

## Optimal Bell Tests Do Not Require Maximally Entangled States

Antonio Acín,<sup>1</sup> Richard Gill,<sup>2</sup> and Nicolas Gisin<sup>3</sup>

<sup>1</sup>*ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

<sup>2</sup>*Mathematical Institute, University of Utrecht, Box 80010, 3508 TA Utrecht, The Netherlands*

<sup>3</sup>*GAP-Optique, University of Geneva, 20, Rue de l'École de Médecine, CH-1211 Geneva 4, Switzerland*

(Received 7 July 2005; published 17 November 2005)

Any Bell test consists of a sequence of measurements on a quantum state in spacelike separated regions. Thus, a state is better than others for a Bell test when, for the optimal measurements and the same number of trials, the probability of existence of a local model for the observed outcomes is smaller. The maximization over states and measurements defines the optimal nonlocality proof. Numerical results show that the required optimal state does not have to be maximally entangled.

DOI: [10.1103/PhysRevLett.95.210402](https://doi.org/10.1103/PhysRevLett.95.210402)

PACS numbers: 03.65.Ud, 03.67.—a

As first shown by Bell [1] in 1964, the correlations among the measurement outcomes of spacelike separated parties on some quantum states cannot be reproduced by a local theory. This fact is often referred to as quantum nonlocality and has been recognized as the most intriguing quantum feature. The fundamental importance of the work by Bell was that it provided conditions for experimentally testing quantum mechanics (QM) versus the whole set of local models, the so-called Bell inequalities. The experimental demonstration [2], up to some loopholes, of a Bell inequality violation definitely closed the Einstein-Podolsky-Rosen program [3] for the existence of a local theory alternative to QM.

The interest on quantum correlations, or entanglement, has dramatically increased during the last two decades due to the emerging field of quantum information science (QIS) [4]. It has been realized that quantum states provide new ways of information processing and communication without analog in classical information. The essential resource for most of these applications are entangled states. This has motivated a strong effort devoted to the characterization and quantification of the entanglement of quantum states. Although many questions remain unanswered, the problem is completely solved for the case of pure states in bipartite systems. For a state  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , its amount of entanglement is specified by the so-called entropy of entanglement [5],  $E(\Psi) = S(\rho_A)$ , where  $S$  is the usual von Neumann entropy and  $\rho_A = \text{tr}_B(|\Psi\rangle\langle\Psi|)$ . In particular, this means that the maximally entangled state in a bipartite system of dimension  $d \times d$  reads

$$|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |ii\rangle, \quad (1)$$

where  $\{|i\rangle\}$  define orthonormal bases in  $\mathcal{H}_A$  and  $\mathcal{H}_B$ .

Apart from their importance for quantum information applications, entangled states provide the only known way of establishing nonlocal correlations among spacelike separated parties. It is meant here by nonlocal those correlations that (i) cannot be explained by a local model but

(ii) do not allow any faster-than-light communication, that is, they are consistent with the no-signaling condition. Indeed, it is a well-established result that a quantum pure state violates a Bell inequality if and only if it is entangled [6]. However, it is also well known that there exist nonlocal correlations that are not achievable by measuring quantum states [7]. In a similar line of thought, it has very often been assumed that  $|\Phi_d\rangle$  represents the most nonlocal quantum state too. However, no precise demonstration of this fact has ever been given and, indeed, it is one of the scopes of this work to raise some doubts about this statement.

In what follows, entangled states constitute a resource for constructing nonlocality proofs. The strength of a Bell experiment has to be computed by means of statistical tools: a Bell test is better than another when, for the same number of trials, the probability that a local model explains the observed outcomes is smaller. Recall that statistical fluctuations on finite samples allow a local theory to predict the possibility of data violating a Bell inequality. The goal is then to identify those states needed in the construction of optimal Bell tests. The importance of constructing optimal nonlocality proofs is twofold. First, from an experimental point of view, they allow improving present Bell experiments, especially in terms of the needed resources. Second, Bell tests also represent an important tool for QIS [8]. In particular, they are useful for testing the quantumness of devices. This is a hardly explored problem in QIS that, for instance, can be especially relevant in cryptographic applications [9]: given some observed correlations among several parties, how can its quantum origin be certified? Could these correlations have alternatively been established by classical means, i.e., shared randomness? Bell inequalities provide an answer to the previous questions.

