Violations of Bell inequalities as lower bounds on the communication cost of nonlocal correlations

Stefano Pironio*
Service de Physique Théorique, CP 225, Université Libre de Bruxelles, 1050 Brussels, Belgium
(Received 5 May 2003; published 4 December 2003)

To reproduce in a local hidden variables theory correlations that violate Bell inequalities, communication must occur between the parties. We show that the amount of violation of a Bell inequality imposes a lower bound on the average communication needed to produce these correlations. Moreover, for every probability distribution there exists an optimal inequality for which the degree of violation gives the minimal average communication. As an example, to produce using classical resources the correlations that maximally violate the Clauser-Horne-Shimony-Holt inequality, $\sqrt{2}-1\simeq 0.4142$ bits of communication are necessary and sufficient. For Bell tests performed on two entangled states of dimension $d\geqslant 3$ where each party has the choice between two measurements, our results suggest that more communication is needed to simulate outcomes obtained from certain nonmaximally entangled states than maximally entangled ones.

DOI: 10.1103/PhysRevA.68.062102 PACS number(s): 03.65.Ud, 03.65.Ta, 03.67.Hk

I. INTRODUCTION

Characterizing the features of quantum mechanics that differentiate it from classical theories is an important issue for quantum information theory, as well as from a fundamental perspective. One such peculiarity is the nonlocal character of quantum mechanics, i.e., the fact that quantum correlations are incompatible with local realistic theories. Apart from being one of the most intriguing aspects of nature, nonlocality is deeply related to several quantum information processing tasks [1,2], and is at the core of quantum communication complexity [3,4].

It was Bell [5] who first showed that correlations obtained by measuring two separated subsystems cannot be explained by a classical realistic theory if no communication between the subsystems is allowed. The question which then follows is how much communication is required to reproduce these correlations? This is a natural way to quantify the nonlocal character of quantum correlations in terms of classical resources. We will show that the inequalities introduced by Bell 40 years ago not only tell us that some communication is necessary to produce the correlations but also how much.

The situation we consider is the one encountered in bipartite Bell scenarios. Two spatially separated parties, Alice and Bob, receive local inputs x and y and subsequently produce outputs a and b. We denote by M_A the number of possible inputs on Alice's side and by M_B the number of inputs on Bob's side and restrict ourselves to the case where a finite number of distinct outcomes is associated to each input. The scenario is completely characterized by the probabilities $p_{ab|xy}$ that Alice outputs a when given x and Bob outputs b when given y. We therefore associate to each Bell scenario a correlation vector $\mathbf{p}$ with entries $p_{ab|xy}$. Note that these entries satisfy the normalization constraints

$$\sum_{a,b} p_{ab|xy}=1 \quad \text{for } x=0,...,M_A-1 \text{ and } y=0,...,M_B-1. \tag{1}$$

In the quantum version of the Bell scenario, Alice and Bob share an entangled quantum state on which they perform local measurements. The inputs x and y then correspond to the possible settings of their measuring apparatus and the outcomes a and b correspond to the results of these measurements.

In the classical version of the Bell scenario, Alice and Bob may use only classical resources, i.e., shared randomness (local hidden variables) and classical communication, to determine their outcomes a and b. If the two parties have unrestricted access to shared randomness, the classical cost of producing the correlations $\mathbf{p}$ is the minimum amount of communication they must exchange in a classical protocol to achieve this goal. Different measures of this amount of communication are possible.

(i) $C_w(\mathbf{p})$. *Worst case communication*: the maximal amount of communication exchanged between Alice and Bob in any particular execution of the protocol. See [6–9].

(ii) $\overline{C}(\mathbf{p})$. *Average communication*: the average communication exchanged between Alice and Bob, where the average is taken over the inputs and the shared randomness. See [10–12].

(iii) $C_\infty(\mathbf{p})$. *Asymptotic communication*: the limit $\lim_{n\to\infty}\overline{C}(\mathbf{p}^n)/n$, where $\mathbf{p}^n$ is the probability distribution obtained when n runs of the Bell scenario are carried out in parallel, that is when the parties receive n inputs and produce n outputs in one go. See [13].

In each of these definitions the costs are defined with respect to the optimal protocol that gives the lowest value for each quantity.

The asymptotic measure C_∞ may be the most appropriate when one is concerned with practical applications that make use of the correlations but is less preoccupied with whether the measurements are performed individually or collectively. On the other hand, the first two measures of communication relate to protocols where the outcomes are determined after each single pair of inputs is chosen. This is in particular the situation encountered in Bell tests. These two measures thus more properly count the communication necessary to simulate classically nonlocality and it could be expected that they

*Electronic address: spironio@ulb.ac.be

are closely connected to Bell inequalities. Relations between the worst case situation and Bell inequalities were examined in [8] where the authors introduced new Bell inequalities that are satisfied by correlations that necessitate at most 1 bit of communication to be simulated.

In the present paper we concentrate on the average communication $\bar{C}$. We first point out that the amount by which the probabilities $\mathbf{p}$ violate a Bell inequality imposes a lower bound on $\bar{C}(\mathbf{p})$. This bound is simply a bound on the amount of communication needed to simulate classically a violation of the inequality. It is *a priori* unclear that one particular manifestation of the nonlocal content of correlations, the violation of a specific Bell inequality, suffices to characterize exactly the communication $\bar{C}(\mathbf{p})$ necessary to reproduce the entire set of correlations (all the less since in general correlations violate more than one inequality). Yet, to each probability distribution $\mathbf{p}$ is associated an optimal inequality such that the bound the violation imposes on $\bar{C}(\mathbf{p})$ is saturated, i.e., it gives the minimal average communication needed to reproduce these correlations. We then investigate in detail the case of the Clauser-Horne-Shimony-Holt (CHSH) inequality [14]. We show that for two-settings and two-outcomes Bell scenarios, the CHSH inequality is optimal for all quantum correlations. This implies in particular that $\sqrt{2}-1 \simeq 0.4142$ bits are necessary and sufficient on average to reproduce classically the correlations that lead to the maximal violation of the inequality. We then apply our approach to the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality [15]. We find that for two-measurements scenarios more communication is needed to reproduce the effect of measuring certain nonmaximally entangled states of two qutrits than is necessary for maximally entangled ones. Our results, combined with those of [16], suggest that this is also the case for qudits with $d \geqslant 3$. Finally we ask whether for quantum correlations the optimal inequalities from the communication point of view are always facet inequalities. We give an example where this is not the case.

This paper is organized as follows. We first describe in Sec. II how the average communication $\bar{C}$ relates to the degree of violation of Bell inequalities. We then apply these ideas to the CHSH inequality in Sec. III and to the CGLMP inequality in Sec. IV. In Sec. V we discuss the relations between optimal and facet inequalities.

