Bell scenarios in which nonlocality and entanglement are inversely related

Giuseppe Vallone,1,2 Gustavo Lima,3 Esteban S. Gómez,3
Gustavo Cañas,3 Jan-Åke Larsson,4 Paolo Mataloni,2,5 and Adán Cabello6

1Department of Information Engineering, University of Padova, I-35131 Padova, Italy
2Dipartimento di Fisica della “Sapienza” Università di Roma, I-00185 Roma, Italy
3Center for Optics and Photonics, MSI-Nucleus on Advanced Optics, Departamento de Física, Universidad de Concepción, 160-C Concepción, Chile
4Institutionen för Systemteknik, Linköpings Universitet, SE-38183 Linköping, Sweden
5Istituto Nazionale di Ottica (INO-CNR), L.go E. Fermi 6, I-50125 Florence, Italy
6Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain

(Dated: October 24, 2018)

We show that for two-qubit chained Bell inequalities with an arbitrary number of measurement settings, nonlocality and entanglement are not only different properties but are inversely related. Specifically, we analytically prove that in absence of noise, robustness of nonlocality, defined as the maximum fraction of detection events that can be lost such that the remaining ones still do not admit a local model, and concurrence are inversely related for any chained Bell inequality with an arbitrary number of settings. The closer quantum states are to product states, the harder it is to reproduce quantum correlations with local models. We also show that, in presence of noise, nonlocality and entanglement are simultaneously maximized only when the noise level is equal to the maximum level tolerated by the inequality; in any other case, a more nonlocal state is always obtained by reducing the entanglement. In addition, we observed that robustness of nonlocality and concurrence are also inversely related for the Bell scenarios defined by the tight two-qubit three-setting $I_{3322}$ inequality, and the tight two-qutrit inequality $I_3$.

PACS numbers: 03.65.Ud,03.67.Bg,42.50.Xa

I. INTRODUCTION

Nonlocality and entanglement are two core concepts in quantum information. If $p_{\rho}(ab)$ is the joint probability that Alice obtains $a = 1$ and Bob $b = 1$ on a system prepared in state $\rho$, nonlocality is the impossibility of expressing $p_{\rho}(ab)$ as

$$
\sum_{\lambda} p_{\rho}(\lambda)p_{\rho}(a, \lambda)p_{\rho}(b, \lambda),
$$

where $\lambda$ are preestablished classical correlations. Entanglement is the impossibility of expressing a quantum state as a convex combination of separable states. Nonlocality and entanglement are related concepts in the sense that, to have nonlocality, entanglement is needed [2]. The difference between both concepts has been pointed out before. First, it was noticed that there are entangled states which do not violate specific Bell inequalities [3]. Then, in Ref. [4], the statistical strength of Bell tests was studied, showing that stronger tests (for a given family of Bell inequalities) require nonmaximally entangled states. Similarly, it was shown in [5] that nonmaximally entangled states allow for larger violations (or equivalently a stronger resistance to noise) of the $I_{3}$ two-qutrit inequality [6]. In [7] it was demonstrated that, for general bipartite Bell inequalities with $n$ inputs, $n$ outputs, and $n$-dimensional Hilbert spaces, the entropy of entanglement of the state is essentially irrelevant in obtaining large violation. Finally, in [8][9], it is shown that, for certain inequalities, weakly entangled states outperform maximally entangled ones of arbitrary dimension.

One difficulty in reaching a general conclusion about the relationship between nonlocality and entanglement is that of finding a general scenario where incontrovertible measures of nonlocality and entanglement can be compared. Bipartite scenarios have the advantage that any of the many measures of entanglement assign zero entanglement to product states and maximum entanglement to maximally entangled states [10][11]. Nonlocality is a more delicate issue since different restrictions on the number of measurement settings usually lead to different measures of nonlocality. This suggests that to study such relationship, one needs to consider a general scenario in which each party can perform an arbitrary number of local measurements.

The structure of the paper is the following: In Sec. II we define a measure of nonlocality called robustness of nonlocality that will be used through all the paper. In Sec. III we discuss a general bipartite scenario in which both parties have the same number of settings and prove that, no matter the number of settings, robustness of nonlocality and entanglement are inversely related. We then study how noise affects this conclusion. In Sec. IV we numerically explore the second simplest tight bipartite Bell inequality $I_{3322}$ [12], which has three settings per party, each with two outcomes. In Sec. V we study a tight two-qutrit Bell inequality $I_3$ [6]. In all cases considered we observe the same behavior, namely, that entanglement and robustness of nonlocality are inversely related.