*The scenario:* We consider the standard scenario for any Bell test. Two spacelike separated parties, called Alice and Bob, share copies of a pure quantum state  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , of dimension  $d \times d$ . They can choose among  $m$  possible measurements, each of  $n$  outcomes, this being denoted by  $m \times n$ .  $A_j^i$  denotes the positive operator corre-

sponding to the outcome  $j$  of the measurement  $i$  for Alice, so  $\sum_j A_j^i = \mathbb{1}$ . Similarly, Bob's measurement operators are denoted by  $B_j^i$ . Here, we restrict most of our considerations to projective measurements, where  $A_j^i$  and  $B_j^i$  are mutually orthogonal one-dimensional projectors. The probability that Alice and Bob obtain the outcomes  $j_A$  and  $j_B$  after applying the measurement  $i_A$  and  $i_B$ , where  $j_A, j_B = 1, \dots, n$  and  $i_A, i_B = 1, \dots, m$ , on  $|\Psi\rangle$  is

$$p_Q(j_A, j_B | i_A, i_B) = \text{tr}(A_{j_A}^{i_A} \otimes B_{j_B}^{i_B} |\Psi\rangle\langle\Psi|). \quad (2)$$

In a Bell experiment, a quantum state is prepared and sent to the parties who measure it. After  $N$  repetitions of the experiment, the frequencies of the results define a  $(m^2 n^2)$ -dimensional vector whose components tend to  $p_Q(j_A, j_B | i_A, i_B)$  when  $N \rightarrow \infty$ . A vector of probabilities is achievable using QM when there exist a state  $|\Psi\rangle$  and measurements  $A_{j_A}^{i_A}$  and  $B_{j_B}^{i_B}$  satisfying (2).

On the other hand, in a local model any observed correlation between measurement results in spacelike separated regions should come from initially shared random data, denoted by  $\lambda$ . QM is nonlocal because some of the vectors (2) do not allow a local description; i.e., they cannot be written as

$$p_L(j_A, j_B | i_A, i_B) = \sum_{\lambda} p(\lambda) p(j_A | i_A, \lambda) p(j_B | i_B, \lambda). \quad (3)$$

Therefore, shared quantum states can be used to establish nonlocal correlations.

The goal of any Bell experiment is to test the hypothesis  $\mathcal{Q}$ , “the observed outcomes are governed by a quantum probability distribution (2)”, against the composite hypothesis  $\mathcal{L}$ , “there exists a local model (3) reproducing the data” [10]. The statistical tool that quantifies the *average amount of support in favor of  $\mathcal{Q}$  against  $\mathcal{L}$  per trial when the data are generated by  $\mathcal{Q}$*  is the so-called relative entropy or Kullback-Leibler (KL) divergence [11],  $D$ . For two probability distributions,  $\vec{p}_1$  and  $\vec{p}_2$  associated with the event  $\{z\}$ , it reads [12]

$$D(\vec{p}_1 || \vec{p}_2) = \sum_z p_1(z) \log(p_1(z)/p_2(z)). \quad (4)$$

We denote by  $\vec{q} = (q(1, 1, 1, 1), \dots, q(m, m, n, n))$  the quantum probabilities for the measurement settings  $i_A$  and  $i_B$ , and outcomes  $j_A$  and  $j_B$ . Using (2), one has

$$q(i_A, i_B, j_A, j_B) = \text{tr}(A_{j_A}^{i_A} \otimes B_{j_B}^{i_B} |\Psi\rangle\langle\Psi|) p_M(i_A, i_B), \quad (5)$$

where  $p_M(i_A, i_B)$  characterizes the choice of measurements by Alice and Bob. Now, the support in favor of  $\mathcal{Q}$  against  $\mathcal{L}$  provided by these quantum data is [10]

$$D(\mathcal{Q} || \mathcal{L}) = \min_{\vec{p} \in \mathcal{L}} D(\vec{q} || \vec{p}), \quad (6)$$

where the minimization runs over all alternative local models,  $\vec{p}$ . The vector  $\vec{p}$  is defined analogously to (5), replacing the quantum term  $p_Q$  (2) by a local model  $p_L$