II. GENERAL FORMALISM

A. Deterministic protocols

To state our results it is necessary to consider particular classical protocols, the deterministic ones which do not use any kind of randomness. These protocols therefore always produce the same pair of outcomes for given inputs x and y. The entries of the associated correlation vectors $\mathbf{d}$ are thus of the form $d_{ab|xy} = \delta^a_{\alpha(x,y)} \delta^b_{\beta(x,y)}$ where $\alpha(x,y)$ and $\beta(x,y)$ specify Alice's and Bob's outcomes for measurements x and y. Since there are a finite number of functions $\alpha(x,y)$ and $\beta(x,y)$, there are a finite number of different deterministic strategies $\mathbf{d}^\lambda$ which we index by λ. Their interest is that any classical protocol can be viewed as a probability distribution $\{q_\lambda\}$ of deterministic protocols $\mathbf{d}^\lambda$. That is any correlation vector $\mathbf{p}$ can be written as $\mathbf{p} = \Sigma_\lambda q_\lambda \mathbf{d}^\lambda$ where $q_\lambda \geqslant 0$ and $\Sigma_\lambda q_\lambda = 1$.

Deterministic protocols for which α and β depend only on the measurements performed locally by each party, i.e., $\alpha = \alpha(x)$, $\beta = \beta(y)$, are local protocols. No communication at all is required to implement them. On the other hand, if $\alpha(x,y)$ or $\beta(x,y)$ depends on the input of the other party, some (deterministic) communication $c(x,y)$ between the parties is necessary to carry out the protocol.

It will be convenient to group in subsets $\mathcal{D}_i$ deterministic strategies that need the same comunication c_i to be implemented. Since in the present paper we are interested in the average communication $\bar{C}$, we will group the deterministic strategies with respect to the minimal average communication needed to implement them, expressed in bits. Indexing strategies in $\mathcal{D}_i$ by λ_i, we thus have $\bar{C}(\mathbf{d}^{\lambda_i}) = c_i \ \forall \lambda_i$. We also arrange the subsets $\mathcal{D}_i$ $(i=0,...,N)$ in increasing order with respect to their communication cost: $c_i < c_{i+1}$. Local deterministic strategies thus belong to $\mathcal{D}_0$ for which $c_0 = 0$, while the maximum communication cost c_N is associated with strategies in $\mathcal{D}_N$. This occurs when both parties need to send the value of their input to the other, so $c_N = \log_2 M_A + \log_2 M_B$ [17]. We will further illustrate this grouping of deterministic strategies in Sec. IV.

With the above notation, a decomposition of $\mathbf{p}$ in term of deterministic strategies can be written

$$\mathbf{p} = \sum_i \sum_{\lambda_i} q_{\lambda_i} \mathbf{d}^{\lambda_i}. \tag{2}$$

It then directly follows that the average communication $\bar{C}(\mathbf{p},\{q_\lambda\})$ associated with the protocol (2) is given by

$$\bar{C}(\mathbf{p},\{q_\lambda\}) = \sum_i \sum_{\lambda_i} q_{\lambda_i} \bar{C}(\mathbf{d}^{\lambda_i}) = \sum_i \sum_{\lambda_i} q_{\lambda_i} c_i = \sum_i q_i c_i, \tag{3}$$

where $q_i = \Sigma_{\lambda_i} q_{\lambda_i}$ is the probability to use a strategy from $\mathcal{D}_i$. The minimum amount of communication $\bar{C}(\mathbf{p})$ necessary to reproduce the correlations $\mathbf{p}$ is the minimum of $\bar{C}(\mathbf{p},\{q_\lambda\})$ over all possible decompositions of the form (2). If there exists a decomposition such that $q_0 = 1$, i.e., if the correlations can be written as a convex combination of local deterministic strategies, then $\bar{C}(\mathbf{p}) = 0$ and the correlations are local. If for every decomposition $q_0 < 1$, the correlations are nonlocal and they violate a Bell inequality.

B. Bell inequalities

A Bell inequality can be viewed as a linear form, represented by a vector $\mathbf{b}$, which associates with each set of probabilities $\mathbf{p}$ the number $B(\mathbf{p}) = \mathbf{b} \cdot \mathbf{p}$. One particular number is the local bound $B_0 = \max_{\lambda_0} \{\mathbf{b} \cdot \mathbf{d}^{\lambda_0}\}$. By convexity, every local probability distribution $\boldsymbol{l} = \Sigma_{\lambda_0} q_{\lambda_0} \mathbf{d}^{\lambda_0}$ satisfies the inequality $B(\boldsymbol{l}) \leqslant B_0$. Correlations $\mathbf{p}$ that violate it, $B(\mathbf{p}) > B_0$, are therefore nonlocal. To extract more information from $B(\mathbf{p})$ than a simple detection of nonlocality it is necessary to consider not only the upper bound B_0 the inequality takes on the local subset $\mathcal{D}_0$, but also on all the other subsets $\mathcal{D}_i$,

$$B_i = \max_{\lambda_i} \{\mathbf{b}\cdot\mathbf{d}^{\lambda_i}\}. \tag{4}$$

Given this extra knowledge, a constraint on the decomposition (2) can be deduced from the amount by which $\mathbf{p}$ violates the Bell inequality. This turns into a bound on $\overline{C}(\mathbf{p})$ which is the basis of the present paper.

C. Main results

Proposition 1. For every inequality $\mathbf{b}$ and probability distribution $\mathbf{p}$, the following bound holds:

$$\overline{C}(\mathbf{p}) \geqslant \frac{B(\mathbf{p})-B_0}{B_{j*}-B_0} c_{j*}, \tag{5}$$

where $j*$ is the index such that $(B_{j*}-B_0)/c_{j*} = \max_{j\neq 0}\{(B_j - B_0)/c_j\}$.

Proof. From Eqs. (2) and (4) we deduce $B(\mathbf{p}) = \mathbf{b}\cdot\mathbf{p} = \Sigma_i \Sigma_{\lambda_i} q_{\lambda_i} \mathbf{b}\cdot\mathbf{d}^{\lambda_i} \leqslant \Sigma_i q_i B_i$. Since $\Sigma_i q_i = 1$ we find

$$B(\mathbf{p}) - B_0 \leqslant \sum_{i\neq 0} q_i (B_i - B_0) \tag{6}$$

or

$$q_{j*} \geqslant \frac{B(\mathbf{p})-B_0}{B_{j*}-B_0} - \sum_{i\neq 0, j*} q_i \frac{B_i - B_0}{B_{j*}-B_0}. \tag{7}$$

We thus obtain

$$\begin{aligned} \overline{C}(\mathbf{p}) &= \sum_i q_i c_i \\ &\geqslant \frac{B(\mathbf{p})-B_0}{B_{j*}-B_0} c_{j*} + \sum_{i\neq 0,j*} q_i \left(c_i - \frac{B_i - B_0}{B_{j*}-B_0} c_{j*} \right) \\ &\geqslant \frac{B(\mathbf{p})-B_0}{B_{j*}-B_0} c_{j*}, \end{aligned} \tag{8}$$

where in the last line we used $(B_{j*}-B_0)/c_{j*} \geqslant (B_i - B_0)/c_i$ which follows from the definition of $j*$. ■

The bound (5) the inequality $\mathbf{b}$ imposes on the average communication $\overline{C}(\mathbf{p})$ is proportional to the degree of violation $B(\mathbf{p})$ times a normalization factor $c_{j*}/(B_{j*}-B_0)$ expressed in units of "communication per amount of violation." This naturally suggests to rewrite Bell inequalities in natural units where $c_{j*}/(B_{j*}-B_0) = 1$ so that Eq. (5) takes a simpler form.

