Local content of all pure two-qubit states

Samuel Portmann,[1] Cyril Branciard,[2] and Nicolas Gisin[1]

[1]*Group of Applied Physics, University of Geneva, Chemin de Pinchat 22, CH-1211 Geneva 4, Switzerland*
[2]*School of Mathematics and Physics, The University of Queensland, St. Lucia, Queensland 4072, Australia*

The local (nonlocal) content in the sense of Elitzur, Popescu, and Rohrlich [Phys. Lett. A **162**, 25 (1992)] is a natural measure for the locality (nonlocality) of quantum states. Its computation is in general difficult, even in low dimensions, and is one of the few open questions about pure two-qubit states. We present a complete solution to this long-lasting problem.

I. INTRODUCTION

Bell showed [1] that the classical metaphysical assumptions used by Einstein, Podolsky, and Rosen (EPR) [2] have empirical consequences in the form of a constraint on correlations of quantum measurement outcomes: the Bell inequalities. [Whether or not Einstein thought of a completion of quantum theory in the form of a hidden-variable theory ruled out by Bell's theorem, as Bell believed, has been disputed (see, e.g., the Appendix of Ref. [3] and Ref. [4]).] Bell later subsumed these assumptions into the notion of local causality [5], today commonly abbreviated to just locality. The experiments done so far give strong evidence that Bell inequalities are violated as predicted by quantum mechanics [6] and thus that there cannot be *any* local model for the corresponding correlations: The world is in this sense nonlocal.

Bell's theorem is a metatheoretical result of, at first sight, rather philosophical concern. That it can be of actual practical interest came as a surprise. [It was a surprise for Bell too (see, e.g., Ref. [7], p. 314).] Starting with an intuition of Ekert [8], it was established later on that nonlocal correlations can be made the basis of so-called device-independent quantum information processing [9–13]. There are many examples in the history of physics where a theoretical result is used for applications not foreseen by the pioneers, but Bell's theorem with its foundational character might be the most striking yet.

Whereas the violation of a Bell inequality implies nonlocality, it does not quantify it. To take the amount of the violation as a measure is problematic. For example, the state that violates it maximally depends on the particular inequality [14]. A more natural measure of nonlocality, proposed three times independently [15–17], quantifies it in terms of communication complexity (for precise mathematical definitions of this generalization of the original notion introduced by Yao [18], see, e.g., Refs. [19,20]): how many bits of communication are minimally needed for the simulation of quantum states. For the simplest case of pure two-qubit states most answers are known. Surprisingly, as shown by Toner and Bacon [21], a single bit is already enough for the maximally entangled state and two bits for all partially entangled ones. The main open question is whether one bit is already enough even for partial entanglement or whether the two bits of Ref. [21] are necessary. Whatever the answer to this question, communication complexity is clearly a very coarse measure for the nonlocality of these states. In that respect a more promising approach starts with the so-called local content.

Consider an experiment where one measures the probability distribution $P_Q(A,B|\boldsymbol{a},\boldsymbol{b})$ of the measurement outcomes A and B for measurements $\boldsymbol{a}$ and $\boldsymbol{b}$ by preparing repeatedly particle pairs in the same quantum state. If P_Q comes from a nonlocal state the whole ensemble of particle pairs cannot be described with a local model. However, as observed by Elitzur, Popescu, and Rohrlich [22], there are states for which it is possible to regard a fraction $p_L \in\,]0,1]$ of the particle pairs as behaving locally and the completing fraction $1 - p_L$ as behaving nonlocally. We will then say that there exists an Elitzur-Popescu-Rohrlich (EPR2) decomposition into a local part P_L with weight p_L and a nonlocal part P_{NL} with weight $1 - p_L$ (see Fig. 1):

$$P_Q = p_L P_L + (1 - p_L)P_{NL}, \tag{1}$$

where the decomposition holds for all von Neumann measurements and all outcomes. EPR2 decompositions are in general not unique. What is unique is the local content p_{LC}, which is defined as the maximal value p_L can take over all EPR2 decompositions. We will refer to a decomposition with $p_L = p_{LC}$ as a maximal EPR2 decomposition and define the nonlocality of a state (or its nonlocal content) as $1 - p_{LC}$. This nonlocality measure can be defined for any probability distribution. For instance, the nonlocality of the nonlocal part of a maximal EPR2 decomposition is 1 and such a nonlocal part is thus maximally nonlocal. We also note that the nonlocal part of any EPR2 decomposition of a quantum state is always more nonlocal than or equally nonlocal as the quantum state it decomposes (see Fig. 1) and is in general not quantum; however, it is still nonsignaling since both P_Q and P_L are.

The local content of pure two-qubit states was already investigated by Elitzur, Popescu, and Rohrlich. By making appropriate choices of the basis, all these states can be written as

$$|\Psi(\theta)\rangle = \cos\left(\frac{\theta}{2}\right)|00\rangle + \sin\left(\frac{\theta}{2}\right)|11\rangle, \tag{2}$$

with $\theta \in [0,\pi/2]$. All entangled states ($\theta > 0$) are nonlocal [23,24]. In spite of this, Elitzur, Popescu, and Rohrlich found an explicit EPR2 decomposition with local weight $\frac{1}{4}(1 - \sin\theta)$ for all these states, thus giving a lower bound on their local content. Furthermore, they gave an argument that the maximally entangled state has a local content of zero and is thus fully nonlocal. Further progress was achieved by Barrett, Kent, and Pironio [25]. Since the local part necessarily
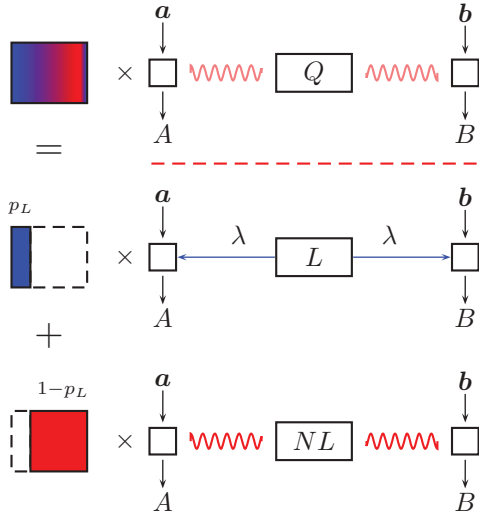
FIG. 1. (Color online) An EPR2 decomposition holds for all measurement directions $\boldsymbol{a}$ and $\boldsymbol{b}$ and measurement outcomes A and B. The red wavy lines indicate nonlocality. The nonlocal part is more nonlocal than the decomposed quantum probability distribution. The outcomes of the local part are determined by λ and the corresponding measurement direction.

obeys Bell inequalities, this places constraints on any EPR2 decomposition and can be used for deriving an upper bound on the local content. With this idea and using a chained Bell inequality [26,27], they were able to rederive (without a technical assumption used by Elitzur, Popescu, and Rohrlich) the result that maximally entangled states are fully nonlocal [28]. Scarani [29] narrowed the set of possible values of p_{LC} further: He derived both a new lower bound of $1 - \sin\theta$ and numerically, using again the same chained Bell inequality as in Ref. [25], an upper bound of $\cos\theta$; he conjectured in a preliminary version [30] of Ref. [29] that this equals the local content, i.e., that $p_{LC} = \cos\theta$. This conjecture was proven to be correct by Branciard, Gisin, and Scarani [31] for the states with $\cos\theta \in [0, \frac{4}{5}]$.