II. ROBUSTNESS OF NONLOCALITY

For an ensemble of entangled particles in a state $|\psi\rangle$ and a given Bell inequality, we define the robustness of nonlocality (RN) against loss of local information as the maximum fraction of random particles per observer that can be lost such that the remaining ones can violate the Bell inequality. The robustness of nonlocality is related to the minimum detection efficiency, $\eta_{\text{crit}}$, required for a loophole-free violation of the Bell inequality [13] as RN = 1 $-$ $\eta_{\text{crit}}$. 

\begin{align}
\text{RN}(\eta_{\text{crit}}) & = \frac{1}{2} \left[ 1 + \frac{1}{\sqrt{2}} \sin(\frac{\pi}{4} - \frac{\theta}{2}) \right] \\
& = 1 - \eta_{\text{crit}}
\end{align}
The idea behind this measure of nonlocality is simple: A violation of a Bell inequality with perfect detection efficiency implies that no local model can reproduce the observed joint probabilities. If the minimum detection efficiency is \( \eta_{\text{crit}} \), this means that no local model exists, even if one locally rejects a fraction \( RN \) of the events. Therefore, the larger \( RN \), the harder it is to reproduce the observed results with local models. Therefore, \( RN \) may be taken as a measure of nonlocality.

In the previous expression, \( p \) in a similar way by using only the first \( d \) outcomes for non-detection gives a simple way to treat nondetections as simply “undefined” [16]. However, it can also be noted that from the experimental viewpoint, this is completely equivalent to relabeling the inputs or outputs of a Bell inequality and using the “−1” outcome in the case of a no-detection event for any observable. Thus, since any no-detection strategy is equivalent to rewrite the Bell inequality, the robustness of nonlocality \( RN \) can be evaluated by optimizing over all possible ways of rewriting the inequality and using the −1 outcome in the case of non-detection (in the case of observable with \( d \) outcomes, the last outcome is typically used in the case of non-detection). In order to violate a Bell inequality written as [5], in the case of detection efficiencies \( \eta_A \) and \( \eta_B \), the following relation must hold:

\[
\eta_A \eta_B \sum_{j=1}^{m_A} \sum_{k=1}^{m_A} c_{jk}p(a_jb_k) + \sum_{j=1}^{m_A} \alpha_j p(a_j) + \sum_{k=1}^{m_B} \beta_k p(b_k) > S_{\text{LHV}}.
\]

Furthermore, if one of the outcomes, different strategies can be used. For instance, a further outcome, corresponding to nondetections, can be added to the observables, [15]. or one can also choose to treat nondetections as simply “undefined” [16]. However, these strategies will require a modification of the Bell inequality. In the present paper we will study the robustness of nonlocality by assigning one of the observable outcomes to inconclusive events.

Each strategy giving a definite output to each observable is completely equivalent to relabeling the inputs or outputs of a Bell inequality and using the “−1” outcome in the case of no-detection for any observable. To give an example, the inequality [2] with Alice giving output +1 only for observable \( A_1 \) in the case of no detection is equivalent to replacing \( p(a_1b_k) \rightarrow p(b_k) - p(a_1b_k) \) and \( p(a_1) \rightarrow 1 - p(a_1) \) and using the −1 outcome in the case of a no-detection event for any observable.

It can also be noted that from the experimental viewpoint, assigning −1 outcomes for non-detection gives a simple way to handle these events. This is because with this assignment, no-detection events do not contribute to the inequality, so that there is no need to distinguish whether there was a pair produced but no detection, or if there was no pair produced. Distinguishing these are sometimes nontrivial, for example in a continuously pumped experiment, but this is not needed with the suggested assignment.

Any bipartite Bell inequality involving \( m_A \) and \( m_B \) dichotomic \((±1)\) observables \( A_j \) and \( B_k \) on Alice’s and Bob’s sides, respectively, can be written in the following form:

\[
\langle S \rangle_{\rho} \leq S_{\text{LHV}},
\]

where \( \langle S \rangle_{\rho} \) is the expectation value of \( S \) in the state \( \rho \) and

\[
S = \sum_{j=1}^{m_A} \sum_{k=1}^{m_A} c_{jk}p(a_jb_k) + \sum_{j=1}^{m_A} \alpha_j p(a_j) + \sum_{k=1}^{m_B} \beta_k p(b_k).
\]

In the present expression, \( p(a_jb_j) = p(A_j = 1, B_k = 1) \) are the joint probabilities of detecting the +1 eigenvectors \( |a_j \rangle \) and \( |b_k \rangle \) of the observables \( A_j \) and \( B_k \). If the observables have \( d_A \) and \( d_B \) outcomes, any Bell inequality can be expressed in a similar way by using only the first \( d_A - 1 \) and \( d_B - 1 \) outcomes.

Let us now evaluate the effect of detection inefficiency. For inequalities involving only +1 outcomes such as [12], it is customary to assume that no-detection events do not contribute to the inequality (they can be seen as detection on the “−1” outcome). However, in order to compute the robustness of nonlocality, it is necessary to optimize over all possible strategies for the no-detection events; for instance, whenever Alice does not get a detection, she can choose to always output +1 for observables \( A_1 \) and output −1 for all other observables [14]. It is worth noting that, instead of grouping inconclusive events with one of the outcomes, different strategies can be used. For instance, a further outcome, corresponding to nondetections, can be added to the observables, [15]. or one can also choose to treat nondetections as simply “undefined” [16]. However, these strategies will require a modification of the Bell inequality. In the present paper we will study the robustness of nonlocality by assigning one of the observable outcomes to inconclusive events.