Proposition 2. Every Bell inequality $\mathbf{b}$ can be rewritten in a normalized form $\mathbf{b}'$ such that $B_i' \leqslant c_i$ $\forall i$. For the normalized inequality the bound (5) becomes

$$\overline{C}(\mathbf{p}) \geqslant B'(\mathbf{p}). \tag{9}$$

Proof. Define the normalized version of the inequality $\mathbf{b}$ as

$$\mathbf{b}' = \frac{c_{j*}}{B_{j*}-B_0} \left(\mathbf{b} - \frac{B_0}{M_A M_B} \mathbf{1} \right), \tag{10}$$

where $j*$ is taken as in Proposition 1. Note that $\mathbf{1}\cdot\mathbf{p} = M_A M_B$ since the entries of all correlations vectors $\mathbf{p}$ satisfy the normalization constraints (1). The effect of the term $-(B_0/M_A M_B)\mathbf{1}$ in Eq. (10) is thus to shift the value the inequality takes on an arbitrary vector $\mathbf{p}$ from $B(\mathbf{p})$ to $B(\mathbf{p}) - B_0$. We therefore get $B_i' = \max_{\lambda_i}\{\mathbf{b}'\cdot\mathbf{d}^{\lambda_i}\} = [c_{j*}/(B_{j*}-B_0)](B_i - B_0) \leqslant c_i$, where the last inequality holds by definition of $j*$.

We then immediately deduce (9) since $B'(\mathbf{p}) = \mathbf{b}'\cdot\mathbf{p} = \Sigma_i \Sigma_{\lambda_i} q_{\lambda_i} \mathbf{b}'\cdot\mathbf{d}^{\lambda_i} \leqslant \Sigma_i q_i B_i' \leqslant \Sigma_i q_i c_i = \overline{C}(\mathbf{p})$. ■

Assuming Bell inequalities are written in this standard way where $B_i \leqslant c_i$, it follows from Eq. (9) that for a given set of probabilities $\mathbf{p}$, the inequality that leads to the strongest bound on $\overline{C}(\mathbf{p})$ is the one for which $B(\mathbf{p})$ takes the greatest value. In fact we have the following.

Proposition 3. Let $\mathbf{b}_*$ be the normalized inequality that gives the maximum violation $B_*(\mathbf{p}) = \max_{\mathbf{b}}\{B(\mathbf{p})\}$ for the correlations $\mathbf{p}$, then

$$\overline{C}(\mathbf{p}) = B_*(\mathbf{p}). \tag{11}$$

Proof. This follows from the duality theorem of linear programming [18]. Indeed $B_*(\mathbf{p})$ is the solution to the following linear programming problem:

$$\begin{aligned} &\max \quad \mathbf{b}\cdot\mathbf{p} \\ &\text{subject to} \quad \mathbf{b}\cdot\mathbf{d}^{\lambda_i} \leqslant c_i \quad \forall \lambda_0,\ldots,\lambda_i,\ldots,\lambda_N \end{aligned} \tag{12}$$

for the variable $\mathbf{b}$. The dual of that problem is

$$\begin{aligned} &\min \quad \sum_i \sum_{\lambda_i} c_i q_{\lambda_i} = \sum_i c_i q_i \\ &\text{subject to} \quad \sum_i \sum_{\lambda_i} q_{\lambda_i} \mathbf{d}^{\lambda_i} = \mathbf{p}, \\ &\qquad q_{\lambda_i} \geqslant 0 \quad \forall \lambda_0,\ldots,\lambda_i,\ldots,\lambda_N \end{aligned} \tag{13}$$

for the variables q_{λ_i}. The solution to the dual problem is $\overline{C}(\mathbf{p})$ since it just amounts to search for the optimal decomposition $\{q_{\lambda_i}\}$ of $\mathbf{p}$ which leads to the lowest average communication (note that the condition $\Sigma_i \Sigma_{\lambda_i} q_{\lambda_i} = 1$ is in fact already implied by the normalization conditions that $\mathbf{d}^{\lambda_i}$ and $\mathbf{p}$ satisfy). Now, the duality theorem of linear programming states that if the primal (dual) has an optimal solution, then the dual (primal) problem also has an optimal solution and, moreover, the two solutions coincide, i.e., $B_*(\mathbf{p}) = \overline{C}(\mathbf{p})$. ■

This last result introduces the concept of an optimal inequality $\mathbf{b}_*$ from the communication point of view for the correlations $\mathbf{p}$. Indeed the bounds (5) and (9) can be interpreted as bounds on the communication necessary to simulate classically a violation of the inequality $\mathbf{b}$ by the amount

$B(\mathbf{p})$. Of course this is also a bound on the average communication $\overline{C}(\mathbf{p})$ necessary to reproduce the entire set of correlations $\mathbf{p}$. In general, however, more communication may be necessary to carry out the latter task than the former. For the optimal inequality $\mathbf{b}_*$, though, the communication is identical in the two cases. If we quantify nonlocality by the amount of communication needed to simulate it classically, a violation of the inequality $\mathbf{b}_*$ by the amount $B_*(\mathbf{p})$ therefore exhibit the complete nonlocality contained in the correlations $\mathbf{p}$.

D. Comparing Bell inequalities

The bound (5) simply expresses that the most efficient strategy to simulate a violation of a Bell inequality uses local deterministic protocols (which do not necessitate any communication) and deterministic protocols from $\mathcal{D}_{j*}$ for which the ratio of violation per communication $(B_{j*}-B_0)/c_{j*}$ is maximal. Indeed, for that strategy a violation by the amount $B(\mathbf{p})=(1-q_{j*})B_0+q_{j*}B_{j*}$ implies

$$q_{j*}=\frac{B(\mathbf{p})-B_0}{B_{j*}-B_0} \qquad (14)$$

and thus a communication

$$\overline{C}=q_{j*}c_{j*}=\frac{B(\mathbf{p})-B_0}{B_{j*}-B_0}c_{j*}$$

which is nothing more than the right-hand side of Eq. (5).