(3). This quantity gives the statistical figure of merit to be maximized in any nonlocality test [10]. It is worth mentioning here that the KL divergence (i) is an asymptotic measure and (ii) appears as the measure of statistical support for the two most commonly used methods for hypothesis testing, frequentist and Bayesian (see Ref. [10] for more details). Moreover, and despite not being symmetric, it can be seen as a measure of statistical distance between probability distributions.

It is often convenient to interpret a Bell test as a game between a quantum and a local player [10,13]. The quantum player has to design an experimental situation for which the local player is unable to provide a model. Thus, the quantum player looks for the experiment that gives him the victory with the minimal number of repetitions, i.e., his task consists of designing the Bell test maximizing (6). In order to do that, he can choose the state to be prepared, the measurements, and the probability governing the choice of measurements,  $p_M(i_A, i_B)$ . Notice that we do not impose  $p_M(i_A, i_B)$  to be product. Indeed, one could think of a configuration where an external referee is sending the choice of measurements to the parties. On the other hand, the local player only assumes the existence of a local model. In particular, he is allowed to change his description according to the observed data.

**Results:** In what follows, the optimality of Bell tests is analyzed according to the KL divergence. The optimization of (6) in full generality is a very hard problem. Here, we mainly consider the standard situation where Alice and Bob apply two projective measurements, i.e.,  $m = 2$ ,  $n = d$ . For a fixed number of measurements, it is possible to search numerically the state and measurements defining the optimal Bell test. In the qubit case,  $d = 2$ , the best nonlocality proof is given by the maximally entangled state  $|\Phi_2\rangle$  (1) and the measurements maximizing the violation of the CHSH inequality [14], as expected. The KL divergence turns out to be equal to 0.046 bits [10] and the optimal choice of settings is completely random,  $p_M(i_A, i_B) = 1/4$ . Actually, the optimal choice of settings turns out to be random for all the situations considered in this work.

Moving to higher dimension, the optimal measurements for the maximally entangled state of two three-dimensional systems are the ones maximizing its violation of the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality [15]. The statistical strength is of 0.058 bits, reflecting the fact that quantum nonlocality increases with the dimension [16]. However, it is known that the largest violation of the CGLMP inequality is given by a nonmaximally entangled state [17]

$$|\Psi_3^{mv}\rangle = \gamma(|00\rangle + |11\rangle) + \sqrt{1 - 2\gamma^2}|22\rangle, \quad (7)$$

where  $\gamma \approx 0.617$ . The measurements maximizing its statistical strength are again those maximizing its Bell violation (which are the same as for  $|\Phi_3\rangle$ ) and give 0.072 bits, larger than the value obtained for the maximally entangled

state. The maximization now over the space of measurements, choices of settings and states gives the same measurements as above but for a different state,

$$|\Psi_3\rangle \approx \delta(|00\rangle + |11\rangle) + \sqrt{1 - 2\delta^2}|22\rangle, \quad (8)$$

where  $\delta \approx 0.642$ , the KL divergence being of 0.077 bits. Therefore, the  $3 \times 3$  state producing the optimal nonlocality test with two projective measurements per site does not have maximal entanglement. Actually, this state has even less entanglement than  $|\Psi_3^{mv}\rangle$ . All these results are summarized in Table I. It is worth mentioning here that the optimal measurements are the same for all three states. Similar results are obtained for  $d = 4$ : (i) the optimal measurements are those maximizing the Bell violation for  $|\Phi_4\rangle$  [15,17] but (ii) the optimal state,  $|\Psi_4\rangle$ , is not maximally entangled. The corresponding KL divergence is of 0.098 bits. The problem in full generality becomes intractable for larger  $d$ , so the following simplifications are considered.