III. ROBUSTNESS OF NONLOCALITY VS CONCURRENCE FOR CHAINED BELL INEQUALITIES

Pearle [13] and Braunstein and Caves (BC) [25, 26] introduced a generalization of the CHSH inequality [17] and Clauser-Horne (CH) [27] Bell inequalities, known as chained Bell inequalities, in which Alice and Bob choose among \( M \geq 2 \) settings. Chained Bell inequalities have some interesting applications: The case \( M = 3 \) fixes a loophole that occurs in some experiments based on the CHSH inequality [28]. Besides, it reduces the number of trials needed to rule out local hidden variable theories [29], and improves the security of some quantum key distribution protocols [30]. In the case in which \( M \) tends to infinity, the inequality allows one to discard nonlocal hidden variable theories with a nonzero local fraction [31]. Chained Bell inequalities have been experimentally tested using pairs of photons, with \( M = 3 \) [22], 4 [33], and 21 [34]. It was re-
The version of the chained Bell inequalities introduced in [33], which is symmetric under the permutation of Alice and Bob, reads (by using the notation of [2])

$$\langle S_M \rangle_\rho \leq 0,$$

where

$$S_M = p(a_M b_M) + \sum_{k=2}^{M} [p(a_k b_{k-1}) + p(a_{k-1} b_k)] - p(a_1 b_1) - \sum_{k=2}^{M} [p(a_k) + p(b_k)],$$

The minimum detection efficiency required for a loophole-free violation of chained Bell inequalities for any $M \geq 2$ using maximally entangled states has been obtained in [36]. The fact that the maximum quantum violation of chained Bell inequalities is always achieved with maximally entangled states [37] might suggest that the minimum detection efficiency occurs for maximally entangled states, but no proof exists of whether the detection efficiency for the chained Bell inequalities can indeed be reduced when one considers more general classes of entangled states. Indeed, for case $M = 2$, corresponding to the CH inequality (that is equivalent the CHSH), the minimum detection efficiency occurs for almost product states [18][21].

In the following we will show that, in absence of noise (e.g., considering pure states), the states with higher robustness of nonlocality (or the minimum detection efficiency) for any chained Bell inequality written in the form of [5] are almost product states for which the robustness of nonlocality tends to

$$\text{RN}_M = \frac{1}{2M - 1}. \quad (6)$$

The important point here is that this value is larger than the maximum value of $\text{RN}_M$ for maximally entangled states [36], namely,

$$\text{RN}_M^{\text{MES}} = \frac{M \cos \left( \frac{\pi}{2M} \right) - M + 1}{M \cos \left( \frac{\pi}{2M} \right) + M - 1}. \quad (7)$$

Moreover, for Bell inequalities of the form [4] with fixed $M$, we will show that the value in (6) is the maximum achievable robustness of nonlocality for any quantum state. This shows that, for all chained Bell inequalities, entanglement and nonlocality of pure states are inversely related.

**Theorem:** The maximum of the robustness of nonlocality of inequality [4] is $\text{RN}_M = \frac{1}{2M - 1}$ and can be obtained by almost product state.

**Proof:** Assuming the same detection efficiency for every party and setting, i.e., $\eta_A = \eta_B = \eta$, the value of $S_M$ becomes

$$\eta^2 \langle S_M \rangle_\rho - \eta (1 - \eta) \sum_{k=2}^{M} [p(\rho(a_k)) + p(\rho(b_k))], \quad (8)$$

where $p_\rho(a_k)$ is the expectation value of $p(a_k)$ in the state $\rho$. Therefore, inequality [4] is violated when $\eta > \eta^{(M)}_{\text{crit}}$, with

$$\eta^{(M)}_{\text{crit}} = \frac{\sum_{k=2}^{M} [p(\rho(a_k)) + p(\rho(b_k))]}{(\langle S_M \rangle_\rho + \sum_{k=2}^{M} [p(\rho(a_k)) + p(\rho(b_k))])}. \quad (9)$$

Since $p_\rho(a_1 b_1) \geq 0$, it is easy to show that

$$\langle S_M \rangle_\rho + \sum_{k=2}^{M} [p_\rho(a_k) + p_\rho(b_k)] \leq p_\rho(a_M b_M) + \sum_{k=2}^{M} [p_\rho(a_k b_{k-1}) + p_\rho(a_{k-1} b_k)].$$

Then,

$$\eta^{(M)}_{\text{crit}} \geq \frac{\sum_{k=2}^{M} [p_\rho(a_k) + p_\rho(b_k)]}{p_\rho(a_M b_M) + \sum_{k=2}^{M} [p_\rho(a_k b_{k-1}) + p_\rho(a_{k-1} b_k)]}. \quad (10)$$