The bound (5) can thus be viewed as the minimal communication needed to produce a given violation of the inequality $\mathbf{b}$. This allows us to compare the amount of violation of different Bell inequalities, possibly corresponding to different Bell scenarios. If the inequalities are normalized so that $B_i\leqslant c_i$, the bound takes the form (9) and the comparison is even more direct: the greater the violation, the greater the nonlocality exhibited by the inequality.

This way of weighing Bell inequalities is correct, however, only if $B(\mathbf{p})\leqslant B_{j*}$. Indeed if this is not the case, the strategy just described no longer works since in Eq. (14) $q_{j*}>1$. Though the bounds (5) and (9) are still valid, it is then, in principle, possible to infer strongest bounds from the violation of the Bell inequality. This should be taken into account when comparing Bell inequalities in this way.

In the remainder of the paper, we will only be concerned with two-settings Bell scenarios. Note that in that case, $B(\mathbf{p})\leqslant B_{j*}$ is always satisfied for quantum correlations. Indeed the minimal possible communication in a (nonlocal) deterministic protocol is 1 bit and is associated with strategies in $\mathcal{D}_1$. However, every quantum correlation of a two-settings Bell scenario can be reproduced with 1 bit of communication (indeed since quantum correlations satisfy the no-signaling conditions, it suffices for one of the parties to send his input to the other so that they are able classically to simulate them). It therefore follows that $B(\mathbf{p})\leqslant B_1\leqslant B_{j*}$.

E. Other measures of communication

The general arguments we presented in this section remain valid independently of the precise way communication is counted and the way determinist strategies are accordingly partitioned. Depending on the physical quantity one is interested in, different measures for the communication cost c_i are thus possible. For example, to obtain bounds on the average communication needed to reproduce quantum correlations in classical protocols that use only one-way communication, the cost of deterministic strategies using two-way communication would be taken to be $c=\infty$. Our results therefore apply to all averaged-type measures of communication.

Note that one can also count the communication using Shanon's entropy if it is assumed that the parties may perform block coding. This is natural, for instance, if the parties perform several runs of the protocol at once as in the definition of the asymptotic communication C_∞. The resulting bound, however, will not be a lower bound on the asymptotic communication C_∞. This is because for Bell scenarios corresponding to n runs in parallel, there are deterministic strategies than cannot be written as the product of n one-run strategies. As n increases, there thus exist new ways of decomposing the correlations in terms of deterministic protocols that can possibly result in lower communication per run but which are not taken into account in the one-run decomposition (2).

Finally, note that computing the communication costs associated with deterministic strategies is in general a difficult task. It is a particular problem of the field of communication complexity for which several techniques have been specially developed [19]. However, in the case of the CHSH and the CGLMP inequality, the bound (5) can easily be deduced.

III. CHSH INEQUALITY

Let us now focus on the simplest inequality, the CHSH inequality [14]. The CHSH inequality refers to two-settings and two-outcomes Bell scenarios. The value the inequality takes on an arbitrary vector $\mathbf{p}$ is

$$\begin{aligned}B(\mathbf{p})=&p(a_0=b_0)+p(b_0\neq a_1)+p(a_1=b_1)+p(b_1=a_0)\\&-[p(a_0\neq b_0)+p(b_0=a_1)+p(a_1\neq b_1)\\&+p(b_1\neq a_0)],\end{aligned} \qquad (15)$$

where $p(a_x=b_y)=p_{00|xy}+p_{11|xy}$ and $p(a_x\neq b_y)=p_{10|xy}+p_{01|xy}$. The local bound of this inequality is $B_0=2$. The maximal violation of the CHSH inequality by quantum mechanics is $2\sqrt{2}$ and is obtained by performing measurements on Bell states. On the other hand, the maximum value it can take for all possible correlations is 4, when the four terms with a plus sign are equal to one.

To derive a bound on $\overline{C}(\mathbf{p})$ from Eq. (15), we need to compute $\max_{j\neq 0}\{(B_j-B_0)/c_j\}$. Note that in a deterministic protocol, either the two parties do not communicate at all, or one of the parties start speaking to the other. In the latter case, the minimum communication he can send is 1 bit. This implies that the minimum possible average communication for nonlocal deterministic strategies is $c_1=1$. The following protocol with entries $d_{ab|xy}=\delta^a_{\alpha(x,y)}\delta^b_{\beta(x,y)}$, where

$$\alpha(x,y)=0 \quad \text{for } x,y=0,1$$

$$\beta(0,0)=0, \quad \beta(1,0)=1, \quad \beta(0,1)=0, \quad \beta(1,1)=0 \tag{16}$$

can be implemented with 1 bit of communication. Indeed it suffices for Alice to send the value of her input to Bob. Moreover, the value $B(\mathbf{d})$ it takes on the inequality (15) is the maximum possible $B(\mathbf{d})=4$. It thus follows that $\max_{j\neq 0}\{(B_j-B_0)/c_j\}=(4-2)/1=2$, so that for the CHSH inequality the bound (5) becomes

$$\overline{C}(\mathbf{p})\geq\frac{1}{2}B(\mathbf{p})-1. \tag{17}$$

This implies, for instance, that to reproduce the optimal quantum correlations at least $\sqrt{2}-1\simeq 0.4142$ bits of communication are necessary. Note that to reproduce all possible von Neumann measurements on a Bell state 1 bit is sufficient [9].

Is it possible to find a protocol that reproduces these correlations with that amount $\overline{C}(\mathbf{p})=\sqrt{2}-1$ of communication? It turns out, in fact, that the CHSH inequality is optimal, i.e., the bound (17) is saturated for all quantum correlations. Indeed, quantum correlations satisfy the no-signaling conditions:

$$\sum_b p_{ab|xy}=\sum_b p_{ab|xy'} \quad \forall y,y'$$

$$\sum_a p_{ab|xy}=\sum_a p_{ab|x'y} \quad \forall x,x' \tag{18}$$

which express that Alice's marginal probabilities are independent of Bob's input and conversely. For correlations that obey these constraints, we have the following.

Proposition 4. $\overline{C}(\mathbf{p})=\frac{1}{2}B(\mathbf{p})-1$ bits of communication are necessary and sufficient to simulate two-settings and two-outcomes correlations $\mathbf{p}$ that violate the CHSH inequality (15) and satisfy the no-signaling conditions (18).

Proof. As the "necessary" part follows from the bound (17), we just have to exhibit a classical protocol that reproduces the correlations with that communication.

First note that when the bound (5) is saturated, it follows from the proof of Proposition 1 that the optimal protocol uses only strategies from $\mathcal{D}_0$ and $\mathcal{D}_{j*}$ and, moreover, in these subsets only strategies that attain the maximal values B_0 and B_{j*} on the inequality $\mathbf{b}$ [there could be more than one subset $\mathcal{D}_{j*}$ if they are several indexes j_* for which $(B_{j*}-B_0)/c_{j*}$ is maximum]. In our case, this implies that the optimal protocol must be built from local strategies $\mathbf{d}^{\lambda_0}$ and from 1-bit strategies $\mathbf{d}^{\lambda_1}$ such that $\mathbf{b}\cdot\mathbf{d}^{\lambda_0}=B_0=2$ and $\mathbf{b}\cdot\mathbf{d}^{\lambda_1}=B_1=4$.