The main result of this paper is an explicit model for the local part that generalizes the model of Ref. [31] and leads to a valid EPR2 decomposition with local weight $p_L = \cos\theta$ for all states (2). Since $\cos\theta$ is an upper bound for the local content (which we prove analytically in Appendix B), this implies that $p_{LC} = \cos\theta$ and thus proves Scarani's conjecture for all pure two-qubit states.

II. PROBLEM

We will parametrize the von Neumann measurements of Alice and Bob by two three-dimensional unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ on the Bloch sphere in the form [32]

$$\boldsymbol{a} = \begin{pmatrix} a_\perp \cos\alpha \\ a_\perp \sin\alpha \\ a_z \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_\perp \cos\beta \\ b_\perp \sin\beta \\ b_z \end{pmatrix}. \quad (3)$$

With the definitions $s := \sin\theta$, $c := \cos\theta$, and $\chi := \alpha - \beta$, the expectation values of Alice's and Bob's measurement

outcomes $A, B \in \{1, -1\}$ for the states (2) are

$$\langle AB \rangle_{Q|a,b} = a_z b_z + s a_\perp b_\perp \cos\chi, \quad (4)$$

$$\langle A \rangle_{Q|a} = c a_z, \quad (5)$$

$$\langle B \rangle_{Q|b} = c b_z. \quad (6)$$

Without loss of generality [33,34], we consider only deterministic models for the local part. That is, we take the outcomes to be functions of some shared random variable λ and of the corresponding measurement directions:

$$A = A(\boldsymbol{a}, \lambda), \quad B = B(\boldsymbol{b}, \lambda). \quad (7)$$

Because of locality, A is independent of $\boldsymbol{b}$ and B is independent of $\boldsymbol{a}$. Moreover, the probability density $\rho(\lambda)$ of the shared random variable is assumed to be independent of both measurement directions. The expectation values of the local part are then given as

$$\langle AB \rangle_{L|a,b} = \int d\lambda \, \rho(\lambda) A(\boldsymbol{a}, \lambda) B(\boldsymbol{b}, \lambda),$$

$$\langle A \rangle_{L|a} = \int d\lambda \, \rho(\lambda) A(\boldsymbol{a}, \lambda), \quad (8)$$

$$\langle B \rangle_{L|b} = \int d\lambda \, \rho(\lambda) B(\boldsymbol{b}, \lambda).$$

Ultimately, our aim is to prove that $p_{LC} = c$. Since $p_{LC} \leqslant c$, as proven numerically in Ref. [29] and analytically in Appendix B, it is enough to find a model for the local part that admits a local weight $p_L = c$.

Barrett, Kent, and Pironio [25] showed that the maximally entangled state, which is fully nonlocal, necessarily has random marginals. We prove in Appendix C that the nonlocal part of a maximal decomposition of the states (2), which is also fully nonlocal, must also have random marginals. As can be seen from Eqs. (5) and (6), this implies the constraints $\langle A \rangle_L = a_z$ and $\langle B \rangle_L = b_z$ for the local part. Note that this agrees with the ansatz used in Ref. [31].

Once one has found candidate functions (7) for the local part, it remains to be checked whether the implied nonlocal part $P_{NL} = \frac{P_Q - c P_L}{1 - c}$ (assuming $p_L = c$ [35]) is a probability for all measurement directions and outcomes. Since the nonlocal part has random marginals, this is the case if the correlation term is between -1 and 1. The problem to be solved reduces to the following task [31] for any given value of c.

Problem. Find functions $A(\boldsymbol{a}, \lambda)$ and $B(\boldsymbol{b}, \lambda)$ taking values in $\{1, -1\}$ and a probability distribution $\rho(\lambda)$ such that

$$\langle A \rangle_{L|a} = a_z, \quad (9)$$

$$\langle B \rangle_{L|b} = b_z, \quad (10)$$

$$|\langle AB \rangle_{Q|a,b} - c \langle AB \rangle_{L|a,b}| \leqslant 1 - c \quad (11)$$

hold for all measurement directions $\boldsymbol{a}$ and $\boldsymbol{b}$.

Note that this formulation of the problem is redundant since, as shown in Appendix C, constraint (11) implies Eqs. (9) and (10).

III. MODEL

Our starting point is the model for the local part introduced in Ref. [31], which we will call the Branciard-Gisin-Scarani (BGS) model. The complete model below will be a generalization of this model.

Branciard-Gisin-Scarani model for the local part. Alice and Bob share the random variable $\boldsymbol{\lambda}$, which is a random three-dimensional unit vector uniformly distributed over the Bloch sphere. For measurement directions $\boldsymbol{a}$ and $\boldsymbol{b}$, Alice and Bob output, respectively,

$$A(\boldsymbol{a},\boldsymbol{\lambda}) = \text{sgn}[a_z - \boldsymbol{a} \cdot \boldsymbol{\lambda}], \qquad (12)$$

$$B(\boldsymbol{b},\boldsymbol{\lambda}) = \text{sgn}[b_z - \boldsymbol{b} \cdot \boldsymbol{\lambda}], \qquad (13)$$

where sgn is the sign function ($\in \{1,-1\}$).

This model gives the right marginals (9) and (10) and the correlation becomes [see Ref. [31] or Eq. (E15) in Appendix E for the explicit form of $E_L^{\text{BGS}}(a_z,b_z,\chi)$]

$$\langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{b}}^{\text{BGS}} = \begin{cases} 1 - |a_z - b_z| & \text{if } \chi = 0, \\ -1 + |a_z + b_z| & \text{if } \chi = \pi, \\ E_L^{\text{BGS}}(a_z,b_z,\chi) & \text{if } 0 < |\chi| < \pi. \end{cases} \qquad (14)$$

This model solves the problem partially: As verified numerically in Ref. [31], Eq. (11) is fulfilled for all states such that $c \in [0,\frac{4}{5}]$.