Clearly,

$$0 \leq p_\rho(a_j b_k) \leq \min[p_\rho(a_j), p_\rho(b_k)], \quad (11)$$

and the lowest possible bound of the right-hand side of (10) is obtained when $p_\psi(a_j b_k) = p_\psi(a_j) = p_\psi(b_k)$ for $j$ and $k$ not both equal to 1. We obtain

$$\eta^{(M)}_{\text{crit}} \geq \frac{(2M - 2)p_\rho(a_1)}{(2M - 1)p_\rho(a_1)} = \frac{2M - 2}{2M - 1}, \quad (12)$$

which cannot be achieved exactly, but arbitrarily close by the following procedure: Any generic two-qubit pure states $\rho = |\psi\rangle \langle \psi|$, can be written (in a suitable basis) as

$$|\psi\rangle = \sin \frac{\gamma}{2} |00\rangle + \cos \frac{\gamma}{2} |11\rangle,$$

with $0 \leq \gamma \leq \pi/2$. Let us consider the following eigenstates:

$$|a_1\rangle = |b_1\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle, \quad (14a)$$
$$|a_k\rangle = |b_k\rangle = |0\rangle, \quad \text{with } k = 2, \ldots, M, \quad (14b)$$
and choose $\theta$ such that $\tan^2 \frac{\theta}{2} = \tan^2 \frac{\gamma}{2}$. Then, $p_\rho(a_1 b_1) = 0$ and the critical efficiency becomes

$$\eta^{(M)}_{\text{crit}} = \frac{2M - 2}{2M - 3 + \frac{2}{1 + \tan \gamma/2}}, \quad (15)$$

which, when $\gamma$ tends to zero (i.e., when the state tends to a product state), tends to

$$\eta^{(M)}_{\text{crit}} \xrightarrow{\gamma \to 0} \frac{2M - 2}{2M - 1} \Rightarrow \text{RN}_M \xrightarrow{\gamma \to 0} \frac{1}{2M - 1}, \quad (16)$$

concluding our proof.

We have numerically obtained, by using the method of conjugate gradient, $\text{RN}_M$ as a function $C$ of the pure state used to violate the inequality and compared it with the corresponding maximal achievable violation of the Bell inequality $S_M$. Moreover, through exhaustive numerical searches, we have obtained that the form [5] gives the maximum $\text{RN}$ for any given state (in the specific case of a maximally entangled state this is analytically demonstrated in the Appendix section). Note that, for nonmaximally entangled states such as
A. Adding noise

How does noise affect this conclusion? In the presence of white noise, the state becomes \( \rho = (1 - q)|\psi\rangle\langle\psi| + \frac{q}{2} \mathbb{1} \) and the threshold detection for the chained Bell inequalities efficiency is changed to

\[
\eta_{\text{crit}}^{(M)} = \frac{\sum_{k=2}^{M} [p_{\rho}(a_k) + p_{\rho}(b_k)] + \frac{q}{1-q} \left( M - 1 \right)}{(\mathbb{S}_M)_{\rho} + \sum_{k=2}^{M} [p_{\rho}(a_k) + p_{\rho}(b_k)]}.
\]  

In Fig. 2 we show, for three different values of noise \( q = 0.01 \), \( q = 0.05 \), and \( q = 0.1 \), the dependence of \( R_{N2} \) and \( R_{N3} \) and the maximum values of \( S_2 \) and \( S_3 \) with the degree of entanglement of the initial pure state. We observe that, when the noise is different from 0, the best quantum state giving the lowest threshold is not an almost separable state, but a nonmaximally entangled state depending on \( q \). However, the lower the noise \( q \), the smaller the entanglement required to obtain the optimal threshold.

Furthermore, in Fig. 2(b) we observe that, the lower \( M \) is, the more resistant to noise is the violation of the Bell inequality. In fact, it is possible to calculate the maximum tolerated noise to violate the chained Bell inequalities. Given \( \gamma \) and the maximal violation of \( S_M \) defined as \( S_M^{\text{max}}(\gamma) \), the maximum tolerated noise is \( q_{\text{max}} = \frac{2S_M^{\text{max}}(\gamma)}{2S_M^{\text{max}}(\gamma) + M - 1} \).

Using the method of conjugate gradient to minimize Eq. (17), it is also possible to obtain the threshold and the required entanglement for any value of the noise \( q \). The results are shown in Fig. 3. We observe that, for chained Bell inequalities, nonlocality and entanglement are simultaneously maximized 

\textit{only in the case of extreme noise}, namely the maximum noise level tolerated by the inequality. A better threshold detection efficiency is obtained by lowering the noise and suitably decreasing the entanglement. From this we conclude that nonlocality and entanglement are synonymous only for extremely noisy scenarios.

IV. ROBUSTNESS OF NONLOCALITY VS CONCURRENCE FOR \( I_{3322} \) BELL INEQUALITY

After the results presented in the previous section, a natural question is whether or not the same behavior occurs for other bipartite Bell inequalities. In this section we present the results for the second simplest tight bipartite Bell inequality, namely, \( I_{3322} \) [12, 38, 39], involving three dichotomic measurements on both \( A \) and \( B \) sides (the simplest tight bipartite Bell inequality is the CHSH inequality or \( S_2 \), studied in the previous section).

