{2k+1}+x_{2k+2})+(k^2+k)/2}. \quad (32)$$

One can readily check that this coincides with Eq. (23) for $m=2k+2$ and $q=k+1$. Equation (23) thus satisfies the recursion relation (19) for $2k$ even. □

Proof of Eq. (24). Equation (17) can be rewritten as

$$M_{2k+1} = \frac{1}{2}\left[M_{2k}\sum_{x_{2k+1}=0}^{1}A_{2k+1}^{x_{2k+1}} + M'_{2k}\sum_{x_{2k+1}=0}^{1}(-1)^{x_{2k+1}}A_{2k+1}^{x_{2k+1}}\right]. \quad (33)$$

Using Eq. (15), this is equivalent to the relation

$$F_{2k+1}(\vec{x}) = \frac{1}{2}[F_{2k}(\vec{x}) + F_{2k}(\vec{\bar{x}})(-1)^{x_{2k+1}}]. \quad (34)$$

Substituting Eq. (29) into Eq. (34) one recovers Eq. (24). □

Proof of Eq. (23) *for* $m=2k+1,k$ *integer*. Inserting Eqs. (29) and (34) into Eq. (22), for m odd, yields

$$\mu_{2k+1}(\vec{x}) = \frac{1}{4}F_{2k}(\vec{\bar{x}})\{1+(-1)^{\Sigma_{i=1}^{2k}(x_i+1)+x_{2k}+k+1} + (-1)^{\Sigma_{i=1}^{2k}x_i+k} \times[1+(-1)^{\Sigma_{i=1}^{2k}x_i+k+x_{2k+1}}]\}. \quad (35)$$

This becomes, by identity (31),

$$\mu_{2k+1}(\vec{x}) = \frac{1}{2}F_{2k}(\vec{\bar{x}})(-1)^{x_{2k+1}(\Sigma_{i=1}^{2k}x_i+k)}. \quad (36)$$

This can be rewritten as

$$\mu_{2k+1}(\vec{x}) = \frac{1}{2^{k+1}}(-1)^{\Sigma_{i<j}^{2k}x_ix_j+x_{2k+1}\Sigma_{i=1}^{2k}x_i} \times(-1)^{(k+2)(\Sigma_{i=1}^{2k}x_i+x_{2k+1})+(k^2+k)/2}, \quad (37)$$

which coincides with Eq. (23). □

VII. CLASSES PP AND TP

The central result of this article is to obtain bounds on the maximum value of S_m attainable in different models. Let us recall what is already known in this respect:

Theorem 1 (Seevinck and Svetlichny [7], Collins et al.[8], Mitchell et al[12]):

(i) Local hidden variable models and separable models satisfy the Svetlichny inequality

$$\langle S_m\rangle_{\text{lhv, separable models}} \leq 2^{m-\lceil m/2\rceil-1}. \quad (38)$$

(ii) The maximum value of S_m attainable by quantum mechanics (reached by carrying out measurements on the GHZ state) is

$$\langle S_m\rangle_{\text{quantum mechanics}} \leq 2^{m-\lceil m/2\rceil-1/2}. \quad (39)$$

(iii) The algebraic maximum of S_m [obtained by taking $A_1^{x_1}\cdots A_m^{x_m}=\mu(\vec{x}_m)/|\mu(\vec{x}_m)|$ and using Eq. (23)] is

$$S_m^{\text{alg}} = 2^{m-\lceil m/2\rceil}. \quad (40)$$

We now go back to the classification of extreme points in terms of communication patterns introduced in Sec. III. We define two classes of graph and formulate two theorems about them.

Definition 1: partially paired (PP) graphs (see Fig. 2). A communication pattern represented by a directed graph in which there exist two (or more) parties i,j such that there is no party with results dependent on x_i,x_j. Graphically, this definition can be rephrased as: Take the subgraph composed only of the vertices receiving arrows originating from i and j and the arrows themselves. This graph is separable: it can be split into two disconnected graphs one including vertex i and the other j.

Examples are Figs. 1(i)-1(iii) since, in these graphs, there is always a pair of vertices i,j in the graph that neither send arrows to each other, nor both send to the same party.