To get a leading idea for our generalization of the BGS model, we will concentrate on a necessary condition for the local part [31]. If Alice measures in a direction $\boldsymbol{a}$ and gets the outcome 1, Bob's qubit collapses to the state $|\boldsymbol{a}(1)\rangle$ [the eigenstate of $\boldsymbol{a}(1) \cdot \boldsymbol{\sigma}$ with eigenvalue 1, where $|\boldsymbol{a}(1)\rangle$ is one step above $|\boldsymbol{a}\rangle$ on the Hardy ladder [36,37] (see Appendix A). The vector $\boldsymbol{a}(1)$ is given by

$$\boldsymbol{a}(1) = \begin{pmatrix} a_\perp(1)\cos\alpha \\ a_\perp(1)\sin\alpha \\ a_z(1) \end{pmatrix} \quad \text{for} \quad a_z(1) := \frac{a_z + c}{1 + ca_z} \qquad (15)$$

and $a_\perp(1) = \sqrt{1 - a_z(1)^2}$. If Bob now measures in the direction $\boldsymbol{a}(1)$, he will always get the outcome 1 and thus $P_Q(1,-1|\boldsymbol{a},\boldsymbol{a}(1)) = 0$. In a similar way, we also have $P_Q(-1,1|\boldsymbol{a},\boldsymbol{a}(-1)) = 0$, where

$$\boldsymbol{a}(-1) = \begin{pmatrix} a_\perp(-1)\cos\alpha \\ a_\perp(-1)\sin\alpha \\ a_z(-1) \end{pmatrix} \quad \text{for} \quad a_z(-1) := \frac{a_z - c}{1 - ca_z}, \qquad (16)$$

and $\boldsymbol{a}(-1)$ is one step below $\boldsymbol{a}$ on the Hardy ladder. Since $P_Q = p_L P_L + (1 - p_L)P_{NL}$, this implies (for $p_L \neq 0$) that the corresponding probabilities for the local part must be zero too:

$$P_L(1,-1|\boldsymbol{a},\boldsymbol{a}(1)) = 0, \qquad (17)$$

$$P_L(-1,1|\boldsymbol{a},\boldsymbol{a}(-1)) = 0. \qquad (18)$$

The BGS model fulfills these conditions for all states, including the ones that violate constraint (11) and thus for which the model does not give a valid EPR2 decomposition. To see this, we note for instance that $\langle A \rangle_{L|\boldsymbol{a}}^{\text{BGS}} = a_z$, $\langle B \rangle_{L|\boldsymbol{a}(1)}^{\text{BGS}} = a_z(1)$, and that, since $\boldsymbol{a}$ and $\boldsymbol{a}(1)$ have the same azimuth angle (i.e., $\chi = 0$)

and $a_z(1) \geqslant a_z$,

$$\langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{a}(1)}^{\text{BGS}} = 1 - |a_z - a_z(1)| = 1 + a_z - a_z(1). \qquad (19)$$

Hence one indeed has

$$P_L^{\text{BGS}}(1,-1|\boldsymbol{a},\boldsymbol{a}(1))$$
$$= \tfrac{1}{4}\big(1 + \langle A \rangle_{L|\boldsymbol{a}}^{\text{BGS}} - \langle B \rangle_{L|\boldsymbol{a}(1)}^{\text{BGS}} - \langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{a}(1)}^{\text{BGS}}\big) = 0. \qquad (20)$$

Let us now see in what way we can alter Alice's output function (12) in the BGS model such that conditions (17) and (18) still hold [we ignore condition (9) for a moment]. Assume that Bob measures in direction $\boldsymbol{b} = \boldsymbol{a}(1)$: Condition (17) still holds if Alice chooses to output $A(\boldsymbol{a}(t),\boldsymbol{\lambda})$ with

$$\boldsymbol{a}(t) = \begin{pmatrix} a_\perp(t)\cos\alpha \\ a_\perp(t)\sin\alpha \\ a_z(t) \end{pmatrix} \qquad (21)$$

and $-1 \leqslant a_z(t) \leqslant a_z(1)$. Similarly, assume that Bob measures in direction $\boldsymbol{b} = \boldsymbol{a}(-1)$: Condition (18) still holds if Alice chooses to output $A(\boldsymbol{a}(t),\boldsymbol{\lambda})$ with $a_z(-1) \leqslant a_z(t) \leqslant 1$. Taken together, since Bob could be measuring in either direction $\boldsymbol{a}(1)$ or $\boldsymbol{a}(-1)$ about which Alice has no information, Alice can choose to output $A(\boldsymbol{a}(t),\boldsymbol{\lambda})$ with $a_z(-1) \leqslant a_z(t) \leqslant a_z(1)$ and still be sure that conditions (17) and (18) hold. We will say that Alice is free to choose $\boldsymbol{a}(t)$ in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$. To use this wiggle room is precisely the idea for the construction of our θ-dependent generalization of the BGS model.

Bob's output will be identical to his output in Eq. (13) but Alice will now output $A(\boldsymbol{a}(t),\boldsymbol{\lambda})$, where $t \in [-1,1]$ parametrizes the vectors in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ such that $\boldsymbol{a}(0) = \boldsymbol{a}$ and $\boldsymbol{a}(1)$ and $\boldsymbol{a}(-1)$ are the two vectors introduced in Eqs. (15) and (16). Alice picks a particular $\boldsymbol{a}(t)$ depending on a local (nonshared) random variable t, which has a probability distribution $\rho_{\boldsymbol{a}}(t)$ that depends on a_z (besides depending on θ) [38]. One obvious constraint on $\rho_{\boldsymbol{a}}(t)$ comes from the condition that Eq. (9) should still be fulfilled. Note that for $\chi = 0$ the model implies the same correlation term (14) as the BGS model if $\boldsymbol{b}$ does *not* lie in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$. Incidentally, for such measurement directions, one can show (see Appendix D) that the BGS model does lead to a valid EPR2 decomposition for all values of θ, as it should be for the following generalization having a chance to be valid for all states and all measurement directions.

Complete model for the local part. Alice and Bob share the random three-dimensional vector $\boldsymbol{\lambda}$, uniformly distributed on the Bloch sphere. Alice can sample an additional local random variable $t \in [-1,1]$ according to the following distribution, which depends on her setting $\boldsymbol{a}$ through its component a_z:

$$\rho_{\boldsymbol{a}}(t) := \frac{s\ln(\gamma)}{4c} \frac{1 + G(t)a_z}{\sqrt{1 - G(t)^2}} \qquad (22)$$

with

$$\gamma := \frac{1 + c}{1 - c}, \quad G(t) := \frac{\gamma^t - 1}{\gamma^t + 1}. \qquad (23)$$

Alice's and Bob's outputs are then defined as

$$A(\boldsymbol{a},\boldsymbol{\lambda},t) = \text{sgn}[a_z(t) - \boldsymbol{a}(t) \cdot \boldsymbol{\lambda}], \qquad (24)$$

$$B(\boldsymbol{b},\boldsymbol{\lambda}) = \text{sgn}[b_z - \boldsymbol{b} \cdot \boldsymbol{\lambda}], \qquad (25)$$

where

$$\boldsymbol{a}(t) := \begin{pmatrix} a_\perp(t)\cos\alpha \\ a_\perp(t)\sin\alpha \\ a_z(t) \end{pmatrix} \quad \text{for} \quad a_z(t) := \frac{a_z + G(t)}{1 + G(t)a_z} \quad (26)$$

and $a_\perp(t) = \sqrt{1 - a_z(t)^2}$.