The \( I_{3322} \) inequality may be written as

\[
\langle I_{3322} \rangle_{\rho} \leq 0,
\]

where \( I_{3322} \) was defined in [12] as \( I_{3322} = p(a_1 b_1) + p(a_1 b_2) + p(a_1 b_3) + p(a_2 b_1) + p(a_2 b_2) + p(a_3 b_1) - p(a_2 b_3) - p(a_3 b_2) - 2p(a_1) - p(a_2) - p(b_1) \). However, this form will not lead to the best RN. We have numerically checked that the
Given a maximally entangled state $|\psi\rangle = \frac{1}{\sqrt{2}} (|a_1\rangle + |a_2\rangle)\frac{1}{\sqrt{2}} (|b_1\rangle + |b_2\rangle)$, to violate the inequality it is necessary that

\[ \eta^2 \langle I_{3322}\rangle_{\rho} + \eta (1 - \eta) [p_{\rho'}(a_2) - 1] + \eta (1 - \eta) [p_{\rho'}(b_1) - 1] - (1 - \eta)^2 > 0. \]  

(21)

Remembering that for a maximally entangled state (MES) $p_{\rho'}(a) = \frac{1}{2}$, we obtain

\[ \eta > 2(\sqrt{2} - 1) \approx 0.828 \implies RN^{MES} \approx 0.172. \]  

(22)

The maximal robustness of non-locality $RN = \frac{1}{4}$ can be achieved for almost product states. If we consider the $I_{3322}$ form of the inequality, the critical efficiency can be written as $\eta_c = \frac{p_{\rho}(a_1) + p_{\rho}(a_2) + p_{\rho}(b_1) + p_{\rho}(b_2)}{4}$. Let us choose $|a_1\rangle = |b_1\rangle = |a_3\rangle = |0\rangle$, $|a_2\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle$, $|b_2\rangle = |b_3\rangle = \sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle$, $|\psi\rangle = \sin \frac{\theta}{2} |00\rangle + \cos \frac{\theta}{2} |11\rangle$. By using $\theta$
such that \( \cos \theta = \frac{1 + \cos \gamma}{2(1 + \sin \gamma)} \) we obtain
\[
\eta_{\text{crit}} = \frac{12(1 + \sin \gamma)}{13 + 3 \cos \gamma + 12 \sin \gamma},
\]
which, when \( \gamma \) tends to zero tends to
\[
\eta_{\text{crit}} \xrightarrow{\gamma \to 0} \frac{3}{4} \quad \Rightarrow \quad \text{RN} \xrightarrow{\gamma \to 0} \frac{1}{4}.
\]
We observe that the maximum RN for the \( I_{3322} \) is greater than the one for the \( S_3 \) inequality, which has the same number of local settings.

V. ROBUSTNESS OF NONLOCALITY VS CONCURRENCE FOR THE \( I_3 \), TWO-QUTRIT INEQUALITY

For the two-qubit Bell inequalities discussed above we have observed that nonlocality and entanglement are inversely related. Here we show that this is also true for other bipartite scenarios. For this purpose we repeat our analysis but now for a tight bipartite inequality maximally violated by two-qutrit states, the \( I_3 \) inequality [6].

The inequality is given by \( I_3 = P(A_1 = B_1) + P(B_1 = A_2 + 1) + P(A_2 = B_2) + P(B_2 = A_1) - P(A_1 = B_3 - 1) - P(B_1 = A_2) - P(A_2 = B_2 - 1) - P(B_2 = A_1 - 1) \leq 2 \), where \( P(A_m = B_n + k) = \sum_{j=1}^{3} P(a_j^m b_j^n + k \mod 3) \). Here, \( n \) and \( m (n, m = 1, 2) \) denote the settings that the parties may choose for the local measurements, and the index \( j \) denotes each measurement outcome \( j = 1, 2, 3 \). The inequality can be rewritten in the form of [2] as \( \langle I_3 \rangle_\rho \leq 0 \), with

\[
I_3 = p(a_1 b_1) + p(a_2 b_2) + p(a_2 b_2) - p(a_2 b_2) - p(a_2 b_2) - p(a_1) - p(a_1) - p(b_1).
\]

and the following measurements \( |a_1 \rangle = |b_1 \rangle = |3 \rangle, |a_2 \rangle = |b_1 \rangle = |2 \rangle, |a_2 \rangle = \cos \frac{\theta}{2} |1 \rangle + \sin \frac{\theta}{2} |3 \rangle, |b_2 \rangle = \sin \frac{\theta}{2} |1 \rangle + \cos \frac{\theta}{2} |3 \rangle, |a_2 \rangle = |b_2 \rangle = |2 \rangle \). In the form [25] the threshold efficiency becomes (non-detection events correspond to 3-eigenvalues and thus does not contribute to the inequality):

\[
\eta_c = \frac{p_r(a_1) + p_r(a_1) + p_r(b_1) + p_r(b_1)}{\langle I_3 \rangle_\rho + p_r(a_1) + p_r(a_1) + p_r(b_1) + p_r(b_1)}.
\]

With the above measurement the critical efficiency becomes

\[
\eta_c = \frac{4(1 - \cos \gamma)}{3 - \cos \theta(2 + \cos \theta) - 4 \cos \gamma + \sin^2 \theta \sin \gamma}. \quad (28)
\]