Definition 2: partially paired correlations. These are the correlations that can be written as convex combinations of the extremal correlations whose associated graph is partially paired. These correlations form the set C_{PP}:

$$C_{\text{PP}} = \cup_{G\in\text{PP}}C_G$$

(where PP is the set of all PP graphs). Note that C_{PP} is the space of correlations associated with all possible PP graphs, not just a single PP graph G_{PP}.

One of our main results is the following.

Theorem 2. All partially paired correlations (i.e., all correlations in the set C_{PP}) satisfy the multi party Svetlichny inequality $S_m \leq 2^{m-q-1}$.

Note that the Svetlichny inequality is violated by some quantum states; see theorem 1. Thus PP correlations cannot reproduce all quantum correlations.

The complementary class to PP graphs are the totally paired graphs.

Definition 3: totally paired (TP) graphs. Any graph which is not PP. Graphically this can be rephrased as the following. For any two parties i,j there always exists a party k such that k receives arrows originating from i and j (k could coincide with i or j).

Examples are Figs. 1(iv) and 1(v) (see [13] for a graph-theoretic analysis of TP graphs).

The definition of TP graphs will allow us to prove the complement of theorem 2. Namely, we will show that for any TP graph G_{TP}, there exist correlations whose associated graph is G_{TP} and which maximally violate the Svetlichny inequality.

But it should be noted that—except in the case of lhv models—an extreme point is a deterministic signaling strategy: one party's results provides information about the settings of other parties. Thus one could argue that such extreme points are unphysical and cannot reproduce the predictions of causal theories, such as quantum mechanics, which do not allow signaling. But we will show that different strategies, with the same associated graph G_{TP}, can be combined to produce no-signaling correlations while continuing to maximally violate the Svetlichny inequality. Thus maximal violation of the Svetlichny inequality by correlations in $C_{G_{\mathrm{TP}}}$, where G_{TP} is an arbitrary TP graph, is compatible with causality. These results are summarized as follows.

Theorem 3. For any totally paired communication graph G_{TP}, there exist correlations in the set $C_{G_{\mathrm{TP}}}$ (the set of correlations obtained by convex combinations of the extremal points $E_{G_{\mathrm{TP}}}$ whose associated graph is G_{TP}) which both attain the algebraic maximum of the Svetlichny polynomial and are no-signaling.

Note that $C_{G_{\mathrm{TP}}}$ is the set of correlations associated with a single TP graph G_{TP}.

Proof of Theorem 2. The Svetlichny inequalities can be written in the form

$$\langle S_m\rangle = \sum_{\vec{x},\vec{a}} \mu_m(\vec{x})(-1)^{\Sigma_i a_i} P(\vec{a}|\vec{x}) \leq 2^{m-q-1}. \quad (41)$$

In the class PP there always exists at least one pair of settings—say, x_1, x_m—such that no party's outcome is dependent on both. As in Fig. 2, the m parties can be divided into the set $\{1,\ldots,k\}$ dependent on x_1 (and not x_m) and $\{k+1,\ldots,m\}$ dependent on x_m (and not x_1). Defining $\Sigma_{i=1}^{k} a_i \bmod 2 \equiv \beta_1, \Sigma_{i=k+1}^{m} a_i \bmod 2 \equiv \beta_2$ and rewriting the left-hand side of Eq. (41) using Eq. (23):

$$\langle S_m\rangle_{\mathrm{PP}} = \frac{1}{2^q} \sum_{x_2,\ldots,x_{m-1}} (-1)^{\Sigma_{i<j,i\neq 1}^{m-1} x_i x_j + (q+1)\Sigma_{i=2}^{m-1} x_i + (q^2-q)/2} \times \left(\sum_{x_1,x_m,\vec{a}} (-1)^{x_1 x_m + (q+1+\Sigma_{i=2}^{m-1} x_i)(x_1+x_m)} \times (-1)^{\beta_1+\beta_2} P(\vec{a}|\vec{x}) \right). \quad (42)$$

By the definition of β_1, β_2,

$$\sum_{x_1,x_m,\vec{a}} (-1)^{x_1 x_m + (q+1+\Sigma_{i=2}^{m-1} x_i)(x_1+x_m)} (-1)^{\beta_1+\beta_2} P(\vec{a}|\vec{x}) \quad (43)$$

has the same form as a Clauser-Horne-Shimony-Holt expression [14] and thus has a modulus ≤ 2. Substituting this into Eq. (42) yields $\langle S_m\rangle_{\mathrm{PP}} \leq 2^{m-q-1}$. □

Thus, at least insofar as the Svetlichny inequality is concerned, the set C_{PP} is only as strong as its subset, the separable correlations considered in [7,8].