The parametrization of the vectors in the Hardy sectors $\mathcal{H}_{\boldsymbol{a}}$ as $\boldsymbol{a}(t)$ is motivated in Appendix A. The probability distribution $\rho_{\boldsymbol{a}}(t)$ is derived in Appendix E from the postulate that constraint (11) is saturated for the case that $\boldsymbol{b}$ lies in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$. This, together with the previous remark on the case where $\boldsymbol{b}$ does not lie in $\mathcal{H}_{\boldsymbol{a}}$, allows us to prove analytically that our local model above defines a valid EPR2 decomposition, with a maximal local weight $p_L = p_{LC} = c$, for all measurement directions $\boldsymbol{a}, \boldsymbol{b}$ with $\chi = 0$. One can show in a similar way that this also works for measurements with $\chi = \pi$.

For measurement settings $\boldsymbol{a}, \boldsymbol{b}$ with $0 < |\chi| < \pi$, we checked numerically with standard numerical optimization tools of a common computational software program that constraint (11) holds within the expected numerical precision (of absolute order 10^{-10}). Note that for any θ, $\langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{b}}$ depends only on the three parameters a_z, b_z, and χ. We also note that the implied nonlocal part is nonquantum for all $c \in \,]0,1[$ (see Appendix F).

IV. DISCUSSION

Our EPR2 decomposition for two-qubit states, inspired by that of Ref. [31] and based on properties of Hardy ladder vectors, thus allows us to show that the local content of all pure two-qubit states (2) equals $c = \cos\theta$. This proves a conjecture formulated in Ref. [30]. Interestingly, the nonlocality $1 - c$ of these states is found to be monotonically related to its entanglement, against the trend of results that suggest that entanglement and nonlocality are of a rather different nature (see, e.g., Refs. [39–41]).

One obvious next step is to consider more general states. For the case of higher-dimensional partially entangled two-part states, we only know of the numerical lower bound on p_{LC} found by Scarani [29] for a class of entangled qutrits. We believe that the methods presented in Appendixes B and C should be generalizable to get analytical upper bounds that might then pave the way for models of the local part. Whether the monotonic connection between entanglement and nonlocality then still holds is a question of particular interest.

We would like to end with one possible application of our decomposition and some ensuing open questions. *Any* EPR2 decomposition can be used as a starting point for a simulation protocol, where only the nonlocal part remains to be simulated [42]. Now, the arguably biggest difficulty in the simulation of partially entangled states is their nonrandom marginals. Using our maximal EPR2 decomposition seems in that respect ideally suited since its nonlocal part has random marginals. We also note that nonlocal correlations with random marginals that can be simulated with one bit of classical communication can be simulated with one nonlocal Popescu-Rohrlich (PR) box [43] too [44]. Thus, *if* the nonlocal part of a maximal EPR2 decomposition for two-qubit correlations

can be simulated with one bit of communication, then the corresponding quantum state can be simulated with one PR box. This however is known to be impossible for weakly entangled two-qubit states [39]; hence their nonlocal part in a maximal EPR2 decomposition cannot be simulated with one bit of communication. This does not yet answer the open question of whether partially entangled states can be simulated with a single bit: Whereas it is always possible to simulate the decomposed state by simulating the nonlocal part, this might not be the optimal way to do it. On the one hand, maximal EPR2 decompositions lead to a natural alternative measure of nonlocality; on the other hand, they thus lead to insights into the quantification of nonlocality in terms of communication complexity.

APPENDIX A: HARDY-LADDER VECTORS

The vectors $\boldsymbol{a}(1)$ and $\boldsymbol{a}(-1)$, one step above and below $\boldsymbol{a} = (a_\perp \cos\alpha, a_\perp \sin\alpha, a_z)$ on the Hardy ladder, are defined such that $P_Q(1, -1|\boldsymbol{a}, \boldsymbol{a}(1)) = 0$ and $P_Q(-1, 1|\boldsymbol{a}, \boldsymbol{a}(-1)) = 0$, respectively, and are given by

$$\boldsymbol{a}(1) = \begin{pmatrix} a_\perp(1)\cos\alpha \\ a_\perp(1)\sin\alpha \\ a_z(1) \end{pmatrix} \quad \text{for} \quad a_z(1) = \frac{a_z + c}{ca_z + 1} \quad (A1)$$

and

$$\boldsymbol{a}(-1) = \begin{pmatrix} a_\perp(-1)\cos\alpha \\ a_\perp(-1)\sin\alpha \\ a_z(-1) \end{pmatrix} \quad \text{for} \quad a_z(-1) = \frac{a_z - c}{-ca_z + 1}.$$

$$(A2)$$

The successive steps of the Hardy ladder are iteratively defined as

$$P_Q(1, -1|\boldsymbol{a}(n), \boldsymbol{a}(n+1)) = 0, \quad (A3)$$

$$P_Q(-1, 1|\boldsymbol{a}(-n), \boldsymbol{a}(-n-1)) = 0 \quad (A4)$$

for all $n \in \mathbb{N}$. To get the explicit form of $\boldsymbol{a}(n)$ for all $n \in \mathbb{Z}$, we note that the maps

$$a_z \mapsto \frac{a_z \pm c}{\pm c a_z + 1} \quad (A5)$$

are (real) Möbius transformations. The composition of Möbius transformations basically reduces to the product of their associated matrices, in our case to the product of

$$\begin{pmatrix} 1 & \pm c \\ \pm c & 1 \end{pmatrix}. \quad (A6)$$

Defining (we assume $c < 1$)

$$\gamma := \frac{1+c}{1-c} \geqslant 1, \quad G(n) := \frac{\gamma^n - 1}{\gamma^n + 1} \in [-1,1], \quad \text{(A7)}$$

one gets the explicit form of $\boldsymbol{a}(n)$ as

$$\boldsymbol{a}(n) = \begin{pmatrix} a_\perp(n)\cos\alpha \\ a_\perp(n)\sin\alpha \\ a_z(n) \end{pmatrix} \quad \text{for} \quad a_z(n) = \frac{a_z + G(n)}{G(n)a_z + 1} \quad \text{(A8)}$$

and $a_\perp(n) = \sqrt{1 - a_z(n)^2}$. This can be independently checked, e.g., by induction.