If we choose \( \cos \theta = -\frac{1}{\sin \gamma + 1} \) and let \( \gamma \) go to zero we get
\[
\eta_c \xrightarrow{\gamma \to 0} \frac{2}{3} \Rightarrow \quad \text{RN} \xrightarrow{\gamma \to 0} \frac{1}{3}. \quad (29)
\]

For maximally entangled states, for which \( p(a_j) = p(b_j) = p(a_1) = p(b_1) = \frac{1}{2} \), the RN is given by
\[
\text{RN}^{\text{MES}} = \frac{I_3^{\text{MES}}}{I_3} + 4/3 = \frac{4\sqrt{3} - 3}{4\sqrt{3} + 15} \simeq 0.1791. \quad (30)
\]

VI. CONCLUSIONS

We would argue that robustness of nonlocality RN is a good measure of nonlocality, since it marks the border where local
hidden variable descriptions become possible: The larger robustness of nonlocality is, the harder it is to express the joint probabilities with local models.

We have shown that, for the two-party $M$-setting chained Bell scenario (for any $M \geq 2$ finite), for a tight two-qubit Bell inequality $I_{3322}$ and a tight two-qutrit Bell inequality $I_3$, robustness of nonlocality and concurrence, are in the absence of noise, inversely related.

The main result of this paper is the observation that, for many distinct types of Bell scenarios, larger nonlocality requires smaller entanglement; in the absence of noise, almost product state are the most nonlocal ones. We analytically showed that the maximal RN can be achieved with almost product state. The maximal values of RN (related to the minimum required detection efficiency as $\eta_c = 1 - \text{RN}$) are given by $\text{RN} = \frac{1}{2M-1}$, $\text{RN} = \frac{1}{4}$ and $\text{RN} = \frac{1}{4}$ for the $S_M$ chained Bell inequality, the $I_{3322}$ inequality and the $I_3$ inequality respectively.

When noise is present, the most nonlocal states acquire some amount of entanglement; however, the smaller the noise is, the lower their entanglement becomes.

Some questions naturally arise: are the nonlocality and entanglement inversely related in any Bell inequality involving $m_A$, $m_B$ observables with $d_A$ and $d_B$ outcomes? If yes, is there some physical mechanism for such counterintuitive behavior? These questions require further research.

Acknowledgments

GL, ESG and GC were supported by the CONICYT, AGCI, FONDECYT 1120067, MilenioP10-030-F and PIA-CONICYT PFB0824. GV was supported by the Strategic-Research-Project QUINTET of the Department of Information Engineering, University of Padova and the Strategic-Research-Project QUANTUMFUTURE of the University of Padova. PM acknowledge the Chistera EU project QUASAR. AC was supported by Project No. FIS2011-29400 (MINECO, Spain).

Appendix A: Optimality of detection strategy for maximally entangled states

In this section we will demonstrate which is the optimal way of rewriting the Bell inequalities analyzed in the main text in case of maximally entangled states. We start by giving the general framework to solve the optimization.

Let us consider a general bipartite Bell inequality involving $m_A$ and $m_B$ observables $A_j$ and $B_k$ on the Alice and Bob side. The observables have $d_A$ and $d_B$ outcomes respectively, $\mu = 1, 2, \cdots, d_A$ and $\nu = 1, 2, \cdots, d_B$. Any Bell inequality can be written as

$$\langle S \rangle_p \leq S_{\text{LHV}}, \quad (A1)$$

with

$$S = \sum_{\mu=1}^{m_A} \sum_{k=1}^{d_A} \sum_{\mu=1}^{m_B} \sum_{k=1}^{d_B} c_{jk}^{\mu\nu} p(a_1^\mu b_1^\nu) + \sum_{\mu=1}^{m_A} \sum_{k=1}^{d_A} a_1^\mu p(a_1^\mu) + \sum_{\mu=1}^{m_B} \sum_{k=1}^{d_B} b_1^\nu p(b_1^\nu). \quad (A2)$$

In the previous expression $p(a_j^\mu b_k^\nu) = p(A_j = \mu, B_k = \nu)$ are the joint probabilities of detecting the $\mu$ and $\nu$ eigenvectors $|a_j^\mu\rangle$ and $|b_k^\nu\rangle$ of the observables $A_j$ and $B_k$. Note that only the first $d_A - 1$ and $d_B - 1$ outcomes are involved in the inequality.