Proof of Theorem 3. Any extreme point in the set of correlations—i.e. any deterministic scenario—has $\vec{a}=\vec{a}(\vec{x})$ and so Eq. (41) becomes

$$\langle S_m\rangle = \sum_{\vec{x}} \mu_m(\vec{x})(-1)^{\Sigma_i a_i(\vec{x})}. \quad (44)$$

In order to reach the algebraic maximum of S_m we require $(-1)^{\Sigma_i a_i(\vec{x})} = \mu(\vec{x})/|\mu(\vec{x})|$. From lemma 1, this means that the algebraic maximum is attained if

$$\sum_{i=1}^{m} a_i = \sum_{i<j}^{m} x_i x_j + (q+1)\sum_{i=1}^{m} x_i + (q^2-q)/2. \quad (45)$$

Let us show that for any totally paired graph G_{TP}, there is an extreme point whose associated graph is G_{TP} and which satisfies Eq. (45). To see this note the following.

(i) $a_i(\vec{x})$ is a function of x_j only when there is an arrow originating at vertex j and ending at vertex i (in addition a_i can always depend on x_i).

(ii) For any pair of vertices i,j in a TP graph there is always a vertex k which either receives edges from both members of the pair or is itself a member of the pair (i.e., k can coincide with either i or j). The output of this vertex $a_k(\vec{x})$ can thus be equal to any function of x_i and x_j (plus possibly other functions depending on the other arrows leading into vertex k). By an appropriate choice of these functions we can reproduce the term $\Sigma_{i<j}^{m} x_i x_j$ on the right-hand side of Eq. (45).

(iii) The output of any vertex i can depend on the input to vertex i. Hence the output of vetex i can contain a term equal to $a_i(x_i) = c x_i + d$. By combining these terms we can reproduce the term $(q+1)\Sigma_{i=1}^{m} x_i + (q^2-q)/2$.

By combining points (ii) and (iii) above, we can satisfy Eq. (45).

Thus certain extreme points in $E_{G_{\mathrm{TP}}}$ (where G_{TP} is any TP graph) define functions $\vec{a}(\vec{x})$ such that Eq. (45) holds. These reach the algebraic maximum of the multiparty Svetlichny polynomials. Convex combinations of these will also obtain the maximum. This is why Eq. (13) has the same form as the exponent of Eq. (25).

We now turn to the second part of theorem 3. We will show that different strategies, with the same associated graph, G_{TP}, can be combined to produce no-signaling correlations while continuing to maximally violate the Svetlichny inequality. Let us first recall that by no-signaling correlations we mean correlations $P(\vec{a}|\vec{x})$ such that one subset of parties—say, parties $1,\ldots,k$—cannot communicate to the other parties $k+1,\ldots,m$ by changing the settings of their measurement device. Mathematically this is expressed by the condition that, for all $a_{k+1},\ldots,a_m$,

$$\sum_{a_1,\ldots,a_k} P(\vec{a}|\vec{x}) = P(a_{k+1},\ldots,a_m|\vec{x}), \quad (46)$$

where the right-hand side is independent of $x_1,\ldots,x_k$.

Let us now consider a particular graph G_{TP} in the set TP. To this graph we can associate at least one deterministic strategy $\vec{a}^0(\vec{x})$ such that Eq. (45) holds. This deterministic strategy is necessarily signaling; i.e., Eq. (46) is not independent of $x_1,\ldots,x_k$. The first step in proving the second part of theorem 3 is to note that $\vec{a}^0(\vec{x})$ is not the only deterministic strategy which has G_{TP} as its associated graph and which obeys Eq. (45). In fact from $\vec{a}^0(\vec{x})$ we can easily construct a set of 2^{m-1} deterministic strategies that all have G_{TP} as their associated graph and obey Eq. (45). To this end define the m component vectors, $\vec{b}^\mu \in \{0,1\}^m$ with the property $\Sigma_i b_i^\mu \bmod 2 = 0$. There are 2^{m-1} such vectors, $\vec{b}^1, \vec{b}^2, \ldots, \vec{b}^{2^{m-1}}$. Now we define $\vec{a}^\mu(\vec{x}) = \vec{a}^0(\vec{x}) + \vec{b}^\mu$. Note that $\Sigma_i a_i^\mu \bmod 2 = \Sigma_i a_i^0 \bmod 2$ hence, Eq. (45) holds for all deterministic strategies $\vec{a}^\mu(\vec{x})$.