In order to get a parametrization of the vectors between $\boldsymbol{a}(-1)$ and $\boldsymbol{a}(1)$, we let n take real values t in the interval $[-1,1]$. Note that $a_z(t)$ is a monotonically increasing function. We then define for every vector $\boldsymbol{a}$ its Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ to be the set of all vectors that can be written as $\boldsymbol{a}(t)$ for some $t \in [-1,1]$:

$$\mathcal{H}_{\boldsymbol{a}} := \{\boldsymbol{a}(t)|t \in [-1,1]\}. \quad \text{(A9)}$$

APPENDIX B: ANALYTIC PROOF THAT c IS AN UPPER BOUND FOR p_{LC}

Let us label N measurement directions of Alice as $\boldsymbol{a}(-N)$, $\boldsymbol{a}(-N+2)$, ..., $\boldsymbol{a}(N-2)$ and N directions of Bob as $\boldsymbol{b}(-N+1), \boldsymbol{b}(-N+3), \ldots, \boldsymbol{b}(N-1)$. For a given probability distribution P with binary outcomes A, B, we define as a correlation measure the expression

$$I_N(P) = P(A_{-N} \neq B_{-N+1}) + P(A_{-N+2} \neq B_{-N+1}) + \cdots$$
$$+ P(A_{N-2} \neq B_{N-1}) + P(A_{-N} = B_{N-1}), \quad \text{(B1)}$$

where $P(A_i \neq B_j)$ is shorthand notation for $P(A \neq B|\boldsymbol{a}(i), \boldsymbol{b}(j))$. As shown in Refs. [26,27], I_N is a chained Bell polynomial, defining the following chained Bell inequality, satisfied by all local correlations:

$$I_N(P_L) \geqslant 1. \quad \text{(B2)}$$

Since by linearity

$$I_N(P_Q) = p_L I_N(P_L) + (1 - p_L)I_N(P_{NL}) \quad \text{(B3)}$$

and $I_N(P_{NL})$ does not have any constraints except that the probabilities are between zero and one and thus $I_N(P_{NL}) \geqslant 0$, $I_N(P_Q)$ is an upper bound for p_L [25]:

$$p_L \leqslant I_N(P_Q). \quad \text{(B4)}$$

We show now that the measurement directions of Alice and Bob can be chosen such that

$$\lim_{N \to \infty} I_N(P_Q) = c. \quad \text{(B5)}$$

For that, we start with an arbitrary unit vector

$$\boldsymbol{v}(0) := \begin{pmatrix} v_\perp(0)\cos(\omega) \\ v_\perp(0)\sin(\omega) \\ v_z(0) \end{pmatrix} \quad \text{(B6)}$$

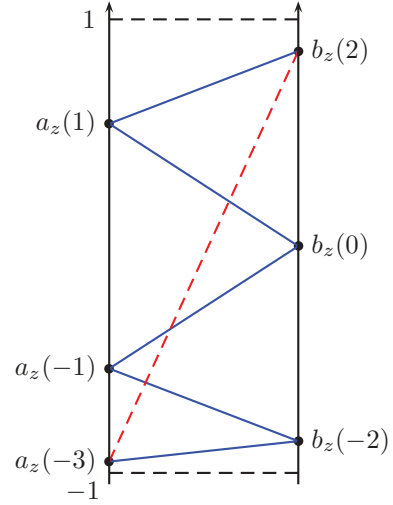
FIG. 2. (Color online) Hardy ladder for $N = 3$. Each colored line represents one term in the chained Bell inequality, the dashed red one being the last one $P(A_{-N} = B_{N-1})$.

for any $v_z(0) \in]-1,1[$, $v_\perp(0) = \sqrt{1 - v_z(0)^2}$ and $\omega \in [0, 2\pi[$. Let us then define, with the notation of Eq. (A8),

$$\boldsymbol{v}(n) \equiv \begin{pmatrix} v_\perp(n)\cos(\omega) \\ v_\perp(n)\sin(\omega) \\ v_z(n) \end{pmatrix}. \quad \text{(B7)}$$

We then choose Alice's settings to be $\boldsymbol{a}(n) = \boldsymbol{v}(n)$ for $n \in \{-N, -N+2, \ldots, N-2\}$ and Bob's settings to be $\boldsymbol{b}(n) = \boldsymbol{v}(n)$ for $n \in \{-N+1, -N+3, \ldots, N-1\}$. An example of the z coordinates of such a set of vectors is drawn in Fig. 2 for $N = 3$. Note that we have

$$\lim_{n \to \infty} \boldsymbol{v}(n) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \lim_{n \to -\infty} \boldsymbol{v}(n) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \quad \text{(B8)}$$

Since

$$P_Q(1, -1|\boldsymbol{a}(n), \boldsymbol{b}(n+1)) = 0 \quad \text{(B9)}$$

for two neighboring Hardy steps $\boldsymbol{a}(n) = \boldsymbol{v}(n)$ and $\boldsymbol{b}(n+1) = \boldsymbol{v}(n+1)$, one has

$$\langle A_n B_{n+1} \rangle_Q := \langle AB \rangle_{Q|\boldsymbol{a}(n), \boldsymbol{b}(n+1)} = 1 + c[v_z(n) - v_z(n+1)]$$
$$\text{(B10)}$$

and hence

$$P_Q(A_n \neq B_{n+1}) = \frac{1}{2}[1 - \langle A_n B_{n+1} \rangle_Q]$$
$$= \frac{c}{2}[v_z(n+1) - v_z(n)]. \quad \text{(B11)}$$

Similarly,

$$P_Q(A_{n+1} \neq B_n) = \frac{c}{2}[v_z(n+1) - v_z(n)]. \quad \text{(B12)}$$

Then $I_N(P_Q)$ becomes

$$
\begin{aligned}
I_N(P_Q) &= \sum_{n=-N}^{N-2} \frac{c}{2}[v_z(n+1) - v_z(n)] \\
&\quad + \frac{1}{2}[1 + \langle A_{-N}B_{N-1}\rangle_Q] \\
&= \frac{c}{2}[v_z(N-1) - v_z(-N)] + \frac{1}{2}[1 + \langle A_{-N}B_{N-1}\rangle_Q].
\end{aligned}
$$
(B13)

Because of Eq. (B8), we then have

$$
\lim_{N \to \infty} I_N(P_Q) = c,
$$
(B14)

which leads, with Eq. (B4), to

$$
p_L \leqslant c.
$$
(B15)

APPENDIX C: PROOF THAT $\langle A \rangle_{NL} = \langle B \rangle_{NL} = 0$ FOR A MAXIMAL EPR2 DECOMPOSITION OF P_Q

As shown in Appendix B, there exist measurement directions for which

$$
\lim_{N \to \infty} I_N(P_Q) = c.
$$
(C1)

Since we have, for a maximal EPR2 decomposition of P_Q (i.e., for $p_L = p_{LC} = c$),

$$
I_N(P_Q) = cI_N(P_L) + (1-c)I_N(P_{NL}),
$$
(C2)

with $I_N(P_L) \geqslant 1$ and $I_N(P_{NL}) \geqslant 0$, then we must have

$$
\lim_{N \to \infty} I_N(P_L) = 1, \quad \lim_{N \to \infty} I_N(P_{NL}) = 0
$$
(C3)

for such a set of measurement directions.