When inefficiencies are present it is necessary to give a strategy for the non-detection events. Let us suppose that the strategy on Alice’s side is the following. If Alice is measuring the observable $A_j$ and the particle is not detected, she assigned, with probability $A_j^{(\mu)}$, the outcome $\mu$. Clearly, $\sum_{\mu=1}^{d_A} A_j^{(\mu)} = 1$. The same happens at Bob’s side, with probabilities $B_k^{(\nu)}$. If we consider Alice and Bob inefficiencies as $\eta_A$ and $\eta_B$, the Bell inequality is violated if

$$\eta_A \eta_B \langle S \rangle_p + (1 - \eta_A) \eta_B T_A + \eta_A (1 - \eta_B) T_B + (1 - \eta_A)(1 - \eta_B) X_{AB} > S_{\text{LHV}}, \quad (A3)$$
with
\[ T_A = \sum_{j,k,\mu,\nu} c^{\mu}_{jk} A_j^{(\mu)} p_\rho(b_k^{\nu}) + \sum_{j,\mu} c_j^{(\mu)} A_j^{(\mu)} + \sum_{k,\nu} \beta_k^{(\nu)} p_\rho(b_k^{\nu}) \]
\[ T_B = \sum_{j,k,\mu,\nu} c^{\mu}_{jk} p_\rho(a_j^{(\mu)}) B_k^{(\nu)} + \sum_{j,\mu} \alpha_j^{(\mu)} p_\rho(a_j^{(\mu)}) + \sum_{k,\nu} \beta_k^{(\nu)} B_k^{(\nu)} \]
\[ X_{AB} = \sum_{j,k,\mu,\nu} c^{\mu}_{jk} A_j^{(\mu)} B_k^{(\nu)} + \sum_{j,\mu} \alpha_j^{(\mu)} A_j^{(\mu)} + \sum_{k,\nu} \beta_k^{(\nu)} B_k^{(\nu)}. \] (A4)

The sum is taken over \( i = 1, \ldots, m_A \) and \( j = 1, \ldots, m_B \) while \( \mu = 1, \ldots, d_A - 1 \) and \( \nu = 1, \ldots, d_B - 1 \): also in the previous expression the outcomes \( d_A \) and \( d_B \) of each observable are not present.

We start with the chained Bell inequalities, and then analyze the \( I_{3322} \) and \( I_3 \) inequalities.

1. Chained Bell inequalities

For the chained Bell inequalities of section III we have dichotomous observables. Then, in the case of non detection on the observable \( A_j \), Alice chooses to output the +1 outcome with probability \( A_j \) and the −1 outcome with probability \( 1 - A_j \). The same happens to Bob. Remembering that, for MES, \( p_\rho(a_j) = p_\rho(b_k) = \frac{1}{2} \) we have:
\[ T_A = T_B = \frac{M - 1}{2} \]
\[ X_{AB} = A_M B_M + \sum_{k=2}^{M} (A_k B_{k-1} + A_{k-1} B_k) + \frac{1}{4} \eta^2 - \frac{1}{2} \eta (1 - \eta) - (1 - \eta)^2 > 0, \] (A8)

solved by
\[ \eta > 2(\sqrt{2} - 1) \simeq 0.828. \] (A9)

The choice of the \( A \)'s and \( B \)'s corresponds to choosing for the non-detection events the outcome −1 for the inequality written as \( I_{3322}^{(2)} \).

2. \( I_{3322} \) inequality

Let us consider the \( I_{3322} \) inequality written in its original form
\[ I_{3322} = p(a_1 b_1) + p(a_1 b_2) + p(a_1 b_3) + p(a_2 b_1) + p(b_2 b_1) + p(b_3 b_1) - p(a_2 b_3) - p(a_3 b_2) - 2p(a_1) - p(a_2) - p(b_1). \]

In the case of inefficiencies with non maximally entangled states we have
\[ T_A = -\frac{1}{2}(A_1 + A_2 + 1) \]
\[ T_B = \frac{1}{2}(B_1 + B_2 - 3) \]
\[ X_{AB} = A_3(B_1 - B_2) + B_3(A_1 - A_2) + A_1 B_1 + A_1 B_2 + A_2 B_1 + A_2 B_2 - 2A_1 - A_2 - B_1. \] (A6)

Since the maximal value of \( \langle I_{3322} \rangle \) with maximally entangled state is 1/4, the Bell parameter in the case of detection inefficiencies \( \eta_A = \eta_B = \eta \) becomes
\[ \frac{1}{4} \eta^2 + \frac{1}{2} \eta (1 - \eta) (B_1 + B_2 - A_1 - A_2 - 4) + (1 - \eta)^2 X_{AB}. \] (A7)

The choice that minimizes the critical efficiency is given by \( B_1 = B_2 = 1 \) and \( B_3 = A_1 = A_2 = A_3 = 0 \), giving \( X_{AB} = -1 \), \( T_A = T_B = -\frac{1}{2} \) and
\[ \frac{1}{4} \eta^2 - \eta (1 - \eta) - (1 - \eta)^2 > 0, \] (A8)