The entries of the vectors $\mathbf{p}$ corresponding to the Bell scenario associated with the CHSH inequality consists of 16 probabilities $p_{ab|xy}$ since *a*, *b*, *x*, and *y* each take two possible values. Half of these probabilities appear with a plus sign in the CHSH expression (15) and half of them with a minus sign. Since entries $d_{ab|xy}=\delta^a_{\alpha(x,y)}\delta^b_{\beta(x,y)}$ of deterministic strategies are either equal to 0 or 1, for a deterministic strategy $\mathbf{d}$ to satisfy $B(\mathbf{d})=2$, it must contribute to Eq. (15) with one $-$ and three $+$. For local strategies, which assign local values $\alpha(x)$ and $\beta(y)$ to Alice's and Bob's outcomes, this leaves eight possibilities. Indeed, if we choose one of the eight entries appearing in Eq. (15) with a $-$ sign to be equal to one, the requirement that three entries appearing with a $+$ sign must also be equal to one fully determines the functions $\alpha(x)$ and $\beta(y)$. The resulting eight possible local strategies $\mathbf{d}^{\lambda_0}$ ($\lambda_0=0,\ldots,7$) are given in Table I. On the other hand, for a deterministic strategy to attain $B(\mathbf{d})=4$, it must contribute to Eq. (15) with four terms weighted by a $+$. The assignment of outcomes of 1-bit strategies $\mathbf{d}^{\lambda_1}$ are either of the form $\alpha(x)$, $\beta(x,y)$ (when Alice sends her input to Bob), or $\alpha(x,y)$, $\beta(y)$ (when it is Bob who sends his input to Alice). For each of the four possible functions $\alpha(x)$, the requirement that all the entries of the deterministic vector equal to one appear with a $+$ in the CHSH inequality fixes the function $\beta(x,y)$ and similarly for the four possible functions $\beta(y)$. There are thus eight protocols in $\mathcal{D}_1$ that attain the bound $B_1=4$. These strategies are given in Table II.

TABLE I. The eight local deterministic strategies for which $B(\mathbf{d}^{\lambda_0})=2$.

	$\mathbf{d}^{0_0}$	$\mathbf{d}^{1_0}$	$\mathbf{d}^{2_0}$	$\mathbf{d}^{3_0}$	$\mathbf{d}^{4_0}$	$\mathbf{d}^{5_0}$	$\mathbf{d}^{6_0}$	$\mathbf{d}^{7_0}$
$d_{00\|00}$	1	1	0	0	1	0	0	0
$d_{10\|00}$	0	0	0	0	0	1	0	0
$d_{01\|00}$	0	0	1	0	0	0	0	0
$d_{11\|00}$	0	0	0	1	0	0	1	1
$d_{00\|10}$	1	0	0	0	0	0	0	0
$d_{10\|10}$	0	1	0	0	1	1	0	0
$d_{01\|10}$	0	0	1	1	0	0	1	0
$d_{11\|10}$	0	0	0	0	0	0	0	1
$d_{00\|01}$	1	1	1	0	0	0	0	0
$d_{10\|01}$	0	0	0	1	0	0	0	0
$d_{01\|01}$	0	0	0	0	1	0	0	0
$d_{11\|01}$	0	0	0	0	0	1	1	1
$d_{00\|11}$	1	0	1	1	0	0	0	0
$d_{10\|11}$	0	1	0	0	0	0	0	0
$d_{01\|11}$	0	0	0	0	0	0	1	0
$d_{11\|11}$	0	0	0	0	1	1	0	1

Having characterized the deterministic strategies from which the protocol is built, it remains to determine the probabilities q_λ with which these strategies are used. These must be chosen so that

$$p_{ab|xy}=\sum_{\lambda_0=0}^{7} q_{\lambda_0}d^{\lambda_0}_{ab|xy}+\sum_{\lambda_1=0}^{7} q_{\lambda_1}d^{\lambda_1}_{ab|xy} \tag{19}$$

holds for the 16 entries $p_{ab|xy}$. Let us focus first on the entries that enter in Eq. (15) with a $-$ sign. For each of these eight entries, the only contribution to the right-hand side of Eq. (19) different from zero comes from a local deterministic

TABLE II. The eight 1-bit deterministic strategies for which $B(\mathbf{d}^{\lambda_1})=4$.

	$\mathbf{d}^{0_1}$	$\mathbf{d}^{1_1}$	$\mathbf{d}^{2_1}$	$\mathbf{d}^{3_1}$	$\mathbf{d}^{4_1}$	$\mathbf{d}^{5_1}$	$\mathbf{d}^{6_1}$	$\mathbf{d}^{7_1}$
$d_{00\|00}$	1	0	1	0	1	0	1	0
$d_{10\|00}$	0	0	0	0	0	0	0	0
$d_{01\|00}$	0	0	0	0	0	0	0	0
$d_{11\|00}$	0	1	0	1	0	1	0	1
$d_{00\|10}$	0	0	0	0	0	0	0	0
$d_{10\|10}$	0	0	1	1	1	0	1	0
$d_{01\|10}$	1	1	0	0	0	1	0	1
$d_{11\|10}$	0	0	0	0	0	0	0	0
$d_{00\|01}$	1	0	1	0	1	1	0	0
$d_{10\|01}$	0	0	0	0	0	0	0	0
$d_{01\|01}$	0	0	0	0	0	0	0	0
$d_{11\|01}$	0	1	0	1	0	0	1	1
$d_{00\|11}$	1	1	0	0	1	1	0	0
$d_{10\|11}$	0	0	0	0	0	0	0	0
$d_{01\|11}$	0	0	0	0	0	0	0	0
$d_{11\|11}$	0	0	1	1	0	0	1	1

strategy $\mathbf{d}^{\lambda_0}$. This therefore fixes the value of the corresponding probability q_{λ_0}. For instance, $q_{0_0}=p_{00|10}$ or $q_{1_0}=p_{10|11}$.

We now have to determine the value of the probabilities q_{λ_1} so that the eight entries $p_{ab|xy}$ that enter (15) with a + sign satisfy Eq. (19). For simplicity let us focus on one of these entries: $p_{00|00}$. Using Tables I and II, Eq. (19) becomes

$$p_{00|00}=q_{0_0}+q_{1_0}+q_{4_0}+q_{0_1}+q_{2_1}+q_{4_1}+q_{6_1} \tag{20}$$

or

$$q_{0_1}+q_{2_1}+q_{4_1}+q_{6_1}=p_{00|00}-p_{00|10}-p_{10|11}-p_{01|01}, \tag{21}$$

where we replaced each of the probabilities q_{λ_0} with their value previously determined. From Eq. (15) and using the no-signaling conditions (18) and the normalization conditions (1), it is not difficult to see that the left-hand side of this equation is equal to $[B(\mathbf{p})-2]/4$. The same argument can be carried for all the seven other entries that contribute to the CHSH inequality with a + sign, each time finding that the sum of four probabilities q_{λ_1} equals $[B(\mathbf{p})-2]/4$. Taking $q_{\lambda 1}=[B(\mathbf{p})-2]/16$ for $\lambda_1=0,...,7$ one therefore obtains a solution to (19).