Note that apart from the last term $P(A_{-N} = B_{N-1})$, a term that is present in $I_N(P_{NL})$ is also present in all subsequent polynomials $I_{N+2k}(P_{NL})$ for $k \in \mathbb{N}$. Since all terms appearing in $I_N(P_{NL})$ are probabilities and $\lim_{N \to \infty} I_N(P_{NL}) = 0$, then all terms have to be zero, except possibly for the last one, which must only *tend to* zero (as $N \to \infty$). Take one term $P(A_n \neq B_{n\pm1}) = 0$, which implies that

$$
\begin{aligned}
P(A_n &= +1, B_{n\pm1} = -1) \\
&= \tfrac{1}{4}(1 + \langle A_n \rangle_{NL} - \langle B_{n\pm1} \rangle_{NL} - \langle A_n B_{n\pm1} \rangle_{NL}) = 0, \\
P(A_n &= -1, B_{n\pm1} = +1) \\
&= \tfrac{1}{4}(1 - \langle A_n \rangle_{NL} + \langle B_{n\pm1} \rangle_{NL} - \langle A_n B_{n\pm1} \rangle_{NL}) = 0,
\end{aligned}
$$

and therefore

$$
\begin{aligned}
P(A_n &= +1, B_{n\pm1} = -1) - P(A_n = -1, B_{n\pm1} = +1) \\
&= \tfrac{1}{2}(\langle A_n \rangle_{NL} - \langle B_{n\pm1} \rangle_{NL}) = 0.
\end{aligned}
$$
(C4)

Hence $\langle A_n \rangle_{NL} = \langle B_{n\pm1} \rangle_{NL}$ and by induction $\langle A_i \rangle_{NL} = \langle B_j \rangle_{NL}$ for any odd value of i and even value of j. In a similar

way, $P(A_{-N} = B_{N-1}) \to 0$ implies that

$$
\begin{aligned}
P(A_{-N} &= B_{N-1} = +1) \\
&= \tfrac{1}{4}(1 + \langle A_{-N} \rangle_{NL} + \langle B_{N-1} \rangle_{NL} + \langle A_{-N} B_{N-1} \rangle_{NL}) \to 0, \\
P(A_{-N} &= B_{N-1} = -1) \\
&= \tfrac{1}{4}(1 - \langle A_{-N} \rangle_{NL} - \langle B_{N-1} \rangle_{NL} + \langle A_{-N} B_{N-1} \rangle_{NL}) \to 0,
\end{aligned}
$$

and therefore

$$
\begin{aligned}
P(A_{-N} &= B_{N-1} = +1) - P(A_{-N} = B_{N-1} = -1) \\
&= \tfrac{1}{2}(\langle A_{-N} \rangle_{NL} + \langle B_{N-1} \rangle_{NL}) \to 0.
\end{aligned}
$$
(C5)

Now, from the previous argument, $\langle A_{-N} \rangle_{NL} = \langle B_{N-1} \rangle_{NL} = \langle A_i \rangle_{NL} = \langle B_j \rangle_{NL}$ for any i, j. Equation (C5) then implies that $\langle A_i \rangle_{NL} = \langle B_j \rangle_{NL} = 0$ for all the measurement directions appearing in $I_N(NL)$.

Since we started the construction in Appendix B with an arbitrary measurement vector [see Eq. (B6)], we must have, for the maximal EPR2 decomposition considered here,

$$
\langle A \rangle_{NL|\boldsymbol{a}} = \langle B \rangle_{NL|\boldsymbol{b}} = 0
$$
(C6)

for all $\boldsymbol{a}$ and $\boldsymbol{b}$. This in turn implies that the marginals of the local part are, as given in Eqs. (9) and (10),

$$
\langle A \rangle_{L|\boldsymbol{a}} = a_z, \quad \langle B \rangle_{L|\boldsymbol{b}} = b_z.
$$
(C7)

APPENDIX D: PROOF THAT THE BGS MODEL GIVES A VALID EPR2 DECOMPOSITION FOR MEASUREMENT SETTINGS SUCH THAT $\chi = 0$ AND $b \notin \mathcal{H}_a$

In this Appendix we prove that when Alice's and Bob's settings are such that $\chi = 0$ but $\boldsymbol{b}$ does not lie in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$ [i.e., $b_z > a_z(1)$ or $b_z < a_z(-1)$], then the BGS EPR2 decomposition [31] is valid for all states. For that we will show that Eq. (11) holds for all values of c.

Suppose indeed that $\chi = 0$ and $b_z > a_z(1) = \frac{a_z+c}{1+ca_z} \geqslant a_z$. Then by omitting the conditional settings we have

$$
\langle AB \rangle_Q - c\langle AB \rangle_L = a_z b_z + s a_\perp b_\perp - c(1 - b_z + a_z).
$$

Remembering that $b_\perp = \sqrt{1-b_z^2}$, we find $\frac{\partial^2}{\partial b_z^2}[\langle AB \rangle_Q - c\langle AB \rangle_L] = -s a_\perp b_\perp^{-3} \leqslant 0$ and therefore that $\frac{\partial}{\partial b_z}[\langle AB \rangle_Q - c\langle AB \rangle_L]$ decreases with b_z. Now observing that $\frac{\partial}{\partial b_z}[\langle AB \rangle_Q - c\langle AB \rangle_L]_{|b_z=a_z(1)} = 0$, we conclude that for $b_z \geqslant a_z(1)$, $\frac{\partial}{\partial b_z}[\langle AB \rangle_Q - c\langle AB \rangle_L] \leqslant 0$. Hence $\langle AB \rangle_Q - c\langle AB \rangle_L$ decreases monotonically for $b_z \in [a_z(1),1]$ and therefore takes values between $[\langle AB \rangle_Q - c\langle AB \rangle_L]_{|b_z=1} = (1-c)a_z \geqslant -(1-c)$ and $[\langle AB \rangle_Q - c\langle AB \rangle_L]_{|b_z=a_z(1)} = 1-c$. It follows that $|\langle AB \rangle_Q - c\langle AB \rangle_L| \leqslant 1-c$, i.e., that Eq. (11) is satisfied.

The case $b_z < a_z(-1)$ can be analyzed in a very similar way. We thus conclude that when Alice's and Bob's settings are such that $\chi = 0$ but with $\boldsymbol{b} \notin \mathcal{H}_{\boldsymbol{a}}$, the BGS model gives a valid EPR2 decomposition for all states.

APPENDIX E: DERIVATION OF $\rho_a(t)$

We consider the case $\chi = 0$ and assume for now that $\boldsymbol{b}$ lies in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$. The correlator reads

$$
\langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{b}}^{\chi=0} = \int_{-1}^{1} dt\, \rho_{\boldsymbol{a}}(t)(1 - |a_z(t) - b_z|).
$$
(E1)

To determine $\rho_{\boldsymbol{a}}(t)$ we will demand that condition (11) is exactly saturated.

Proposition 1. If $\rho_{\boldsymbol{a}}(t) := \frac{s\ln(\gamma)}{4c}\frac{1+G(t)a_z}{\sqrt{1-G(t)^2}}$, then $\langle AB\rangle_{Q|\boldsymbol{a},\boldsymbol{b}}^{\chi=0} - c\langle AB\rangle_{L|\boldsymbol{a},\boldsymbol{b}}^{\chi=0} = 1-c$.