3. Two-qutrit \( I_3 \) inequality

For this two-qutrit inequality Alice has three outcomes for each observable \( A_j \). In the case of non-detection she assigns with probability \( A_j^{(1)} \) the outcome 1, with probability \( A_j^{(2)} \) the outcome 2, and with probability \( 1 - A_j^{(1)} - A_j^{(2)} \) the outcome 3. The same applies to Bob. For maximally entangled states \( p(a_j) = p(b_j) = p(\pi_j) = p(\bar{b}_j) = \frac{1}{2} \), and we have
\[ T_A = T_B = -\frac{2}{3} \]
\[ X_{AB} = A_1^{(1)} B_1^{(1)} + A_1^{(1)} B_2^{(1)} + A_2^{(1)} B_1^{(1)} - A_2^{(1)} B_2^{(1)} + A_1^{(2)} B_1^{(2)} + A_1^{(2)} B_2^{(2)} + A_2^{(2)} B_1^{(2)} - A_2^{(2)} B_2^{(2)} + A_1^{(3)} B_1^{(3)} + A_1^{(3)} B_2^{(3)} + A_2^{(3)} B_1^{(3)} - A_2^{(3)} B_2^{(3)} - A_1^{(4)} - A_2^{(4)} - B_1^{(4)} - B_2^{(4)}. \] (A10)

The optimal choice of \( A_k^{(\mu)} \)'s and \( B_k^{(\mu)} \)'s is the one that maximizes \( X_{AB} \). This term is clearly upper bounded by 0 (it corresponds to the Bell inequality). Then the choice \( A_j^{(\mu)} = B_k^{(\nu)} = 0 \) saturates the bound. This choice corresponds to choosing for the non-detection events the outcome 2 for the inequality written as [28].
[1] J. S. Bell, Physics 1, 195 (1964).
[2] N. Gisin, Phys. Lett. A 154, 201 (1991).
[3] R. F. Werner and M. M. Wolf, Phys. Rev. A 61, 062102 (2000).
[4] A. Acín, R. Gill, and N. Gisin, Phys. Rev. Lett. 95, 210402 (2005).
[5] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Phys. Rev. A 65, 052325 (2002).
[6] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. 88, 040404 (2002).
[7] M. Junge and C. Palazuelos, Comm. Math. Phys. 306, 695 (2011).
[8] T. Vidick and S. Wehner, Phys. Rev. A 83, 052310 (2011).
[9] Y.-C. Liang, T. Vértesi and N Brunner, Phys. Rev. A 83, 022108 (2011).
[10] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998).
[11] P. Rungta, V. Bužek, C. M. Caves, M. Hillery, and G. J. Milburn, Phys. Rev. A 64, 042315 (2001).
[12] D. Collins and N. Gisin, J. Phys. A: Math. Gen. 37, 1775 (2004).
[13] P. M. Pearle, Phys. Rev. D 2, 1418 (1970).
[14] N. Brunner and N. Gisin, Phys. Lett. A 372, 3162 (2008).
[15] S. Massar, S. Pironio, J. Roland, and B. Gisin, Phys. Rev. A 66, 052112 (2002).
[16] J.-Å. Larsson, Phys. Rev. A 57, 3304 (1998).
[17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).
[18] P. H. Eberhard, Phys. Rev. A 47, R747 (1993).
[19] A. Cabello and J.-Å. Larsson, Phys. Rev. Lett. 98, 220402 (2007).
[20] N. Brunner, N. Gisin, V. Scarani, and C. Simon, Phys. Rev. Lett. 98, 220403 (2007).
[21] J.-Å. Larsson and J. Semitecolos, Phys. Rev. A 63, 022117 (2001).
[22] T. Vértesi, S. Pironio, and N. Brunner, Phys. Rev. Lett. 104, 060401 (2010).
[23] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. 98, 140402 (2007).
[24] G. Vallone, Phys. Rev. A 87, 020101 (2013).
[25] S. L. Braunstein and C. M. Caves, in Bell’s Theorem, Quantum Theory, and Conceptions of the Universe, edited by M. Kafatos (Kluwer, Dordrecht, 1989), p. 27.
[26] S. L. Braunstein and C. M. Caves, Ann. Phys. (N.Y.) 202, 22 (1990).
[27] J. F. Clauser and M. A. Horne, Phys. Rev. D 10, 526 (1974).
[28] S. Aerts, P. Kwiat, J.-Å. Larsson, and M. Żukowski, Phys. Rev. Lett. 83, 2872 (1999).
[29] A. Peres, Fortschr. Phys. 48, 531 (2000).
[30] J. Barrett, L. Hardy, and A. Kent, Phys. Rev. Lett. 95, 010503 (2005).
[31] J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. 97, 170409 (2006).
[32] G. Vallone, I. Gianani, E. B. Inostroza, C. Saavedra, G. Lima, A. Cabello, and P. Mataloni, Phys. Rev. A 83, 042105 (2011).
[33] D. Boschi, S. Branca, F. De Martini, and L. Hardy, Phys. Rev. Lett. 79, 2755 (1997).
[34] M. Barbieri, F. De Martini, G. Di Nepi, and P. Mataloni, Phys. Lett. A 334, 23 (2005).
[35] C. Dhara, G. Prettico, and A. Acín, Phys. Rev. A 88, 052116 (2013).
[36] A. Cabello, J.-Á. Larsson, and D. Rodríguez, Phys. Rev. A 79, 062109 (2009).
[37] S. Wehner, Phys. Rev. A 73, 022110 (2006).
[38] M. Froissart, Nuovo Cimento B 64, 241 (1981).
[39] C. Śliwa, Phys. Lett. A 317, 165 (2003).