The second step in proving the second part of theorem 3 is to note that, while staying constrained by the graph G_{TP}, the parties need not use a deterministic strategy. Instead, before the protocol starts, they can choose one value of μ at random, according to some probability distribution $p(\mu) \geqslant 0, \Sigma_\mu p(\mu) = 1$. They then carry out the deterministic strategy $\vec{a}^\mu(\vec{x})$. Since μ is chosen at random the resulting correlations have the form

$$P(\vec{a}|\vec{x}) = \sum_\mu p(\mu) P^\mu(\vec{a}|\vec{x}), \tag{47}$$

where $P^\mu(\vec{a}|\vec{x})$ are the correlations obtained by using the deterministic strategy $\vec{a}^\mu$. Thus the parties have formed a convex combination of deterministic strategies, all with the same associated graph G_{TP}.

Let us now show that if $p(\mu) = 2^{-(m-1)}$ is the uniform distribution, then the correlations defined by Eq. (47) are no-signaling. This follows from the fact that the correlations associated with $\vec{a}^\mu$ have the form $P^\mu(\vec{a}|\vec{x}) = \delta(a_1 - (a_1^0(\vec{x}) + b_1^\mu))\delta(a_2 - (a_2^0(\vec{x}) + b_2^\mu)) \cdots \delta(a_m - (a_m^0(\vec{x}) + b_m^\mu))$. We can now show that $P(a_{k+1}, \ldots, a_m|\vec{x})$ is independent of $x_1, \ldots, x_k$:

$$\begin{aligned} P(a_{k+1},\ldots,a_m|\vec{x}) &= \frac{1}{2^{m-1}} \sum_{a_1,\ldots,a_k} \sum_\mu P^\mu(\vec{a}|\vec{x}) \\ &= \frac{1}{2^{m-1}} \sum_{a_1,\ldots,a_k} \sum_\mu \delta(a_1 - [a_1^0(\vec{x}) + b_1^\mu]) \cdots \delta(a_m - [a_m^0(\vec{x}) + b_m^\mu]) \\ &= \frac{1}{2^{m-1}} \sum_\mu \Bigg(\delta(a_{k+1} - [a_{k+1}^0(\vec{x}) + b_{k+1}^\mu]) \cdots \delta(a_m - [a_m^0(\vec{x}) + b_m^\mu]) \\ &\quad \times \sum_{a_1,\ldots,a_k} \delta(a_1 - [a_1^0(\vec{x}) + b_1^\mu]) \cdots \delta(a_k - [a_k^0(\vec{x}) + b_k^\mu]) \Bigg) \\ &= \frac{1}{2^{m-1}} \sum_\mu \delta(a_{k+1} - [a_{k+1}^0(\vec{x}) + b_{k+1}^\mu]) \cdots \delta(a_m - [a_m^0(\vec{x}) + b_m^\mu]), \end{aligned} \tag{48}$$

where we use the fact that

$$\sum_{a_1,\ldots,a_k} \delta(a_1 - [a_1^0(\vec{x}) + b_1^\mu]) \cdots \delta(a_k - [a_k^0(\vec{x}) + b_k^\mu]) = 1, \tag{49}$$