First, the  $2 \times d$  measurements are taken to be equal to those maximizing the Bell violation for  $|\Phi_d\rangle$ : the parties apply a unitary operation with only nonzero terms in the diagonal,  $e^{i\phi_a(j)}$  for Alice and  $e^{i\phi_b(j)}$  for Bob, with  $j = 0, \dots, d-1$  and  $a, b = 1, 2$ . These phases read [15]

$$\begin{aligned} \phi_1(j) &= 0, & \phi_2(j) &= \frac{\pi}{d}j, \\ \varphi_1(j) &= \frac{\pi}{2d}j, & \varphi_2(j) &= -\frac{\pi}{2d}j. \end{aligned} \quad (9)$$

Then, Alice carries out a discrete Fourier transform,  $U_{FT}$ , and Bob applies  $U_{FT}^*$ , and they measure in the computational basis. Thus, it is assumed in what follows that these measurements define the optimal  $2 \times d$  Bell test. This is known to be the case for qubits, and our numerical results indicate that this also happens for  $d = 3, 4$ .

Once the settings are fixed, the problem is cast in a formulation very similar to a standard Bell inequality. The goal is now to obtain the  $d \times d$  state maximizing (6) for the given settings. Let  $\vec{q}^s$  and  $\vec{p}^s$  denote a pair forming a solution to this problem, i.e.,

$$\max_{\vec{q}} \min_{\vec{p}} D(\vec{q}||\vec{p}) = D(\vec{q}^s||\vec{p}^s) = D^s. \quad (10)$$

For small deviations from this solution one has  $D(\vec{q}^s + \delta\vec{q}||\vec{p}^s) \leq D^s$  and  $D(\vec{q}^s||\vec{p}^s + \delta\vec{p}) \geq D^s$ . Therefore, all vectors of quantum and local probabilities close to the

previous solution satisfy

$$\sum_i \log\left(\frac{q_i^s}{p_i^s}\right) q_i \leq D^s, \quad (11)$$

$$\sum_i \frac{q_i^s}{p_i^s} p_i \leq 1. \quad (12)$$

The values  $D^s$  and 1 are found by substituting  $q_i = q_i^s$  and  $p_i = p_i^s$ , respectively. Actually this has to be true for all  $\vec{q}$  and  $\vec{p}$ . If this were not the case, using convexity arguments one could construct a point arbitrarily close to  $\vec{p}^s$  or  $\vec{q}^s$  violating these conditions. Indeed, assume there exists a vector of quantum probabilities  $\vec{q}'$  violating (11). Then,  $(1 - \epsilon)\vec{q}^s + \epsilon\vec{q}'$  would also violate the same condition for arbitrarily small  $\epsilon$ . Note that the quantity on the left hand side of (11) can be seen as the mean value of a Bell operator, while (12) defines a Bell inequality. Then,  $\vec{q}^s$  maximizes (11) over all  $\vec{q}$ , while  $\vec{p}^s$  does it for (12).

After inspection, one can see that the Bell inequality (12) corresponding to the optimal solution for  $d = 2, 3, 4$  is of the CGLMP form, up to taking a linear combination with the normalization condition  $\sum_i p_i = 1$ . Actually, Eq. (12) can be rewritten in these three cases as

$$\begin{aligned} \langle [A_1 - B_1] + [B_1 - A_2] + [A_2 - B_2] \\ + [B_2 - A_1 - 1] \rangle \geq d - 1, \end{aligned} \quad (13)$$

where  $[X]$  stands for  $X$  modulo  $d$  and  $\langle X \rangle = p(X = 1) + 2p(X = 2) + \dots + (d-1)p(X = d-1)$ . This inequality easily follows from the identity

$$[A_1 - B_1 + B_1 - A_2 + A_2 - B_2 + B_2 - A_1 - 1] = d - 1, \quad (14)$$

and the fact that  $[X] + [Y] \geq [X + Y]$ . One can see that Eq. (13) represents an extremely compact way of writing all CGLMP inequalities for arbitrary dimension. Then, it is assumed that the inequality (12), derived from Eq. (10), has the CGLMP form (13), up to linear combination with  $\sum_i p_i = 1$ , also for  $d > 4$ . Thus the  $q_i^s/p_i^s$  terms in (11) and (12) are known functions of one parameter.