The communication associated with this protocol is thus $\overline{C}=\Sigma_\lambda q_\lambda \overline{C}(\mathbf{d}^\lambda)=\Sigma_{\lambda_1=0}^{7} q_{\lambda_1}=\frac{1}{2}B(\mathbf{p})-1.$ ■

IV. MORE DIMENSIONS: THE CGLMP INEQUALITY

The CGLMP inequality [15] generalizes the CHSH inequality for d-dimensional systems. This inequality refers to measurement scenarios where Alice's and Bob's local settings take two values $x,y=0$, 1 and each measurement gives d possible outcomes $a,b=0,...,d-1$. The value the CGLMP inequality takes on an arbitrary vector $\mathbf{p}$ is

$$\begin{aligned}B^d(\mathbf{p})=\sum_{k=1}^{[d/2]-1}\left(1-\frac{2k}{d-1}\right)\{&P(a_0=b_0+k)+P(b_0=a_1+k\\&+1)+P(a_1=b_1+k)+P(b_1=a_0+k)\\&-[P(a_0=b_0-k-1)+P(b_0=a_1-k)+P(a_1=b_1\\&-k-1)+P(b_1=a_0-k-1)]\},\end{aligned} \tag{22}$$

where $P(a_x=b_y+k)=\Sigma_{b=0}^{d-1}p_{(b+k)b|x,y}$ is the probability that Alice and Bob results satisfy $a=b+k$ mod d when measuring x and y. As shown in [15], the local bound of the inequality is $B_0^d=2$.

When $d=2$ we recover the CHSH inequality and in that case the maximal quantum violation is $B_{ME}^2\simeq 2.828$. For $d>2$, the (conjectured) maximal violations obtained from maximally entangled qudits are given in [15]. For qutrits the maximum is $B_{ME}^3\simeq 2.8729$ and this value increases with d. This suggests that the CGLMP inequality exhibits stronger nonlocal correlations for larger d. This has been made more precise by connecting the violation of the CGLMP inequality to the resistance of the correlations to the admixture of noise [15]. It has, however, been argued in [16] that the resistance to noise is not a good measure of nonlocality. Quite surprisingly it was also found in [16] that for $d>2$ the strongest violation of the CGLMP inequality is obtained using certain nonmaximally entangled states. For qutrits, for instance, the maximal violation obtained from a nonmaximally entangled state is $B_{NME}^3\simeq 2.9149$ which is higher than the maximum $B_{ME}^3\simeq 2.8729$ for the maximally entangled one. Moreover, this discrepancy between maximally and nonmaximally states grows with the dimension. This raises several questions on how one should interpret and compare these manifestations of nonlocality.

A natural answer is through the bound (5). The derivation of the bound for the CHSH inequality in the previous section can directly be applied to the CGLMP inequality. This yields

$$\overline{C^d}(\mathbf{p})\geqslant \tfrac{1}{2}B^d(\mathbf{p})-1. \tag{23}$$

This bound is the same for all the inequalities of the family (22), and the strength of these different inequalities can therefore simply be measured by the degree by which they are violated. This confirms the intuition that the nonlocality displayed by the CGLMP inequality grows with the dimension.

On the other hand, the fact that for $d>2$ the CGLMP inequality is maximally violated for nonmaximally entangled states translates into more severe constraints on the average communication necessary to reproduce correlations obtained by measuring certain nonmaximally entangled states than maximally entangled ones. For instance, for qutrits (23) implies that $\overline{C_{ME}^3}\geqslant 0.4365$ while $\overline{C_{NME}^3}\geqslant 0.4575$. It could, however, be that for these particular correlations the CGLMP inequality is not optimal and that another inequality will impose stronger bounds for maximally entangled states.

To verify that assertion, we numerically solved the linear programming problem (13) for the correlations that maximally violate the CGLMP inequality both on maximally and nonmaximally entangled states for $d \leqslant 8$. There exists many different algorithms for linear programming and the only difficulty in solving Eq. (13) is to characterize the sets $\mathcal{D}_i$ of deterministic strategies and their corresponding communication costs c_i. A deterministic strategy assigns a definite value $\alpha(x,y)$ to Alice's outcomes and $\beta(x,y)$ to Bob's outcomes for each of the four possible pair of inputs $\{x,y\}$. To simplify the notation we write $\alpha_x(y) = \alpha(x,y)$ and $\beta_y(x) = \beta(x,y)$. There are two possibilities for α_x: either α_x is constant (const), i.e., $\alpha_x(0) = \alpha_x(1)$, and given input x Alice does not need any information from Bob to determine her output; or $\alpha_x \neq \text{const}$, that is $\alpha_x(0) \neq \alpha_x(1)$, and Alice's outcome depends not only on her local setting x but also on Bob's one. In that case Alice needs one bit of information from Bob to output her result. The situation is similar for Bob. This leads to four possible sets of deterministic strategies.

(i) $\mathcal{D}_0$: the set of local deterministic strategies for which $\alpha_x = \text{const}$ and $\beta_y = \text{const}$ for $x = 0,1$ and $y = 0,1$. These do not need any communication to be implemented: $c_0 = 0$.

(ii) $\mathcal{D}_1$: the strategies where $\alpha_x = \text{const}$ for $x = 0,1$ and at least one of the $\beta_y \neq \text{const}$. These strategies necessitate 1 bit of communication from Alice to Bob. This set also contains the reverse strategies which need 1 bit of communication from Bob to Alice. The communication cost associated to $\mathcal{D}_1$ is therefore $c_1 = 1$.

(iii) $\mathcal{D}_2$: the protocols where $\alpha_x = \text{const}$ for one of the two values $x = 0$ or $x = 1$, $\alpha_{\bar{x}} \neq \text{const}$ for the other value $\bar{x}$ and at least one of the $\beta_y \neq \text{const}$. These strategies can be implemented by Alice sending 1 bit to Bob, the value of her input, and then Bob sending back to Alice the value of his input if Alice's input equals $\bar{x}$. The average communication exchanged is 3/2 bits so that $c_2 = 3/2$. This set also contains the strategies where Alice's and Bob's positions are permuted.

(iv) $\mathcal{D}_3$: $\alpha_x \neq \text{const}$ and $\beta_y \neq \text{const}$ for $x = 0,1$ and $y = 0,1$. To implement these strategies both parties need to know the input of the other, so $c_3 = 2$.