Proof. We will use the ansatz for $t \in [-1,1]$,

$$\rho_{\boldsymbol{a}}(t) = [1 + G(t)a_z]\rho(t), \tag{E2}$$

where $\rho(t)$ is a (normalized) even probability distribution, and we further assume that

$$\int_0^{\pm 1} dt\, \rho(t)G(t) = \frac{1-s}{2c}\left[=\frac{1}{2}G\left(\frac{1}{2}\right)\right]. \tag{E3}$$

The ansatz (E2) ensures that $\rho_{\boldsymbol{a}}(t)$ is a well-defined probability distribution and $\langle A\rangle_L = a_z$ as desired [which follow from the facts that $1 + G(t)a_z \geqslant 0$ and $\rho(t)G(t)$ is an odd function], while (E3) ensures that the proposition is true for $\boldsymbol{a} = \boldsymbol{b}$. At the end we will have to check that condition (E3) holds and that $\rho(t)$ is indeed a probability distribution.

Since $\boldsymbol{b}$ is assumed to lie in $\mathcal{H}_{\boldsymbol{a}}$, b_z can then be written as

$$b_z = a_z(T) \quad \text{for some} \quad T \in [-1,1]. \tag{E4}$$

We have

$$\langle AB\rangle_{L|\boldsymbol{a},\boldsymbol{b}}^{\chi=0} = 1 - \int_{-1}^{1} dt\, \rho_{\boldsymbol{a}}(t)|a_z(t) - b_z| \tag{E5}$$

$$= 1 + \int_{-1}^{T} dt\, f(t) - \int_{T}^{1} dt\, f(t), \tag{E6}$$

with $f(t) = \rho_{\boldsymbol{a}}(t)[a_z(t) - b_z] = \rho(t)[a_z - b_z + (1-a_zb_z)G(t)]$. By decomposing the two integrals we find

$$\langle AB\rangle_{L|\boldsymbol{a},\boldsymbol{b}}^{\chi=0} = 1 + \int_{-1}^{0} dt\, f(t) + 2\int_{0}^{T} dt\, f(t) - \int_{0}^{1} dt\, f(t)$$

$$= 1 - \int_{0}^{1} dt\,[f(t) - f(-t)] + 2\int_{0}^{T} dt\, f(t)$$

$$= 1 - (1-a_zb_z)\frac{1-s}{c} + 2\int_{0}^{T} dt\, f(t), \tag{E7}$$

where we made use of the parity of $\rho(t)$ and of assumption (E3).

Demanding that $\langle AB\rangle_Q^{\chi=0} - c\langle AB\rangle_L^{\chi=0} = 1-c$ with

$$\langle AB\rangle_Q^{\chi=0} = a_zb_z + s\,a_\perp b_\perp \tag{E8}$$

leads to

$$2c\int_{0}^{T} dt\, f(t) = sa_\perp b_\perp - s(1-a_zb_z). \tag{E9}$$

Expressing now b_z as $\frac{a_z+G(T)}{1+G(T)a_z}$, one gets

$$a_\perp b_\perp = \frac{a_\perp^2}{1+G(T)a_z}\sqrt{1-G(T)^2},$$

$$a_z - b_z = \frac{-a_\perp^2}{1+G(T)a_z}G(T), \quad 1-a_zb_z = \frac{a_\perp^2}{1+G(T)a_z}. \tag{E10}$$

Using these expressions, Eq. (E9) leads to the integral equation

$$\int_{0}^{T} dt\, \rho(t)[G(T) - G(t)] = \frac{s}{2c}[1 - \sqrt{1-G(T)^2}], \tag{E11}$$

which can be solved easily: e.g., after substituting

$$t \to G(t), \tag{E12}$$

one gets an equation with a linear kernel. The unique solution reads

$$\rho(t) = \frac{s\ln(\gamma)}{4c}\frac{1}{\sqrt{1-G(t)^2}}. \tag{E13}$$

As can be checked, $\rho(t)$ is an even, normalized probability distribution on the interval $[-1,1]$ and condition (E3) holds.

The choice $\rho_{\boldsymbol{a}}(t) := \frac{s\ln(\gamma)}{4c}\frac{1+G(t)a_z}{\sqrt{1-G(t)^2}}$ thus ensures that our maximal EPR2 decomposition, with $p_L = p_{LC} = c$, is valid for all measurement directions $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\chi = 0$ when $\boldsymbol{b}$ lies in the Hardy sector $\mathcal{H}_{\boldsymbol{a}}$ of $\boldsymbol{a}$. If $\chi = 0$ but $\boldsymbol{b} \notin \mathcal{H}_{\boldsymbol{a}}$, our local model gives the same correlation as the BGS model, which we proved in Appendix D to also give a valid EPR2 decomposition. Note that if $\chi = \pi$, the situation is quite similar to $\chi = 0$ and one can prove again that our EPR2 decomposition is still valid.

The question now is whether our maximal EPR2 decomposition remains valid for all other settings when $0 < |\chi| < \pi$. The difficulty in checking that Eq. (11) indeed holds is that one needs to calculate $\langle AB\rangle_{L|\boldsymbol{a},\boldsymbol{b}}$, given by

$$\langle AB\rangle_{L|\boldsymbol{a},\boldsymbol{b}} = \int_{-1}^{1} dt\, \rho_{\boldsymbol{a}}(t)\, E_L^{\text{BGS}}(a_z(t),b_z,\chi), \tag{E14}$$

with

$$E_L^{\text{BGS}}(a_z,b_z,\chi) = 1 - \frac{2|\chi|}{\pi}$$
$$+ \frac{2}{\pi}a_z\arctan\left(\frac{a_\perp b_z - a_zb_\perp\cos\chi}{b_\perp\sin|\chi|}\right)$$
$$+ \frac{2}{\pi}b_z\arctan\left(\frac{a_zb_\perp - a_\perp b_z\cos\chi}{a_\perp\sin|\chi|}\right) \tag{E15}$$

from the BGS model [31].

Unfortunately, we were not able to calculate the integral (E14) explicitly. However, we carried intensive numerical checks to convince ourselves that Eq. (11) always holds (up
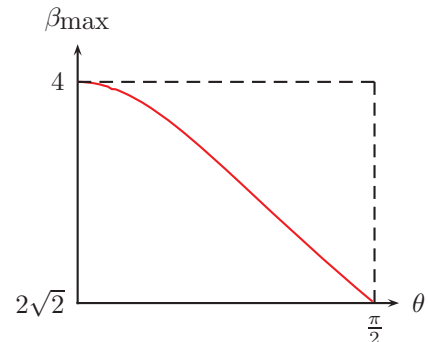

FIG. 3. (Color online) Maximum violation $\beta_{\max}$ of the Clauser-Horne-Shimony-Holt inequality by the nonlocal part P_{NL} of our EPR2 decomposition.

to machine precision of absolute order 10^{-10}) and therefore that our maximal EPR2 decomposition is valid for all possible measurement settings.