The problem has now been hugely simplified. Under the mentioned assumptions, the state for an optimal  $2 \times d$  Bell test is given by the eigenvector of largest eigenvalue of the Bell operator (11), where the measurements are fixed as before, and where the coefficients of the Bell operator are determined (up to one unknown parameter, over which we also optimize) by the CGLMP inequality. The associated eigenvalue gives the optimal KL divergence. This computation can be done up to very large dimension; the results can be found in Fig. 1.

*Discussion:* Fig. 1 shows that, beyond the qubit case, none of the optimal Bell tests requires a maximally entangled state. In all the studied situations, the Schmidt basis for the optimal state was the computational one. Assuming this is always the case, we can compute the convec-

TABLE I. Nonlocal content and entanglement for the maximally entangled state  $|\Phi_3\rangle$ , the state maximally violating the CGLMP inequality  $|\Psi_3^{mv}\rangle$ , and the optimal state  $|\Psi_3\rangle$ .

State	KL Divergence (bits)	Entanglement (bits)
$ \Phi_3\rangle$	0.058	1.585
$ \Psi_3^{mv}\rangle$	0.072	1.554
$ \Psi_3\rangle$	0.077	1.495

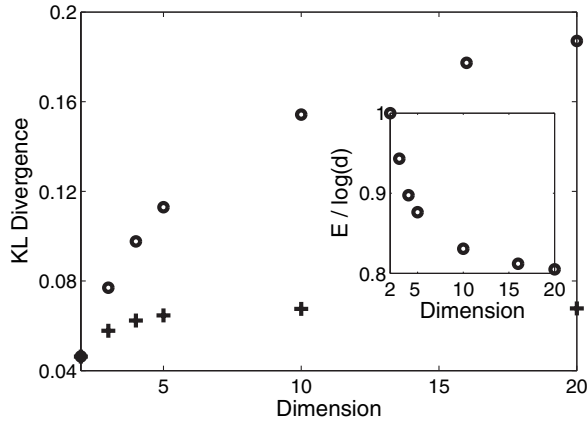


FIG. 1. KL divergence for Bell tests using two projective measurements under the mentioned assumptions. Circles correspond to the optimal state  $|\Psi_d\rangle$ , while crosses to the maximally entangled state  $|\Phi_d\rangle$ . In the inset, the entropy of entanglement for  $|\Psi_d\rangle$  is shown.

tured optimal state for large  $d$ , say  $d = 1000$ , finding  $E(\Psi_d) \rightarrow \log_e d \approx 0.69 \log d$  bits [12]. The gap between the KL divergence for the optimal and the maximally entangled state, under our assumptions, increases with the dimension. Actually, the increase of the KL divergence for  $|\Phi_d\rangle$  is negligible for small  $d$ .

It also follows from Fig. 1 that two measurements per site may not be optimal for large  $d$ . For instance, when  $d = 16$  the conjectured optimal  $2 \times 16$  test is worse than the  $4 \times 16$  test consisting of two independent realizations of the optimal  $2 \times 4$  Bell test for two copies of  $|\Psi_4\rangle$ . Indeed, it is always possible to interpret two independent realizations of this Bell test as a “new”  $2^2 \times 4^2$  Bell test for  $|\Psi\rangle^{\otimes 2}$ . Using that (i) the KL divergence is additive,  $D(\tilde{q}^2 || \tilde{p}^2) = 2D(\tilde{q} || \tilde{p})$ , and (ii) the closest local model to two independent realizations of the same Bell test corresponds to two independent realizations of the best local model for the single-copy case, the KL divergence for this test is twice the initial one.