With this assignment of communication costs to deterministic strategies and for the correlations considered ($d \leqslant 8$), it turns out from the results of the numerical optimization (13) that the CGLMP inequality is optimal, i.e., the bound (23) is saturated. For these particular measurements, those that gives rise to the maximal violation of the CGLMP inequality, more communication is thus necessary to reproduce outcomes obtained on nonmaximally entangled states than on maximally entangled ones.

It is nevertheless possible that these measurements are not optimal to detect the nonlocality of maximally entangled states. We performed numerical searches for $d = 3$, optimizing the two von Neumann measurements the parties carry out on the maximally entangled state. We found that the measurements that necessitate the maximal communication to be simulated are the ones that maximize the CGLMP inequality.

These results therefore suggest that two measurement settings on each side do not optimally detect the nonlocality of maximally entangled states for $d \geqslant 3$. It is still possible that the simulation of positive-operator-valued measures would necessitate further communication. However, concurring with [16], we believe that more settings per site and a corresponding new Bell inequality are needed [20].

V. OPTIMAL INEQUALITIES AND FACET INEQUALITIES

The CHSH and the CGLMP inequalities are special inequalities: they are facet inequalities. Local correlations $\boldsymbol{l}$ are convex combinations of a finite number of points, the local deterministic strategies: $\boldsymbol{l} = \sum_{\lambda_0} q_{\lambda_0} \mathbf{d}^{\lambda_0}$. The set of local correlations thus forms a convex polytope. Every polytope can be characterized either by its vertices (the local deterministic strategies) or by its facets which are a finite set of inequalities $\mathbf{b}^i$ $(i = 1,\ldots,M)$

$$\boldsymbol{l} = \sum_{\lambda_0} q_{\lambda_0} d^{\lambda_0} \quad \Leftrightarrow \quad \mathbf{b}^i \cdot \boldsymbol{l} \leqslant B_0^i \quad i = 1,\ldots,M. \tag{24}$$

Facet inequalities thus form a minimal set of inequalities that fully characterize the local correlations. They can therefore be viewed as tight detectors of non-locality. Complete sets of facet inequalities are known in some cases [21–24]. For two-settings, two-outcomes Bell scenarios, the CHSH is the unique (up to symmetries and besides trivial inequalities that are always satisfied by quantum correlations) facet inequality. It turns out that it is also optimal with regard to the average communication $\bar{C}$ for all quantum correlations. For Bell scenarios involving more outcomes, we have seen that the CGLMP inequality is optimal for certain correlations.

Is it the case that for quantum correlations, optimal inequalities are always facet inequalities? Consider, for instance, the following correlations belonging to a two-settings, three-outcomes Bell scenario: Alice and Bob share the maximally entangled state of two qutrits $|\psi\rangle = 1/\sqrt{3}(|00\rangle + |11\rangle + |22\rangle)$. The measurements they perform consist of each carrying out the transformation $|i\rangle \rightarrow e^{i\phi(i)}|i\rangle$, followed by a Fourier transform U_{FT} for Alice and $U_{\text{FT}*}$ for Bob and then a measurement in the computational basis. The settings of their measuring apparatus are thus determined by the three phases they use. For Alice's setting $x = 0$ and $x = 1$ the phases are (0, 0, 0) and (0, 0, $\pi/2$), while for Bob's settings $y = 0$ and $y = 1$ they are (0, 0, $\pi/4$) and (0, 0, $-\pi/4$). This results in the probabilities

$$\begin{aligned} p(a_x = b_y) &= [5 + (-1)^{f(x,y)} 2\sqrt{2}]/9, \\ p(a_x = b_y + 1) &= [2 - (-1)^{f(x,y)} \sqrt{2}]/9, \\ p(a_x = b_y + 2) &= [2 - (-1)^{f(x,y)} \sqrt{2}]/9, \end{aligned} \tag{25}$$

where $f(x,y) = x(y+1)$.

These correlations violate the CGLMP inequality by the amount $B^3(\mathbf{p}) = 2/3(1 + 2\sqrt{2}) \simeq 2.5523$. On the other hand, consider the inequality (15), which has to be viewed now as a three-outcomes inequality, i.e., $p(a_x = b_y) = \sum_k p_{kk|xy}$ and $p(a_x \neq b_y) = \sum_{k \neq l} p_{kl|xy}$, where the sum over k and l runs from 0 to 2. The above correlations violate this straightforward generalization of the CHSH inequality to more outcomes by the amount $B^{3c}(\mathbf{p}) = 2/9(1 + 8\sqrt{2}) \simeq 2.7364$. Since for both inequalities $\overline{C}(\mathbf{p}) \geqslant \frac{1}{2}B(\mathbf{p})-1$, the generalized CHSH inequality is stronger than the CGLMP ones for these particular correlations. Moreover, numerically solving the linear problem (13) we found $\overline{C}(\mathbf{p})=0.3682$ so that the bound $\overline{C}(\mathbf{p}) \geqslant 0.3682$ implied by the generalized CHSH is saturated, i.e., the inequality is optimal.

The generalized CHSH inequality, however, is not a facet inequality. Indeed, for an inequality to be a facet, the local deterministic strategies that attain the local bound B_0 (i.e., the vertices that belong to the facet) must generate a space of dimension one less than the dimension of the polytope, since they form its boundary. It is shown in [25] that the two-settings, three-outcomes polytope lies in a hyperplane of dimension 24. For the inequality (15), it is easily checked that there are only 21 local deterministic strategies that attain the limit $B_0=2$. They thus generate at best a space of dimension 21 which is less than the expected value of 23 for (15) to be a facet.

Does there exist a facet inequality that imposes the same bound $\overline{C}(\mathbf{p}) \geqslant 0.3682$ as the generalized CHSH inequality? There exist algorithms that compute all the facets of a polytope given its vertices. Using both the reverse search vertex enumeration algorithm [26] and the double description method [27] we obtained the complete set of facet inequalities of the two-settings, three-outcomes local polytope which consists of 1116 inequalities. The correlations described above violate 23 of these inequalities.