whatever the value of $\vec{a}^0(\vec{x}) + \vec{b}^\mu$ and for all μ and $\vec{x}$. Now note that for any given $m-k$ element bit string $(a_{k+1}, \ldots, a_m)$ and $(m-k)$-element bit string $(a_{k+1}^0(\vec{x}), \ldots, a_m^0(\vec{x}))$ (for given $\vec{x}$), there are 2^{k-1} vectors $\vec{b}^\mu$ such that $(a_{k+1}, \ldots, a_m) = (a_{k+1}^0(\vec{x}), \ldots, a_m^0(\vec{x})) + (b_{k+1}^\mu, \ldots, b_m^\mu)$. Thus, upon summing over μ, one finds, from Eq. (48), that $P(a_{k+1}, \ldots, a_m|\vec{x}) = 2^{k-m}$ for all $a_{k+1}, \ldots, a_m$ and for all $\vec{x}$. The result is true for any $1 \leqslant k \leqslant m$. Thus no nontrivial subgroup of the parties can signal to any other. The correlations in Eq. (47) are therefore non-signaling. □

VIII. CONCLUSION

Svetlichny [6] and then others [7,8] demonstrated that classical models which allow superluminal communication within subsets of parties cannot reproduce all multipartite quantum correlations. We have extended this approach and showed that much more general classical communication patterns than those considered in [7,8] cannot reproduce all multipartite quantum correlations. We have shown how to describe such communication patterns in terms of directed graphs. (For instance, the correlations considered in [6–8] are described by separable graphs.) Our main result is to prove that the correlations described by partially paired graphs (see definition 1) cannot reproduce all quantum correlations. PP graphs are much more general than separable graphs, and therefore our result shows that, in the multipartite setting, quantum correlations are much more nonlocal than previously thought. To obtain this result we carried out a detailed analysis of the properties of the multiparty generalization of the Svetlichny inequality for which the bounds attained by lhv models, quantum mechanics, and completely nonlocal models were previously known. We showed that the correlations associated with PP graphs attain the same bound as the lhv models and therefore cannot reproduce all quantum correlations. While for purposes of exposition we described the four-party case in Sec. V, it should be noted that there are three-party nonseparable PP correlations—for example, $a_1(x_1,x_2,\wedge),a_2(x_1,x_2,\wedge),a_3(\wedge,x_2,x_3)$. In other words, the phenomenon we have described in this paper—namely, that quantum mechanics is stronger than some inseparable correlations—appears even for three-party states. We have also found that another class of correlations which are convex combinations of some of the extreme points associated with totally paired graphs (see definition 3) can both maximally violate the Svetlichny inequality and be nosignaling. However, this does not necessarily mean that any TP graph has associated extreme points which can reproduce all multipartite quantum correlations. Indeed the above results give an essentially complete characterization of how much different classical communication patterns violate the Svetlichny inequality. But there are many other Bell inequalities which can be used to probe the nonlocality of quantum correlations. It may be that—using another Bell inequality as test—one can show that some TP graphs represent correlations that cannot reproduce all quantum correlations. On the other hand, it may be that the extreme points associated with any TP graph can reproduce all quantum correlations. We leave this as an open question for future research. Indeed the present work shows that the nonlocality present in multipartite quantum correlations is stronger, and structurally richer, than previously thought.

ACKNOWLEDGMENTS

We are grateful for support from the EPSRC, from Project RESQ IST-2001-37559 of the IST-FET program of the EC, from the Communauté Française de Belgique under Grant No. ARC00/05-251 and from the IUAP program of the Belgian government under Grant No. V-18.

[1] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[2] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).
[3] M. Steiner, Phys. Lett. A **270**, 239 (2000).
[4] S. Massar, D. Bacon, N. J. Cerf, and R. Cleve, Phys. Rev. A **63**, 052305 (2001).
[5] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).
[6] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[7] M. Seevinck, and G. SvetlichnyPhys. Rev. Lett. **89**, 060401 (2002).
[8] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002).
[9] R. F. Werner and M. M. Wolf, Phys. Rev. A **64**, 032112 (2001).
[10] M. Zukowski and C. Brukner, Phys. Rev. Lett. **88**, 210401 (2002).
[11] Effectively there are two hidden variables. In addition to λ, the variable $i\in\{1,2,3\}$ defines which pair in the triple can communicate.
[12] P. Mitchell, S. Popescu, and D. Roberts, Phys. Rev. A **70**, 060101 (2004).
[13] G. Gutin, N. S. Jones, A. Rafiey, S. Severini, and A. Yeo, e-print math.CO/0411653.
[14] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).