*A priori*, one would have expected the maximally entangled state to be always the optimal state. A thorough numerical search of Bell tests for the maximally entangled state  $|\Phi_3\rangle$  using three projective measurements per site or two general measurements and with equal setting probabilities has been performed. No improvement over the optimal  $2 \times 3$  case was obtained. Actually, it is remarkable that Bell tests with two projective measurements per site are so good for low dimensional systems. Thus, all the previous numerical results show that beyond qubits and for the same resources, the optimal state for a Bell test is not maximally entangled.

**Conclusions:** Nonlocal correlations constitute an information theoretic resource *per se* [18], that can be distributed by means of quantum states. It is known that there are nonlocal correlations that cannot be established by measuring quantum states [7]. Moreover, the nonlocal correla-

tions obtained from the maximally entangled state  $|\Phi_3\rangle$  seem to be less robust against noise than those from  $|\Psi_3^{mv}\rangle$  [17]. Actually, the communication cost of simulating the nonlocal correlations for  $|\Psi_3^{mv}\rangle$  seems to be higher than for  $|\Phi_3\rangle$  [19]. More recently, it has been shown that the so-called nonlocal machine [7,18] is sufficient for the simulation of the correlations in a singlet state [20], but it fails for some nonmaximally entangled states of two qubits [21]. All these result suggest that, despite the fact that all pure entangled states contain nonlocal correlations [6], the relation between entanglement and nonlocality is subtler than firstly expected.

In this work, entangled states are analyzed as a tool for the construction of Bell tests. For all the studied scenarios and beyond the qubit case, the states needed for an optimal Bell test are not maximally entangled.

This work is supported by the ESF, an MCYT “Ramón y Cajal” grant, the Generalitat de Catalunya, the Swiss NCCR “Quantum Photonics,” and OFES within the EU project RESQ (IST-2001-37559).

- 
- [1] J. S. Bell, *Physics* (Long Island City, N.Y.) **1**, 195 (1964).
  - [2] A. Aspect, P. Grangier, and G. Roger, *Phys. Rev. Lett.* **47**, 460 (1981); W. Tittel *et al.*, *ibid.* **81**, 3563 (1998); G. Weihs *et al.*, *ibid.* **81**, 5039 (1998); M. Rowe *et al.*, *Nature* (London) **409**, 791 (2001).
  - [3] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
  - [4] See, for instance, M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge, England, 2000).
  - [5] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, *Phys. Rev. A* **53**, 2046 (1996).
  - [6] N. Gisin, *Phys. Lett. A* **154**, 201 (1991).
  - [7] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
  - [8] A. Acín, N. Gisin, L. Masanes, and V. Scarani, *Int. J. Quantum Inform.* **2**, 23 (2004).
  - [9] D. Mayers and A. Yao, *Quantum Inf. Comput.* **4**, 273 (2004).
  - [10] W. van Dam, R. Gill, and P. Grünwald, *IEEE Trans. Inf. Theory* **51**, 2812 (2005).
  - [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley Interscience, New York, 2000).
  - [12] Unless specified, all the logarithms are in base two.
  - [13] A. Peres, *Fortschr. Phys.* **48**, 531 (2000).
  - [14] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
  - [15] D. Collins *et al.*, *Phys. Rev. Lett.* **88**, 040404 (2002).
  - [16] D. Kaszlikowski *et al.*, *Phys. Rev. Lett.* **85**, 4418 (2000).
  - [17] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, *Phys. Rev. A* **65**, 052325 (2002).
  - [18] J. Barrett *et al.*, *Phys. Rev. A* **71**, 022101 (2005).
  - [19] S. Pironio, *Phys. Rev. A* **68**, 062102 (2003).
  - [20] N. J. Cerf, N. Gisin, S. Massar, and S. Popescu, *Phys. Rev. Lett.* **94**, 220403 (2005).
  - [21] N. Brunner, N. Gisin, and V. Scarani, *New J. Phys.* **7**, 88 (2005).