Note that there are various ways of writing a Bell inequality which are equivalent for local and quantum correlations. Indeed local and quantum correlations satisfy the normalization (1) and no-signaling conditions (18) which we express as the constraints

$$\mathbf{g}^j \cdot \mathbf{p} = G^j, \quad j=1,\ldots,J. \tag{26}$$

For probabilities that satisfy these conditions, the inequality $\mathbf{b}\cdot\mathbf{p} \leqslant B$ can be rewritten in the equivalent form

$$\left(\mu_0 \mathbf{b} + \sum_j \mu_j \mathbf{g}^j\right) \cdot \mathbf{p} \leqslant \mu_0 B + \sum_j \mu_j G^j. \tag{27}$$

In particular, with that rewriting, a facet inequality will remain a facet inequality and an inequality which is violated by correlations satisfying (26) will still be violated. This can be geometrically understood as follows. Probabilities that satisfy the constraints (26) lie in a hyperplane $\mathcal{G}$ of dimension less than the total dimension of the space $\mathcal{P}$ of all vectors $\mathbf{p}$. An inequality $\mathbf{b}\cdot\mathbf{p} \leqslant B$ defines a half-space in $\mathcal{P}$. The fact that for probabilities in $\mathcal{G}$, Bell inequalities can be written in different equivalent ways corresponds to the fact that they are different half-spaces of $\mathcal{P}$ that have the same intersection with the hyperplane $\mathcal{G}$. It is shown in [25] that the dimension of the two-settings, three-outcomes polytope (the set of all local correlations) is the same as the hyperplane $\mathcal{G}$ defined by the conditions (26) of normalization and no-signaling. It therefore follows that the rewriting (27) based on these constraints is the unique way to rewrite Bell inequalities in an equivalent form for local correlations.

However, for probabilities which do not satisfy the no-signaling conditions, such as nonlocal deterministic strategies in $\mathcal{D}_{i\neq 0}$, the rewritten inequalities (27) are not equivalent to the original one. They will thus lead to different bounds on $\overline{C}(\mathbf{p})$. Since an inequality can always be written in the normalized form of Propostion 2 using the normalization constraints in Eq. (26) (which are satisfied by all correlations), the strongest bound on $\overline{C}(\mathbf{p})$ a facet inequality $\mathbf{b}$ will impose on the correlations $\mathbf{p}$ is the solution to the following linear programming problem for the variables μ_j:

$$\begin{aligned} \max \quad & \mu_0 B(\mathbf{p}) + \sum_j \mu_j G^j \\ \text{subject to} \quad & \left(\mu_0 \mathbf{b} + \sum_j \mu_j \mathbf{g}^j\right) \cdot \mathbf{d}^{\lambda_i} \leqslant c_i . \end{aligned} \tag{28}$$

We numerically solved this linear problem for the correlations described above and each of the 23 facet inequalities they violate. The highest bound obtained was given by the CGLMP inequality and is $\overline{C}(\mathbf{p}) \geqslant 0.2764$.

This example shows that there exist quantum correlations for which the strongest bound on $\overline{C}(\mathbf{p})$ deduced from facet inequalities is lower than the (optimal) bound given from a nonfacet inequality. This is contrary to the common view according to which facet inequalities are ''optimal'' tests of nonlocality [25].

VI. CONCLUSION

In summary, we have shown that the average communication necessary to simulate classically a violation of a Bell inequality is proportional to the degree of violation of the inequality. Moreover, to each set of correlations is associated an optimal inequality for which that communication is also sufficient to reproduce the entire set of correlations. The key ingredient was to compare the amount of violation of Bell inequalities not only with the maximum value they take on local deterministic strategies, but also on nonlocal ones that necessitate some communication to be implemented.

Part of the interest of this work is that it gives a physical meaning to the degree of violation of Bell inequalities and thus provides an objective way to compare violation of different inequalities. It also gives a way to view and understand Bell inequalities that could shed light on some of their aspects. For instance, it was commonly assumed that facet inequalities are optimal tests of nonlocality because they are tight ''detectors'' of nonlocality. However, if we measure nonlocality by the communication needed to reproduce it, in certain situations nonfacet inequalities are better ''meters'' of nonlocality than are facet ones.

This work also provides a tool to characterize and quantify the nonlocality inherent in quantum correlations. As a result, for instance, for two measurements on each side it seems that the correlations that necessitate the most communication to be reproduced are obtained on nonmaximally entangled states rather than on maximally entangled ones for $d>2$. It would be interesting to know whether this is still the case for more settings and if not, what is the corresponding Bell inequality.

ACKNOWLEDGMENTS

I would like to thank Jonathan Barrett and Serge Massar for stimulating discussions and useful comments on this manuscript. I acknowledge financial support from the Communauté Française de Belgique under Grant No. ARC 00/05-251, from the IUAP program of the Belgian government under Grant No. V-18, and from the EU under project No. RESQ (IST-2001-37559).

[1] V. Scarani and N. Gisin, Phys. Rev. A **65**, 012311 (2002).

[2] A. Acín, V. Scarani, and M. M. Wolf, Phys. Rev. A **66**, 042323 (2002).

[3] G. Brassard, e-print quant-ph/0101005.

[4] C. Brukner, M. Żukowski, J. Pan, and A. Zeilinger, e-print quant-ph/0210114.

[5] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).

[6] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[7] J. A. Csirik, Phys. Rev. A **66**, 014302 (2002).

[8] D. Bacon and B. F. Toner, Phys. Rev. Lett. **90**, 157904 (2003).

[9] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).

[10] T. Maudlin, in *PSA 1992*, edited by D. Hull, M. Forbes, and K. Okruhlik (Philosophy of Science Association, East Lansing, 1992), Vol. 1, pp. 404–417.

[11] M. Steiner, Phys. Lett. A **270**, 239 (2000).

[12] A. A. Méthot, e-print quant-ph/0304122.

[13] N. J. Cerf, N. Gisin, and S. Massar, Phys. Rev. Lett. **84**, 2521 (2000).

[14] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[15] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).

[16] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Phys. Rev. A **65**, 052325 (2002).

[17] Plausibly for quantum correlations, it would be sufficient to consider deterministic strategies with communication cost $c'_N \leqslant \log_2[\min(M_A,M_B)]$. Indeed, since they satisfy the no-signaling conditions (18), they can be simulated by one of the parties sending his input to the other. However, it is logically possible that allowing strategies with higher communication cost than c'_N would result in lower communication on *average*. To avoid losing full generality, we thus consider the maximal communication cost to be c_N.

[18] A. Schrijver, *Theory of Linear and Integer Programming* (John Wiley & Sons, New York, 1986).

[19] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, New York, 1997).

[20] A candidate might be the inequality introduced by Bechmann-Pasquinucci and Gisin [Quant. Inf. Comput. **3**, 157 (2003)] which is maximally violated for maximally entangled states.

[21] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[22] I. Pitowsky and K. Svozil, Phys. Rev. A **64**, 014102 (2001).

[23] R. F. Werner and M. M. Wolf, Phys. Rev. A **64**, 032112 (2002).

[24] M. Żukowski and C. Brukner, Phys. Rev. Lett. **88**, 210401 (2002).

[25] L. Masanes, Quant. Inf. Comput. **3**, 345 (2003).

[26] D. Avis and K. Fukuda, Discrete Comput. Geom. **8**, 295 (1992), URL http://cgm.cs.mcgill.ca/\%7Eavis/C/lrs.html

[27] K. Fukuda and A. Prodon, Lect. Notes Comput. Sci. **1120**, 91 (1996), URL http://cgm.cs.mcgill.ca/~fukuda/soft/cdd_home/cdd.html