APPENDIX F: MAXIMAL VIOLATION OF THE CHSH INEQUALITY OF THE NONLOCAL PART

The correlator of the nonlocal part P_{NL} of our EPR2 decomposition is given by

$$E_{NL}(\boldsymbol{a},\boldsymbol{b}) := \langle AB \rangle_{NL|\boldsymbol{a},\boldsymbol{b}} = \frac{\langle AB \rangle_{Q|\boldsymbol{a},\boldsymbol{b}} - c \langle AB \rangle_{L|\boldsymbol{a},\boldsymbol{b}}}{1-c}.$$

We denote by $\beta_{\max}$ the maximum violation of the Clauser-Horne-Shimony-Holt inequality [45]

$$\beta_{\max} := \max_{\substack{\boldsymbol{a}_1,\boldsymbol{a}_2 \\ \boldsymbol{b}_1,\boldsymbol{b}_2}} |E_{NL}(\boldsymbol{a}_1,\boldsymbol{b}_1) + E_{NL}(\boldsymbol{a}_1,\boldsymbol{b}_2)$$
$$+ E_{NL}(\boldsymbol{a}_2,\boldsymbol{b}_1) - E_{NL}(\boldsymbol{a}_2,\boldsymbol{b}_2)|. \quad \text{(F1)}$$

In Fig. 3 the numerically determined values of $\beta_{\max}$ are plotted for all states $\theta \in [0,\pi/2]$. All states with $\theta < \frac{\pi}{2}$ violate Tsirelson's bound $2\sqrt{2}$ [46], approaching the algebraic maximum of 4 for $\theta \to 0$; this proves that the corresponding nonlocal parts P_{NL} are nonquantum.

[1] J. S. Bell, Physics **1**, 195 (1964), reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed. (Cambridge University Press, Cambridge, 2004), pp. 14–21.

[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[3] J. S. Bell, in *Einstein-Podolsky-Rosen Experiments*, Proceedings of the Symposium on Frontier Problems in High Energy Physics, Pisa, 1976 (Scuola Normale Superiore, Pisa, 1976), reprinted in *Speakable and Unspeakable in Quantum Mechanics* (Ref. [1]), pp. 52–62.

[4] J. Stachel, in *From Quarks to Quasars: Philosophical Problems of Modern Physics*, edited by R. G. Colodny (University of Pittsburgh Press, Pittsburgh, 1986), p. 349, reprinted in J. Stachel, *Einstein from 'B' to 'Z,'* Einstein Studies Vol. 9 (Birkhäuser, Boston, 2002), p. 367.

[5] J. S. Bell, CERN Report No. TH-2053-CERN, 1975 (unpublished), reprinted in *Speakable and Unspeakable in Quantum Mechanics* (Ref. [1]), pp. 52–62.

[6] A. Aspect, Nature (London) **398**, 189 (1999).

[7] L. Gilder, *The Age of Entanglement: When Quantum Physics Was Reborn* (Vintage, New York, 2009).

[8] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[9] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. **95**, 010503 (2005).

[10] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[11] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New J. Phys. **11**, 045021 (2009).

[12] S. Pironio, A. Acín, S. Massar, A. Boyer de La Giroday, D. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).

[13] R. Colbeck and A. Kent, J. Phys. A **44**, 095305 (2011).

[14] Y.-C. Liang, T. Vértesi, and N. Brunner, Phys. Rev. A **83**, 022108 (2011).

[15] T. Maudlin, *Proceedings of the Biennial Meeting of the Philosophy of Science Association* (University of Chicago Press, Chicago, 1992), p. 404.

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

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

[18] A. Yao, in *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing* (ACM, New York, 1979), pp. 209–213.

[19] J. Roland and M. Szegedy, *Automata, Languages and Programming*, Lecture Notes in Computer Science, edited by S. Albers, A. Marchetti-Spaccamela, Y. Matias, S. Nikoletseas, and W. Thomas (Springer, Berlin, Heidelberg, 2009), Vol. 5555, pp. 738–749.

[20] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).

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

[22] A. C. Elitzur, S. Popescu, and D. Rohrlich, Phys. Lett. A **162**, 25 (1992).

[23] V. Capasso, D. Fortunato, and F. Selleri, Int. J. Theor. Phys. **7**, 319 (1973).

[24] N. Gisin, Phys. Lett. A **154**, 201 (1991).

[25] J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. **97**, 170409 (2006).

[26] P. Pearle, Phys. Rev. D **2**, 1418 (1970).

[27] S. Braunstein and C. Caves, Ann. Phys. (NY) **202**, 22 (1990).

[28] It was also shown in Ref. [25] that this result is true for maximally entangled two-part states of any dimension.

[29] V. Scarani, Phys. Rev. A **77**, 042112 (2008).

[30] V. Scarani, arXiv:0712.2307v1 [quant-ph].

[31] C. Branciard, N. Gisin, and V. Scarani, Phys. Rev. A **81**, 022103 (2010).

[32] On a technical note, in order for the expectation values $\langle AB \rangle_{Q\boldsymbol{a},\boldsymbol{b}}$ to take the form of Eq. (4), the y directions of Alice and Bob's Bloch spheres should be defined with opposite directions (in other words, with standard definitions, one needs to change, for instance, throughout the paper, $\boldsymbol{b}$ to $\boldsymbol{b}'$, its reflection with respect to the xz plane; see, e.g., Ref. [31]).

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

[34] R. Werner and M. Wolf, Quantum Inf. Comput. **1**, 1 (2001).

[35] For technical reasons we assume from now on that $c < 1$; note that for $c = 1$, the state (2) is a product state and therefore is fully local: $p_{LC} = 1 = c$.

[36] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).

[37] D. Boschi, S. Branca, F. De Martini, and L. Hardy, Phys. Rev. Lett. **79**, 2755 (1997).

[38] Colbeck and Renner [R. Colbeck and R. Renner, Phys. Rev. Lett. **101**, 050403 (2008)] investigated models in which at least one party uses such a local nonshared random variable. The methods

used therein can be adapted to show that the nonlocal part of the following maximal decomposition cannot have any such local nonshared random variables.

[39] N. Brunner, N. Gisin, and V. Scarani, New J. Phys. **7**, 88 (2005).

[40] A. Méthot and V. Scarani, Quantum Inf. Comput. **7**, 157 (2007).

[41] M. Junge and C. Palazuelos, Commun. Math. Phys. **306**, 695 (2011).

[42] N. Brunner, N. Gisin, S. Popescu, and V. Scarani, Phys. Rev. A **78**, 052111 (2008).

[43] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[44] J. Barrett and S. Pironio, Phys. Rev. Lett. **95**, 140401 (2005).

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

[46] B. S. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).